NYC OpenData

Getting ready for your agency's annual Open Data Report

Each City agency is required by law to report annually on their efforts to make their public datasets available and accessible. This report is shared with the Mayor and City Council, and made public on NYC Open Data. To make this process as easy as possible, the Open Data Team shares a workbook for each agency to complete, reviews these workbooks, and consolidates them into a single citywide report and series of associated datasets.

Here are some steps to get ready for your reporting and successfully completing this workbook:

1. **Check in with the core divisions/teams in your agency to explore any new data or systems for potential Open Data**

   As an Open Data Coordinator, you should be talking regularly with your agency's teams and divisions to better understand the data that they manage, and ensure your agency is fulfilling its legal requirement to publish public datasets. Some of the key questions to consider here include:
   - Are the core services your agency provides reflected in the data on Open Data?
   - Is the data that is used by or collected with new applications, systems, or programs shared on Open Data?
   - Are your Mayor's Management Report (MMR) indicators reproducible with or reflected on Open Data?

   An important reminder here is that while agencies are required to report on the existence of new datasets as part of this annual reporting process, these new datasets don't need to be published immediately. Datasets should be prioritized based on demand and team capacity and can be scheduled for publication into the future.

2. **Talk with your agency's legal counsel…**

   …about FOIL requests

   An essential part of the annual reporting process is inventorying your agency's data that is already being shared publicly in response to Freedom of Information Law (FOIL) requests. The intention here isn't to publish exactly what's shared through FOIL, but to use those as a guide to where a dataset might exist. For example, a FOIL request for a specific set of addresses, statuses, or dates should result in the underlying system or database used to fulfill that request being evaluated for publication. If your agency does not track the data sources used for answering FOIL requests, you will also need to identify those sources to complete the annual report.

   …about the legal definition of Open Data

   One of the key outcomes of the Open Data annual reporting process is a schedule for when newly identified datasets are being released. While the exact contents of a dataset do not need to be determined for a release date to be set, understanding NYC's law about what is and is not a public dataset will be helpful in evaluating what your agency is required to publish.

   …about certifying your workbook

Before we can accept your Open Data annual reporting workbook, we require that your agency's general counsel or legal representative certify the workbook's accuracy and completeness. Identifying who is doing this for your agency, and informing them of this requirement, will help you to quickly move through this process.

3. **Connect with the people responsible for your agency's website**

Similar to FOIL requests, data that is being shared on public websites managed by or on behalf of your agency is a good indicator of what needs to be published on Open Data. As with FOIL, the intention here is not to reproduce the specific datasets as they're presented on the internet, but to use those datasets as indicators of underlying systems or databases that need to be evaluated for publication. Talking to the people responsible for your agency's website to identify any new datasets (whether they're shared as a downloadable file, a link, or even a table in a report) or data-powered tools (like a dashboard or map) is a good step to prepare your agency to complete this section of the report.

4. **Review your agency's past publication commitments**

The ultimate goal of the check-ins with your agency's divisions, examination of FOIL request sources, and investigation of the origins of your agency's website data is a schedule of what datasets will be published when. While you might be adding new datasets to this schedule as part of this year's reporting process, you should also review the list of to-be-published datasets that your agency has already committed to sharing.

As part of this review, you should explore whether the datasets your agency listed make sense as originally structured (Are some individually listed datasets better shared as a collection of interrelated datasets? Can multiple tables in separate entries that are listed be flattened into a single dataset?) and ensure that the deadlines your agency committed to are still viable (Is more time necessary to restructure a dataset? Create a comprehensive plain-language data dictionary? Review with your Agency Privacy Officer?).

5. **Review your agency's inventory of already published datasets**

While the primary focus of the annual reporting exercise is identifying new datasets and sharing updates on datasets that are yet to be published, this is also a good time to review your agency's inventory of existing open data. You can filter NYC Open Data's asset inventory for your agency or use the Open Data dashboard to review this inventory. Are your datasets being updated as scheduled? Do the update frequencies make sense? Are there new fields that need to be added? Should there be any changes to the dataset titles or other metadata? Reach out to our team (opendatateam@oti.nyc.gov) with any edits or questions.