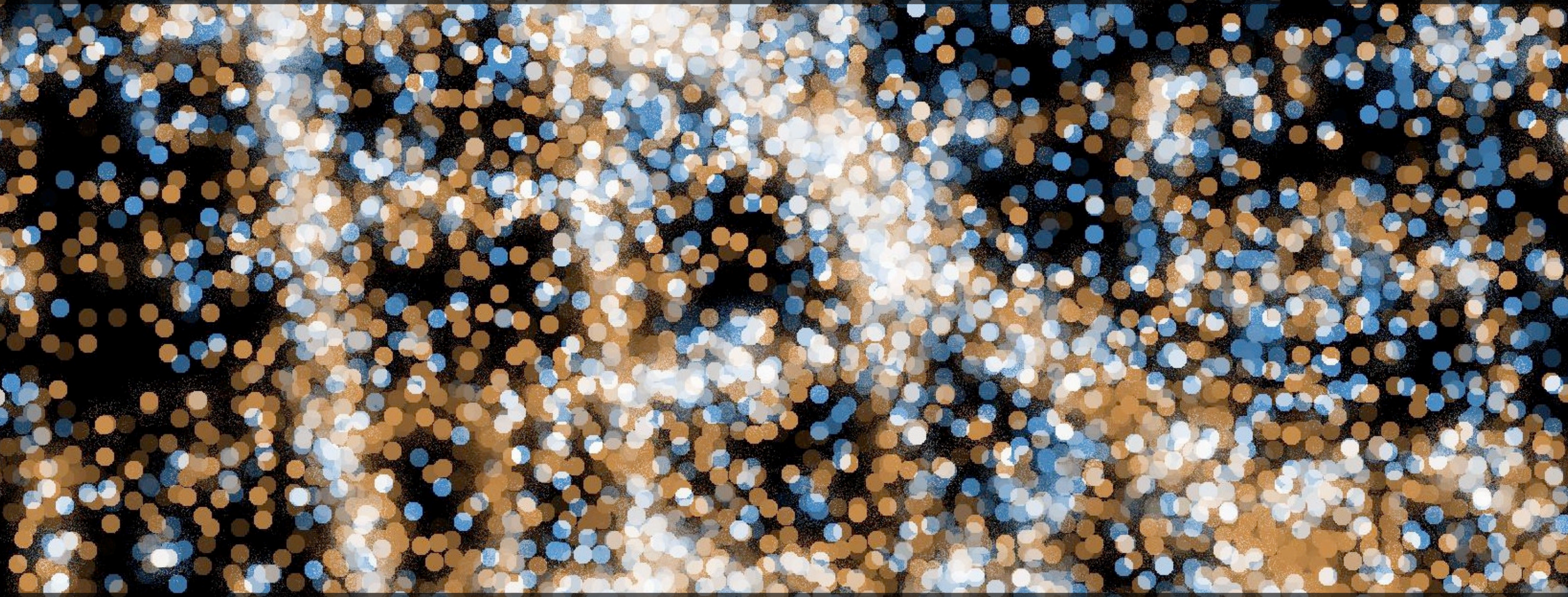
Mapping millions of tweets



Eric Fischer @enf Mapbox

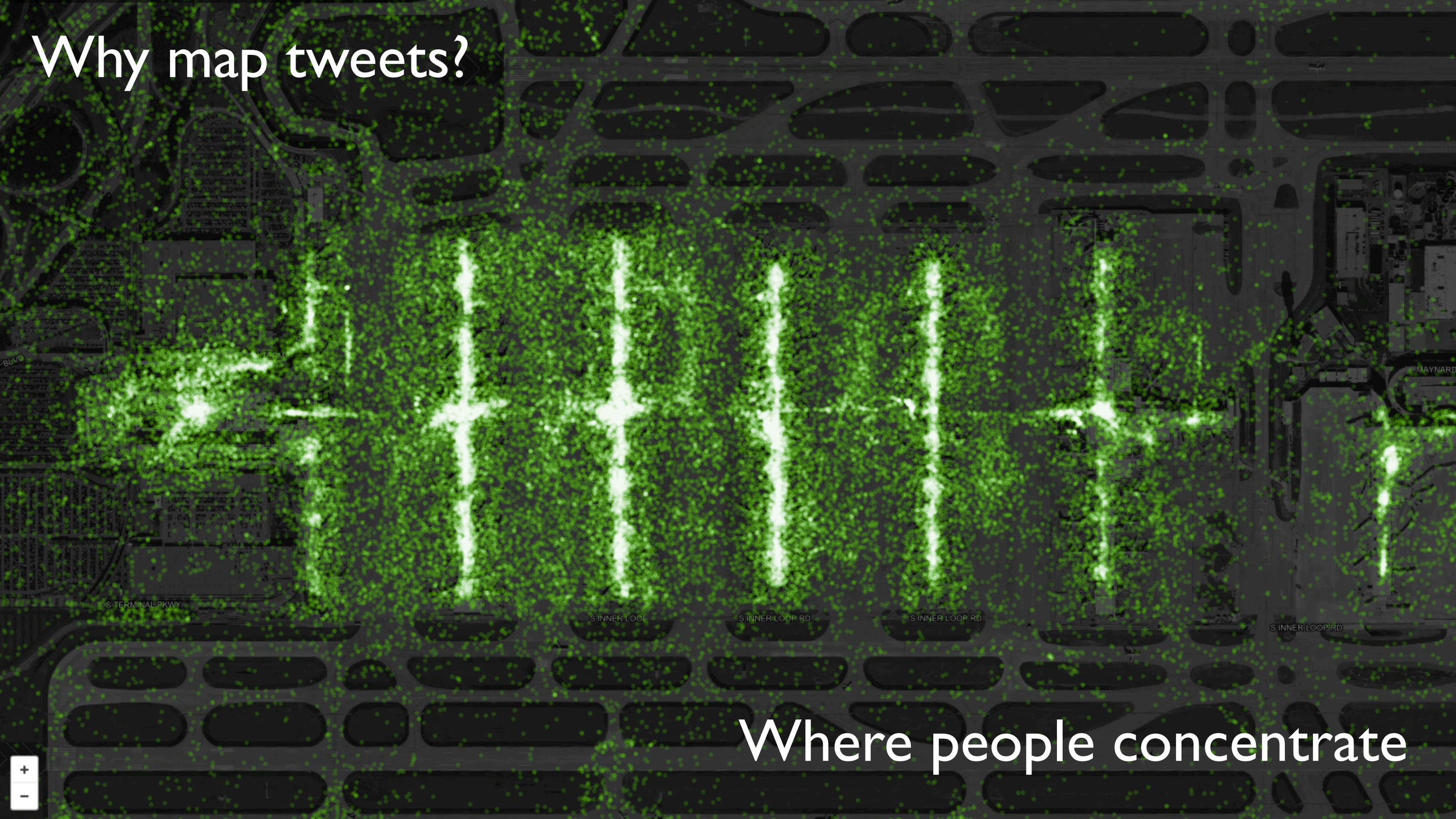
Why map tweets?

Where people travel

Why map tweets?

Who communicates with whom

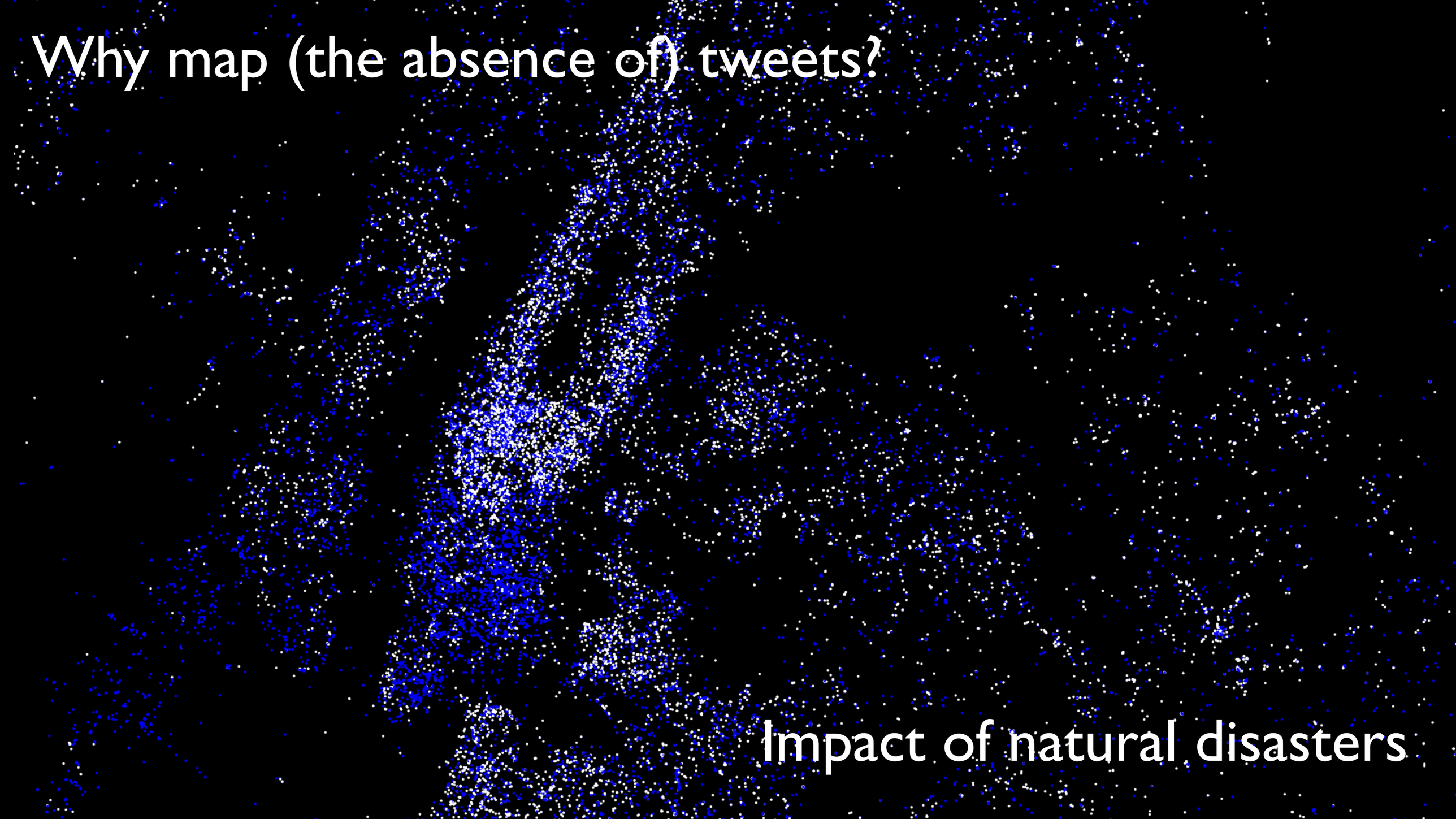
Why map tweets?



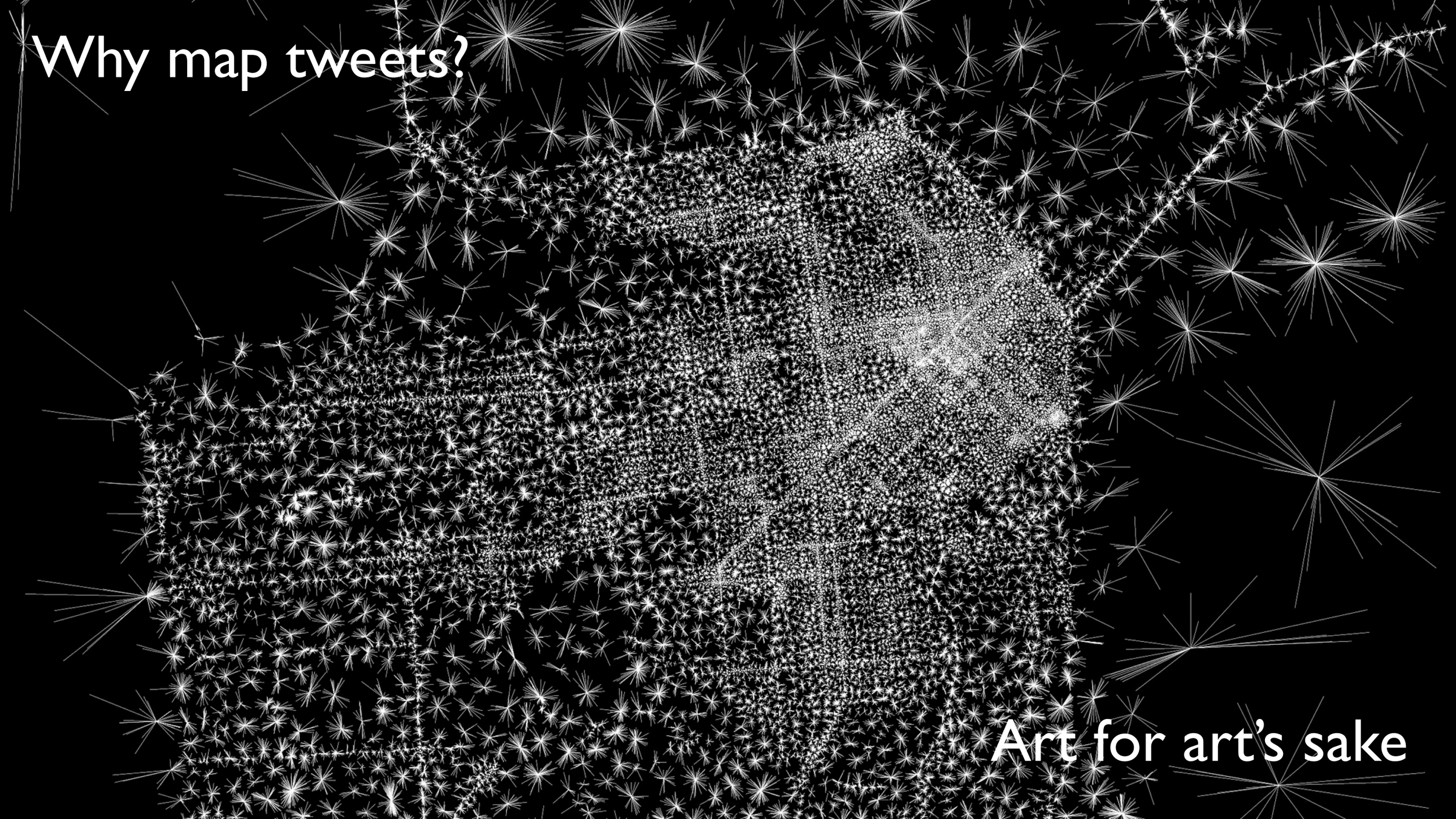
Where people concentrate


FERGUSON, MO TWEETS
Local #ferguson tweets
Visitor #ferguson tweets
All Tweets
LANG DR
COPPER CREEK RD
CANFIELD DR
CLARION DR
ELLISON DR
FORGE DR
GADDIEFIELD RD
STONEFIELD RD
SHARONDALE CIR
FERGUSON AVE
Insight into news

Why map (the absence of) tweets?

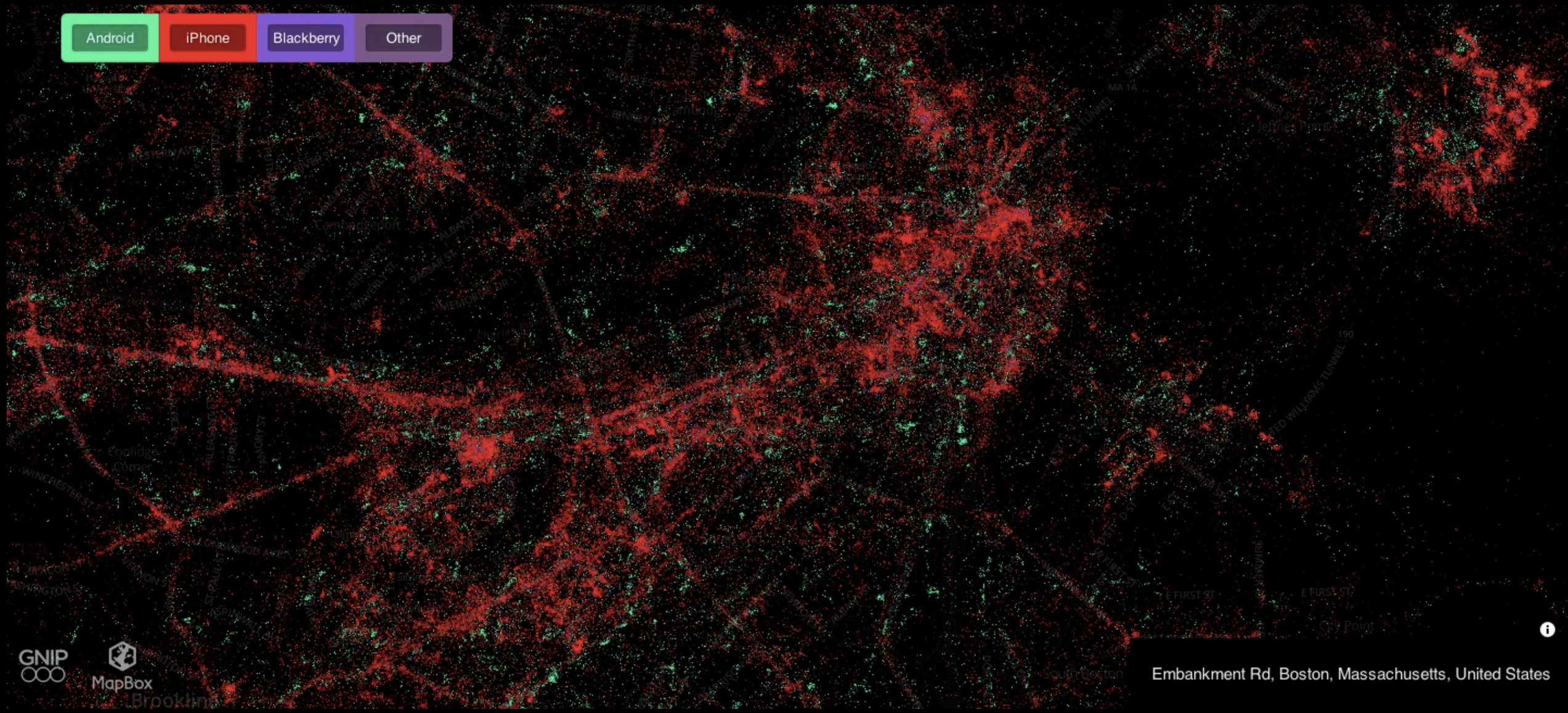


Why map tweets?



Art for art's sake

Mapping Twitter's archives with Gnip



Locals and Tourists, Languages, Phone brands

Public tweets from the “filter” stream

```
curl --verbose --compress --silent
 --header ‘Authorization: OAuth …’
 ’https://stream.twitter.com/1.1/statuses/
 filter.json?locations=-180,-90,180,90’
```

```
{"created_at": … }
{"created_at":"Mon Sep 07 21:44:21 +0000 2015",
 "text":"@mwichary Oh, because it's optically condensed.",
 "source":"… Twitter for Android …",
 "user":{"name":"Eric Fischer","screen_name":"enf"},
 "coordinates":{"coordinates":[-122.251561,37.826818]}},
 "lang":"en"}
{"created_at": … }
```

In-reply-to

Date and time

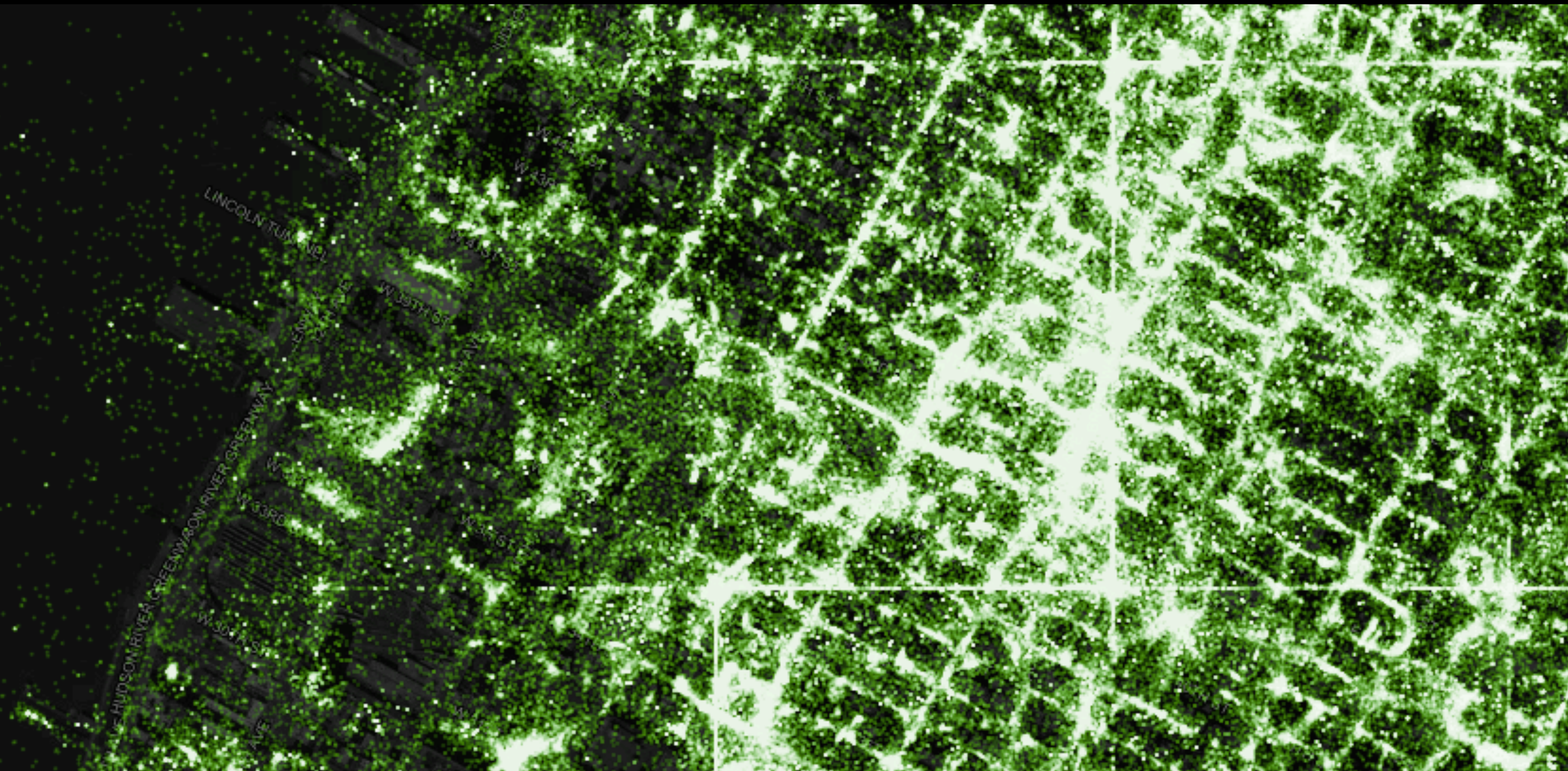
Phone type

Identity

Location

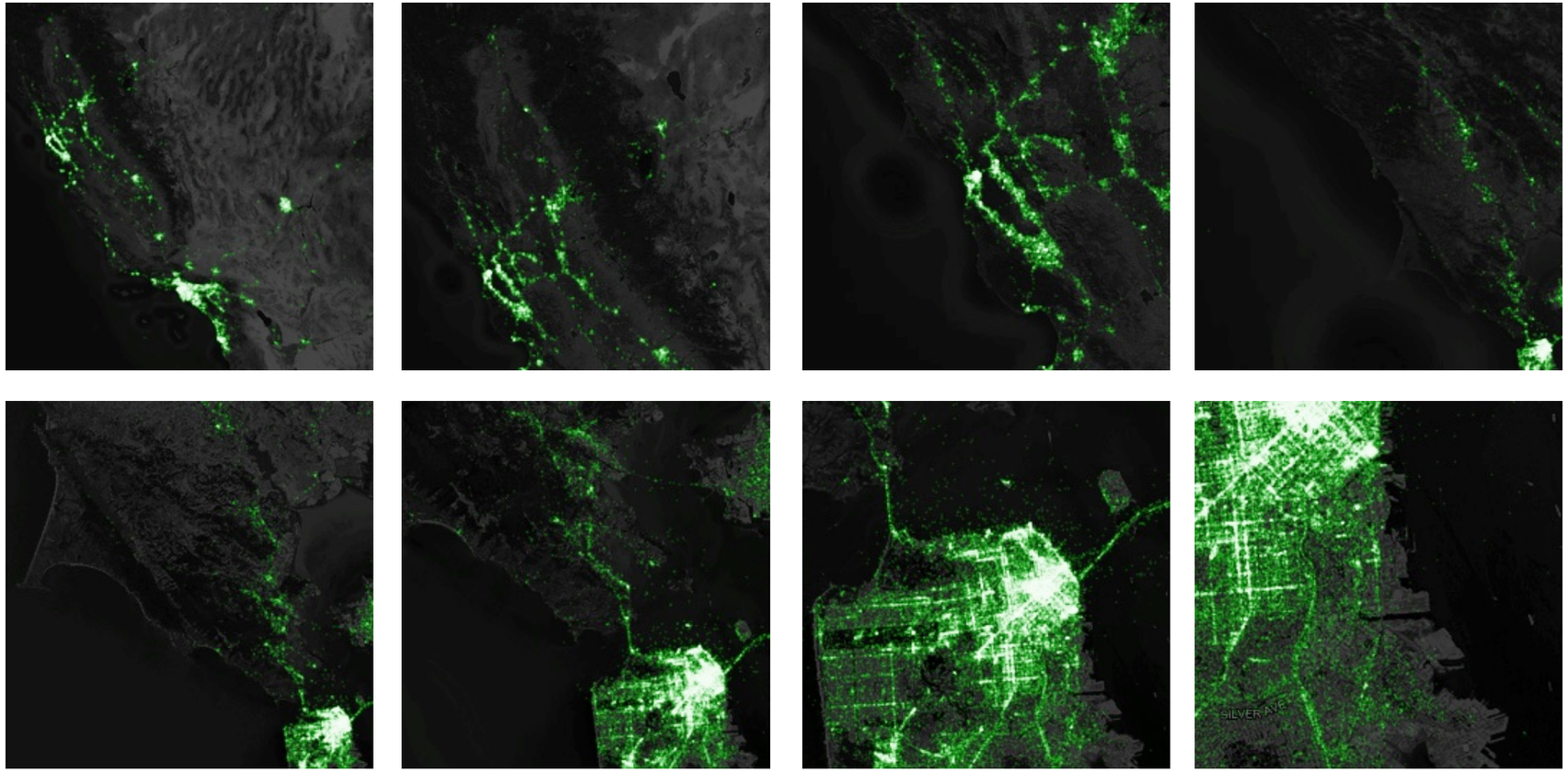
Language

Debanding and despeckling



Vector tiles

At zoom level *n*, each tile contains instructions to draw $1/4^n$ of the Earth



Making vector tiles with Tippecanoe

mapbox/tippecanoe

GitHub, Inc. [US] https://github.com/mapbox/tippecanoe

tile.h	Change some defaults that keep tripping people up:	11 days ago
vector_tile.proto	Add vector tile boilerplate	a year ago
version.h	Add tippecanoe -v	3 months ago

README.md

tippecanoe

Builds vector tilesets from large collections of GeoJSON features. This is a tool for making maps from huge datasets.

Intent

The goal of Tippecanoe is to enable making a scale-independent view of your data, so that at any level from the entire world to a single building, you can see the density and texture of the data rather than a simplification from dropping supposedly unimportant features or clustering or aggregating them.

If you give it all of OpenStreetMap and zoom out, it should give you back something that looks like "All Streets" rather than something that looks like an Interstate road atlas.

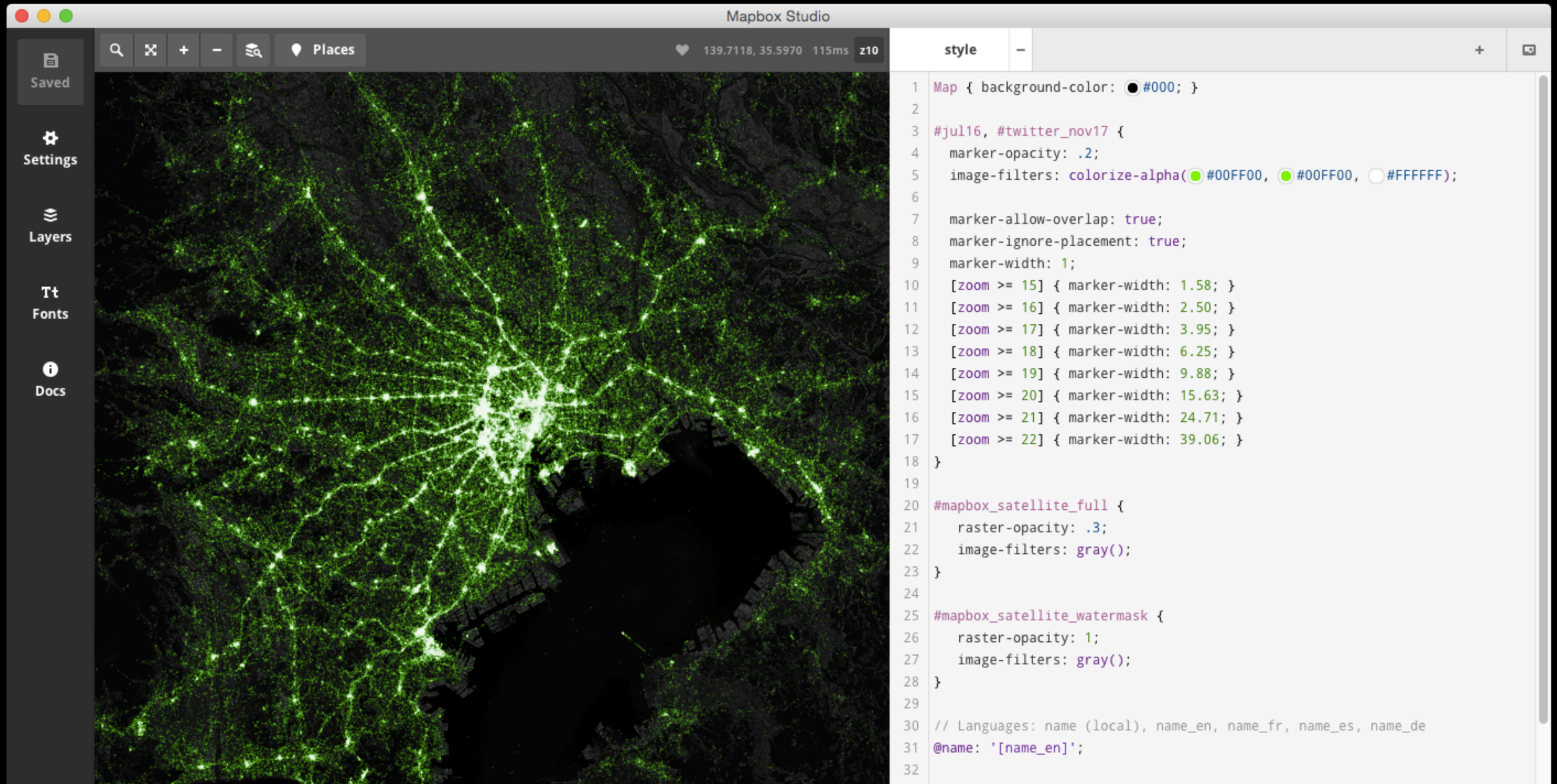
If you give it all the building footprints in Los Angeles and zoom out far enough that most individual buildings are no longer discernable, you should still be able to see the extent and variety of development in every neighborhood, not just the largest downtown buildings.

If you give it a collection of years of tweet locations, you should be able to see the shape and relative popularity of every point of interest and every significant travel corridor.

Key Tippecanoe features

- Exponential reduction of dot density at low zoom levels
- Zoom level-sensitive tile detail
- Optional thinning of near-duplicate features

Styling the data in Mapbox Studio Classic

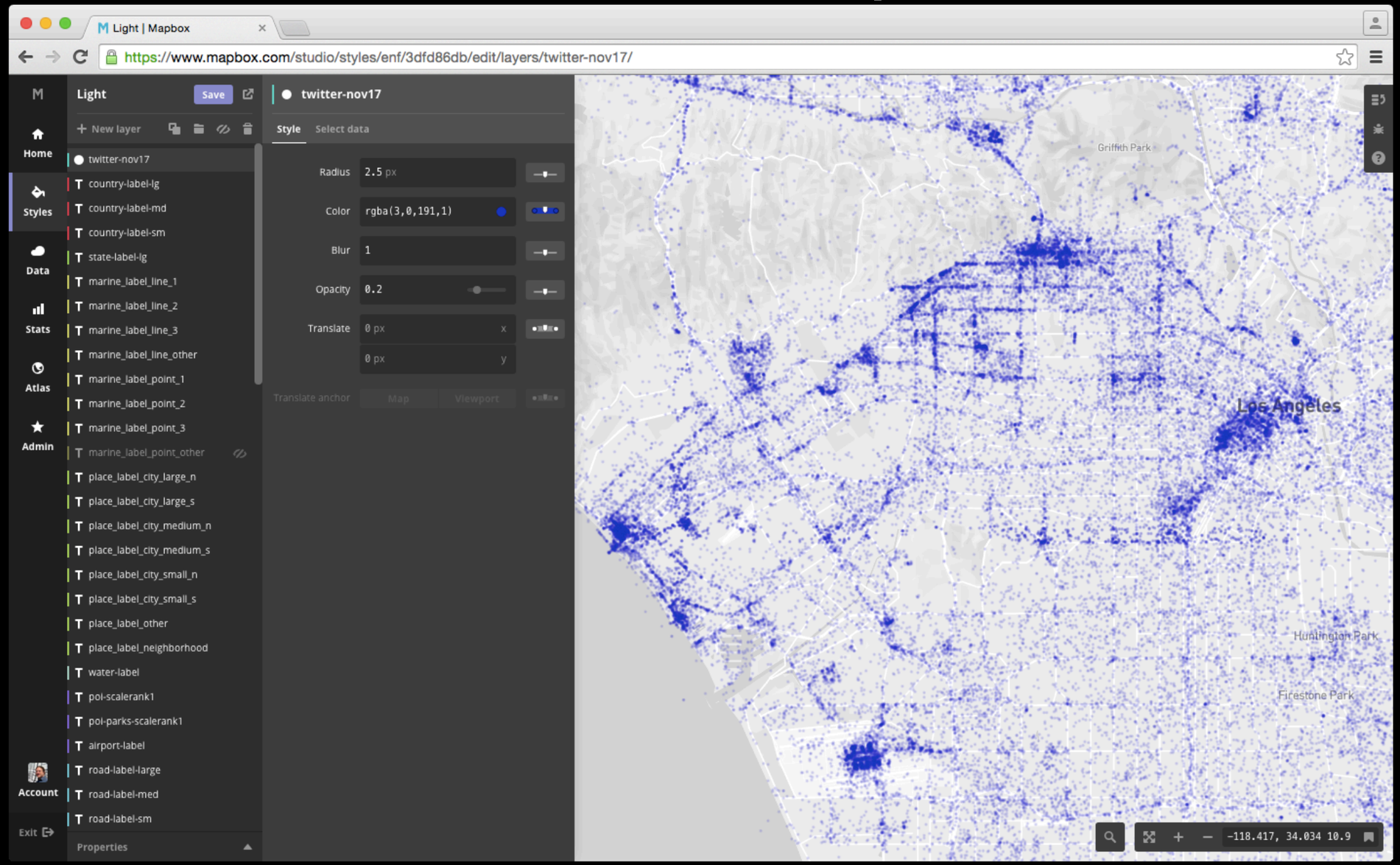


Vectors on the server, bitmaps to the browser

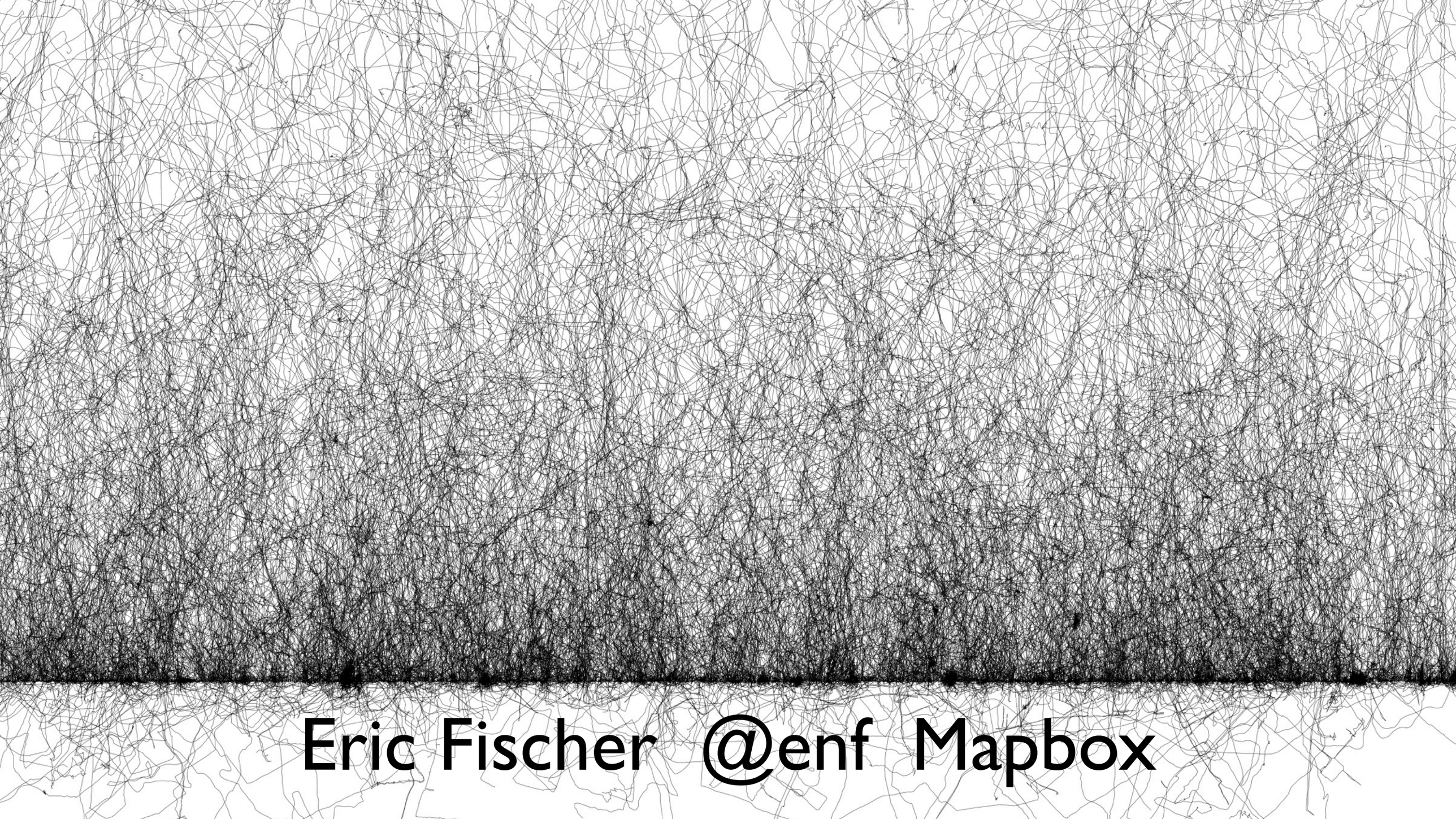
- Styling is only semi-dynamic
- Scales are only integer zoom levels
- No client-side analysis is possible

But ultimately, data is better than pictures of data

Live data in the browser with Mapbox Studio and GL JS



Dynamic scale and style on the client



Eric Fischer @enf Mapbox