



# First Things First

*Laying the foundation  
for a Smart City*

**GovEx**



## *First Things First*

### *Laying the Foundation for a Smart City*

[Introduction](#)

[Imagining a Smart City](#)

[Baltimore's Integrated Approach](#)

[Start with the people](#)

[Understand the local context](#)

[Strengthen Internal Data Management](#)

[Data governance](#)

[Open data policy](#)

[Ensure Security and Accountability of Public Data](#)

[Working with vendors](#)

[Assessing sensitivity](#)

[Considering cybersecurity](#)

[Opening and sharing data](#)

[Develop portal requirements](#)

[Determining dataset publication priority](#)

[Ensure accessibility of data](#)

[Communicating with residents](#)

[Conclusion](#)

## Special thanks to:

**Kate Bender**, Senior Management Analyst, Kansas City

**Ed Blayney**, Innovation Project Manager, Louisville

**Brandi Braun**, Deputy Innovation Office, Columbus

**Seema Iyer**, Associate Director and Research Assistant Professor, Baltimore  
Neighborhood Indicators Alliance

**Shonte Eldridge**, Deputy Chief of Operations, Office of Mayor Catherine Pugh,  
Baltimore City

**Kevin Martin**, Smart City pdx/tech Services Manager, City of Portland, Bureau of  
Planning & Sustainability

**Gerrit Knaap**, Professor of Urban Studies and Planning and Executive Director of the  
National Center for Smart Growth Research and Education at the University of  
Maryland

**Eric Roche**, Chief Data Officer, Kansas City

**Michael Schnuerle**, Data Officer, Louisville

**Julie Steenson**, Deputy Performance Officer, Kansas City

# Introduction

This guide is the product of a [National Science Foundation \(NSF\)](#) funded research collaborative composed of [University of Maryland](#), [Morgan State University](#), the [Center for Government Excellence \(GovEx\)](#) at Johns Hopkins University, and [University of Baltimore](#). Since October 2017, these research institutions have worked together to determine how data and technology can meet the needs of residents in West Baltimore communities. This guide captures the experience of these partners in Baltimore and other cities around the country in laying the foundation for a Smart Cities project that meets the needs of residents and improves city service delivery. In addition to participating in the Baltimore project, GovEx interviewed key stakeholders from Portland, OR; Columbus, OH; and Louisville, KY, to gather their insight and expertise into key aspects of Smart Cities planning and implementation. The guide is divided into sections based on the necessary steps to establish the strong data management and use practice that are necessary for any Smart City.

## Imagining a Smart City

For the purposes of this guide, think of Smart Cities as cities that respond to the needs of residents by using data and technology to transform government services. As this field evolves, GovEx will update this guide to reflect the current definitions, examples, and practices. Although discussions of Smart Cities often focus on technology, it's also important to think before the project starts about how residents lives will be improved. For instance, imagine a city that leverages data and technology to improve infant mortality outcomes. That's Mayor Ginther's vision for Columbus, OH, which won of \$40 million in grant funding from the federal [Department of Transportation \(USDOT\)](#) to improve mobility for residents. [Reducing infant mortality](#) by 40 percent by 2020 is an urgent priority for the mayor, as county-wide infant mortality among black infants (16 per 1,000 live births) is twice as high as white infants (9 per 1,000 live births). Mayor Ginther views mobility as the "great equalizer," giving citizens power to improve their lives through access to transportation. In Columbus, the Mayor believes equitable access to transportation is an effective way to address infant mortality by providing reliable and flexible trips for low-income mothers.

City staff worked with academic and nonprofit partners through techniques like surveys and community conversations to consult on the changes they implemented. Through this study, staff learned from expectant mothers that the current system of providing transportation is cumbersome and unreliable. This lack of reliable transportation in Columbus is a [health factor](#) that can contribute to stress and missed appointments, which affects health outcomes like infant mortality. Based on this research, the city included prenatal trip assistance for expectant mothers who are eligible for Medicaid as a key project for [Smart Columbus](#). This new funding will support initiatives such as the Columbus-Franklin County [Celebrate One campaign](#) and the

[Greater Columbus Infant Mortality Task Force](#) led by Mayor Ginther, which have identified transportation as a social determinant of health that affects infant mortality.

Columbus’s plan illustrates how a city can harness data and technology in the transportation sector to improve quality of life. And the City is not alone in its desire to capitalize on the Smart Cities movement. According to a [2016 U.S. Conference of Mayors survey](#), cities of all sizes from around the country are planning or implementing Smart Cities initiatives including energy, physical infrastructure, governance, security, and healthcare solutions.

Population	# Cities	# Implemented Projects	# Planned Projects
>1M	6	69	103
150K to 1M	25	168	225
<150K	23	98	131

The rest of this guide will explore Baltimore’s people first approach to smart cities, followed by examples and recommendations to ensure that data is managed as an asset internally and kept secure, and that stewards of the data are accountable to the public.

## Baltimore’s Integrated Approach

The NSF proposal to work in Baltimore City focused on West Baltimore, which is a low-income section of the city where lack of access to transportation as well as education, jobs, and safe and healthy housing lead to poor health outcomes. For instance, in the West Baltimore Study Site neighborhoods of [Druid Heights/Upton](#), 63.7 percent of children live below the poverty line, and the median household income is \$19,038. Nearly 60 percent of households pay more than one third of their total household income on rent. A higher percentage of residents in these neighborhoods use public transportation to go to work (32.5) compared with the citywide average (18.4). Life expectancy is five years lower than the Citywide average (68 years compared with 73.2).

The purpose of the research grant is to determine whether and how Smart Cities efforts can improve equity in these neighborhoods. Equity means equal access to the same opportunities in education, employment, transportation, and housing as other neighborhoods in Baltimore. While the following sections of this guide lay out the practices that are necessary to ensure strong data management and use, this section reviews a key first step from the Baltimore experience -- start with the people.

### Start with the people

The first step in any successful Smart Cities project is to ask residents about their problems. The Baltimore team began its Smart Cities project by meeting with city residents to identify

those needs and using the results to jointly design appropriate smart city solutions. [Faculty](#) and students from the University of Maryland conducted community conversations and surveys with residents of West Baltimore to understand the challenges and priorities of the communities, how residents use technology, and the vision of residents for improving the neighborhood. Through community conversations in locations that represent different demographics of the neighborhoods, and 116 written surveys, the Engagement team learned that job access and neighborhood safety and security are clear neighborhood priorities.

In addition, the team identified four main findings on the needs and wants of residents regarding data and technology:

- Residents in West Baltimore lack access to computers and do many computer-based activities on phones, such as searching for jobs and housing, and staying connected with friends.
- To address access to jobs and training, residents would like to have more ownership and control of technological resources that are provided in their neighborhood. For instance, if fiber optic cable is installed, there may be an opportunities to train and employ West Baltimore residents to work on these sites.
- In community conversations, residents repeatedly mentioned that they want and need access to data that be collected through Smart Cities sensors, in order to improve neighborhood conditions. For instance, sharing crime data collected on city cameras could facilitate greater relationships between police and residents, and potentially lead to enhanced community safety.
- Residents are diverse. Technology should address the limitations and opportunities for specific vulnerable groups, while meeting common needs across groups.

## Understand the local context

When getting started with a Smart Cities project, it's important to take inventory of existing and planned Smart Cities initiatives in your city. During the course of the research grant in Baltimore, in March 2018, Baltimore City released its [Inclusive Digital Transformation Strategic Plan](#), which includes plans for a Smart City Council and a Data Analytics Hub. The Data Analytics Hub will house the City's operational data as well as data collected through Smart Cities initiatives such as Citiwatch crime cameras. According to the City's Deputy of Operations, "If we don't have data collection and storage infrastructure, and analytics, we don't have a Smart Cities strategy." As noted by Seema Iyer, Associate Director of the Jacob France Institute at University of Baltimore, "New technology should be adding to a synthesis of knowledge of how a city already works." This flips the traditional Smart Cities approach of solutions looking for problems on its head.

# Strengthen Internal Data Management

Before getting started on any initiative that involves collecting new data, like a Smart Cities initiative, it's necessary to strengthen internal data management practices. The data that you will collect through a Smart Cities project must be treated as a strategic asset, just like any other data. The steps in this section will provide guidance on implementing foundational data management and use practices in the Smart Cities context. Although not exhaustive, links to resources are included for additional reading. This section will help you follow the necessary steps, which are:

- Create a data governance team and plan
- Inventory your existing data assets
- Develop a meaningful open data policy
- Understand the importance of investing in people

## Data governance

*In order to explore how data governance supports Smart Cities projects, consider the City of Louisville, KY, and its [IoT Smoke Detectors project](#), in which low-cost wireless smoke detectors were installed in vacant and abandoned properties in order to alert firefighters of a fire before flames spread to surrounding homes. To develop this project, the City first identified a problem that the community already faced -- a higher prevalence of fires in impoverished neighborhoods. Next, the City leveraged its relationship with the Fire Department and Department of Community Services to create a cross-functional team that also included the City's Office of Performance Improvement and Innovation. This team used data and analytics to create a plan to provide fire prevention education and smoke detectors to at-risk residents. Regarding residents as experts, starting with the community's needs, and leveraging existing relationships leads to the best and more impactful projects. Now, the city's [Smart Louisville](#) initiative houses projects online in order to increase transparency with its community members.*

## Convening governance committees

In Louisville, the data governance group lays the foundation for all current and future Smart Cities work. The governance group is building up skill sets to address the data and technology that will be part of future Smart Cities efforts, and to ensure that data and analysis skills and thinking are embedded into the culture of city government. According to a conversation with Ed Blayney and Michael Schnuerle from the City of Louisville, "There will be so much more data and tech in the future, if Smart Cities is really about using more data and tech in government. It won't even be 'Smart Cities' at that point, it will just be what's expected; data-driven government service delivery."

Data governance is critical to treat data as a strategic asset and a [data governance committee](#) is necessary for these practices. Key topics for a data governance committee may include data management and access, organizational strategy, and problem solving. Convening a governance body composed of representatives from deep within the organization with clear roles and responsibilities is an important first step.

*The City of Louisville worked with staff to establish a data governance structure, team, and policy with an underlying goal of shifting its culture around data usage. In Louisville, there are about 50 people from almost every department and division that work with data that make up the Data Governance team. The projects, working groups, and training around data are in place to build culture of innovation, analysis, and data sharing. The data governance team's projects started initially as a way of getting data to open data portal, but over time, its scope expanded. Now, the team's goal is to take action and write policy on key components of data management like establishing data standards, conducting new inventories, and publishing findings on its open data portal. In Louisville, the data governance committee is laying the foundation for all current and future Smart Cities work by equipping city employees with the skills and tools needed to be able to be more involved.*

Before taking action on developing a data governance committee, first get a holistic understanding of the major data governance milestones and determine which structure is best suited for your department. The GovEx resources linked here can help you and your core team think through key [data governance milestones](#) and the [data governance structure](#) that is best for your organization. Although these resources are more comprehensive, below are some key reminders.

### **Support open data leadership team in creating a collaborative environment that serves the City as a whole**

*With the Mayor's executive order driving them, the City of Louisville's department directors were asked to appoint two people from each department to be a part of the governance committee, which was initially tasked with creating an open data portal for the city. The team wanted to ensure that it was not working in silos, but rather with the intention of building relationships with the veteran city employees in order to change the culture around data and make progress in becoming a Smart City. The governance team knew that building relationships would be the most important factor in achieving those goals. The City also continuously reevaluated the makeup of its governance team. In order to both encourage employees to learn data skills and to identify the employees best fit for a role on the governance team, the City created a [badge incentive program](#) through which city workers could earn achievement badges in recognition of reaching data-related milestones.*

*During Louisville's monthly data governance team meetings, a volunteer from the committee presents or leads a session. Other teams in the City, like within the health department, are following suit. Many of the data governance committee practices, such as leading presentations*

*and the badge program, are becoming institutionalized throughout other departments, and they work interdepartmentally to rely on each other's goals. Because they're not working in silos and instead are collaborating in order to change the current culture and make progress on goals, Louisville is simultaneously developing the supports needed for the open data leadership team.*

Below are key questions to consider when developing a data governance team:

- Who is involved in the data governance team? List team members, including names and position titles.
- What are the responsibilities of these team members?
- What if members aren't able to attend? Who will serve as a secondary representative?
- Will members of the community be allowed to serve on the organization's data governance team? If so, how will they be involved in decision making?
- Who must sign-off on decisions made by the committee?

### **Hold regular data governance meetings**

*Throughout its first two years of monthly meetings, Louisville's committee expanded its original scope and took action to write data standards and policies, execute new data inventories, and finally publish the data on the City's open data portal.*

Your team should specify how many times it plans to meet each year. GovEx recommends the data governance team meets at least once a quarter. The frequency of the meetings are dependent on the types of projects adopted by your data governance team. For example, if the team is launching a city's open data program, it may make sense to meet once a month, or even once a week. A schedule for data governance meetings should be established at the beginning of the calendar year and distributed to the team.

### **Work with your governance committee to create a timeline, plan, and data collection instrument for your city's data inventory**

A data inventory is a fully described record of the data assets maintained by your city. The inventory records basic information about a data asset including its name, contents, update frequency, use license, owner/maintainer, privacy considerations, data source, and other relevant details. The details about a dataset are known as metadata. Because your city may have thousands of datasets across multiple servers, databases, and computers, it's helpful to narrow down which datasets should be included in the inventory overall and how to plan for inventory updates in the future. The datasets worth inventorying are those which are considered assets to employees, departments, executive leadership, and the general public. Data assets can range from individual datasets that are connected to forms that people fill out, to integrated databases that track a city's operations in any given field (building permits, public safety responses, etc.). Remember to identify and manage risks related to new data collection efforts. Most importantly, as your team begins to look to implement a new data integration (bringing together data from several disparate, heterogeneous data sources to provide a unified view of the data) product, include someone who understands the architectures of differing platforms. It

is crucial that the team also builds relationships with database administrators across the city. As your team begins or continues the process of completing data inventories, it may be helpful to reference the [GovEx data inventory guide](#).

To cities that are just getting started on creating a data governance team, the City of Louisville recommends:

- **Bringing together a core group of people who are leading the effort, who are motivated to get things started.** If you have a small 2-4 person team, then you can start building events, trainings, and programs. A motivated group of people would be best to sell it to the internal directors, chiefs, and employees.
- **Building the community and the network.** Work interdepartmentally in order to gain momentum behind necessary culture changes and increase buy-in.
- **Having a stable, consistent team.** Make sure your team has clarity of focus; don't chase all the shiny things that come up as ideas or possibilities. Instead, have focused goals and team members who are consistently working towards them.
- **Being tenacious.** The most fun part of any project is when you sit down and talk about what the project could do. Know that the projects might take a long time to get done, and be willing to see them through the finish.

## Open data policy

According to the Sunlight Foundation, an open data policy is critical to address "[what data should be public and how to make data public](#)." An open data policy is a hidden but very important part of the work towards building a Smart City. The ability of the public to have more information available and accessible to use in meaningful ways is predicated on a standard open data policy internally. An open data policy is a city's declaration of its intention to be more transparent and accessible by explicitly outlining what it needs to do to open its data to the public to increase communication and collaboration.

*The City of Colorado Springs, CO, established an online initiative called SpeakUP, which is an online platform that allows community members to engage directly with the local government through providing feedback and input on areas and issues for which the City solicits. The City used the SpeakUP platform to upload its draft open data policy, and provided 2-3 questions for the public to respond to. Many of the comments were more about the open data program in general, which helped shape the policy as well as open data program development. The comments from the public affirmed that the City was headed in the right direction.*

*In Colorado Springs, OpenDataCOS is one of the City's Smart Cities projects, which are captured on the City's SmartCOS platform. Although its development was underway before the SmartCOS strategy was constructed, OpenDataCOS and a formal open data management program for internal City data were identified as crucial elements of any future-learning strategy*

*for City government. As the City looks to multiply the numbers of connected devices in its community, it is crucial that it establishes an intelligent stance toward data management.*

Cities like Colorado Springs that successfully developed open data policies have followed the tenants discussed below. Because this list is not exhaustive, we recommend you review the [Sunlight Foundation Open Data Guidelines](#) and [GovEx's Policy Components For Open Data Program](#).

### **Identify formal and informal standards for opening data**

*When Colorado Springs was positioned to pursue open data, the city began by reaching out to others who had already done so. Using resources like Twitter, the What Works Cities network, and other city leaders, Colorado Springs was able to balance information from subject matter experts and data stewards with its own circumstances in order to best determine how to contextualize findings and follow suit.*

Because open data bears a strong relationship to public records, organizations or departments should specifically consider how it relates to their general public records access laws. At the same time, it is useful to talk through with internal stakeholders how open data is also different from public records: it is the proactive publication of information to everyone at once, which means that your team should discuss how it might distinguish the data it commits to publish online from data which can be requested through public records request. For example, maybe voter registry can be requested through a public records request, but your team might also develop a privacy policy which restricts it from publishing it in unaltered form on the department's website.

### **Review existing policy measures and consider alternatives**

Review other open data policies before drafting. Note what your department or team especially likes about existing open data policies and what it would like to bring over into its own. Try cutting-and-pasting from existing policies to collect all of the components that you like into a single draft.

### **Work closely with city officials**

*The data governance team in Colorado Springs formed a relationship with the deputy chief of staff early on. Getting the deputy involved in the open data process while in its beginning stages made the employee a natural internal champion in the political realm. Because of this, the committee was able to message to the internal stakeholders and gatekeepers the more tangible objectives and potential real-world results of an open data policy, like the reduction of repetitive work and better measurements to achieve better outcomes. Working closely with city officials helped the committee ensure their investment, and thus the success of the open data policy.*

Whatever form of policy your team ends up developing, most open data policies contain:

- Definitions (what is open data?)
- Commitment to open data
- General governance structure & responsibilities
- Location for open data (portal or site)
- Commitment to review and improve

### **Revise publication prioritization to reflect new levels of internal interest and shift some internal sharing to open data portal**

*In constant collaboration with numerous community stakeholders, Colorado Springs identified a number of outcomes that the city would like to achieve, leveraging Internet of Things, gigabit speeds, prolific sensors, and other technology. The treatment of the data generated by these sensors, controllers and devices will be useful for the City, and much of it can be a valuable asset for the community. As Colorado Springs moves forward with other SmartCOS projects, it will be useful to 1) have a formal data management program, specifically the publication of Open Data, and 2) have the OpenDataCOs program managed in the same team that PMs the other SmartCOS efforts. This will make sure that the City can “default to open” as much as possible as it leans into its SmartCOS projects, and also that it can depend on an intentional data management program that has been constructed to treat data with the appropriate sensitivity and accountability.*

Make sure your data governance committee has rights of final sign-off on the draft before it goes to the approving authority to ensure that it lines up with their capacities and established areas of responsibility. Finally, make the policy announcement the kick-off of your department’s open data program. Plan internal trainings and external events to follow on from the policy’s launch in order to take advantage of the interest and momentum.

## **Ensure Security and Accountability of Public Data**

The sections above have addressed data management practices, but smart cities initiatives require additional data collection, sometimes at the individual level. So how does a city go about collecting and managing data in a manner respectful to civil liberties and rights? How does a city guarantee that external actors won’t use their data for inappropriate and nefarious purposes? Portland, Oregon’s practice of collecting data from sensors can be a case study for protecting public data. Through careful and deliberate contracting with vendors, transparency with the public, and a steadfast commitment to citizen privacy, the City offers an example for other Smart Cities to emulate. This section will help you:

- Work with vendors
- Assess sensitivity

- Consider cybersecurity risk mitigation strategies

## Working with vendors

Vendors and city governments have different goals, and while working with a vendor can be a helpful way to acquire a technical service that the city cannot provide, it's important to consider whether your problem needs technology to be solved. For instance, would adding a million dollars worth of sensors in a high-need neighborhood provide more value than investing a million dollars in home repair? This is a question that can only be answered by clearly articulating the outcomes that you wish to achieve. Smart Cities should be about solving problems, not purchasing technology. If moving forward with a vendor makes sense, consider building those outcome based goals into a vendor contract in order to ensure accountability to the same goal, which goes beyond simply installing sensors.

Through its [Traffic Safety Sensor Project](#), Portland has been using sensors to understand how and when vehicles, pedestrians and bicycles use street infrastructure; monitor and analyze vehicle speeds; and track supply and demand of parking spaces, in order to design better streets. In order to protect the privacy of residents, the City worked with the project vendors to ensure that photos are not saved and are recorded over, and that any information is anonymized. Portland City Council must approve any changes to these terms.

### **Anticipate challenges and address use and ownership of data**

The City of Portland worked with the American Civil Liberties Union (ACLU) and vendors on the traffic safety sensor project - which included AT&T, Current by GE, Intel and Portland General Electric - to ensure that data would not be used in ways that would encroach on citizen privacy. For instance, the City's contract with AT&T included language around data ownership and use. Under the agreement, all rights to the data, including all intellectual property rights, belong solely to the City, and the City retains that ownership at the end of the project. An excerpt of the language on Use and Ownership of Data from this contract is below:

#### *Use of Data.*

City Data may be used and distributed by the City with no restrictions. Contractor has a limited, non-exclusive license to Access and Use Contractor Data and City Data solely for performing its obligations under the Contract. Unauthorized Use of City Data or Contractor Data by Contractor or third parties acting on behalf of Contractor is prohibited. For the purposes of this requirement, the phrase "unauthorized use" means the data mining or processing of Data, stored or transmitted by the Service, for unrelated commercial purposes, advertising or advertising-related purposes, or for any other purpose other than security or Service delivery analysis that is not explicitly authorized by the Contract.

#### *Ownership of Data.*

Notwithstanding anything to the contrary elsewhere in the Contract, the Parties acknowledge and agree that Contractor and/or the respective Alliance Members own all Contractor Data.

Contractor will provide the City access to City Data. City Data is owned by the City, and all rights, including all intellectual property rights, remain the exclusive property of the City. The City will retain ownership of all City Data following the conclusion of the Project. Any and all Data other than City Data is Confidential Information subject to Master Terms and Conditions. In addition, the Parties acknowledge and agree that any Data in the form of live video is and will remain owned by Contractor, and is not being made available to the City under this Project.

In engaging with the vendors and addressing these concerns from the start, the City is setting a precedent on how to approach this technology and methodology going forward. The contract is public, and formed the beginning of the City's open data policy around IOT.

### **Address personally identifiable information**

Their contracts help ensure that Portland and its partners use the data legally and responsibly, and protect confidentiality and personally identifiable information (PII). Their contracts also stipulate policies, requirements, and specifications for edge processing of data for its sensors. After the City draws the data from the monitors, the video is deleted after approximately one second, maintaining privacy and assuaging concerns that the City might keep the surveillance. Any capability to connect to that video connects to that sensor, but it deletes video after it is converted to a data format, so none of it is stored in the long run. An excerpt of language from the video storage and access section of the AT&T contract is below:

*"The Parties acknowledge and agree that live video captured by the Nodes will be retained only for the time needed to process the video and convert it to City Data, a period of approximately one (1) second. After the video is processed, it will be permanently deleted. The Parties also acknowledge and agree that Access to the live video via the Situational Awareness API will be disabled, and that the City, the Contractor and/or the respective Alliance Members or any other parties will not, under any circumstances, be granted Access to live or recorded video."*

The City knows that transparency is crucial for the process to engineer trust, and incorporates this principle into their practices. The contracts also provide that hardware, peripherals, and software upgrades can be made as technology changes and improves, and continues to meet privacy and security requirements.

### **Assessing sensitivity**

When opening your data to the public, you have to be cognizant of security and privacy risks as you prepare it for release. The City should classify its data according to its sensitivity based on risk factors, potential impacts, and legal/compliance requirements. Typically there are three levels: "Public/Open," "Sensitive," and "Restricted." Classifying data as Public/Open automatically qualifies it as being suitable for opening. Data that is not at this level yet should have thorough analysis and scrutinization.

Your data needs a complete overhaul and double-checking before you publish it. You should properly aggregate your data, anonymize it (make any personal data anonymous), and redact data when necessary. This ensures sensitive data isn't released or reconstructible. While classifying your datasets, you should aim to be as granular as possible (e.g., making sheets/tabs within spreadsheets documents, tables within databases, or groups of tables in enterprise systems).

Protecting citizen privacy in the release of your data should be one of your highest priorities. When protecting privacy, consider whether you are sharing data that at least some citizens might consider private. Don't share anything that puts your citizens at risk from unnecessary targeting of any kind. Always present citizen data in terms of being part of a group or demographic so as to make their shared data less personal. If you experience a breach, you must work immediately to secure your information and prevent further breaches.

While you're collecting and storing your data, you'll need to focus on a couple of factors: Data and documentation ownership, standardizing processes, and how long it'll take to hold on to data of different types. How much storage, redundancy, and security is necessary for your data? Is degrading quality (e.g., for visual, geographic, temporal, categorical, and other types of data) for the sake of space acceptable? Should the highest level of access be maintained for data in perpetuity? Who should maintain this access and how are those decisions made?

On the question of whether to collect or store data and which data to collect or store, what level of privacy is expected by the community in exchange for what benefits? Is some data more risky in the event of a breach than it's worth for analyses? Is some sensitive data permissible to collect, but not to store? When protecting your data, have you considered privacy concerns and back-ups? Addressing these sorts of questions will bolster your security and prioritization. Oakland's City Council adopted a [resolution](#) establishing a [Privacy Advisory Committee](#) to advise the City on best practices to protect the rights of residents as the City purchases and uses surveillance equipment and other technology that collects or stores data.

## Considering cybersecurity

Networked computers are everywhere these days, from our cars to our thermostats. However they are vulnerable to cyber attacks if not configured properly. As Smart City initiatives deploy new sensors and networked devices, they have to plan for security in ways they may not have considered previously. Morgan State University has launched a [research lab](#) focused on security for the Internet of Things to help communities navigate this discussion. Through this program MSU is able to educate cities and residents about common ways that these devices are compromised, and strategies for overcoming cybersecurity challenges while managing an expansive network.

### **Cyber security challenges in Smart Cities: safety, security, and privacy**

The evolution of Smart Cities creates new economic and social opportunities but also [challenges in security and expectations of privacy](#). These challenges deal with illegal access of information and attacks resulting in service disruption. Particularly, privacy protecting systems face technological challenges when it comes to gathering data and triggering emergency responses. Here we provide some general considerations for you and your city going forward.

Locational data, privacy, and a citizen's relationship with the state are key cybersecurity concerns that require deep legal and technical security measures. However the benefits of having a Smart City generally outweigh the risks. As long as you can find an acceptable paradigm for your security apparatus, you can mitigate the risks.

One applicable security paradigm for a Smart City, according to IBM, is the [IN3](#), which entails how staffers should be using their devices in an acceptably secure way for the City:

- Instrumented
  - Give devices to city officials that respond to sensor network
- Interconnected
  - Devices must pass information into a network
- Intelligent
  - Have information available for decision-making
- Areas that need to be secured with computing systems
  - The “privacy” and confidentiality of the information
  - The integrity and authenticity of the information
  - The availability of the information for its use and services

How should a city start developing a cybersecurity strategy? Datasmart at Harvard University offers [six ways cities](#) can make cybersecurity a top priority. These recommendations address funding, attracting and fostering talent, partnering with the public sector, collaborating with an open data program, developing a strong response policy, and proactively creating a culture that values and supports risk management.

## Opening and sharing data

As discussed above, it's critical to work with vendors to understand who owns the data and where it will live. Data collected from communities should be accessible by residents of those communities, and one way to do this is to host it on an open data portal. This section will help you:

- Develop open data portal requirements
- Determine dataset publication priority
- Ensure accessibility of data

## Develop portal requirements

It is crucial to keep track of key roles in the open data process. A **consumer** is anyone who visits the portal to find, access, and use data. A **publisher** is a government employee who is responsible for publishing or updating data and other content on the portal. Finally, an **administrator** is a government employee who manages the technical and configuration options of the portal, such as visual themes, metadata requirements, etc. Generally, a publisher can do anything a consumer can do, and an administrator can do anything a publisher can do. For technical details on procuring or developing an open data portal, see the [GovEx guide](#). Below are some tips:

### Community Features

Consumers should be able to provide comments and feedback on individual data resources, nominate data resources for public release, and provide comments and feedback on the data catalog. Publishers must be able to view consumer-nominated data resources, may be able to update the status of consumer-nominated data resources, may be able to respond to consumer-nominated data resources, may be able to view and respond to consumer comments and feedback, and may be able to review and approve visibility of consumer comments and feedback.

Administrators should consistently track and analyze consumer feedback. With accounts and profiles, consumers should be permitted to register for an account with an email address and create, edit, or remove profile information which may be accessible to other consumers. Publishers should be able to grant specific, registered consumers permission to access specific data resources that they own. Administrators must be able to manage permissions for other administrators and publishers and be able to disable consumer accounts (if consumer accounts are a feature of the portal).

Publishers should be able to add, edit, or remove catalog entries and mark them as private. They must be able to create new data resources to replace or update existing data resources. Administrators must be able to create, edit, and retire metadata categories, and configure global settings for the data catalog (see below), including color theme, branding/logo, titling, custom Cascading Style Sheets (CSS), homepage display elements and layout, displayed catalog elements, and more. Additionally, they must be able to manage permissions for other administrators and publishers.

Consumers must be able to search for datasets by one or more terms contained in dataset metadata, browse or explore data by category, download the catalog in a machine-readable format compatible with the [common core metadata schema](#) from a well-known address (typically /data.json), view a summary page for each dataset which details data resources, metadata, and other relevant documentation, and be permitted to search for datasets by terms contained in the

data. They must be allowed to download stored data in bulk and may access stored data through a REST API.

### **Data Catalog Features**

A data catalog is a listing of available data with each catalog entry corresponding to one or more data resources. It usually includes high-level information for each dataset, such as Title, Description, Release Date, Category(-ies), and Keyword(s).

### **Data Presentation Features**

Consumers must be able to preview data resources through their internet browser, visualize data resources in a graphical way (including maps, graphs, charts, etc.) and be able to save visualizations with configuration settings for future viewing. Publishers may be able to create data visualizations and make them accessible through the data catalog or data resource page(s). Administrators should be able to control whether consumer-created data presentations are visible to other consumers.

### **General Technical Requirements**

The following requirements are general technical requirements which either don't require a specific role or apply to all roles.

- Data resources must be available for download in the format in which they were originally published on the portal.
- API calls which are invalid or fail must return the appropriate HTTP response status code (404, 500, etc) and status message.
- The portal must be accessible through a city-provided internet domain name (e.g., data.somecity.gov).
- Transport Layer Security (TLS) must be used for account registration, logging in, and all actions performed by publishers and administrators.
- Web pages must be accessible on large-screen devices (desktop computers, laptops, etc.).
- Web pages may be accessible on small-screen devices (mobile/smart phones, tablets, etc.).

## **Determining dataset publication priority**

Governments should set a process for reviewing data for priority and quality prior to release. Additionally, they need to focus on incorporating public feedback into their data quality review and release schedules. Cities may want to develop a rubric or scorecard to add rigor to this process, categorizing criteria into quality, value, and cost. They may also want to consider how the number of criteria in each category will affect scoring, and weigh their criteria accordingly. After identifying datasets and their technical characteristics, the next step is to determine the relative value of publishing a given dataset.

When evaluating data, there is a set of measures to consider: completeness, accuracy, timeliness, and the degree to which it is automated. These factors will affect how useful your data is to your citizenry, and encourage staffers to adopt better practices when dealing with their data. All are interrelated, but remain distinct. Here we discuss how you should conceptualize these measures.

A dataset is **complete** if it has all the required information present, including in its metadata. This makes it easily machine readable and easily transformed as a result. Completeness is one of the most basic measurements of your data quality, ensuring you avoid gaps in information and give residents a complete breakdown of your data.

**Accuracy** refers to how well the dataset reflects the real world, and how effectively it avoids making mistakes about the information it means to convey. You should verify your dataset's accuracy as thoroughly as possible before publishing it, lest it accidentally misleads citizens. Does the dataset need an additional edit or revamp in its current format before publishing?

**Timeliness** is related to accuracy, but only deals with how relevant the dataset is at the time of its publication. Does the data represent the situation as it currently is and is it relevant to present issues and concerns? Is it related to a citywide goal and/or priority? Does it relate to department's mission, core function, and/or goals? Is there low effort involved in keeping the data up-to-date? Having a regular series of processes for ensuring that your data is released at a time relevant for the populace.

When we talk about how data is **automated**, we mean gauging if your organization can deliver the data to the public in an automatic process. Instead of requiring manual input every time, having an automated process will save time, increase efficiency, and make delivery of your data easier. It will also institutionalize efficient open data practices, making them integral to your organization. This will allow present and future staffers to learn these conceptual and technical best practices far more easily, reducing time and effort spent on teaching them.

Below we provide some key questions on how to assess data quality, value, and accessibility:

**Data quality** (complete, accurate, timely, automated):

- Is all of the required information present, including dataset metadata?
- Is the data valid and true?
- Does the publication/capture of the data represent the actual situation?
- Is the delivery of the data a manual or automated process?

**Value** (level of public interest, sensitivity, relevance):

- Is the data requested often?
- Does the data have a direct impact on the public?
- Is the data appropriate for publishing?
- Is there cause for concern for any reason?
- Is the data related to a citywide goal and/or priority?

- Is the data related to department's mission, core function, and/or goals?
- Can the data impact economic outcomes in the city?
- Does the data allow citizens to measure City performance?
- Does the data make government information more accessible?

**Accessibility** (machine readable, publishable, low cost)

- Is the dataset in a format that is machine-readable or can be easily transformed?
- Does the dataset require an additional step to publish due to their current format?
- Is there a low level of effort or cost to keep the data up-to-date?

## Ensure accessibility of data

When increasing access to data, you have to be cognizant of making the navigation experience on your site as user easy as possible. People should be able to find and access the data in an intuitive and straightforward way. Making it publicly accessible means not only that it is open, but that its formatting and interpretability (ability to connect with other machines) is high level. As you develop your portal, consider how the community wishes to share updates with the public, and how you can consider to engage the public. A few tips are below:

- Published data should conform to appropriate [open standards](#) (e.g., [GTFS](#) or [BLDS](#)).
- Data should be as accessible as possible to people in the city, including those who don't speak English. Having different language options is key.

Here are some resources you may want to use to open your portal:

- [How to Run an Open Data Bootcamp in Your City](#)
- [GovEx Open Data Standards Directory](#)
- [Open Data Getting Started Guide](#)

Present your data in such a way that lets citizens feel like they are part of the city's decision making process. Aim to have them trust you to make the right decisions by increasing your transparency. Being transparent about how well your programs are doing boosts trust and therefore confidence in city programming. Always aim to make accessing your data a user-friendly experience for your citizens. Incorporating public feedback into data quality review and release schedules is a good tactic for this. As you release useful and timely data it'll be more useful for your citizens, but protecting privacy will become more of an issue so be sure to have your safeguards in place.

### Measuring success

A performance management system can be helpful to identify and manage key indicators in order to understand whether the initiatives are meeting the needs of the community and

achieving Citywide goals. The indicators should be available at the city and [project level](#). To develop these indicators, consider the goals for your open data project. What do you want to accomplish? How will you know when you get there?

## **Analytics**

Kansas City, Missouri notes that smart cities infrastructure has the potential to support city operations, but it's important to start with a use case and then determine the best way to collect and analyze that data. Before speaking with vendors, city staff should start with high-need neighborhoods, and understand the challenges and priorities of people living there. With that information, it can then be helpful to see if sensors or other smart cities solutions would help. There are many problems facing residents, and not all of them are best suited to a technological solution.

Big data is not always necessary to take the pulse of a community, and in some cases it can be more confounding. How helpful is it to gather information on the number of pedestrians on a certain walk way if there is no information on how many are unique counts, and where people are coming from and where they are going? Working with vendors during the RFP phase to ensure that meaningful and timely metrics are shared with the City can help to address this challenge. Sometimes, low-tech solutions and "small data" can give more insight. A resident survey can give detailed information about resident satisfaction with city services in varying categories and jurisdictions across an entire city. These surveys also ask about resident priorities, which Kansas City successfully used to make the case for a program to demolish 800 dangerous buildings. With these baseline and historical data points, a next step might be to deploy sensors to augment it with real-time information. According to Eric Roche in KCMO, "Sensors can (ideally) tell you what is really happening. While surveys give you perception information. You can measure the gap between the two to determine if the problem is best addressed via operational changes, or just requires better communication."

For instance, asking about satisfaction with frequency or comfort of busses and then deploying sensors to measure whether behavior matches what people reported can be helpful information for transportation planners and decision makers. Balance the reality that can be collected through sensors with ground-truthing, and collect data that matters. Sens

Consider the timing of this data collection - is there a strategic planning process that this data could support? Where are you in the city budget cycle? Are there legislative considerations at the state or federal level? Are there opportunities to use data that might be collected in departments, but is not regularly analyzed? Transportation systems, bike shares, 311 and 911 systems, inspection reports are all great examples of potential sources of data. Map out these processes and see if there are unmet opportunities to use existing data and tools. If a city has existing data, but is not using it to make decisions it may indicate a culture unsupportive of data driven governance. In this case, the procurement of smart city technologies is unlikely to have

the desired impact without a sustained program to shift city leadership and staff to a data driven culture.

It's also critical to think about how resources might be deployed as a result of this analysis. For instance, deploying sensors in a high-need neighborhood can help collect valuable data points, but if the investment in sensor infrastructure means there is less funding to invest in solutions for those communities, who is really better off?

## Communicating with residents

During the community conversations described at the beginning of this guide, participants noted repeatedly that there is a lack of trust between the community and city government (as well as researchers, police, and some community-based organizations). The researchers on this project plan to return to the community to reflect the findings from the original community conversations, as well as the City's smart cities plans.

Any City that is involved in a smart cities initiative, and especially those with historical systemic segregation, should be aware of these dynamics. To continue to engender trust, there are a few steps that cities can take:

### **Acknowledge the priorities that residents have identified**

In discussing smart cities initiatives, city leaders and staff should reflect the concerns and priorities shared by residents, rather than focusing only on the technology. For instance, in Colorado Springs, the City uses its SpeakUpCOS platform to solicit input from the community on [focus areas](#) including energy and utilities, buildings and sustainability, transportation and mobility, and city services. When Mayor Suthers speaks about smart cities in Colorado Springs, he talks about the problems that will be addressed through this [initiative](#), rather than focusing on the technology. "We look forward to working with Panasonic in a collaborative process that leads to ways to make Colorado Springs more appealing for residents and tourists alike. The technological solutions include items such as LED streetlights that can detect snow levels, building efficiency technologies, and advanced security technologies to improve public safety." This is in line with Colorado Springs' definition of Smart Cities, which is "using [technology](#) to improve public services, drive economic development and solve issues facing the city."

### **Communicate clearly and effectively**

In addition to opening data, as discussed above, cities should think about developing clear measures to illustrate progress on smart cities initiatives. For instance, Boston's "[Imagine Boston](#)" plan includes smart cities infrastructure as a strategy to achieve priorities like affordability and inequality. The plan includes metrics related to 311 call center performance.

Cities should also build applications that are responsive to community needs. The finding by the University of Maryland team that residents in low-income communities primarily use cell phones to access the internet should guide Baltimore City government in developing or encouraging development of phone-based applications. Baltimore can look to other cities that have developed applications to help residents stay connected with governmental service delivery, such as [sanitation services and weather closures](#); [report blight and track those reports](#); [track progress on Mayoral initiatives](#); [find lost pets](#); and more. Kansas City, MO has a [dashboard](#) that residents can use to access information on parking availability, traffic downtown, wifi usage, and streetcar location in the central business corridor. The City also has interactive [kiosks](#) that are used to share information with the public.

## Conclusion

This guide lays out the foundation for a smart cities program, or any data initiative. As this space evolves and cities increasingly use data and technology to respond to resident concerns and proactively address problems, there will be more examples and guidance, and hopefully a communities of cities will form to share experiences and work through challenges and lessons learned together. We hope that you will share your smart cities experiences with us, and with each other.