



Image/Map Credit: Stamen Design

OPEN DATA: WHY AND WHERE

@abhinemani | CAPP 30130 | University of Chicago



CIVIC TECH: KEYS TO SUCCESS

Empathy

- ▶ Build with, not for
- ▶ Talk to users
- ▶ Understand behavior
- ▶ Learn what's out there

Lean

- ▶ It's more than just a technical process
- ▶ Expect failure
- ▶ Only build what's necessary
- ▶ Test, test, test

Data

- ▶ Data "lives" everywhere
- ▶ Be ready to scrape
- ▶ Open data is your friend
- ▶ ETLs are your best friend

Users

- ▶ ~~If you build it they will come~~
- ▶ Listen and learn
- ▶ Define metrics
- ▶ Find partners

Sustainability

- ▶ Consider your options: startup, open source, non-profit, academic, etc
- ▶ Tech -> Policy
- ▶ Expect to train everyone
- ▶ Document everything

CIVIC TECH: KEYS TO SUCCESS

Empathy

- ▶ Build with, not for
- ▶ Talk to users
- ▶ Understand behavior
- ▶ Learn what's out there

Lean

- ▶ It's more than just a technical process
- ▶ Expect failure
- ▶ Only build what's necessary
- ▶ Test, test, test

Data

- ▶ Data "lives" everywhere
- ▶ Be ready to scrape
- ▶ Open data is your friend
- ▶ ETLs are your best friend

Users

- ▶ ~~If you build it they will come~~
- ▶ Listen and learn
- ▶ Define metrics
- ▶ Find partners

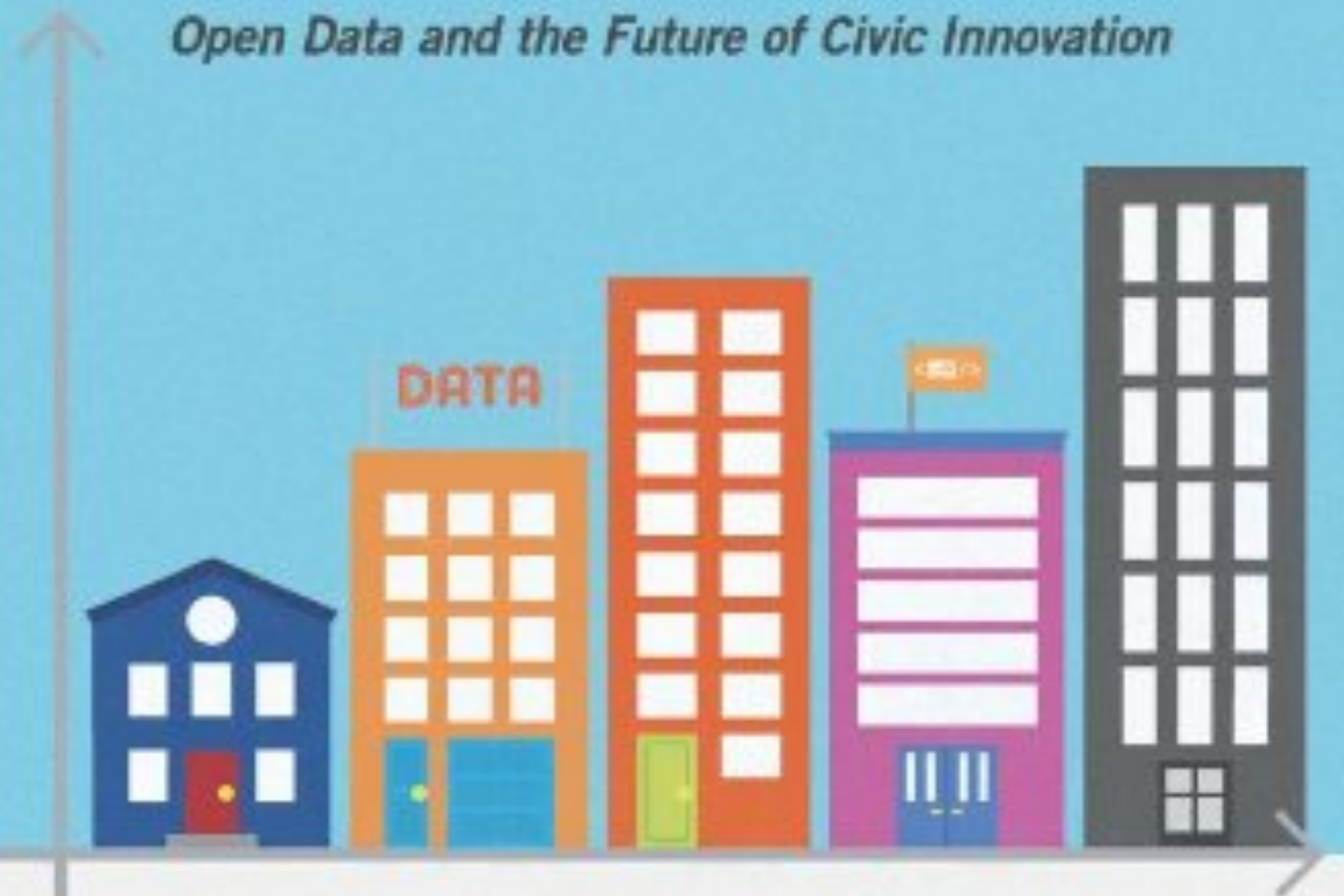
Sustainability

- ▶ Consider your options: startup, open source, non-profit, academic, etc
- ▶ Tech -> Policy
- ▶ Expect to train everyone
- ▶ Document everything

Edited by Brett Goldstein with Lauren Dyson

Beyond Transparency

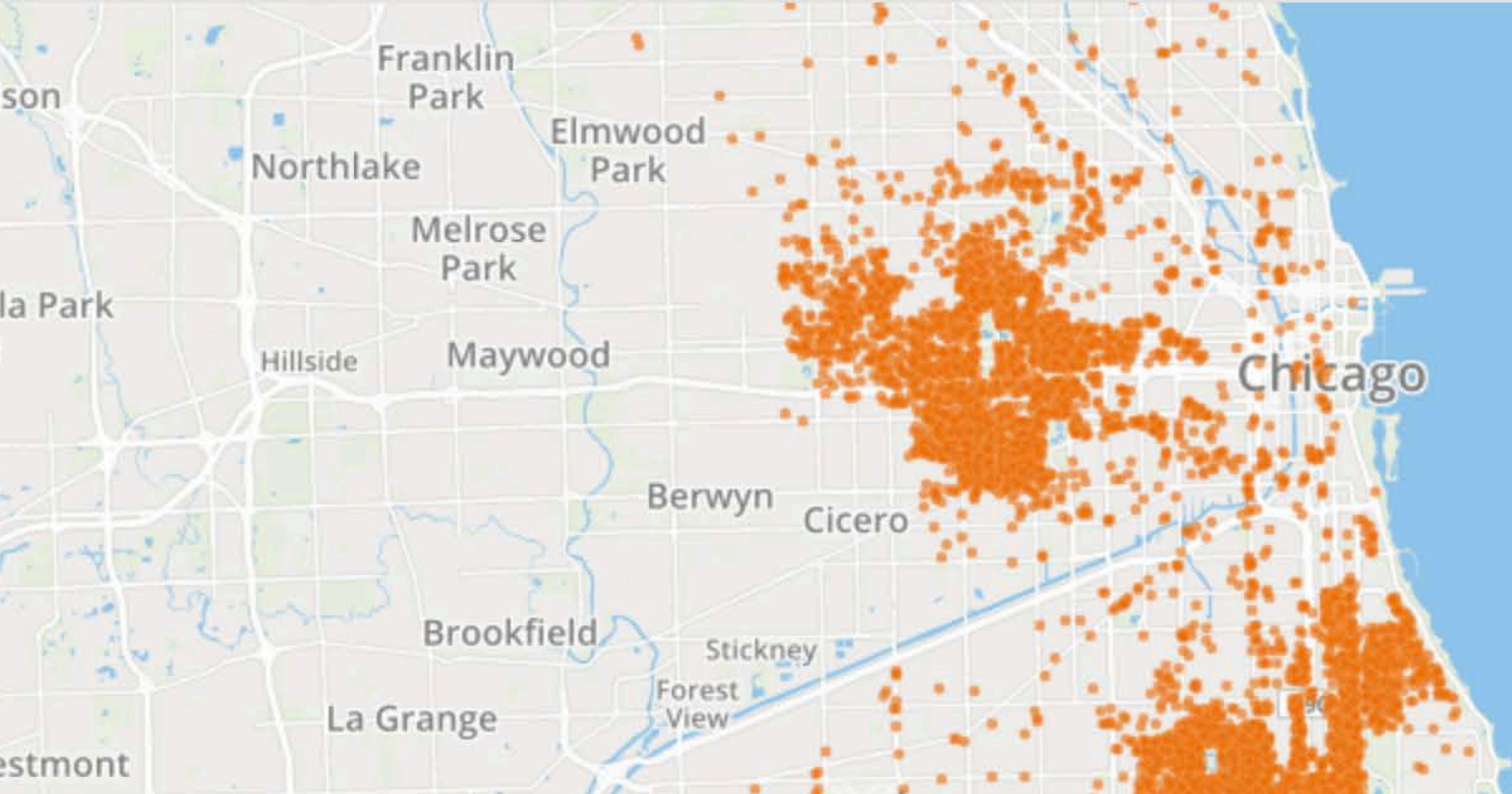
Open Data and the Future of Civic Innovation



"A must-read for anyone who is passionate about what open data can do to transform city living."
—Boris Johnson, Mayor of London

AGENDA

- Case Study: LargeLots
- Why open data?
 - Benefits to the government
- Where to find open data?
 - Ways residents can access governmental data



@abhinemani

CASE STUDY: LARGE LOTS

Open data helping deliver policy to people

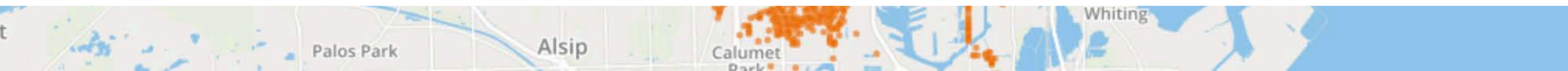


Image credit: ChicagoMag



City of Chicago designed a program to sell unused city-owned property for \$1, per lot in key areas where redevelopment was a priority.

@abhinemani



You could figure it out on your own, but you have to download this spreadsheet from the city data portal with every city owned lot, compare it to this pdf map and then go to the city zoning map... All the pieces are there, but having to go through the steps is a challenge.

- Derek Eder

LARGE LOTS: OPEN DATA CASE STUDY

- Policy: Neighborhood stabilization through sale of city-owned parcels
- Problem: Difficult to find which properties available and apply for purchase
- Approach: Civic technology organization (DataMade) used open data to build LargeLots.org
- Datasets
 - Chicago - City Owned Land Inventory
 - Chicago - Community Areas
 - Cook County - Address Point Chicago
 - Cook County - 2012 Parcels
- Outcomes
 - Hundreds of parcels sold
 - Program expanded across the city



The first challenge was to get all the data in place... the team needed a map of the area, a list of city-owned property and zoning information for Chicago. **Luckily most of the sets (city-owned property, zoning maps, Cook County parcel maps) were already available.**

-WBEZ



City-Owned Land Inventory

View Data	Visualize ▾	Export	API	...
------------------	-------------	--------	-----	-----

Community & Economic Development

Property owned and managed by the City of Chicago. Information provided in the database, or on the City's website generally, should not be used as a substitute for title research, title evidence, title insurance, real estate tax exemption or payment status, environmental or geotechnical due diligence, or as a substitute for legal, accounting, real estate, business, tax or other professional advice. The City assumes no liability for any damages or loss of any kind that might arise from the reliance upon, use of, misuse of, or the inability to use the database or...

[More](#)

Updated
April 18, 2019

Data Provided by
Chicago Department of Planning and Development

About this Dataset

Updated April 18, 2019		Metadata	
Data Last Updated April 18, 2019	Metadata Last Updated February 20, 2019	Data Owner Chicago Department of Planning and Development	
Date Created January 7, 2013		Topics	
		Category Community & Economic Development	
		Tags This dataset does not have any tags.	

@abhinemani

The City of Chicago sells vacant residential lots for **\$1** through the Large Lot Program.


Total lots sold

1248

For information on the next Large Lots application round, please check back in early 2019.

Explore the map below of lots sold, available, and under review.

Enter your address Sold Available Sale Pending



Address	315 W 73RD ST
PIN	20-28-217-001-0000 (info)
Community	GREATER GRAND CROSSING
Ward	6
Status	Sold


 [@abhinemani](#)

Image credit: ChicagoMag



Over 5,000 visitors in the first 2 weeks.



After debuting in Englewood and East Garfield Park in 2014, more than 550 homeowners have so far taken advantage of the program. Now, thanks to its recently expanded scope, **Large Lots will extend to 33 Chicago communities on the West and South sides, offering 4,000 empty properties at the extremely discounted rate.**

-Chicago Tribune



@abhinemani

WHY OPEN DATA

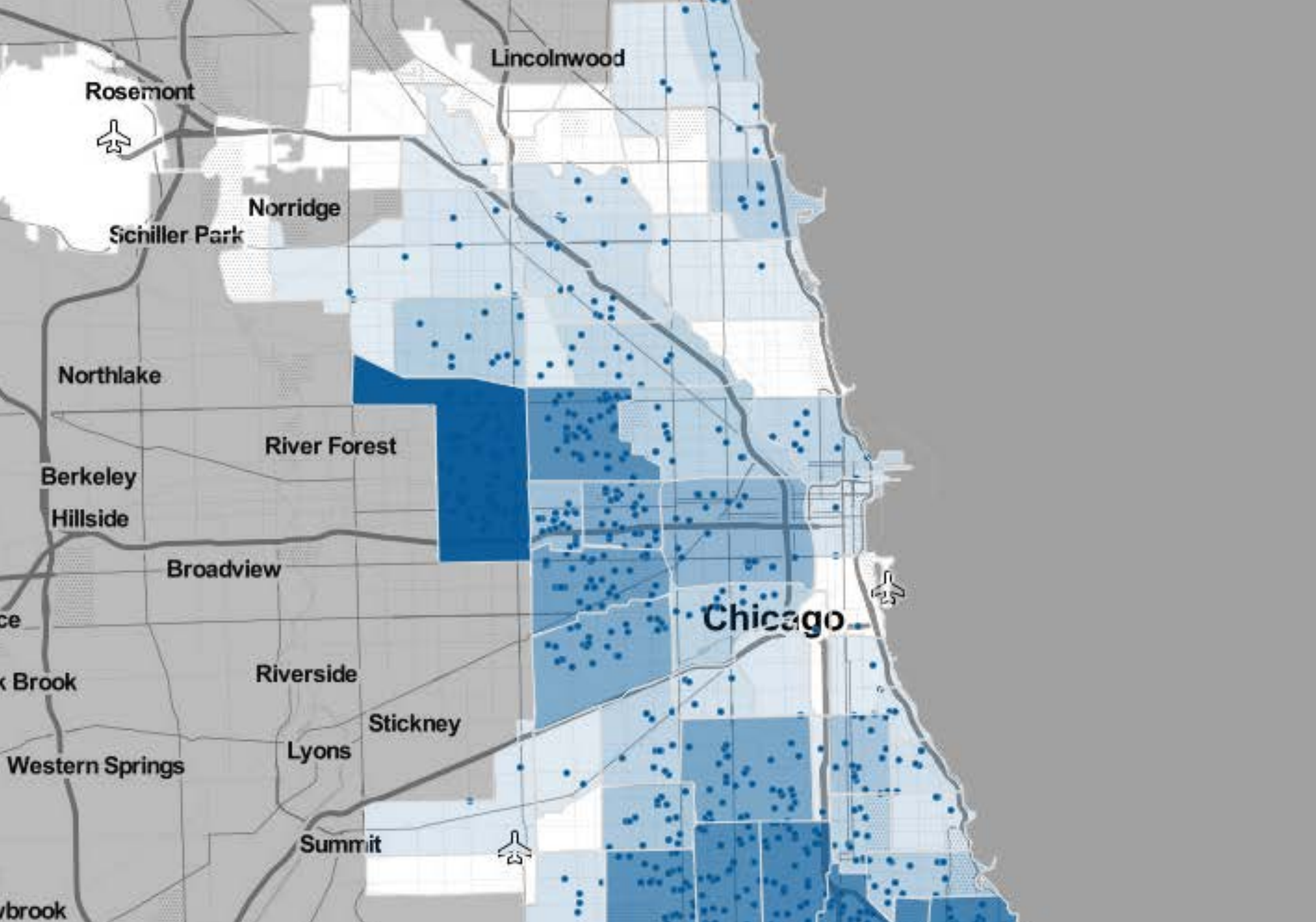
The benefits of opening civic / government data



BENEFITS OF OPEN DATA

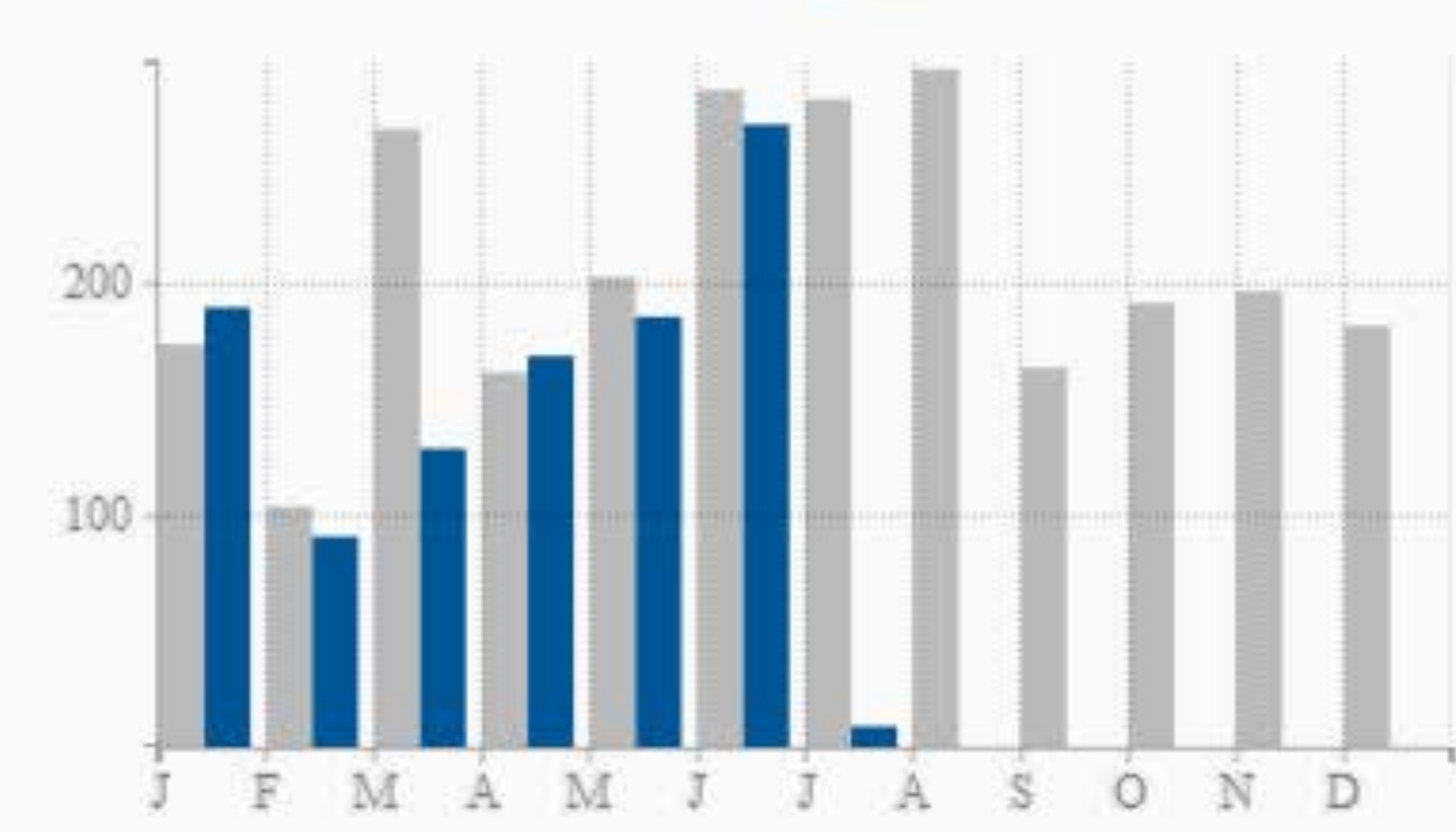
- Transparency
 - Increased accountability (eg Crime maps)
- Internal efficiency
 - Department-to-department communication (eg LA GeoHub)
- Better decision-making
 - Data science / analysis (eg VisionZero)
- Community-supported innovation / engagement
 - Better gov/citizen communication (eg ClearStreets, transit apps, etc)
- Civic entrepreneurship
 - New businesses / jobs (eg BrightScope, DataMade)





1,043 shooting victims in 2013

VICTIMS BY MONTH • 2013 • 2012



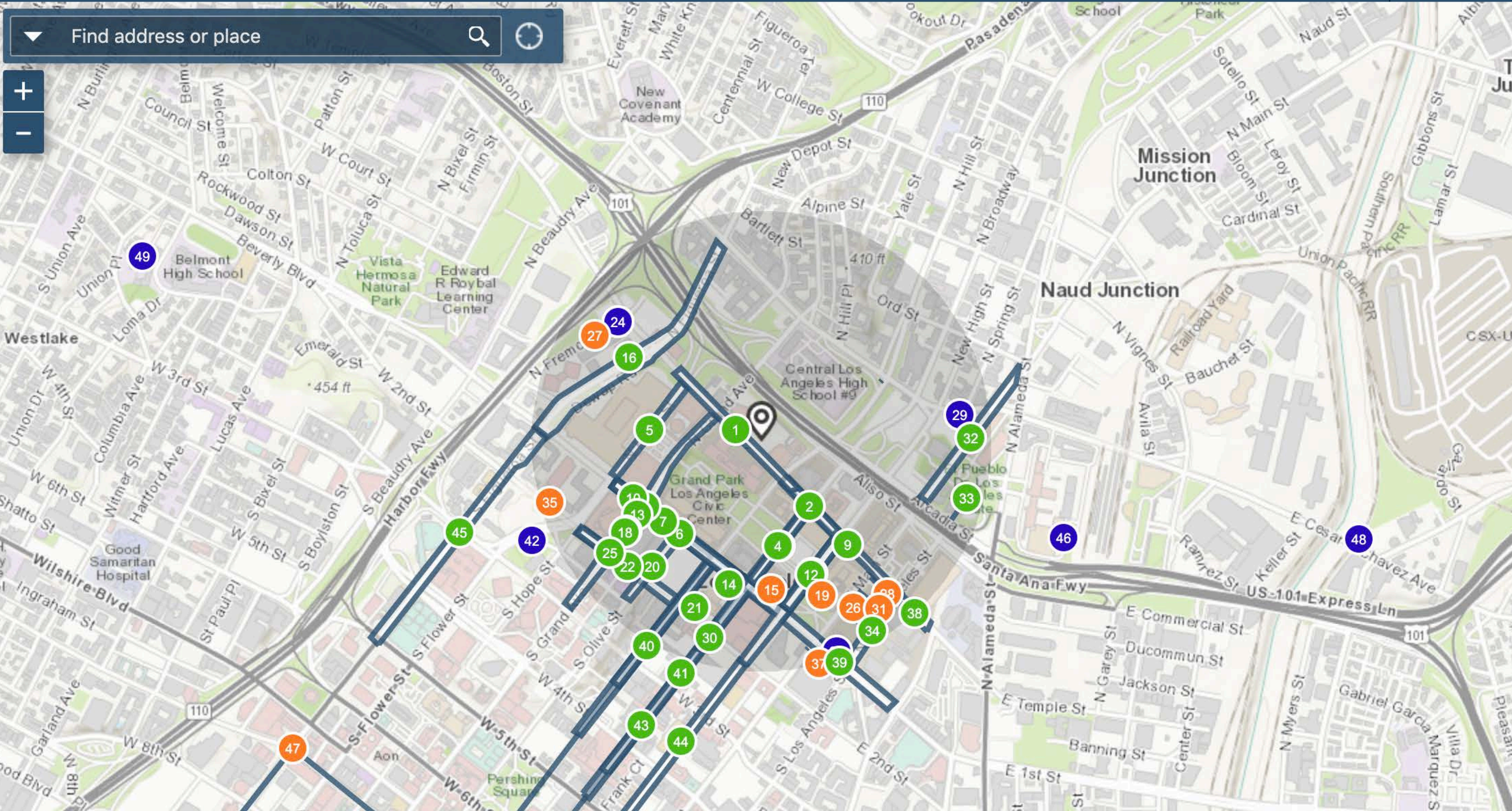
MOST RECENT VICTIMS LAST 30 DAYS

- Male, 23, on July 2 in Washington Heights
- Male, 15, on July 1 in Chicago Lawn
- Male, 22, on July 1 in Washington Park
- Male, 18, on July 1 in Washington Heights

Image credit: Chicago Tribune

Transparency: Using open data journalists, NGOs, and other groups can track and visualize city activity — crime map — boosting accountability.

Find address or place



63
Projects

Search Radius (0.5 mi)



Earliest Date: Apr 21, 2019

Latest Date: Oct 21, 2019

1	Special Events	0.06 MILES	Starts In: 5 day
2	Special Events	0.17 MILES	Starts In: 3 week
3	Special Events	0.23 MILES	Starts In: 5 day
4	Special Events	0.23 MILES	Starts In: 3 week
5	Special Events	0.24 MILES	Starts In: 2 day

Internal Efficiency: LA GeoHub, hosts open data, and powers apps like StreetWize that helps agencies coordinate (and communicate) street closures

@abhinemani

e, coated with bicycle collisions

Bicycle Collisions per Street, 2010 to 2015



VERMONT AV

VENICE BL

FIGUEROA ST

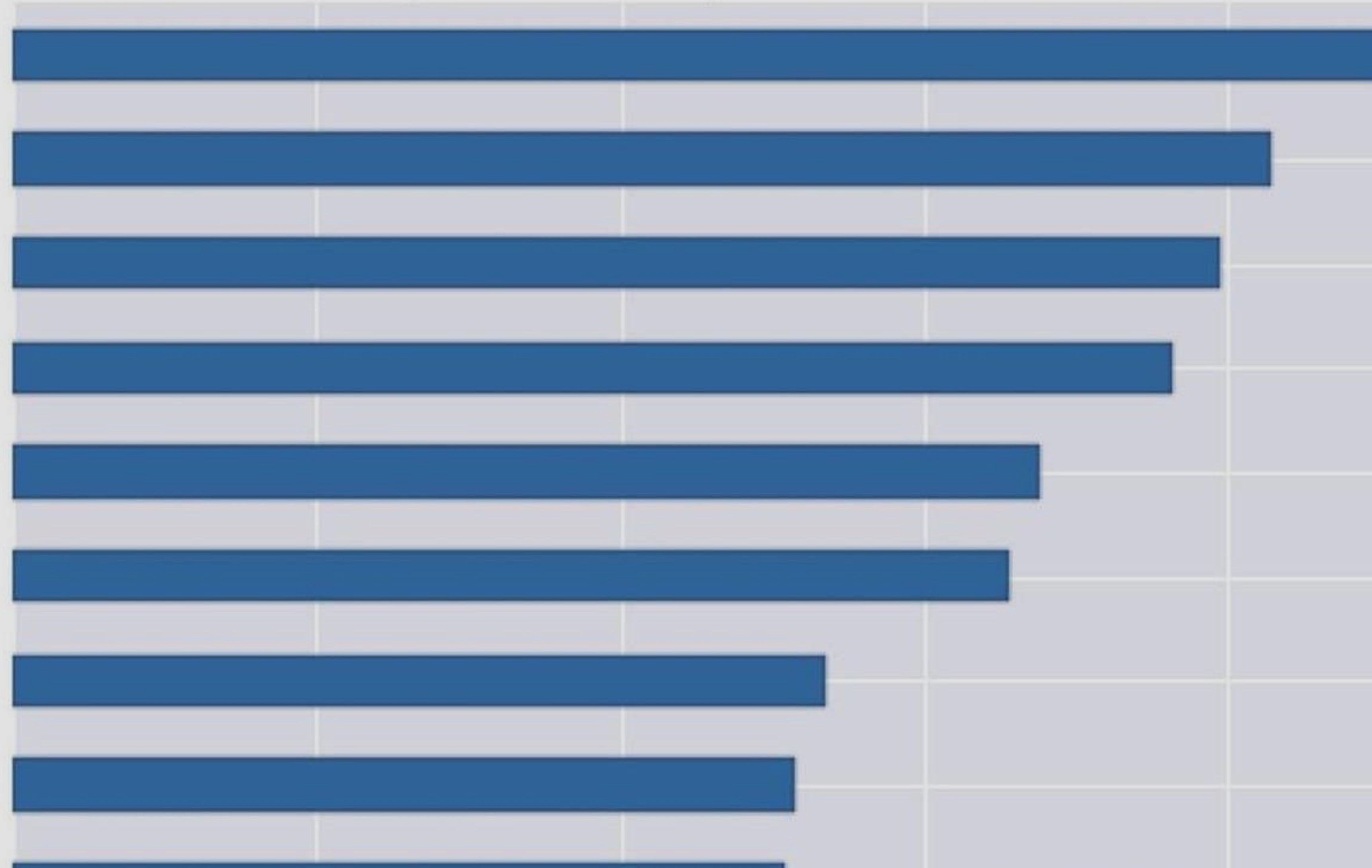
WESTERN AV

SUNSET BL

PICO BL

OLYMPIC BL

VAN NUYS BL



Better decision-making: Open data enables NGOs, universities and companies to perform data science/analysis to support policy initiatives such as VisionZero.

Jan 18, 2019 The Chicago area is expected to get [between 6-10 inches of snow](#).

See which streets have been plowed and when.

Did my street get plowed? (find me)

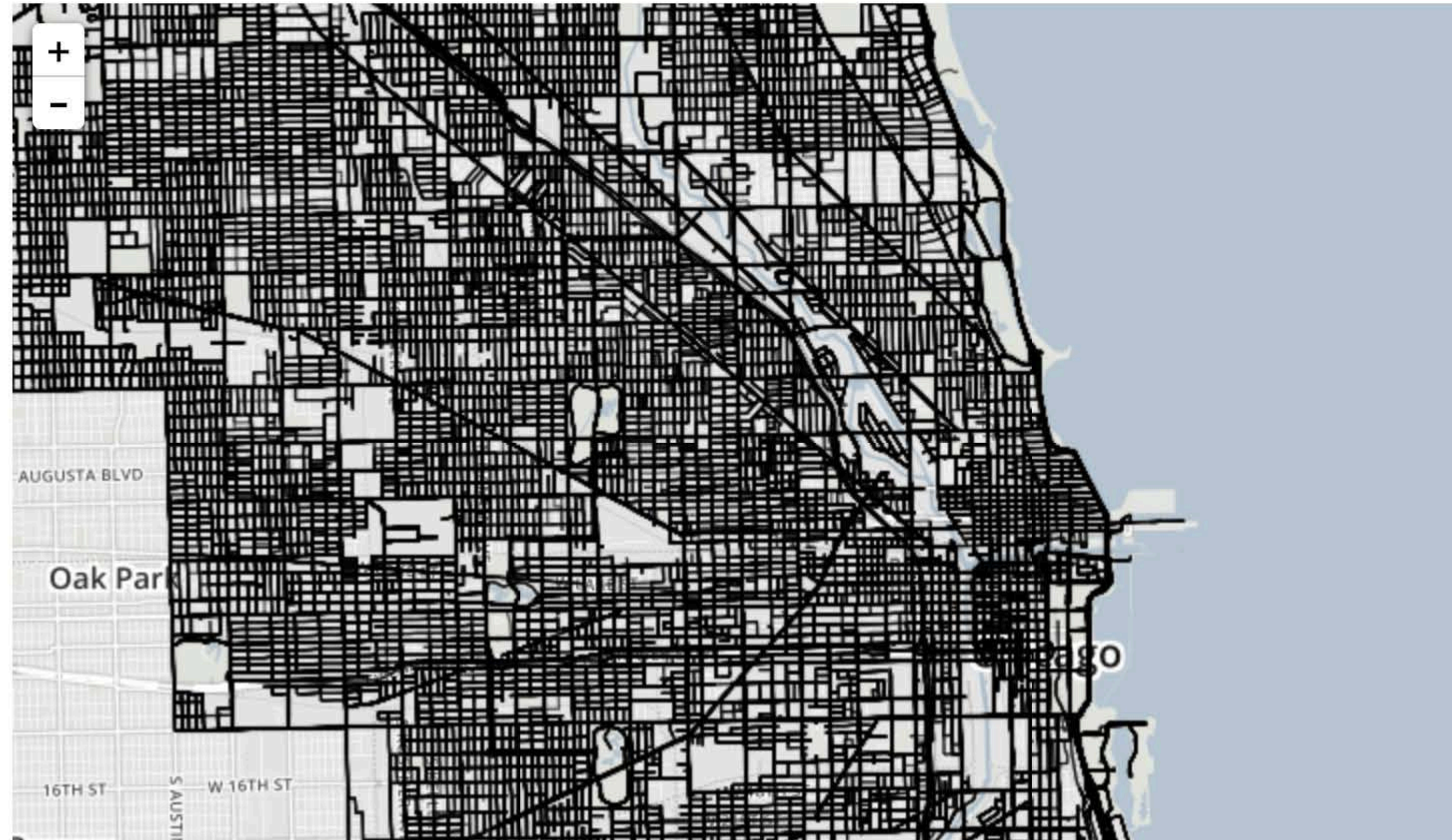
Search

Reset

Plow fleet: **Hibernating**

Clear Streets tracks Chicago's snow plows in real time by using data from the [City of Chicago's Plow Tracker](#). By knowing where the plows have been, we've figured out which streets have been plowed. [Read more »](#)

Did your street get skipped?



Community-supported innovation: Civic hackers can use open data to communicate city activity to residents, like this snow plow tracker.



Shining light on

FINANCIAL ADVISORS

Entrepreneurship: Open data can power new companies, such as BrightScope, which create jobs and economic activity.

@abhinemani



BrightScope Acquired by Institutional Shareholder Services, Inc.

[Click Here to Learn More About ISS](#)



@abhinemani

GETTING “OPEN DATA”

Sometimes easy, sometimes not...

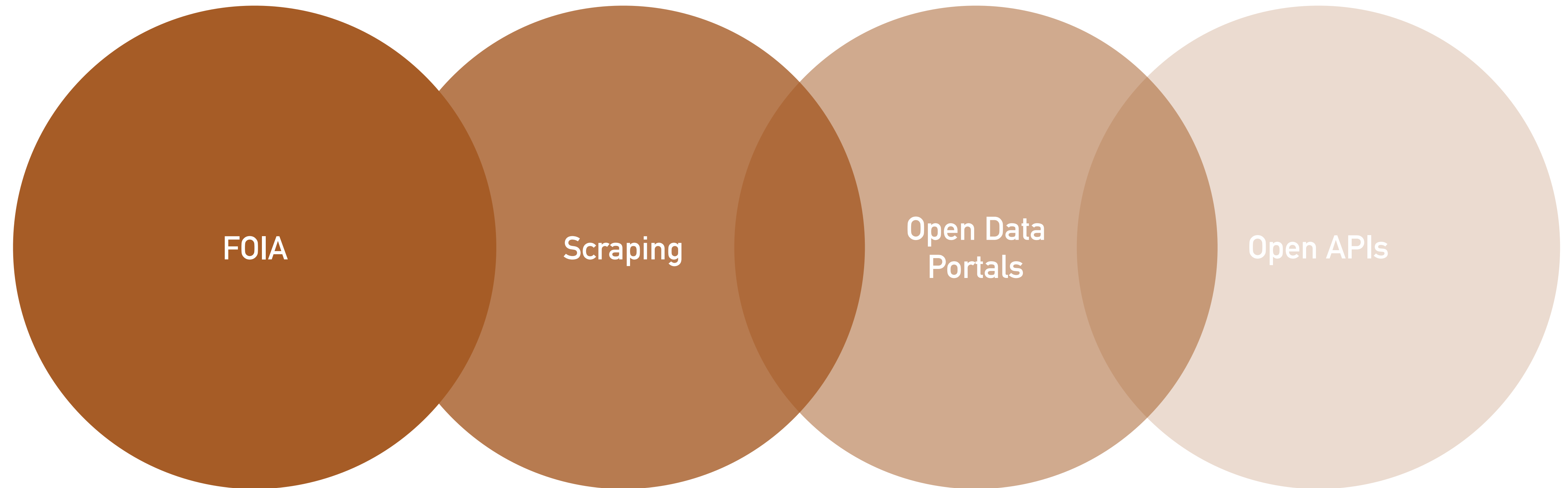




WAYS TO ACCESS GOVERNMENT DATA

- Pull
 - FOIA / PRA
 - Scraping
- Push
 - Open data portals
 - Open APIs

HOW DO YOU GET THE “OPEN” DATA?



MORE

Data available (decreasing)

LESS

HOW DO YOU GET THE “OPEN” DATA?



FOIA

Legally, the government must give you access to any non-“confidential” information it has. The federal law is known as the “Freedom of Information Act” (FOIA); states and cities have their own “Public Records Act(s)” (PRA) or “Sunshine Laws” as well.

FOIA EXAMPLE: SMARTPROCURE.US

HOME

ABOUT US

We have created a community of transparency by pulling together all purchase orders we can into an easily searchable database. Agencies don't just add to the database, but also utilize it for their own purchasing needs.

SmartProcure

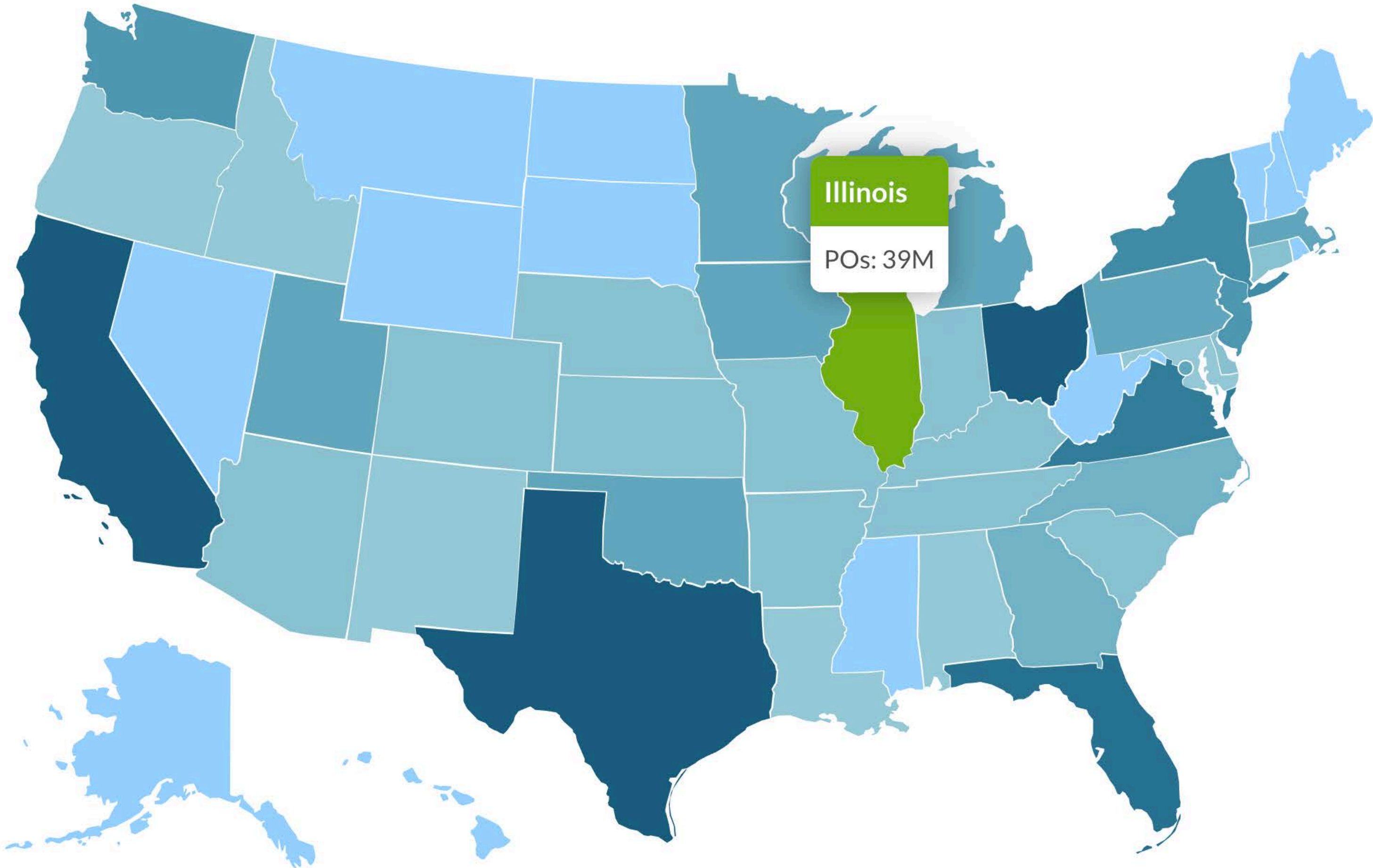
BECOME A MEMBER

419M
Purchase Orders

908M
Line Items

24.9K
Total Agencies

452K
Key Contacts



FOIA GETTING EASIER WITH NEW TOOLS: NEXTREQUEST

Update: SOC 2 security audit complete! - [Learn More](#)

 **NextRequest**

[Case Studies](#)

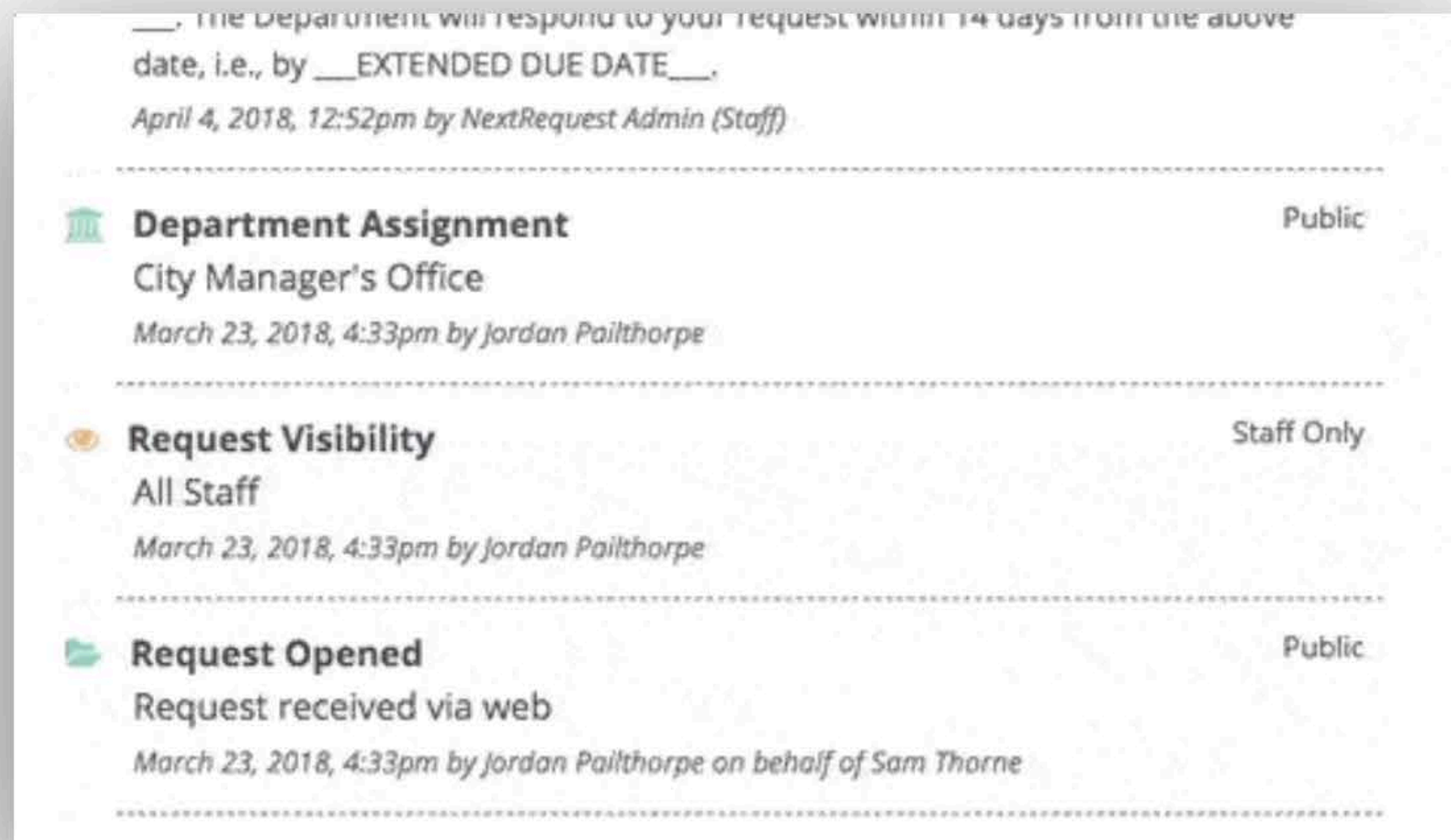
[Features](#)

[Contact Us](#)

[Start A Free Trial](#)

Public Records Is All We Do

As the experts in public records request management, our FOIA software will transform how your government manages records requests.

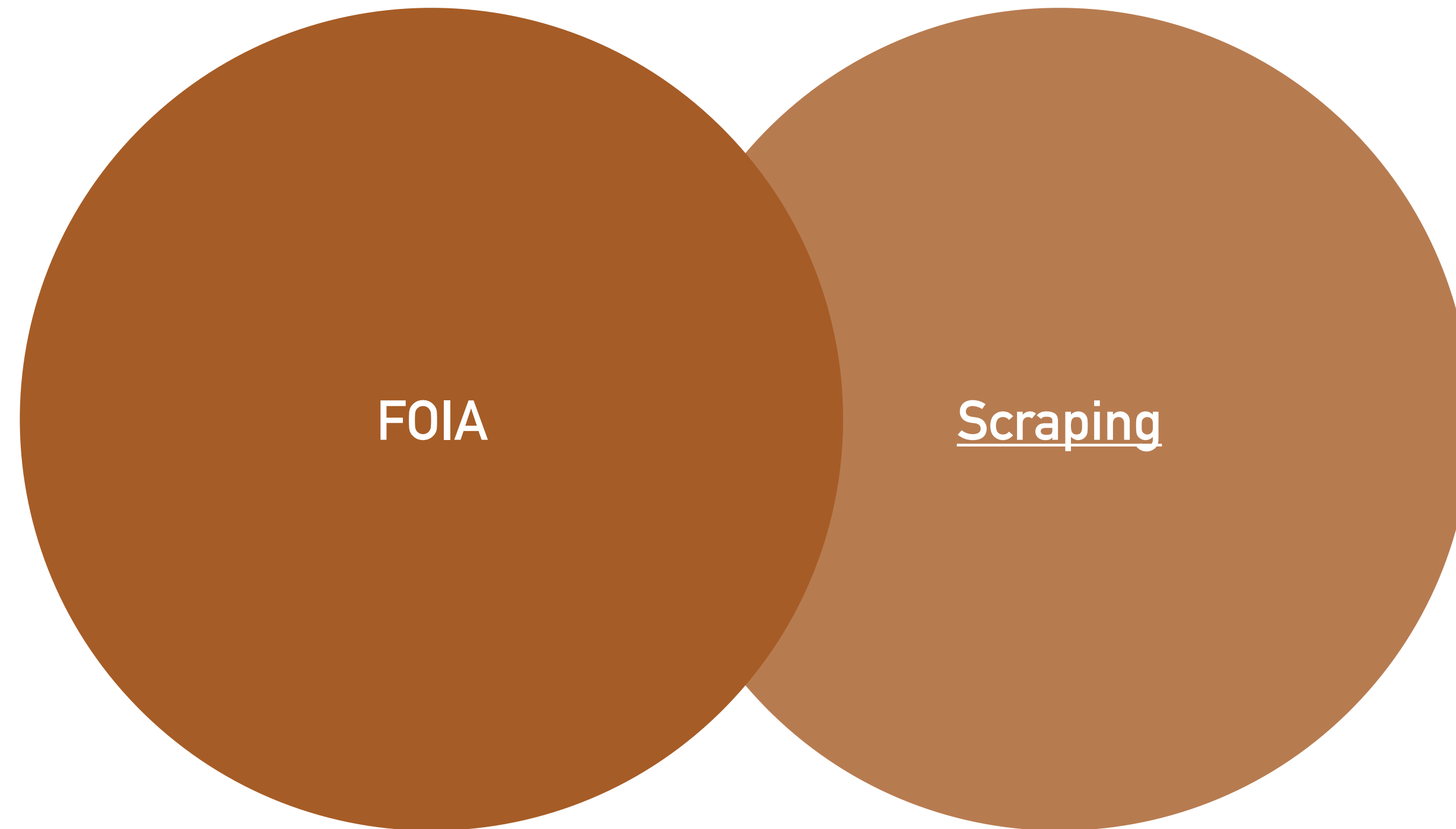


Increase Your Compliance

Give your agency a boost on compliance regulations and laws. With NextRequest, you can rapidly respond to audits, examinations, and legal inquiries with our suite of compliance features.

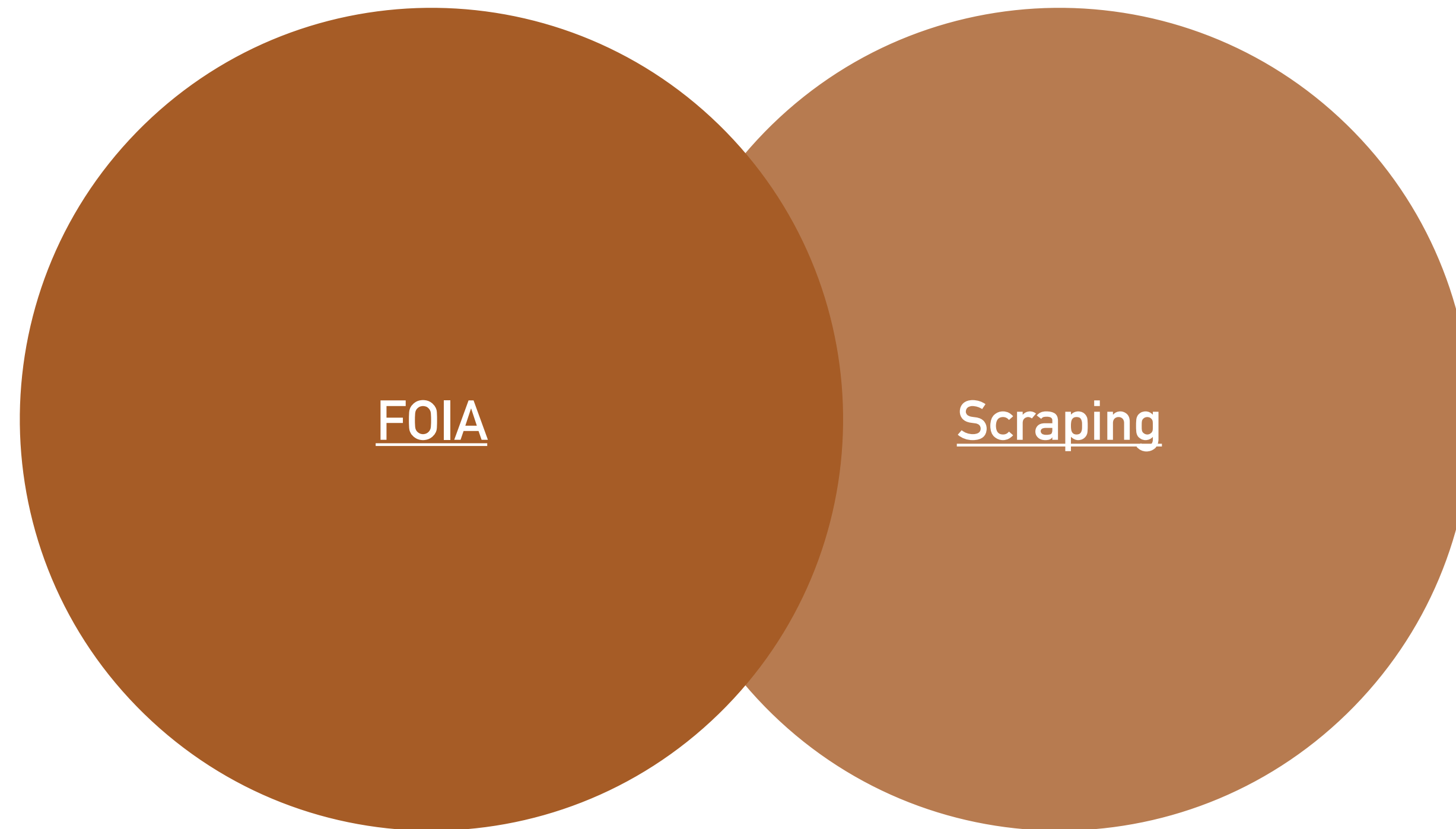
[Learn more](#)

HOW DO YOU GET THE “OPEN” DATA?

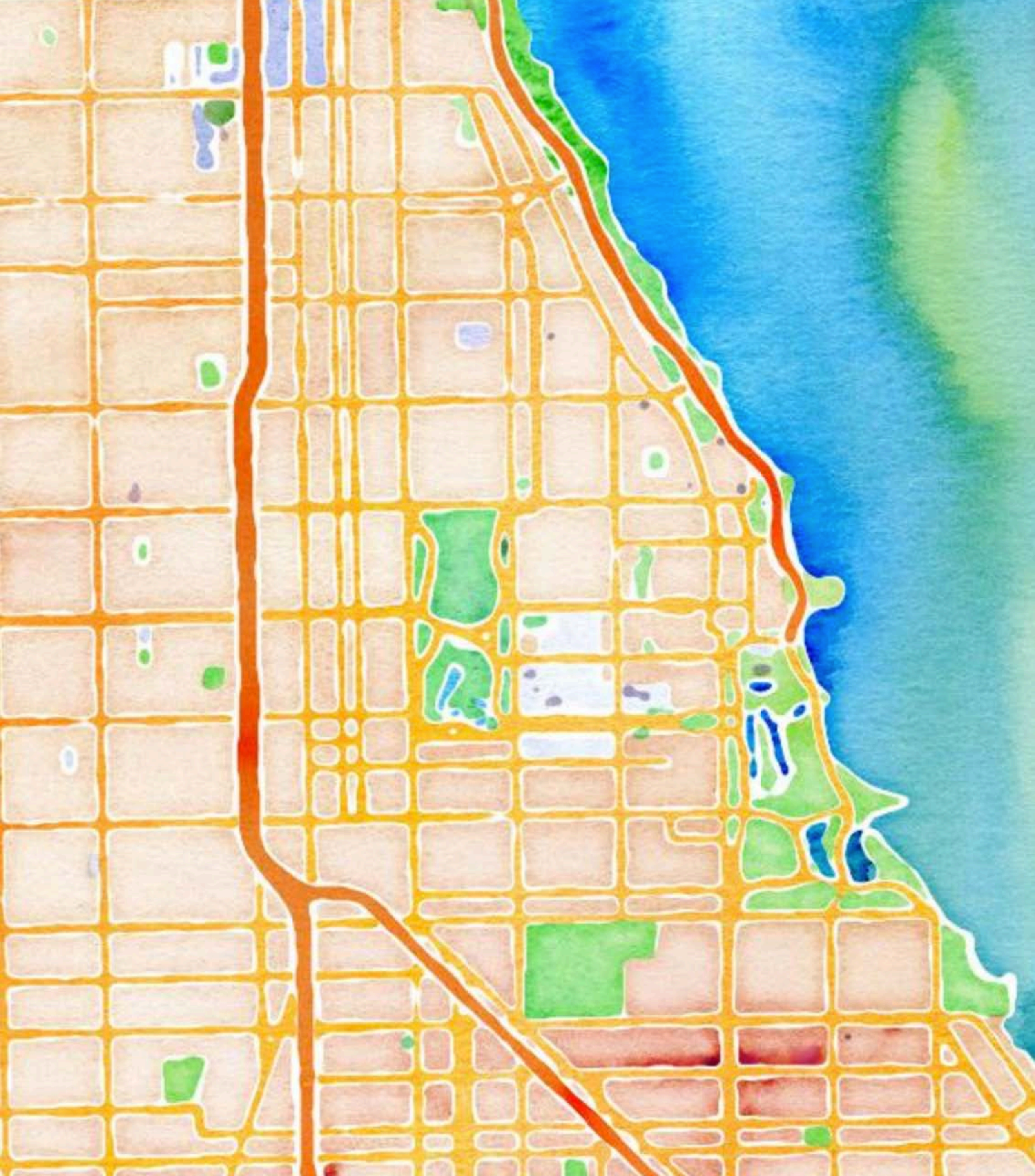


For years, governments have posted “data” online either as downloadable PDFs (e.g. annual/quarterly reports) or as tables/lists on their websites. (Scraping will be the focus of a later lecture)

HOW DO YOU GET THE “OPEN” DATA?



There are few requirements on the manner in which governments complete FOIA requests, so often you will need to run PDF scrapers or an OCR script (object/content recognition) on the FOIA'd files to make the data machine-readable / useful.



CHALLENGES WITH FOIA OR SCRAPING

- ▶ FOIA
 - ▶ “Static” file
 - ▶ Requires ongoing requests to maintain relevance
 - ▶ Proprietary data formats
 - ▶ Limited/no documentation
 - ▶ Can annoy government officials
 - ▶ Can generate additional work
- ▶ Scraping
 - ▶ Typically requires programming
 - ▶ Scripts can become irrelevant with any change by government
 - ▶ Can annoy government officials
 - ▶ Can tax technical infrastructure

PUSH VS. PULL

Open Data Portals
Open APIs

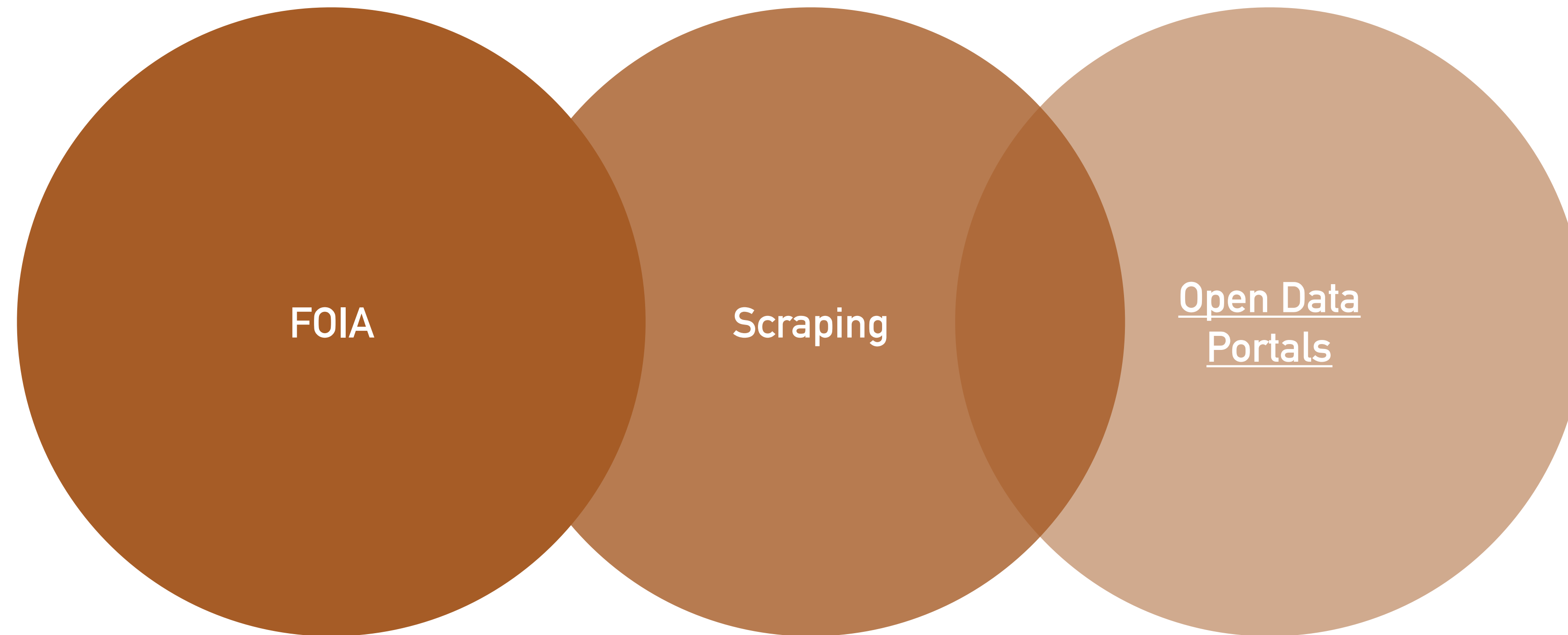
FOIA
Scraping

“

It's easier to throw stones at a house
than to build one.

*- The West Wing
(and many other places)*

HOW DO YOU GET THE “OPEN” DATA?



Governments have begun to pass open data policies / sign open data executive orders, mandating a pro-active release of data, typically through online datastores (ie open data portals).



@abhinemani

My Administration is committed to creating an unprecedented level of openness in Government. We will work together to ensure the public trust and establish a system of transparency, public participation, and collaboration. Openness will strengthen our democracy and promote efficiency and effectiveness in Government.

– PRESIDENT OBAMA, 01/21/09

Categories ▼

Administration & Finance

Buildings

Community & Economic Development

Education

Environment & Sustainable Development

[Show All...](#)

View Types ▼

Calendars

Charts

Data Lens pages

1319 Results

Sort by Most Accessed ▼

Crimes - 2001 to present - Dashboard

Public Safety

 Data Lens

This dataset reflects reported incidents of crime (with the exception of murders where data exists for each victim) that occurred in the City of Chicago from 2001 to present, minus the most recent seven days. Data is extracted...

[More](#)

Updated
April 21, 2019

Views
1,681,412

Tags *No tags assigned*

Current Employee Names, Salaries, and Position Titles

Administration & Finance

 Dataset

This dataset is a listing of all current City of Chicago employees, complete with full names, departments, positions, employment status (part-time or full-time), frequency of hourly employee -where applicable—and annual salaries...

[More](#)

Updated
April 3, 2019

Views
1,311,969

Tags [personnel](#)

[API Docs](#)

Building Permits

Buildings

 Dataset

Permits issued by the Department of Buildings in the City of Chicago from 2006 to the present. The dataset for each

Updated
April 21, 2019

DATA.GOV: A (SOMEWHAT) CENTRAL DATASTORE



DATA TOPICS ▾ IMPACT APPLICATIONS DEVELOPERS CONTACT

The home of the U.S. Government's open data

Here you will find data, tools, and resources to conduct research, develop web and mobile applications, design data visualizations, and [more](#).

GET STARTED

SEARCH OVER 241,835 DATASETS

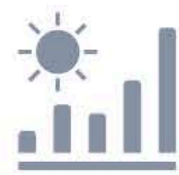
Health Care Provider Charge Data



BROWSE TOPICS



Agriculture



Climate



Consumer



Ecosystems



Education



Energy



Finance

Dataset Counts by Organization Type

Organization Type	Datasets
Federal Government	216,646
Other Federal	3,458
City Government	6,814
County Government	2,005
Education	762
Local Government	9,479
Non-Profit	219
Other Non-Federal	33
Private	235
Regional	823
State Government	17,674
Tribal	8
University	8,209
Total	266,365

Google Dataset Search Beta



Try [boston education data](#) or [weather site:noaa.gov](#)

[Learn more](#) about including your datasets in Dataset Search.

100+ results found

kaggle Crimes in Chicago

www.kaggle.com

Updated Jan 28, 2017



Crimes - 2001 to present

catalog.data.gov
data.wu.ac.at

Updated Apr 5, 2019

kaggle Data from: Chicago Crime

www.kaggle.com

Updated Apr 17, 2018



Crimes - 2001 to present

data.cityofchicago.org

Updated Nov 30, 2013

kaggle Chicago Crime Data

www.kaggle.com

Updated Mar 9, 2017

Crimes - 2001 to present

 data.cityofchicago.org**Dataset created** Sep 30, 2011**Dataset updated** Nov 30, 2013**Dataset published** Sep 30, 2011

@abhinemani

Available download formats from providers

rssxml, xml, csv, json, rdfxml

Description

This dataset reflects reported incidents of crime (with the exception of murders where data exists for each victim) that occurred in the City of Chicago from 2001 to present, minus the most recent seven days. Data is extracted from the Chicago Police Department's CLEAR (Citizen Law Enforcement Analysis and Reporting) system. In order to protect the privacy of crime victims, addresses are shown at the block level only and specific locations are not identified. Should you have questions about this dataset, you may contact the Research & Development Division of the Chicago Police Department at 312.745.6071 or RandD@chicagopolice.org. Disclaimer: These crimes may be based upon preliminary information supplied to the Police Department by the reporting parties that have not been verified. The preliminary crime classifications may be changed at a later date based upon additional investigation and there is always the possibility of mechanical or human error. Therefore, the Chicago Police Department does not guarantee (either expressed or implied) the accuracy, completeness, timeliness, or correct sequencing of the information and the information should not be used for comparison purposes over time. The Chicago Police Department will not be responsible for any error or omission, or for the use of, or the results obtained from the use of this information. All data visualizations on maps should be considered approximate and attempts to derive specific addresses are strictly prohibited. The Chicago Police Department is not responsible for the content of any off-site pages that are referenced by or that reference this web page other than an official City of Chicago or Chicago Police Department web page. The user specifically acknowledges that the Chicago Police Department is not responsible for any defamatory, offensive, misleading, or illegal conduct of other users, links, or third parties and that the risk of injury from the foregoing rests entirely with the user. The unauthorized use of the words "Chicago Police Department," "Chicago Police," or any colorable imitation of these words or the unauthorized use of the Chicago Police Department logo is unlawful. This web page does not, in any way, authorize such use. Data is updated daily Tuesday through Sunday. The dataset contains more than 65,000 records/rows of data and cannot be viewed in full in Microsoft Excel. Therefore, when downloading the file, select CSV from the Export menu. Open the file in an ASCII text editor, such as Wordpad, to view and search. To

About this Dataset

Updated

November 30, 2013

Data Last Updated
November 29, 2013

Metadata Last Updated
November 30, 2013

Date Created
September 30, 2011

Views
50

Downloads
0

Data Provided by
Chicago Police Department

Dataset Owner
cocadmin

Contact Dataset Owner

Metadata

Data Owner	Police
Time Period	2001 to present, minus the most recent seven days
Frequency	Data is updated daily, Tuesday through Sunday
Last Updated Date via Automated Load	2013/10/29 05:00:01

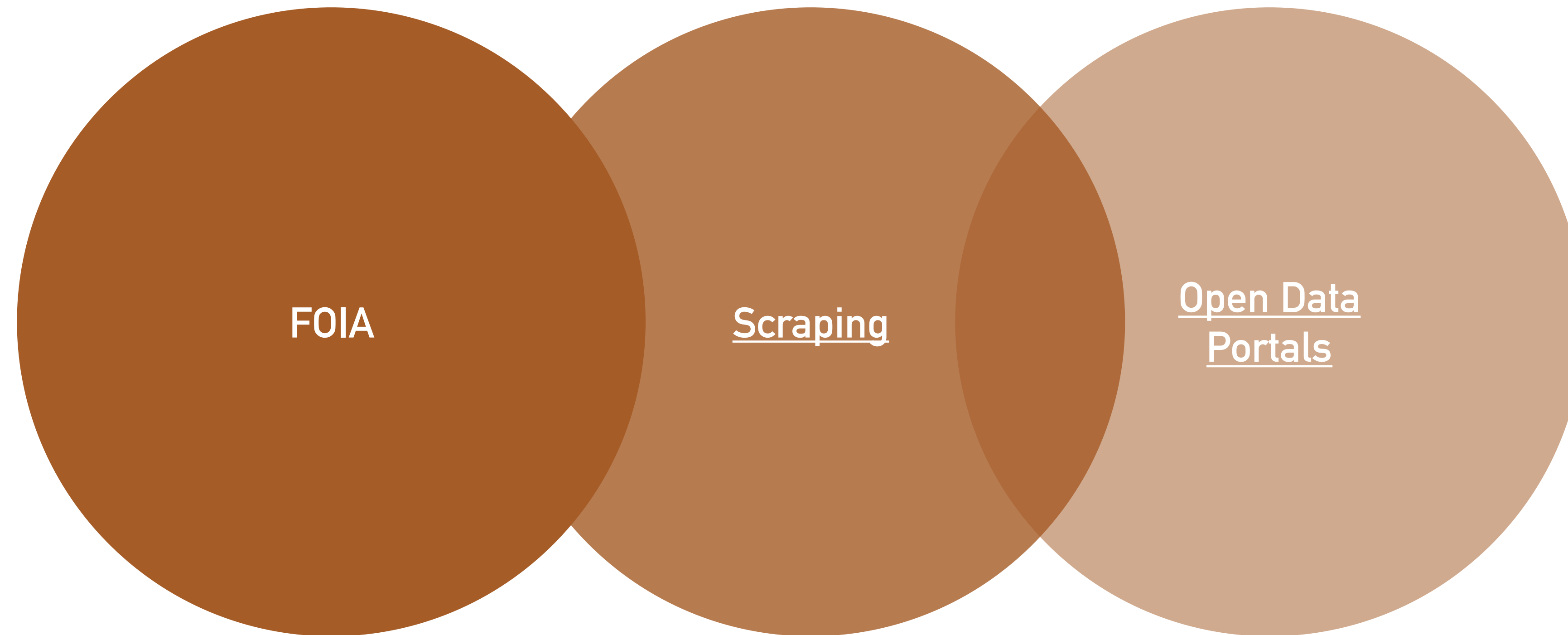
Topics

Category	Public Safety
Tags	crime , police

Licensing and Attribution

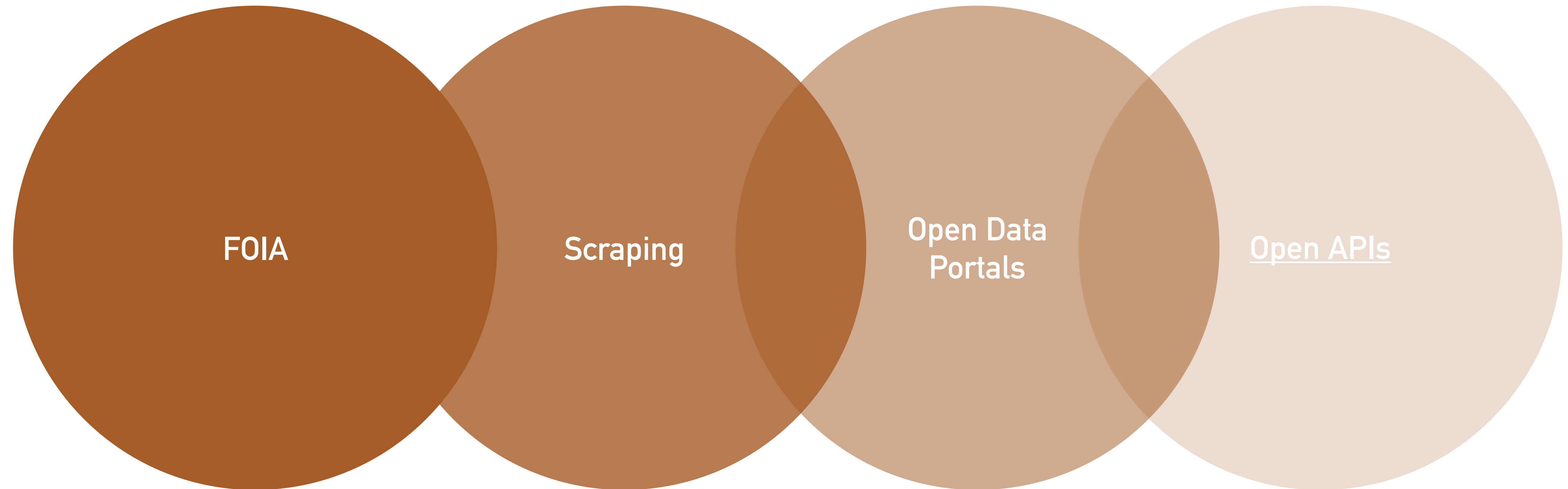
License	<i>The license for this dataset is unspecified</i>
Source Link	https://portal.chicagopolice.org/portal/page/portal/ClearPath ↗

HOW DO YOU GET THE “OPEN” DATA?



When governments prioritize which data to publish to the open data portal, often they will turn to the data they know has been scraped most often, to reduce the burden scraping places on their servers.

HOW DO YOU GET THE “OPEN” DATA?



Some government applications/systems — usually more modern ones — have their own open APIs, which provide direct, programmatic access to the data they contain.

EXAMPLE: SECLICKFIX – OPEN311 API

Open311

Our Open311 endpoint provides two avenues for interacting with our data. First, we provide effective Open311 endpoints for all of our client cities, only containing requests and services utilized by that city. Second, we provide our own unscoped top-level interface, enabling API consumers to query by latitude and longitude to get any relevant requests and service codes at that location, even from multiple jurisdictions. SeeClickFix's Open311 endpoint thus allows you to interact with individual cities and any zones within, or to interact with our system as a whole for performing queries of larger scope.

SeeClickFix's Open311 endpoint is based on the [Open311 GeoReport v2](#) standard.

We have two endpoints available, [production](#) and [test](#).

We currently only support JSON response formats.

Get started

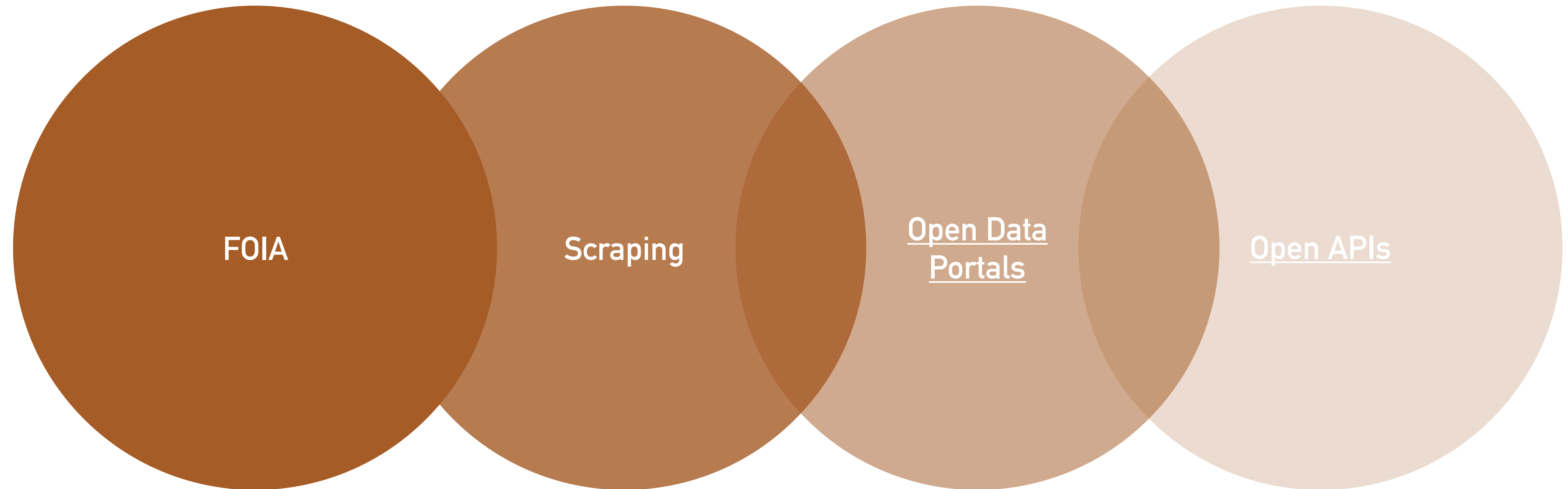
Try querying by latitude and longitude:

```
https://seeclickfix.com/open311/v2/requests.json?lat=41.307153&long=-72.925791  
https://seeclickfix.com/open311/v2/services.json?lat=41.307153&long=-72.925791
```

Or by account ID:

```
https://seeclickfix.com/open311/v2/29/requests.json
```

HOW DO YOU GET THE “OPEN” DATA?



Most open data portals — more specifically, data portals that have a data warehousing functionality — have open APIs enabling standardized, programmatic access to all published data.

EXAMPLE: SOCRATA OPEN DATA API

Code Snippets

The following are grab-and-go code code samples you can use with popular programming languages and data science tools.

jQuery

[Python Pandas](#)

[PowerShell](#)

[RSocrata](#)

[SAS](#)

[soda-ruby](#)

[SODA.NET](#)

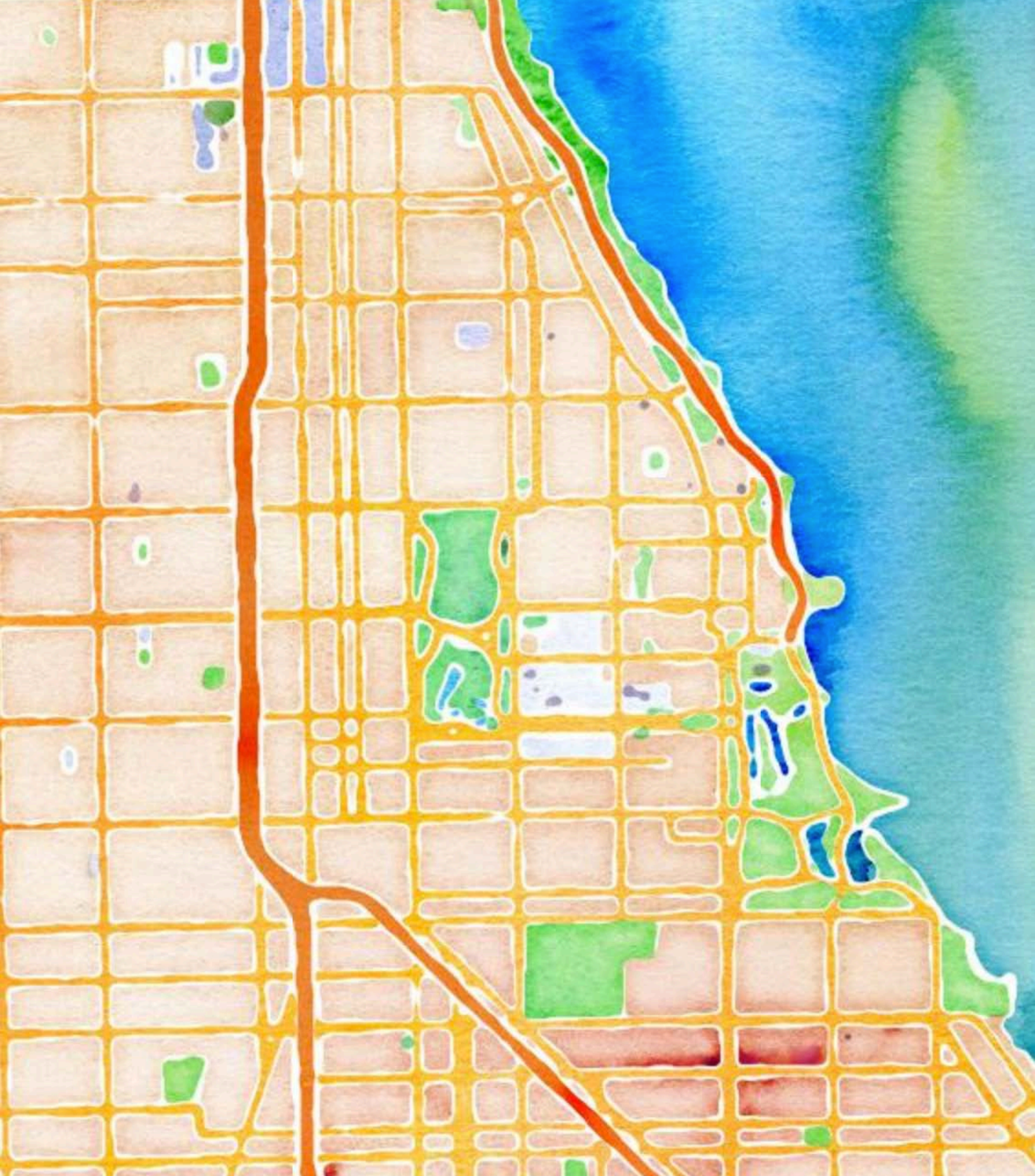
[Stata](#)

jQuery makes it super simple to fetch and parse JSON from an API endpoint.

```
$.ajax({
  url: "https://data.cityofchicago.org/resource/xbin-hpdm.json",
  type: "GET",
  data: {
    "$limit" : 5000,
    "$$app_token" : "YOURAPPTOKENHERE"
  }
}).done(function(data) {
  alert("Retrieved " + data.length + " records from the dataset!");
  console.log(data);
});
```

The following resources might also be helpful:

- [Using a jQueryUI date slider to build a SODA Query](#)
- [Google Maps Mashup](#)
- [chrismetcalf/shady-eats on GitHub](#)



CHALLENGES WITH PORTALS AND APIS

- Portals
 - Select datasets, decided by government / agencies
 - Management and publishing are difficult (more next lecture)
 - Data refresh rate not always consistent / reliable
 - General concern with data extracted from legacy systems
 - Limited/no documentation
- Open APIs
 - Digital services still rare
 - Limited/no documentation
 - Providers can be unreliable



@abhinemani

OPEN DATA PORTALS

A new standard for “smart cities”



LECTURE 7 ASSIGNMENT (ASSIGNED 4/22)

➤ Readings

- *Let's Move Beyond Open Data Portals* by Abhi Nemani
- *My failures: Civic technology ideas that didn't quite work* by Joshua Tauberer
- *Changing Role of the Government CDO* by Mark Headd
- *Los Angeles Open Data Policy*
- Optional: *Beyond Transparency Part II*

➤ Project Work

- Continue customization of JKAN instance