



# 5 Steps to Start Sharing Procurement Data

**Authors:** Sam Elghanayan, Elena Hoffnagle, and Laura Merryfield

**Contributors:** Kailey Burger Ayogu, and Hope Patterson

---

Working in procurement, you probably spend a disproportionate amount of time answering questions from colleagues, vendors, and other key stakeholders. Just as you're starting to get into the flow of strategizing for an upcoming solicitation, a prospective vendor calls to ask about whether a current contract will be renewed this month or be rebid. Or maybe a colleague from another department stops by your office to ask whether your government already holds a blanket contract for office supplies with a specific company.

Now, imagine that instead of pulling up the information yourself for the third time today, you could simply point the question asker to a public website that holds commonly requested procurement and contract information. Once you start publishing contracting data online, you'll have fewer calls, happier vendors and colleagues, and more time for your actual strategic work!

This article will help procurement teams that have not yet published open data understand what open data is, how it benefits governments, purchasing staff and vendors, and how to get started. You'll learn whom to involve, how to choose what data to share, how to prepare the data before making it public, and how to publish it.

## What is open data? And what does it look like?

Open data refers to data or content that is publicly accessible, timely, and free for anyone to use, manipulate, and share for any purpose and without legal restriction. The data should be available on a public website without a password, easy to find, downloadable, and machine readable. (Machine

readable refers to data that can be used and manipulated on a computer with widely available software like Excel or R—rather than a PDF!)

What data gets shared depends, of course, on the agency, department, or government unit. Perhaps a public health department would decide to make hospitalization or immunization data available on its open data portal, or perhaps a parks department might share visitation figures. Social services providers, in turn, might share data about the numbers of residents accessing services or the demographics of the populations they serve. Ultimately, there's no one "ideal" picture of what open data might look like — but the principles of accessible, timely, and manipulable data hold true regardless of a government's particular mission or role.

## Why is open data important?

Open data can improve government transparency and increase public engagement by raising awareness of how tax dollars are spent. It enables innovation and private sector solutions like third-party apps and can increase government efficiency by making data in one department easier to access by another department.

In procurement, open data benefits governments, procurement staff, and prospective and current vendors in several different ways:

- **Reducing duplicative purchasing:** When contracting datasets are easily accessible, government staff can often piggyback off an existing contract or consolidate similar purchases, rather than procuring separately or embarking on a lengthy investigation of what existing or upcoming contracts are held by other departments or agencies.
- **Freeing up staff time:** Many questions commonly directed at buyers (e.g., Does the government have a contract for \_\_\_? Who is the current vendor on \_\_\_ contract? When are you next procuring for \_\_\_?) can be answered quickly without procurement staff's input if relevant data are easy to access and understand. Jim Campion, Purchasing Manager for the Town of Gilbert, AZ, explains that his staff gets many requests from outside agencies and the public about when a contract starts, when another one ends, what the terms are, etc.: "Now, we can point them to our open data page. It saves staff so much time."
- **Elevating the status of procurement:** Proactively sharing data can demonstrate to colleagues that the purchasing function of government is innovative, modern, data-driven, collaborative, and strategic rather than just administrative. It also elevates the status of procurement by showing that procurement data are just as important as other public data, such as 311 calls or restaurant inspection grades.
- **Benefiting vendors:** Publishing upcoming opportunities to do business with the government (often called a "procurement forecast") is particularly useful for vendors. For example, the

# How can open data increase government transparency?

The Organization for Economic Co-operation and Development (OECD) cites three ways that increase transparency and accountability:

- Allows residents and civil society to track how budgets are being spent and to whom money is flowing, which increases accountability of public sector officials and the government as a whole
- Makes information easier to analyze, combine, and compare, increasing the speed and ease of public scrutiny
- Facilitates collaboration between government, civil society, and businesses to improve transparency efforts.<sup>1</sup>

---

1. OECD Directorate of Public Governance Reform of Public Sector Division. "Compendium of good practices on the use of open data for Anti-corruption: Towards data-driven public sector integrity and civic auditing." <https://www.oecd.org/gov/digital-government/g20-oecd-compendium.pdf>

City of Tulsa, OK publishes information on contract renewals, sharing which contracts have exhausted their renewal limit and will soon be re-bid. Sharing this information allows prospective vendors to flag upcoming solicitations they may want to bid on and plan staffing accordingly. In general, open data signals to vendors that the government wants a wide pool of firms to be aware of upcoming contracting opportunities, which can increase trust in government and participation in procurement, especially among small, local, and minority-owned businesses.

## How should I start?

1. **Identify who should be involved.** To turn your open data dreams into reality, you will need to identify who owns or holds the data you want (e.g., the Chief Procurement Officer) and in what systems (e.g., ERP software), who already collects some aspect of the data (e.g., the City Clerk's Office), and who will publish it and where. Governments who are further along on their data journey may have a person or department in charge of data and performance analytics who can analyze, relabel, or summarize existing data.
2. **Find out what data you have and what to publish first.** Don't let big aspirations keep you from getting started! You can start small by publishing the data that's easier to find and upload, then fill in gaps by prioritizing the most important data to collect and publish.

First, map out what data you have. When Tulsa decided to start sharing contracting data publicly, they made a list of all the contracting data they already had and a "wish list" of all the

data they might want to publish in the future.

Next, determine which data to publish first. For each item on their list of data fields, Tulsa answered five questions.

- *How much value would be created by publishing these datasets?* For example, the Town of Gilbert first examined what public records were most often requested, which turned out to include information on contract dollar value, terms, and start and end dates.
  - *Where are the datasets?* You might find that some of this data is in paper files, Excel spreadsheets, or rests within an E-procurement or ERP system.
  - *Are the datasets accessible in their starting form?* Consider how long it will take you to access the data, whether special reports must be run, or whether data fields must be merged or combined to be useful.
  - *Are the datasets easy to keep updated?* Ideally, these data are regularly kept updated already. If not, doing so should be easily worked into an existing process or role.
  - *Is it “clean” or easily cleanable?* Data with standard wording and formatting, few typos and jargon, and without confidential information will be easier to clean for public view.
3. **Make the data accessible, understandable, and usable.** Make sure the data are easy to access, understand, and edit/manipulate/use for different purposes.
- *Accessible:* To make its contract data easy to access, Tulsa adds it to the “Bid Opportunities and Bid Results” page of its public website, which vendors already regularly check.

## What contracting data fields do governments commonly publish?

- Solicitation title or subject (e.g., “Central business district road repaving”)
- Brief solicitation description
- Category of goods or services (e.g., construction, professional services, IT)
- Solicitation or purchase type (e.g., Request for Proposals, Request for Quotes, Request for Information)
- Commodity code category
- Date solicitation released
- Contract start date
- Contract end date
- Length of contract
- Dollar amount
- Renewal date
- Agency or department owning contract
- Contract administrator or point of contact
- Awarded vendor

- *Understandable*: All content should be easily understood by a non-expert and someone who is not familiar with your government’s procurement jargon or acronyms. This might mean renaming category headers, spelling out acronyms, and adding more details to short-hand notes. A “data dictionary” can provide users with detail about what each data field does (and does not) contain.
- *Machine readable*: As discussed in the box below, data that can be manipulated by common software like Excel is much more useful than information that needs humans to interpret it (like PDFs, scans, and images). Machine readable data can be used for a variety of different purposes, such as sorting solicitations by expiration date or filtering the data to only showing a specific category of goods or services. Excel or CSV files are great formats to start with.

## What does machine readable mean?

Machine readable refers to data that can be used and manipulated on a computer with widely available software like Excel or R. Common machine-readable formats include XLS files (Excel), CSV, JSON, XML and RDF. By contrast, a PDF with a table of upcoming solicitations may be easily “readable” (understood) by a human but not by a computer. If the data were in a machine-readable format instead, the public could use software like Excel to, for example, list solicitations in order of dollar value or filter the data to show only janitorial solicitations. This ability to manipulate data is incredibly powerful. For example, a startup company might use machine readable transit data to create a tool for residents to find the best way to commute to work.

4. **“Clean” and verify the accuracy of data.** You can manually “clean” data or use simple Excel formulas or coding to confirm all dates and dollar amounts seem reasonable and are formatted consistently, typos are corrected, acronyms are replaced with complete words (e.g. “Parks and Recreation Department” instead of “Parks & Rec”), product and service categories are consistent (e.g. choose either “IT” or “Information Technology”), the distinction between a value of zero and null or no data is clear, and capitalization and punctuation are consistent. Then, enlist someone with good attention to detail and knowledge of the information being shared to review the data (line by line if feasible or spot checks if the amount of information is too great) to foster accuracy before publishing.

“The more accurately staff enter data into their e-Procurement systems, the less time we spend verifying and correcting data later,” Champion notes.

5. **Publish by starting small and regularly updating.** The City of Tulsa took a low-tech yet effective approach to publishing their data. Every four months, a staff member updates, cleans, and verifies an Excel file with contracting data, which IT then adds to the city’s [website](#). Website visitors can download the Excel file with a simple click.

The Town of Gilbert has developed a slightly higher-tech process over their years working with open data. Twice a month, a staff member downloads a spreadsheet of all open contracts from MUNIS. The Purchasing team verifies that the information is clean and correct. They then use an “extract, transfer, and load” process (ETL) that aggregates the original data to show totals (e.g., “total value of construction contracts”) and standardizes dates to reduce manual cleaning. Finally, IT loads the resulting data to their paid publishing platform, which makes it available on the city’s main [open data portal](#).

**“The more accurately staff enter data into their e-Procurement systems, the less time we spend verifying and correcting data later.”**

Be sure to develop a system with tasks and associated roles and timelines to regularly update your open data. Working this into existing processes makes it easier on staff. Derek Konofalski, the Town of Gilbert’s Data and Technology Analyst, also recommends that governments getting started “use the tools you already have. If you don’t have access to a paid open data publishing platform, check if any of your existing software includes a free open data tool.”

## A Note About Confidential Data

Data containing personal identifiable information (PII) such as staff/resident names, phone numbers, addresses, and social security numbers should NEVER be published. Even if you initially believe your dataset doesn’t contain confidential information, double-check that unique identifiers aren’t inadvertently confidential. For example, a vendor database may use the social security number of the owner as their unique vendor identifier if they are a sole proprietor — and making this database public would also make those individuals’ social security numbers public. Demographic data can also be identifiable in smaller populations. For example, if just one women-owned construction firm exists in a small county, publishing vendor survey results showing opinions broken out by demographic groups could inadvertently reveal that firm’s response.

## Conclusion

Campion advises governments considering publishing open data to “Do it: the sooner the better. It took upfront investment to set it up at the start but has saved us so much time since.” Without hesitation, he concludes, “It’s been a game changer for us.”

### How can you level up?

**Publish more data.** After you publish the “low hanging fruit” of data that is useful to share, easier to access, and relatively clean, you can return to your initial data map to identify data that are highly useful, but harder to access or clean. You might even create a “wish list” of information you would like to begin tracking (e.g., cycle times) and identify how you could collect this data, who would need to input it, and into what database it would be stored.

**Automate the process.** Over time, you can automate steps in the process of collecting and publishing data and get more sophisticated in how you present data. You can add procurement-related open data to your government’s existing larger open data portal, which might contain many more publicly accessible data sources from across your department, agency, or entire government. New York City’s “Citywide Procurement Indicators” [presentation](#) is a good example of how, in addition to open data, your government can analyze data and discuss key takeaways in a narrative report and create visualizations like graphs and maps that make it faster and easier for people (especially those less familiar with procurement or less comfortable with raw data) to glean insights from. [Little Rock](#) and [Gilbert](#) let visitors do simple analyses and visualizations of data straight through the web browser. New York City even [provides APIs](#) that allow individuals, researchers, businesses, etc. to automatically access and repurpose data to, for example, create apps or other software based on the city’s data.

**Apply the OCDS.** The Open Contracting Data Standard (OCDS) may provide you with additional ideas for data fields and help facilitate more advanced and varied uses of data. Additionally, the Open Contracting Partnership [provides](#) valuable guidance and resources that you might find helpful in your journey to share additional procurement data.



The **Procurement Excellence Network** is an initiative of the Government Performance Lab designed to help public sector leaders use government procurement as a tool to improve resident outcomes and advance equity. The **Government Performance Lab**, housed at the Taubman Center for State and Local Government at the Harvard Kennedy School, conducts research on how governments can improve the results they achieve for their citizens. An important part of this research model involves providing hands-on technical assistance to state and local governments. Through this involvement, we gain insights into the barriers that governments face and the solutions that can overcome these barriers. By engaging current students and recent graduates in this effort, we are able to provide experiential learning as well.

The Government Performance Lab is grateful for support from Bloomberg Philanthropies.

What Works Cities Certification provides cities at all points in their data journey with a standard of excellence that shows how investing in data and evidence practices can lead to better and more equitable results for residents. This guide includes strategies in alignment with the following What Works Cities Certification criteria:

- Results-Driven Contracting (RDC) 7: Open and Shared Procurement Data

**Learn more** about how to get your city Certified.

© Copyright 2022 Harvard Kennedy School Government Performance Lab