

NEW YORK STATE OPEN DATA PROGRAM OVERVIEW

data.ny.gov



Open NY History and Introduction

Emerging technologies have dramatically transformed how government does business and changed the ways in which the public interacts with government. To give New Yorkers the access they deserve to government data and information, Governor Andrew M. Cuomo launched the Open NY initiative in March 2013 and signed [Executive Order 95](#). It directs state agencies to identify, catalog, and publish their data on [open.ny.gov](#), the state's open data website administered by the [Office of Information Technology Services \(ITS\)](#).

The award-winning website is a hub for one-stop access to information that helps the public better understand what is happening in government at the state, local, and federal levels. Open NY increases transparency, improves government performance, empowers New Yorkers to participate in government, and encourages research and economic opportunities statewide.

ITS leads the Open NY initiative and is guided by its mission to ensure New York State:

- Proactively releases publishable state data and makes it freely available in accessible formats for the public to re-use and consume.
- Publishes high-quality data sets with metadata and documentation to promote understanding.
- Continuously publishes and updates state data to foster discovery and trust.

“As it expands and evolves over time, [open.ny.gov](#) will spark innovation, improve efficiency, promote accountability, and bring the people back into government.”

– Governor Andrew M. Cuomo

Q&A

What do you mean by “open data”?	2
Where can we find published data?	3
What can I expect to find on the Open NY platform?	3
Who is responsible for publishing and updating data?	4
How do agencies determine which data it publishes?	4
How do agencies determine the priority level of data to publish?	5
Will the data be updated?	5
When can we expect Open NY program updates?	6
Where can I find more information?	6

Q. What do you mean by “open data”?

A. “Open data” refers to data that can be released without restrictions in a user-friendly format, and can be retrieved, downloaded, indexed, and searched by commonly used web search applications. [Open.ny.gov](https://open.ny.gov) gives researchers, citizens, and businesses direct, centralized access to government data to search, explore, download, and share.

According to [Executive Order 95](#), “publishable data” means information that a state entity is able to make public according to all applicable laws, rules, regulations, ordinances, resolutions, policies, or other restrictions, like contractual or legal orders. Data is not publishable if it would violate statute or regulation, invade personal privacy, endanger public health or welfare, hinder government operation or investigations, or impose undue burdens on the state.

Q. Where can we find published data?

A. There are two main platforms for finding data published by state entities.

OPEN.NY.GOV

- Open NY's main platform, also accessible by visiting data.ny.gov
- Established in 2013 by Executive Order 95
- Main website for published data from all covered entities, including state, local, and federal governments
- Maintained by ITS with information published and updated by covered entities

HEALTH.DATA.NY.GOV

- Companion platform that includes health data owned and published by the [New York State Department of Health \(DOH\)](https://www.health.ny.gov)
- Established in 2013 in support of Open NY
- The only known open data site in the U.S. devoted solely to state health data accompanied by targeted public health messaging, extensive metadata, and visualizations
- Data published here is also discoverable on open.ny.gov

Q. What can I expect to find on the Open NY platform?

A. Open NY is a global platform, which makes it valuable not only to New Yorkers, but also for citizens outside of New York State and individuals who are unfamiliar with our state government. Those who are new to exploring open data shouldn't let the technical jargon get in the way of exploring Open NY. Both the platform and the supplemental documentation that agencies are required to submit prior to publication are designed to help the audience understand the who, what, when, where, how, and why of the data. The following are a few elements citizens can expect to find with each data set on open.ny.gov.

Data Set: The way the data is published has a great effect on how consumers interact with, and use, the platform. Participating agencies are required to present the data in a machine-readable format so software, applications, and systems can process it. This also gives users the opportunity to manipulate, search, and customize the raw data for their individual needs.

Metadata: Entities are required not only to standardize the data provided, but also include metadata within their data sets and adhere to metadata standards across the platform, so users can search for and discover common data sets and related information.

Data Dictionary: This piece explains the individual columns in each data set, providing a consistent, standardized vocabulary to help the user understand the information provided.

Overview Document: This separate one- to two-page document explains the data set in detail, provides background and context of the data, explains the data collection process (methodology), highlights statistical and analytical issues, and notes any limitations to data use. The document helps users interpret the data and inspires them to discover new insights and observations.

Entities must also submit a signed approval form with the data. Additional documentation may also be required when it helps users better understand the data (for instance, supplemental links, studies and reports specific to the data, and more).

Q. Who is responsible for publishing and updating data?

A. Each state agency or entity has a data coordinator who is responsible for ensuring the agency is in compliance with [Executive Order 95](#), as well as understanding their agency's data and resources, Open NY guidelines, and directives. The data coordinator is the liaison between their agency and the ITS team, and in larger agencies, receives support from subject matter experts and divisions that help identify, prioritize, and update data that can be published. Data coordinators are also responsible for obtaining agency approvals for publishing, alteration, and removal of data. When an agency has prepared and organized a data set, its data coordinator works with ITS to publish it.

ITS maintains open.ny.gov and health.data.ny.gov, provides technical support and assistance to agencies and entities, and promotes Open NY to the public.

Q. How do agencies determine which data it publishes?

A. When creating a data catalog, ITS advises agencies to determine not only which data sets comply with “publishable state data” under [Executive Order 95](#), but high-quality and complete data that is of value to the public. The information, for example, could increase an agency's accountability or responsiveness, add to the public's knowledge about an agency and its functions, further the mission of an agency, create economic opportunity, or respond to public need or demand the information.

ITS also advises agencies to understand the platform's users when determining which data sets would represent discrete, usable information. Users can include the general public, journalists, academics, nonprofits, and professionals from a wide range of fields and specialties who may find uses for raw data that an agency's data coordinators haven't envisioned. Rather than limiting data sets to only those the agency believes might be understood or useful, ITS encourages agencies to instead ensure that its metadata and overview documents are complete and comprehensive. The supplemental documents should, for example, describe the data, data collection methods, data fields, and presentation of research questions to make the data useful to as wide an audience as possible.

Q. How do agencies determine the priority level of data to publish?

A. Each covered state entity should create and maintain timetables for publishing data. To create a schedule, an agency should balance the data's value and usefulness, quality, availability, and readiness, in addition to demand for the data and the time it will take to prepare it for public consumption.

The amount of time to prepare high-quality data sets can vary depending on the set's complexity and size. To determine a data set's priority and set a publishing schedule, agencies should consider whether the data:

- Has a direct impact on the public.
- Supports decision-making at the state, local, or agency level, or contains information that informs public policy.
- Aligns with new state or agency initiatives or is similar to recently released federal data.
- Creates specific economic opportunity.
- Is frequently requested, has strong demand from constituents, or needed by requestors after-hours.
- Highlights agency performance or benefit the public by setting higher standards.
- Improves agency-to-agency communications.
- Fulfills legal requirements for data publication.
- Addresses regulatory or grant requirements.
- Has been published or made available in machine-readable format so it can be processed, analyzed, or re-used.
- Could be useful for the creation of third-party apps, mobile apps, and services.

If data sets are partial or incomplete, but would be of great value to the public, agencies should work to complete them and set a timetable so they can eventually be published. Agencies should ensure that data is organized or formatted to be machine-readable and capable of being digitally transmitted or processed.

Q. Will the data be updated?

A. Yes! Agencies are required to keep their data sets published on open.ny.gov up to date. To refresh a data set, entities can either replace the data by removing all existing records and adding new ones or append the data by inserting new records into the existing data set. Agencies assist users by indicating the date that the data sets are updated and the posting frequency in the metadata provided.

Q. When can we expect Open NY program updates?

A. Technology is changing every day, and Open NY is evolving with it. The following is a timeline of ITS milestones carried out in accordance with [Executive Order 95](#). Keep an eye out for the latest [Open NY Quarterly Report](#) to stay updated on the progress of the program.

March 2013	Governor Cuomo issues Executive Order 95 , and ITS launches open.ny.gov website
April 2013	The state appoints data coordinators for agencies and establishes a data working group
November 2013	ITS releases first Open Data Handbook
April 2014 to present	ITS publishes Quarterly Reports
June 2016	ITS issues Dataset Submission Guide
May 2020	ITS updates Open Data Handbook with version 2.0
June 2020	ITS updates Dataset Submission Guide with version 4.0
First Quarter 2023	ITS anticipates next update to Open Data Handbook

Q. Where can I find more information?

A. **New York State Open Data Handbook:** This in-depth handbook, updated in May 2020, provides more explicit detail regarding directives and guidelines for participating agencies.

Open NY Quarterly Reports: Find up-to-date information and news about newly published data sets and platform statistics in Open NY's reports released each quarter.

Dataset Submission Guide: This document for agencies and data coordinators includes best practices and lessons learned regarding data set formatting and documentation.