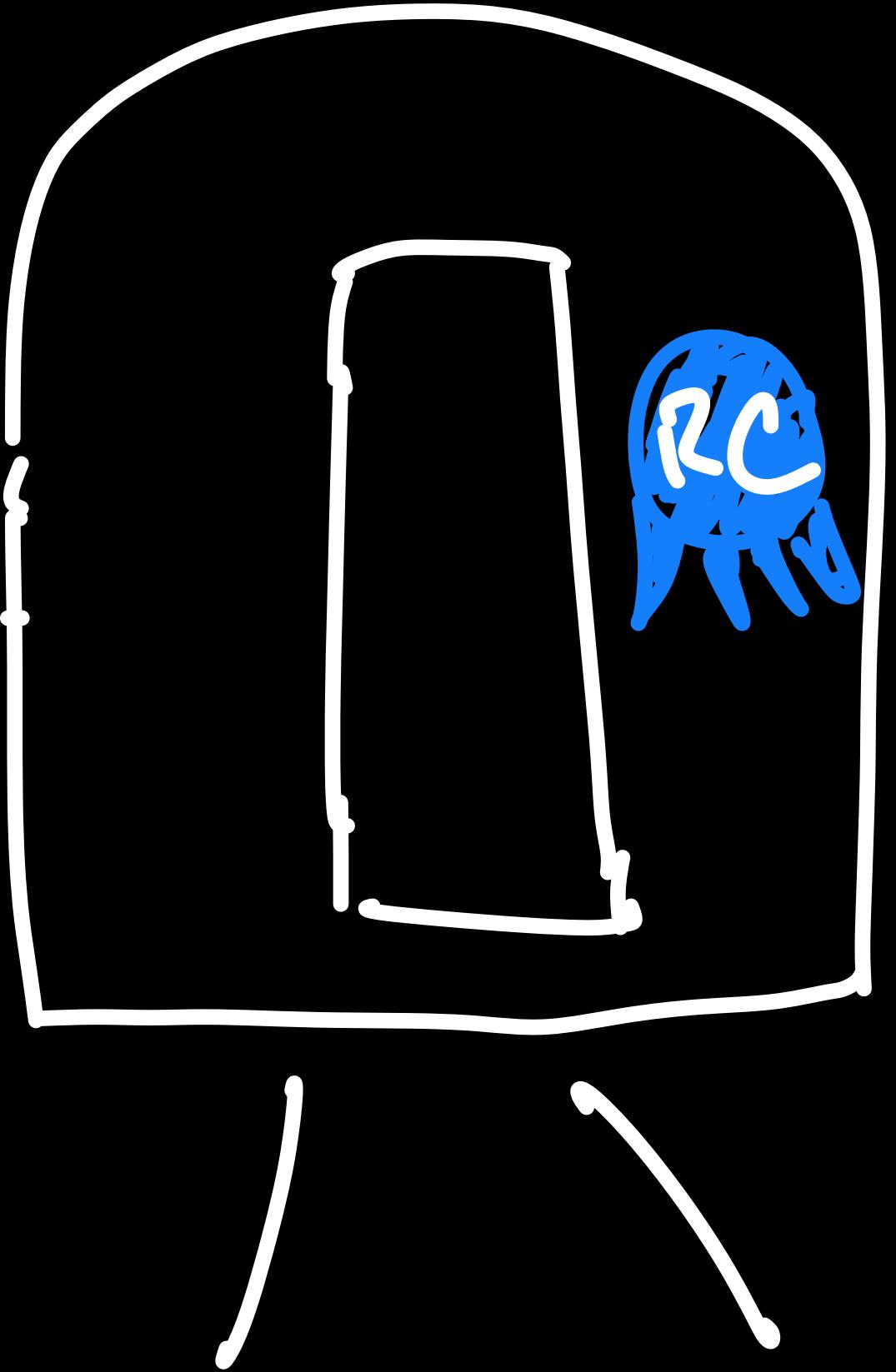


subway surfing

or, my initial adventures with the MTA API

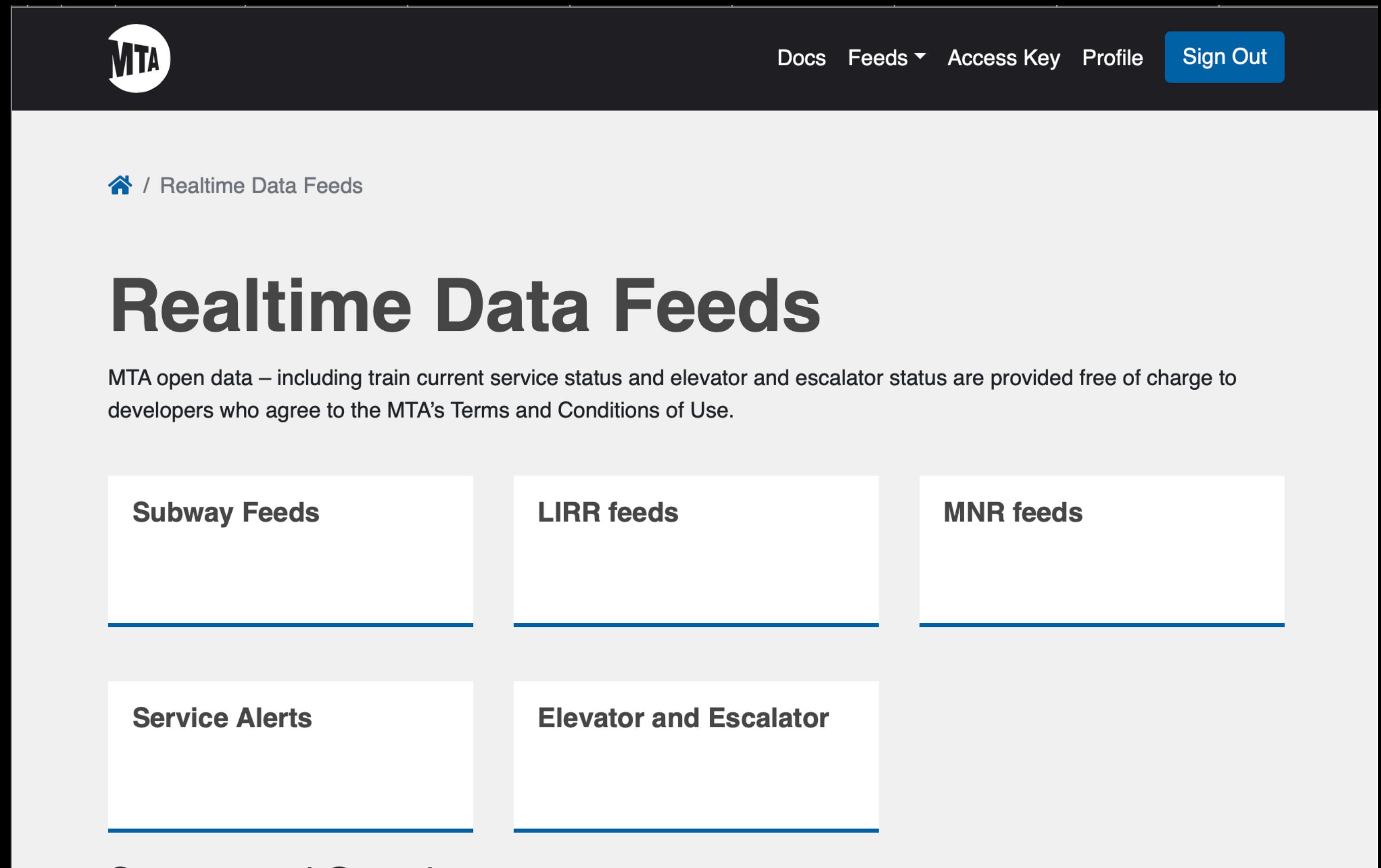
P.B. To
The Recurse Center
Jan 25 2024



“countdown clocks”

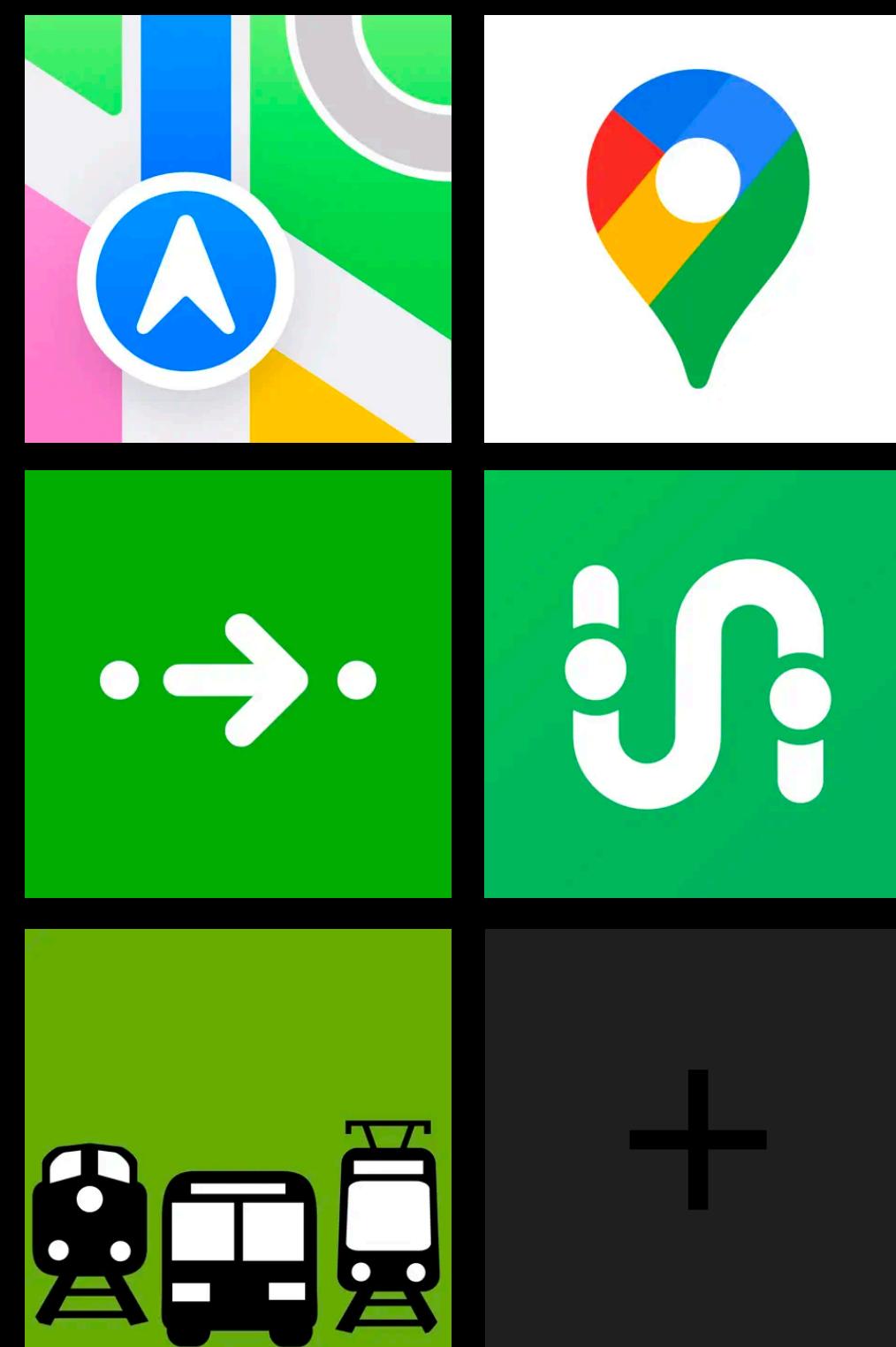


[https://medium.com/good-service/
new-feature-detecting-
delays-48d29df9ba54](https://medium.com/good-service/new-feature-detecting-delays-48d29df9ba54)



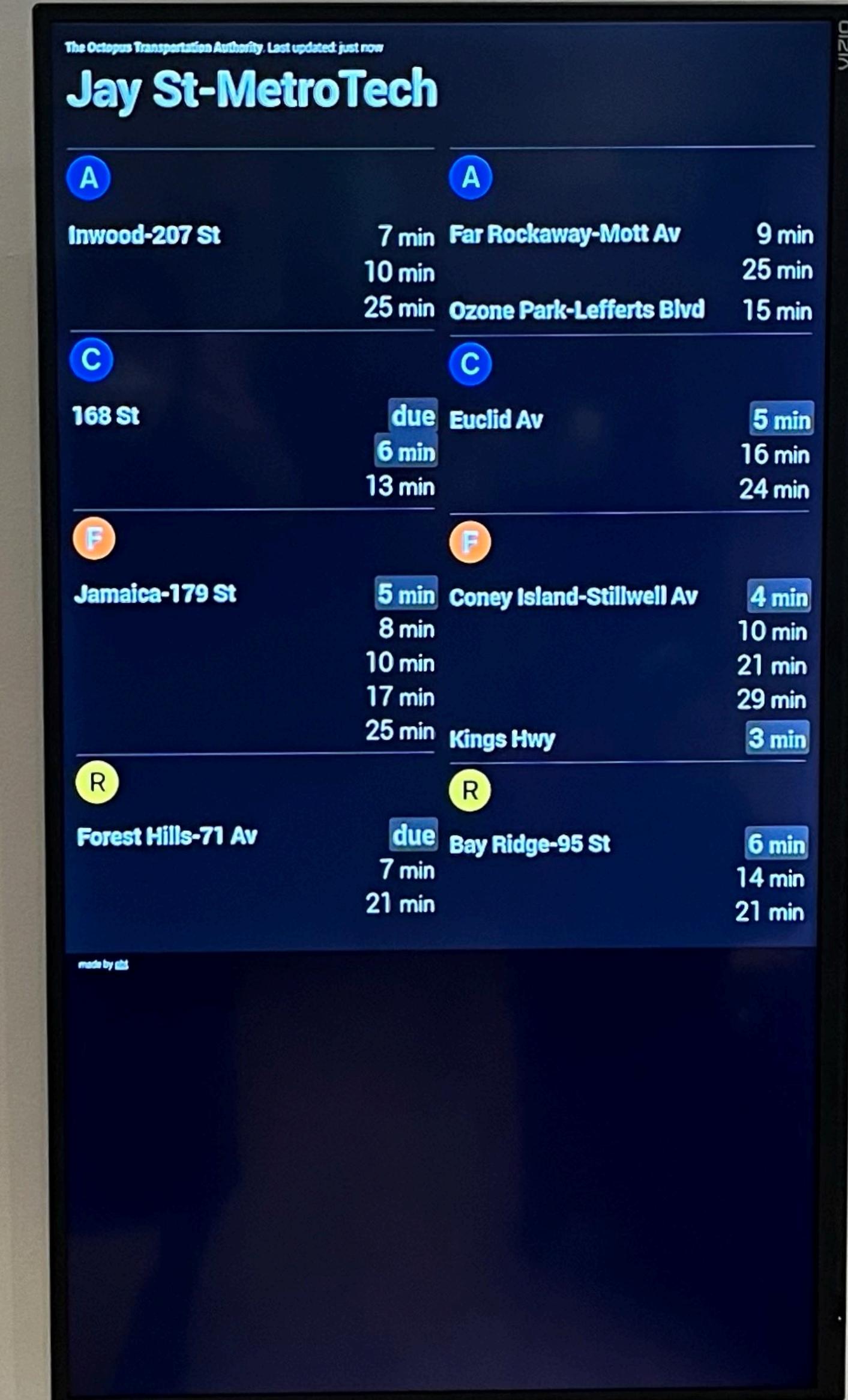
The screenshot shows the MTA Realtime Data Feeds homepage. At the top, there's a dark header with the MTA logo, navigation links for 'Docs', 'Feeds ▾', 'Access Key', 'Profile', and a 'Sign Out' button. Below the header, a breadcrumb trail shows the user is at the 'Realtime Data Feeds' page. The main section features a large heading 'Realtime Data Feeds' and a paragraph explaining that MTA open data is provided free of charge to developers who agree to the Terms and Conditions of Use. There are five rectangular buttons arranged in two rows: 'Subway Feeds' (selected), 'LIRR feeds', 'MNR feeds'; 'Service Alerts', 'Elevator and Escalator'. A 'Get Started' button is located at the bottom left.

<https://api.mta.info/#/landing>



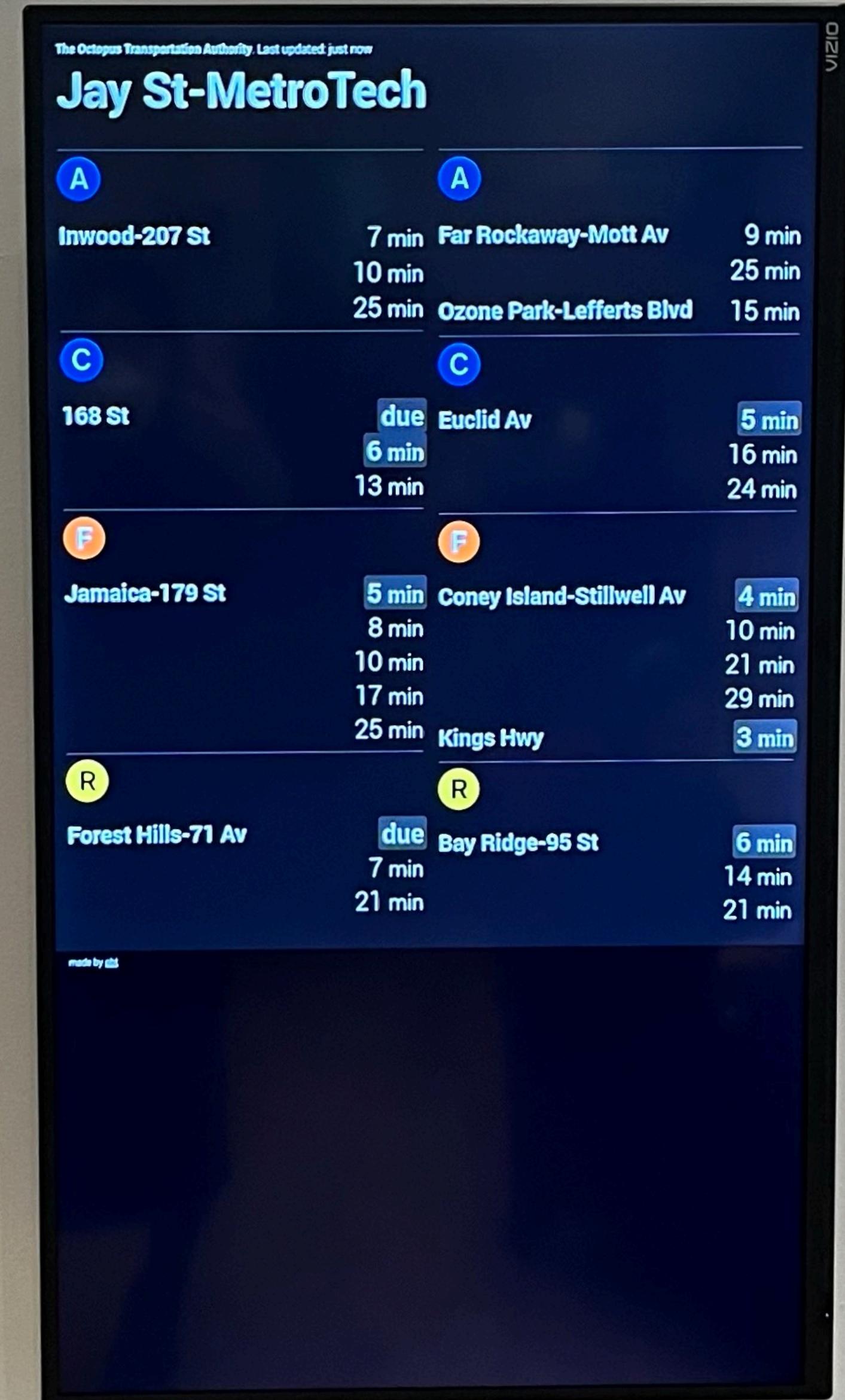
so i made this
it's a departure board/
countdown clock for the
MTA running on RCTV

<https://ota.pbt.dev/>



Goals

1. Learn about the MTA API
2. Make it readable
3. Make it run fast



What I learned about the MTA's API

1. This protobuf file has most of the documentation

```
// Realtime update for arrival and/or departure events for a given stop on a
// trip. Updates can be supplied for both past and future events.
// The producer is allowed, although not required, to drop past events.
message StopTimeUpdate {
    // The update is linked to a specific stop either through stop_sequence or
    // stop_id, so one of the fields below must necessarily be set.
    // See the documentation in TripDescriptor for more information.

    // Must be the same as in stop_times.txt in the corresponding GTFS feed.
    optional uint32 stop_sequence = 1;
    // Must be the same as in stops.txt in the corresponding GTFS feed.
    optional string stop_id = 4;

    optional StopTimeEvent arrival = 2;
    optional StopTimeEvent departure = 3;

    // The relation between this StopTime and the static schedule.
    enum ScheduleRelationship {
        // The vehicle is proceeding in accordance with its static schedule of
        // stops, although not necessarily according to the times of the schedule.
        // At least one of arrival and departure must be provided. If the schedule
        // for this stop contains both arrival and departure times then so must
        // this update.
        SCHEDULED = 0;

        // The stop is skipped, i.e., the vehicle will not stop at this stop.
        // Arrival and departure are optional.
        SKIPPED = 1;

        // No data is given for this stop. The main intention for this value is to
        // give the predictions only for part of a trip, i.e., if the last update
        // for a trip has a NO_DATA specifier, then StopTimes for the rest of the
        // stops in the trip are considered to be unspecified as well.
        // Neither arrival nor departure should be supplied.
        NO_DATA = 2;
    }

    optional ScheduleRelationship schedule_relationship = 5
        [default = SCHEDULED];

    // The extensions namespace allows 3rd-party developers to extend the
    // GTFS Realtime Specification in order to add and evaluate new features
    // and modifications to the spec.
    extensions 1000 to 1999;
}
```

Thankfully, Python bindings exist

```
✓ Trains arriving at Jay Street - MetroTech

[2] from IPython.display import display
    from nyct_gtfs import NYCTFeed
    from nyct_gtfs.gtfs_static_types import Stations
    from itertools import groupby

    from datetime import datetime, date, time, timezone, tzinfo, timedelta

    ApiKey = ...
    nyctime = timezone(-timedelta(hours=5))

    feedACE = NYCTFeed("A", api_key=ApiKey)
    feedF = NYCTFeed("F", api_key=ApiKey)
    feedR = NYCTFeed("R", api_key=ApiKey)
    stations = Stations()

    trains = []
    stops = ["A41N", "A41S", "R29N", "R29S"]

    for stop in stops:
        for feed in [feedACE, feedF, feedR]:
            trains += feed.filter_trips(headed_for_stop_id=stop, underway=True)

    arrivals = [[train.route_id, train.headsign_text if train.headsign_text else train.shape_id, [update.arrival for update in train.stop_time_updates if any
relative_arrivals = [[route, dest, arrival_time, (arrival_time - datetime.now()).seconds // 60] for [route, dest, arrival_time] in arrivals]

display(relative_arrivals)
[('A', 'Far Rockaway-Mott Av', datetime.datetime(2024, 1, 25, 20, 6, 42), 10),
 ('A', 'Ozone Park-Lefferts Blvd', datetime.datetime(2024, 1, 25, 20, 17, 20), 21),
 ('A', 'Ozone Park-Lefferts Blvd', datetime.datetime(2024, 1, 25, 20, 32, 20), 36),
 ('A', 'Far Rockaway-Mott Av', datetime.datetime(2024, 1, 25, 20, 21, 43), 25),
 ('A', 'Far Rockaway-Mott Av', datetime.datetime(2024, 1, 25, 20, 36, 3), 40),
 ('C', 'Euclid Av', datetime.datetime(2024, 1, 25, 19, 59, 52), 3),
 ('C', 'Euclid Av', datetime.datetime(2024, 1, 25, 20, 8, 50), 12),
 ('C', 'Euclid Av', datetime.datetime(2024, 1, 25, 20, 21, 50), 25),
 ('C', 'Euclid Av', datetime.datetime(2024, 1, 25, 20, 27, 20), 31),
 ('C', 'Euclid Av', datetime.datetime(2024, 1, 25, 20, 40, 30), 44),
 ('F', 'Coney Island-Stillwell Av', ...]
```

<https://github.com/Andrew-Dickinson/nyct-gtfs>



[https://flask.palletsprojects.com/
en/3.0.x/](https://flask.palletsprojects.com/en/3.0.x/)

What I learned about the MTA's API

2. There are a lot of feeds

```
@cachetools.func.ttl_cache(maxsize=2, ttl=15)
def get_arrivals(stations):
    print(str(arrow.utcnow()), "getting arrivals")

    feeds = (
        NYCTFeed("1", api_key=ApiKey),
        NYCTFeed("A", api_key=ApiKey),
        NYCTFeed("F", api_key=ApiKey),
        NYCTFeed("G", api_key=ApiKey),
        NYCTFeed("J", api_key=ApiKey),
        NYCTFeed("L", api_key=ApiKey),
        NYCTFeed("N", api_key=ApiKey),
    )
```

What I learned about the MTA's API

3. A "stop" is
just part of a
station

A		A	
Inwood-207 St	5 min	Far Rockaway-	6 min
	7 min	Mott Av	20 min
	23 min	Ozone Park-	14 min
C		Lefferts Blvd	
168 St	5 min	C	
	12 min	Euclid Av	7 min
	19 min		18 min
F			26 min
Jamaica-179 St	8 min	F	
	17 min	Coney Island-	3 min
	26 min	Stillwell Av	15 min
			17 min
			24 min

A41

R		R	
Forest Hills-71 Av	2 min	Bay Ridge-95 St	2 min
	15 min		12 min
	26 min		21 min
			25 min

R29

+

=



What I learned about the MTA's API

4. If the train
hasn't actually
left, it doesn't
exist*

The Octopus Transportation Authority. Last updated: just now

Myrtle-Wyckoff Avs

L

8 Av

L

3 min Canarsie-Rockaway
7 min Pkwy

2 min
5 min

13 min
18 min
23 min

11 min
17 min
22 min

M

57 St

M

due Middle Village-
Metropolitan Av

7 min
18 min
27 min

made by pbt

*in this feed anyway



What I learned about the MTA's API

5. There are exactly two (2) directions



Hoyt-Schermerhorn St



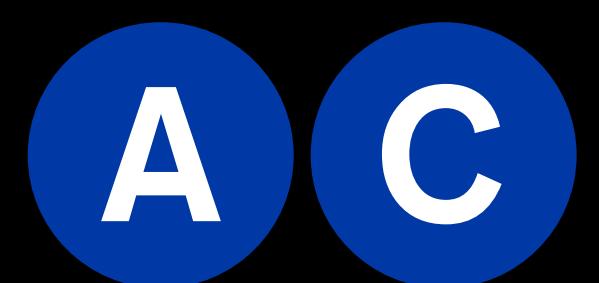
Manhattan



Church Av (Brooklyn)



Court Sq (Queens)



Queens



Hoyt-Schermerhorn St



Manhattan



N



Church Av (Brooklyn)



S



Court Sq (Queens)



N



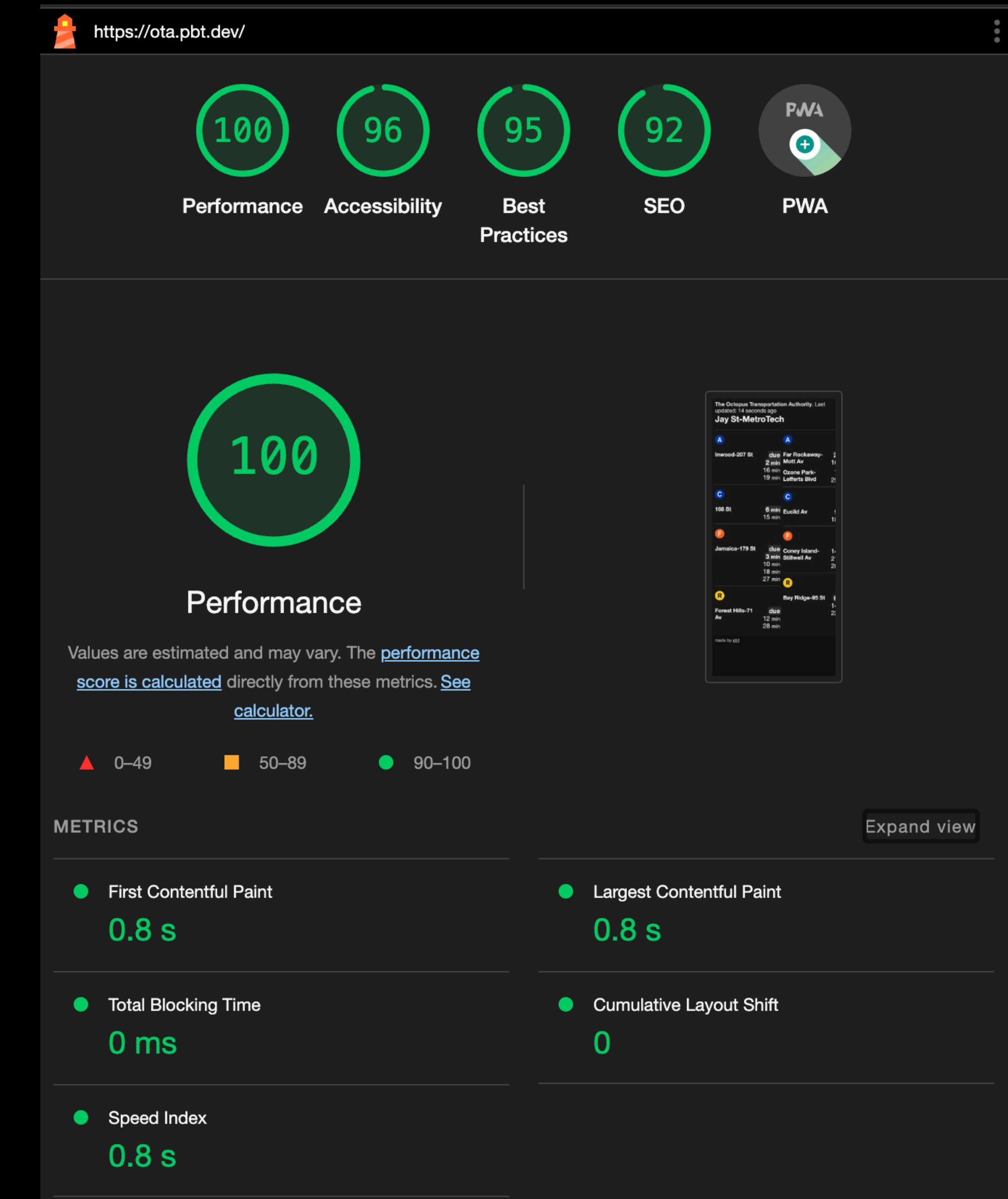
Queens

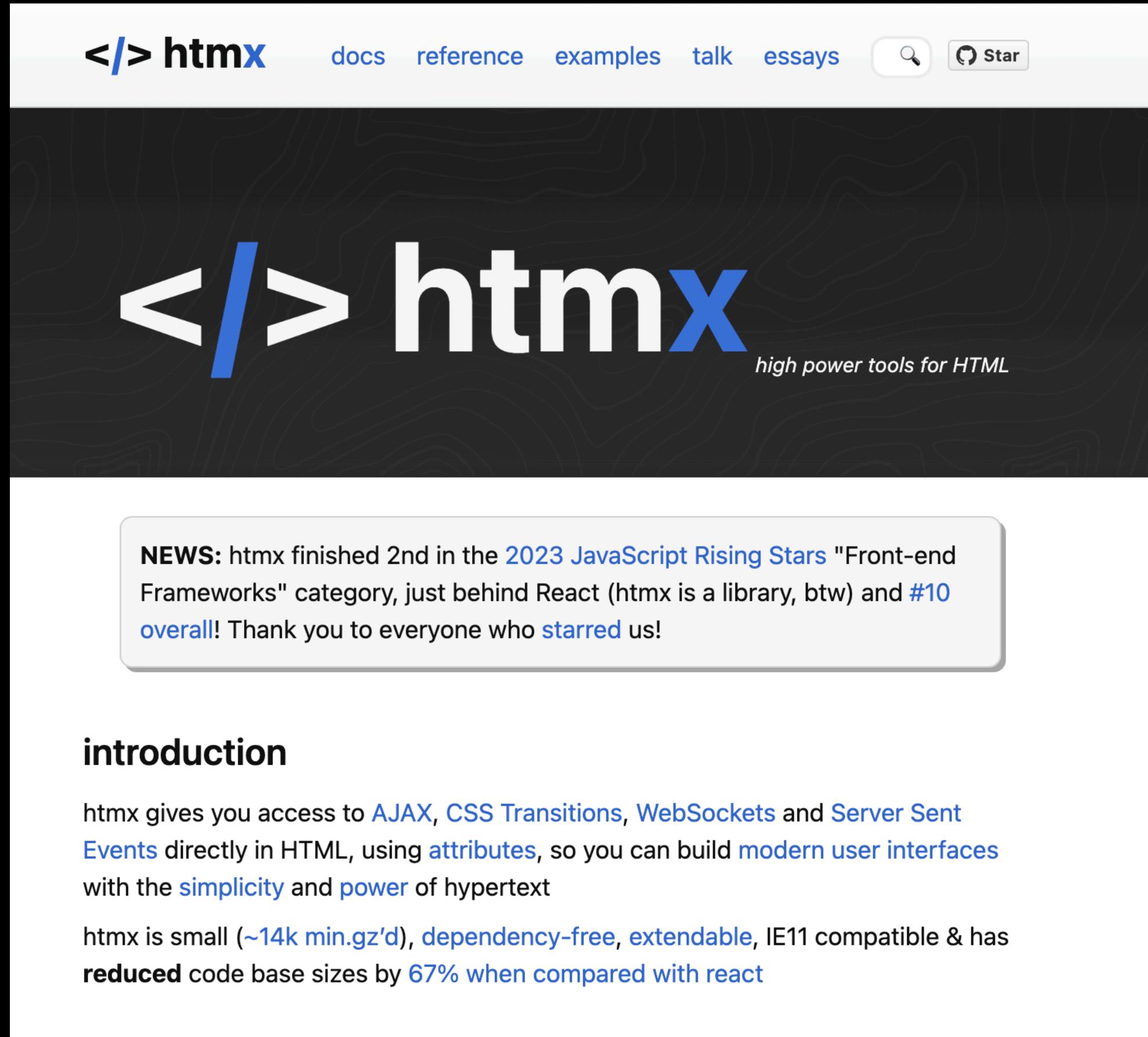


S

making it run fast*

* if it's warmed up...





htmx.org

```
  
```

```
</header>
<main hx-get="{{url_for('countdown', mode=mode, stop_ids=stop_ids)}}> hx-trigger="every 15s" hx-swap="morph:outerHTML"
  | hx-select="main > *">
```

Future Work

1. Making it cache better
2. Adding a UI to navigate stations? Trains?
3. Adding support for Alerts
4. Schedules
5. What do you want to see?



help me with...

1. Ideas
2. Pairing

tysm <3

Matthew Long (W'14)

For giving me a primer
to the MTA data

**Countless, countless
people I've talked to in
the hub**

If you'd like to talk to
me, say hi :)

shoutout to:

<https://wheresthefuckingtrain.com/>

