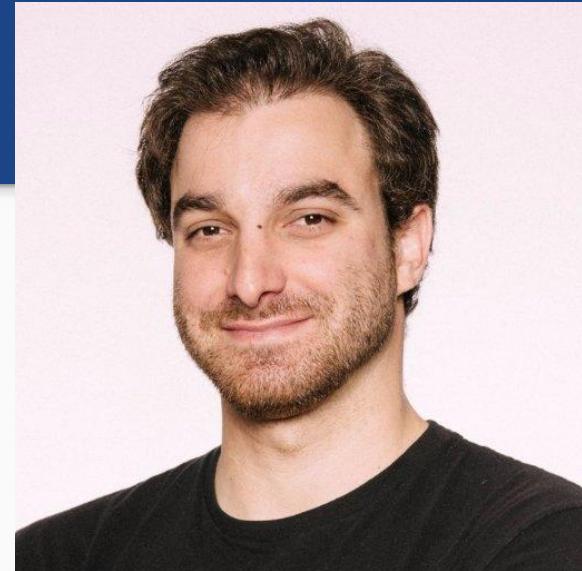




March, 11, 2021 - Open Data Week

Hi, I'm Devin Balkind

- A software producer who organizes teams to deliver solutions for startups, nonprofits and government agencies.
 - Sarapis: nonprofit
 - Pruvn: business
 - BetaNYC: contractor for city agencies (past)
- I was born & raised in NYC.
- Plan to do WeGov work for 30+ years.



@devinbalkind on Twitter
devin@sarapis.org

Let's Make NYC the World's Best Run City

New York City is literally the world's "favorite" city.

We're the richest city in the richest country. We have the resources to be the best.

We aren't. And that's bad for us and bad for the world.

To billion of people around the world, NYC represents the concept of urban life, a global city, a diverse "melting-pot", immigration and opportunity.

We must do better.

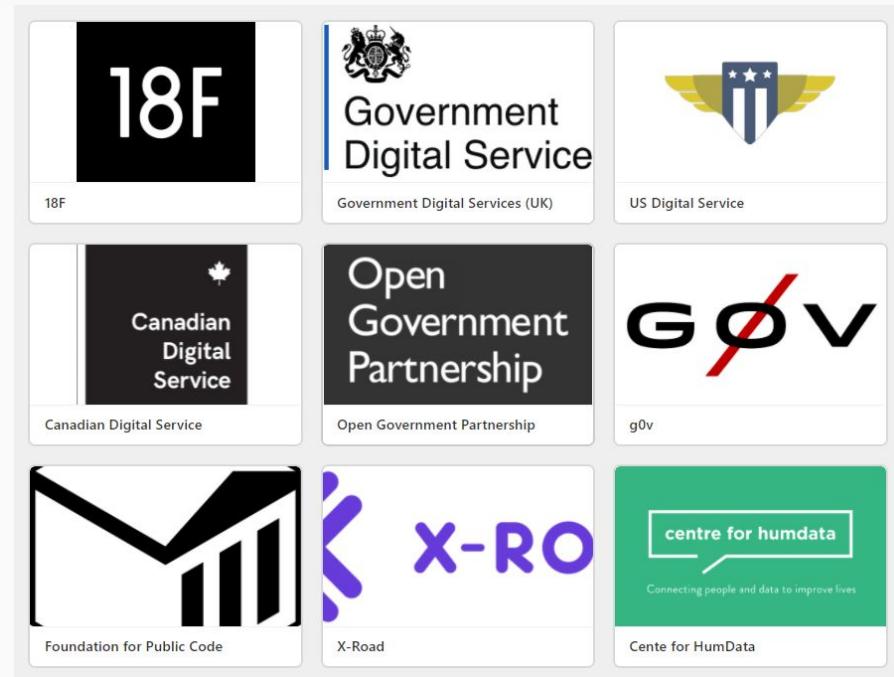


What Does Better Look Like?

Over the last decade, people have figured out how to go government tech better.

New York City doesn't have a general purpose digital service organization applying best practices in digital transformation.

MODA does great work but it's scope is limited.



How do we help?

Phase 1: Tools

We create and support open source software & data projects.

Phase 2: Community

We offer news, knowledge and networking opportunities to like minded people.

Phase 3: Advocacy

We build public awareness and support for best practices and solutions.

Tools

We help technology projects that making NYC more open, effective & participatory.

We've build almost a dozen experimental apps: a City Agency Directory, Capital Project Explorer, City Services API, Laws on Your Phone, Political Directory & Data Maps.

The DataBook is our first real offering.

We also offer resources to projects that align with our values, such as Reported App.

And we provide long-term maintenance for projects, like Capital Projects Directory.

Resources for Reported App

“Reported” helps New Yorkers report dangerous drivers via an information collection app that sends reports to 311, TLC, Twitter and more.

Volunteer developers got stuck and couldn’t get to 3.0 version and the app was broken for half its users.

We gave them \$3,000 worth of a Sarapis developer’s time to get 3.0 launched and are now helping them raise money to keep that developer onboard indefinitely.



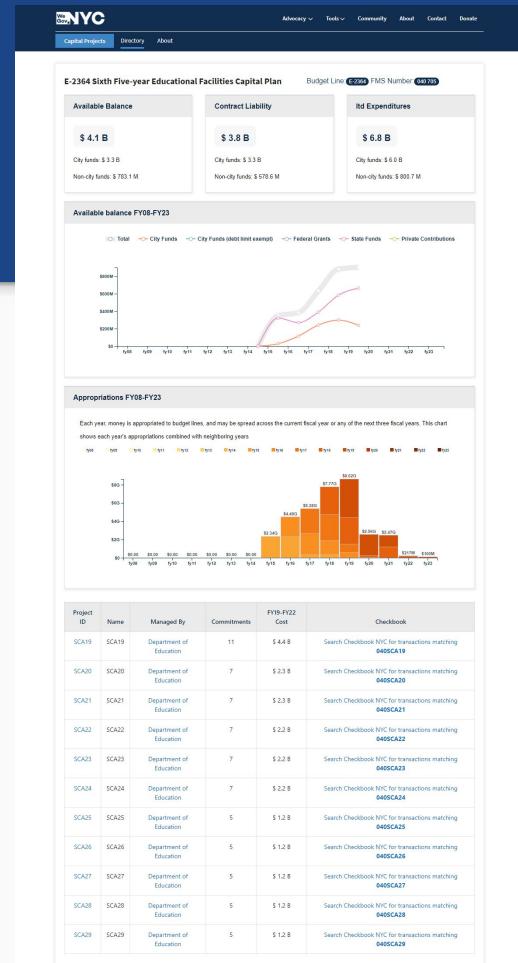
Maintenance for Capital Projects Directory

Chris Whong famously scrapped the capital projects PDFs in 2017 and put up a site.

With his blessing, we forked the site, restyled it and added some features.

This allowed Chris to focus on newer projects, like his newer geocoded capital projects map.

We'll help ensure that is maintained long term too.



Open for Business

Know of tool that aligns with our vision?

Tell us about it.

Maybe we can help!

Plus an insight: we ask too much of volunteer software developers! They can't be expected to launch and maintain a project indefinitely. Better is for them to start and structure projects, do as much work as they want, and then bring in a paid (outsourced) software developer to do the annoying stuff, keep things maintained and the project rolling. We love to fund that!

The WeGov.NYC Databook

Open Data is Good

- One of NYC's only technology success stories is NYC Open Data.
- Congrats to MODA for this. It's a significant and genuine achievement.
- Great data tables, published regularly, with useful metadata. Awesome.

Normalizing the data will unlock its true potential to power more useful and usable public app.

What to Normalize?

- Data Clinic's Scout Tool helps us answer this question.

Agency	124
Agency Name	69
AGENCY	33
Agency Code	14
AGENCY CODE	12
Agency Number	11
Issuing Agency	10
AGENCY NAME	10
Agency Comment	10
agency_name	5

Council District	251
City Council Districts	211
Council District	48
City Council District	33
Community Council	27
COUNCIL DISTRICT	24
COUNCIL_DISTRICT	14
CouncilDistrict	11
CouncilDist	7
Council Member	4

Community Board	341
Community Districts	212
Community School District	28
Community Council	27
BOROUGH / COMMUNITY	23
COMMUNITY BOARD	21
Strong Family-Community Ties - Percent	20
Positive	
CommunityBoard	18
Strong Family-Community Ties Rating	16
Community District	12

General Proposal

NYC Open Data

Official UIDs

- Agency
- Neighborhood (NTA)
- Community District

Normalized Datasets

NYC should:

- Establish UID indexes
- Build a data pipeline that applies those UIDs to open data sets.
- Republish those normalized datasets

We've Done a Little Bit of It

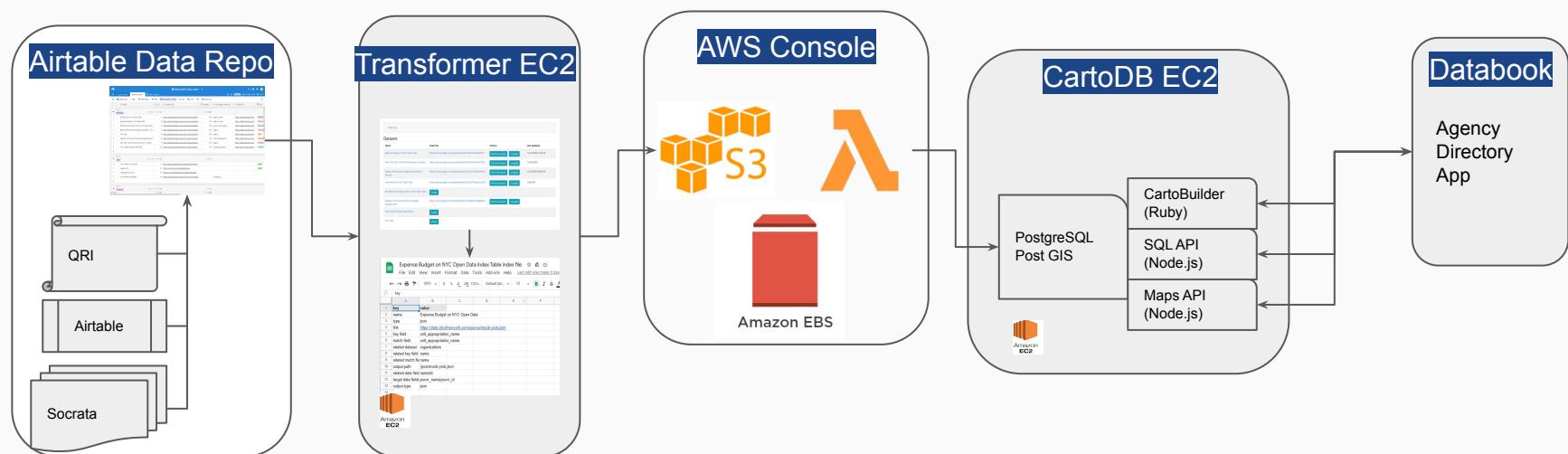
Find data open data sets, scrapers, data entry.

Normalize the data by adding consistent UIDs over all the datasets.

Archive the data in S3 so it's easy to share in bulk and in AWS environment.

Upload data into Carto so its accessible via the Carto SQL API.

Develop apps powered by the SQL API.



How the Transformer Works

1. Log Datasets in Airtable

Datasets are logged in the [WeGovNYC Data Index](#) Airtable's data_sources table. Fields:

- Name: Title of the dataset
- ID: Internal identifier
- Citation URL: Where dataset links to.
- Update Schedule: Frequency of data update in hours.
- API Agency Field Name: name of field in the dataset containing a city agency.
- Data URL: Link to JSON data API
- Description: Text to appear in apps describing dataset.
- Source: Location of dataset
- Status: What to do with it
- Index Table ID: ?
- Last Updated: When the data was last transformed.
- Index Filling Data: When data was last indexed.
- Output URL: Location of transformed data.

The screenshot shows the 'data_sources' table in Airtable. The table has three main sections based on the 'STATUS' field: 'Develop', 'Spec', and 'Evaluate'. Each section contains a list of datasets with columns for Name, ID, Citation URL, API Agency Field Name, Data URL, and a color-coded category. The categories are People (blue), Money (orange), Services (green), Indicators (red), Requests (yellow), and Jobs (purple). The 'Develop' section has 8 records, the 'Spec' section has 4 records, and the 'Evaluate' section has 9 records. The total count is 25 records.

WeGovNYC Data Index						
organizations	data_sources	Add or import	Grid view	...	Hide fields	Filter
Grouped by 1 field						
	A Name	ID	Citation URL	Sort	Color	Share view
	STATUS	Count 8	Sum 53		Sum 4322	
	▼ Develop					
1	Greenbook on NYC Open Data	1	https://data.cityofnewyork.us/City-Government/Greenbook-on-NYC-Open-Dat...	168	agency_name	https://data.cityofnewyork.us/
2	Expense Budget on NYC Open Data	2	https://data.cityofnewyork.us/City-Government/Expense-Budget-on-NYC-Open-Dat...	730	agency_name	https://data.cityofnewyork.us/
3	Benefits and Programs API on NYC Open Data	3	https://data.cityofnewyork.us/Social-Services/Benefits-and-Programs-API-on-NYC-Open-Dat...	168	government_agency	https://data.cityofnewyork.us/
4	Agency Performance Mapping Indicators – An...	4	https://data.cityofnewyork.us/City-Government/Agency-Performance-Mapping-Indicators---An...	730	agency	https://data.cityofnewyork.us/
5	NYC Jobs	5	https://data.cityofnewyork.us/City-Government/NYC-Jobs	336	agency	https://data.cityofnewyork.us/
6	Register of Community Board Budget Request...	6	https://data.cityofnewyork.us/City-Government/Register-of-Community-Board-Budget-Request...	730	responsible_agency	https://data.cityofnewyork.us/
7	New York City Council Discretionary Funding	7	https://data.cityofnewyork.us/City-Government/New-York-City-Council-Discretionary-Funding	730	agency	https://data.cityofnewyork.us/
8	NYC Capital Project Detail Data	25	https://api.gri.cloud/chriswhong/nyc_capital_project_detail_data	730	managing_agency	https://api.gri.cloud/get/ch...
	+					
	STATUS	Count 4	Sum 50		Sum 0	
	▼ Spec					
9	NYC Charter, Code Rules	9	http://library.amlegal.com/nxt/gateway.dll/NewSection?SectionID=10000000000000000000000000000000			Legal
10	Legistar API	10	https://council.nyc.gov/legislation/api/			Legal
11	Checkbook NYC API	15	https://www.checkbooknyc.com/data-feeds/api			Legal
12	NYC Facilities Database	16	https://data.cityofnewyork.us/City-Government/NYC-Facilities-Database		overagency	
	+					
	STATUS	Count 9	Sum 167		Sum 0	
	▼ Evaluate					
25 records			Sum 325			Sum 4322

2. Create, Fill & Prematch in Transformer App

The Transformer pulls data sets from the Data Index and allows the user to:

- **Create** an “Index file” Google Sheets for the file.
- **Fill & Pre-match** the file. This pulls the Airtable Agency Index, matches it to the agencies in the sheet, and allows the user to manually change those matches.
- **Compile** then applies the matched fields to the new dataset and publishes it at the URL accessible via the “Last Updated” field.

Script Log			
Datasets			
Name	Index file	Actions	Last Updated
Expense Budget on NYC Open Data	https://docs.google.com/spreadsheets/d/198rv5odU0H47Tt1T...	<button>Fill & Pre-match</button> <button>Compile</button>	12/10/2020 12:51:34
New York City Council Discretionary Funding	https://docs.google.com/spreadsheets/d/1FAoNkwWShvTrRH...	<button>Fill & Pre-match</button> <button>Compile</button>	12/09/2020
Agency Performance Mapping Indicators – Annual	https://docs.google.com/spreadsheets/d/1nZctLitC4KAzHWw...	<button>Fill & Pre-match</button> <button>Compile</button>	12/10/2020 09:49:19
Greenbook on NYC Open Data	https://docs.google.com/spreadsheets/d/1QUKTGssplAccyND...	<button>Fill & Pre-match</button> <button>Compile</button>	2/6/2018
Benefits and Programs API on NYC Open Data	<button>Create</button>		
Register of Community Board Budget Requests API	https://docs.google.com/spreadsheets/d/1QdARtxUajj9gD5x...	<button>Fill & Pre-match</button> <button>Compile</button>	
NYC Capital Project Detail Data	<button>Create</button>		
NYC Jobs	<button>Create</button>		

3. Match in Google Sheet

A Google Sheets is generated with three sheets:

- Configuration Info from Airtable>Data_Sources
- Organization List from Airtable>Organizations, which is the canonical list of organizations used for matching.
- **Dataset Match**, which displays a list of all agencies in the dataset and an editable fuzzy match with all our official organizations.

A1	Source UID	Source text	Core text	Core UID
1	ACS	ACS	Administration for Childrens Services	Administration for Childrens Services
2	BKBP	BKBP	Brooklyn Borough President	Brooklyn Borough President
3	BKCB	BKCB	skip	skip
4	BPL	BPL	Brooklyn Public Library	Brooklyn Public Library
5	BXBP	BXBP	Bronx Borough President	Bronx Borough President
6	BXCB	BXCB	skip	skip
7	CCRB	CCRB	Civilian Complaint Review Board	Civilian Complaint Review Board
8	CFB	CFB	Campaign Finance Board	Campaign Finance Board
9	CLERK	CLERK	City Clerk	City Clerk
10	COIB	COIB	Conflicts of Interest Board	Conflicts of Interest Board
11	CUNY	CUNY	PSC CUNY	PSC CUNY
12	City Clerk	City Clerk	City Clerk	City Clerk
13	Clerk	Clerk	City Clerk	City Clerk
14	DABK	DABK	District Attorney - Kings	District Attorney - Kings
15	DABX	DABX	District Attorney - Bronx	District Attorney - Bronx
16	DANY	DANY	District Attorney - N.Y.	District Attorney - N.Y.
17	DAON	DAQN	District Attorney - Queens	District Attorney - Queens
18	DASI	DASI	District Attorney - Richmond	District Attorney - Richmond
19	DCA	DCA	Department of Consumer Affairs	Department of Consumer Affairs
20	DCAS	DCAS	Department of Citywide Administrative Services	Department of Citywide Administrative Services
21	DCLA	DCLA	Department of Cultural Affairs	Department of Cultural Affairs
22	DCP	DCP	Department of City Planning	Department of City Planning
23	DDC	DDC	Department of Design and Construction	Department of Design and Construction
24	DEP	DEP	Department of Environmental Protection	Department of Environmental Protection
25	DFTA	DFTA	Department for the Aging	Department for the Aging
26	DHMH	DHMH	Department of Health and Mental Hygiene	Department of Health and Mental Hygiene
27	DHS	DHS	Department of Homeless Services	Department of Homeless Services
28	DJJ	DJJ	Department of Juvenile Justice	Department of Juvenile Justice
29	DOE	DOE	Department of Education	Department of Education
30	DOHMH	DOHMH	Department of Health and Mental Hygiene	Department of Health and Mental Hygiene
31	DOITT	DOITT	Department of Information Technology and Telecommunications	Department of Information Technology and Telecommunications
32	DOP	DOP	Department of Probation	Department of Probation
33	DOT	DOT	Department of Transportation	Department of Transportation
34	DPR	DPR	Department of Parks and Recreation	Department of Parks and Recreation
35	DSBS	DSBS	Department of Small Business Services	Department of Small Business Services
36	DSNY	DSNY	Department of Sanitation	Department of Sanitation
37	DSS	DSS	Social Services Administration	Social Services Administration
38	DSS/HRA	DSS/HRA	Social Services Administration	Social Services Administration

4. Publish to S3 and Carto

S3 data in our public data repo Airtable as “Output URL”

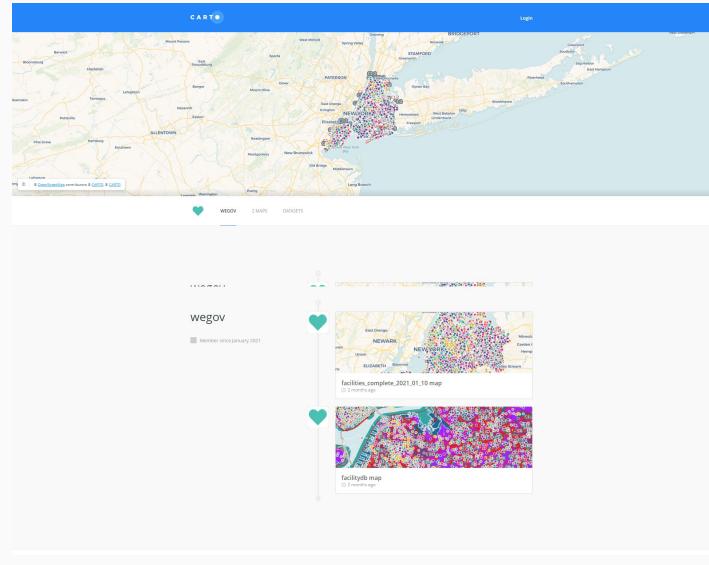
Airtable

1 hidden field Filter Grouped by 1 field Sort ...

Count 16

Name	Description	Citation URL	Data URL	Output Url	Last Updated
102 Greenbook on NYC Open Data	The Green Book Online is a fully searchable dat...	https://data.cityofnyc.org/	https://data.cityofnyc.org/datasets/greenbook-on-nyc-open-data	https://wegov-r-125.s3.amazonaws.com/greenbook-on-nyc-open-data	2/27/2021 08:49
103 Expense Budget on NYC Open Data	Operations/Expense budget for all NYC agencies	https://data.cityofnyc.org/	https://data.cityofnyc.org/datasets/expense-budget-on-nyc-open-data	https://wegov-r-125.s3.amazonaws.com/expense-budget-on-nyc-open-data	2/29/2021 18:23
104 Benefits and Programs API on NYC Open Data	This dataset provides benefit, program, and res...	https://data.cityofnyc.org/	https://data.cityofnyc.org/datasets/benefits-and-programs-api-on-nyc-open-data	https://wegov-r-125.s3.amazonaws.com/benefits-and-programs-api-on-nyc-open-data	2/27/2021 08:56
105 Agency Performance Mapping Indicators – Annual	This dataset provides key performance indicat...	https://data.cityofnyc.org/	https://data.cityofnyc.org/datasets/agency-performance-mapping-indicators-anual	https://wegov-r-125.s3.amazonaws.com/agency-performance-mapping-indicators-anual	2/27/2021 08:45
106 NYC Jobs	This dataset contains current job postings avail...	https://data.cityofnyc.org/	https://data.cityofnyc.org/datasets/nyc-jobs	https://wegov-r-125.s3.amazonaws.com/nyc-jobs	2/27/2021 08:04
107 Register of Community Board Budget Requests API	This dataset reports each of the projects and p...	https://data.cityofnyc.org/	https://data.cityofnyc.org/datasets/register-of-community-board-budget-requests-api	https://wegov-r-125.s3.amazonaws.com/register-of-community-board-budget-requests-api	2/27/2021 09:00
108 New York City Council Discretionary Funding	The dataset reflects applications for discretion...	https://data.cityofnyc.org/	https://data.cityofnyc.org/datasets/new-york-city-council-discretionary-funding	https://wegov-r-125.s3.amazonaws.com/new-york-city-council-discretionary-funding	2/27/2021 08:44
109 NYC Facilities Database	The Facilities Database (FacDB) captures the lo...	https://data.cityofnyc.org/	https://data.cityofnyc.org/datasets/nyc-facilities-database-facdb	https://wegov-r-125.s3.amazonaws.com/nyc-facilities-database-facdb	2/27/2021 08:48
110 NYC Capital Project Detail Data	A dataset of capital projects data scraped from...	https://api.cloud.chicago.gov/	https://api.cloud.chicago.gov/datasets/nyc-capital-project-detail-data	https://wegov-r-125.s3.amazonaws.com/nyc-capital-project-detail-data	2/27/2021 09:01
111 Expense Financial Plan - Exec	This dataset contains agency summary level da...	https://data.cityofnyc.org/	https://data.cityofnyc.org/datasets/expense-financial-plan-exec	https://wegov-r-125.s3.amazonaws.com/expense-financial-plan-exec	3/05/2021 17:46
112 WeGov Organization Index		https://writable.co/			
113 Government Publications Listing	Metadata for documents submitted to the Dep...	https://data.cityofnyc.org/	https://data.cityofnyc.org/datasets/government-publications-listing	https://wegov-r-125.s3.amazonaws.com/government-publications-listing	04/2021 23:26
114 Government Publication- Required Reports	Metadata for documents submitted to the Dep...	https://data.cityofnyc.org/	https://data.cityofnyc.org/datasets/government-publication-required-reports	https://wegov-r-125.s3.amazonaws.com/government-publication-required-reports	03/2021 18:51
115 Local Law 251 of 2017: Published Data Asset Invent...	As per Local Law 251 of 2017, the Open Data P...	https://data.cityofnyc.org/	https://data.cityofnyc.org/datasets/local-law-251-of-2017-published-data-asset-invent	https://wegov-r-125.s3.amazonaws.com/local-law-251-of-2017-published-data-asset-invent	05/2021 17:45
116 NYC Open Data Release Tracker	A list of all datasets that were included on all t...	https://data.cityofnyc.org/	https://data.cityofnyc.org/datasets/nyc-open-data-release-tracker	https://wegov-r-125.s3.amazonaws.com/nyc-open-data-release-tracker	05/2021 17:44
117 OneNYC Indicators	Annual Agency Performance Metrics	https://data.cityofnyc.org/	https://data.cityofnyc.org/datasets/onenyc-indicators	https://wegov-r-125.s3.amazonaws.com/onenyc-indicators	05/2021 17:44

Lamda'd into carto.wegov.nyc



5. Displayed in the Databook

Our first display is a City Agency Directory, with the following datasets for each:

- People - Greenbook on NYC Open Data
- Projects - NYC Capital Project Detail Data
- Services - Benefits and Programs API on NYC Open Data
- Requests - Register of Community Board Budget Requests API
- Indicators - Agency Performance Mapping Indicators – Annual
- Finances - Expense Budget on NYC Open Data
- Jobs - NYC Jobs
- Facilities - NYC Facilities Database
- Organizations - WeGov Organization Index

The screenshot shows a grid of agency profiles from the NYC WeGov Databook. Each profile includes the agency's logo, name, a brief description, and tags. The agencies listed are:

- NYC Administration for Children's Services**: Tags: Social Services
- NYC Administrative Justice Coordinator**: Tags: Civic Services
- NYC Board of Correction**: Tags: Civic Services
- NYC Board of Education Retirement System**: Tags: Education
- NYC Board of Standards and Appeals**: Tags: Civic Services, Housing & development
- Bklyn Public Library**: Tags: Culture & Recreation
- New York City Campaign Finance Board**: Tags: Business
- NYC Center for Innovation through Data Intelligence**: Tags: Business
- The City Clerk**: Tags: Business
- The Council**: Tags: Business

Marginally Very Cool

Timing is Everything:

- 15-30 minutes to normalize a dataset, no coding involved.
- 1 hour to add a new dataset using very popular Laravel and DataTables.net tools.
- Design follows USDS Federal Design System
- Datasets are updated automatically. If an agency name changes or isn't matched, we get an alert to update the matching table.

The screenshot shows a web page from the We Gov. NYC website. The header includes the NYC logo, news links, and a search bar. The main content area is titled "Administration for Childrens Services / Requests". It features a table of budget requests with columns for Publication Date, Borough, C Board, Priority, Tracking Code, and Request description. The table has 10 entries, filtered from 244 total. A note at the bottom states: "This data comes from Register of Community Board Budget Requests API" and "Last updated 01/27/2021".

Publication Date	Borough	C Board	Priority	Tracking Code	Request
2020-07-01	1	1	03	101202103E	Provide, expand, or enhance funding to support higher education and/or workforce development opportunities for youth who are leaving foster care
2020-07-01	1	6	08	106202108E	Provide, expand, or enhance preventive services and community based alternatives for youth
2020-07-01	2	14	19	214202119E	Other foster care and child welfare requests
2020-07-01	2	7	04	207202104E	Provide, expand, or enhance child welfare preventive services slots
2020-07-01	2	7	12	207202112E	Provide, expand, or enhance preventive services and community based alternatives for youth
2020-07-01	2	8	17	208202117E	Provide, expand, or enhance preventive services and community based alternatives for youth
2020-07-01	2	8	39	208202139E	Provide, expand, or enhance primary prevention services to strengthen families
2020-07-01	2	13	41	213202141E	Provide, expand, or enhance housing assistance for youth that are leaving foster care
2020-07-01	2	13	48	213202148E	Provide, expand, or enhance funding to support higher education and/or workforce development opportunities for youth who are leaving foster care
2020-07-01	3	3	11	303202111E	Provide, expand, or enhance housing assistance for youth that are leaving foster care

What's Next?

- 10+ more agency datasets
- Adding charts and maps
- Normalizing geographies: council districts, community districts, NTAs
- Adding interactivity with comments and other user generated content
- Integrating/incorporating other projects

Then hopefully:

- Partnerships with good government orgs
- Advocacy for more, better, normalized open data and API-driven web apps

The screenshot shows a screenshot of the NYC Open Data website. At the top, there is a navigation bar with links for 'Add About', 'Subscribe', 'Write For Us', and 'Contact'. Below the navigation bar, there is a search bar and a 'We Gov. NYC' logo. The main content area has a dark blue header with the text 'Administration for Childrens Services / Requests'. Below the header, there is a sub-header 'Administration for Childrens Services' with a small NYC logo. The main content area contains a table with data from the 'Register of Community Board Budget Requests API'. The table has columns for Publication Date, Borough, C Board, Priority, Tracking Code, and Request. The data in the table includes various budget requests from different community boards across the city.

Publication Date	Borough	C Board	Priority	Tracking Code	Request
2020-07-01	1	1	03	101202103E	Provide, expand, or enhance funding to support higher education and/or workforce development opportunities for youth who are leaving foster care
2020-07-01	1	6	08	106202108E	Provide, expand, or enhance preventive services and community based alternatives for youth
2020-07-01	2	14	19	214202119E	Other foster care and child welfare requests
2020-07-01	2	7	04	207202104E	Provide, expand, or enhance child welfare preventive services slots
2020-07-01	2	7	12	207202112E	Provide, expand, or enhance preventive services and community based alternatives for youth
2020-07-01	2	8	17	208202117E	Provide, expand, or enhance preventive services and community based alternatives for youth
2020-07-01	2	8	39	208202139E	Provide, expand, or enhance primary prevention services to strengthen families
2020-07-01	2	13	41	213202141E	Provide, expand, or enhance housing assistance for youth that are leaving foster care
2020-07-01	2	13	48	213202148E	Provide, expand, or enhance funding to support higher education and/or workforce development opportunities for youth who are leaving foster care
2020-07-01	3	3	11	303202111E	Provide, expand, or enhance housing assistance for youth that are leaving foster care

Showing 1 to 10 of 17 entries (filtered from 244 total entries)

Previous 1 2 Next

ⓘ This data comes from [Register of Community Board Budget Requests API](#)

Last updated 01/27/2021

Zooming Out: We're Ready to Build Community

We want to create spaces where people creating technology in the public interest can raise awareness about issues, showcase their projects and onboard collaborators.

To do that, we're gearing up to offer the following:

- News: stay up to date with our blog, social media, newsletter and Slack.
- Events: meetup to learn and share new things, and find collaborators.
- Knowledge: a dynamic directory of relevant people, places and things.
- Services: a fiscal sponsor enabling projects to access nonprofit benefits including tax-deductible contributions, credits, and more.

Open for Business

We're hiring a data manager & community builder.

We're also looking for sponsors that believe in our mission.

Get Involved: <https://tinyurl.com/wegovnyc>

Contact Me: Devin Balkind, devin@sarapis.org, @devinbalkind on Twitter

Notes

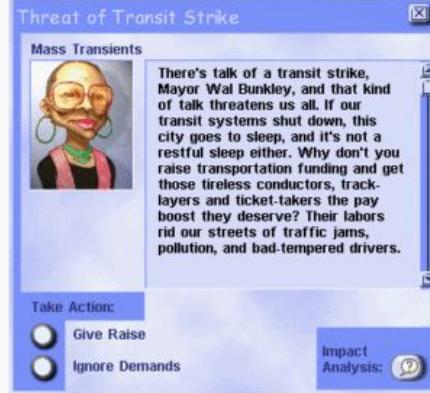
Eventually: SIMNYCity

Imagine an interactive “game” where you could see city data on real maps, follow agency activities, give feedback on issues, make your own budget and more.

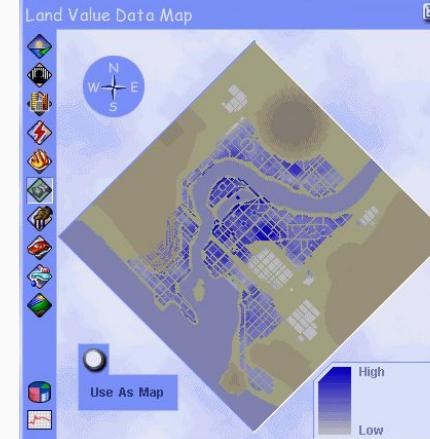
Citizen-Driven Budgets



Decision-Making



Data Maps



Geospatial Awareness

