Getting Open Data through the Western Pennsylvania Regional Data Center or:

How I Learned to Stop Downloading CSVs and Love the API

David Walker Steve Saylor Twitter Bot

With special guest:
Digalot Sirius Mischief

What is “Open Data”?

What is the WPRDC?

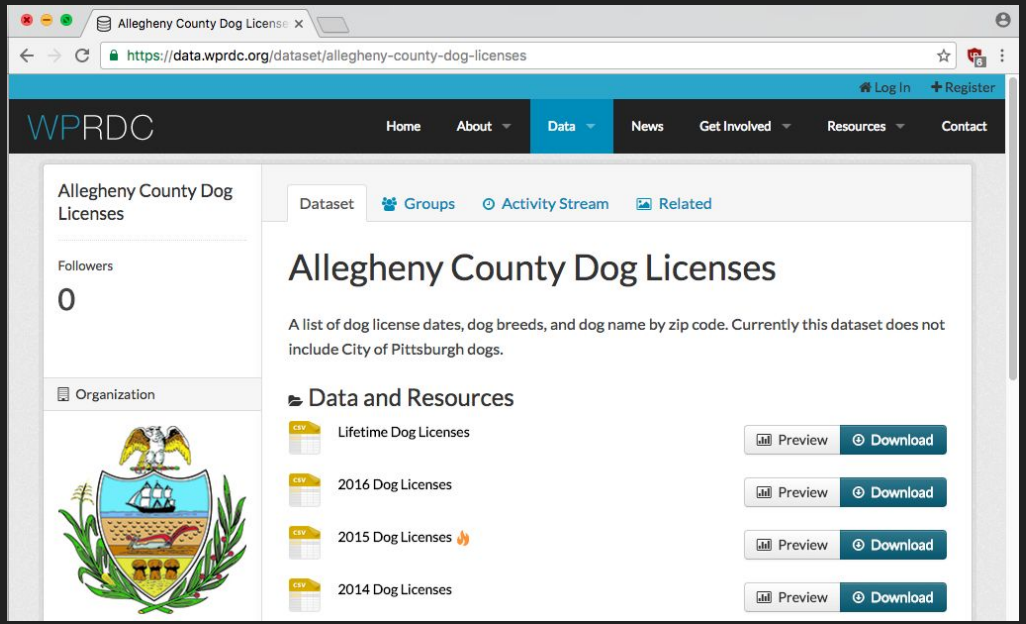
Data? What kind of data?

Civic data. Megatons of it! (155 datasets currently.)

- Updated roughly hourly
  - 311 Requests
  - Air Quality Measurements
  - Snow Plow Tracking
- Updated daily
  - Public Safety (crime blotter, arrests, citations)
  - PLI (Permits, License, and Inspections) Violations
- Updated monthly
  - Restaurant Inspections
  - Property Assessments
  - Property Sales
  - Accidental Drug Overdoses
- Popular yet somewhat static (updated yearly)
  - Car Crashes (Now covering 5 counties!)
  - Dog Licenses

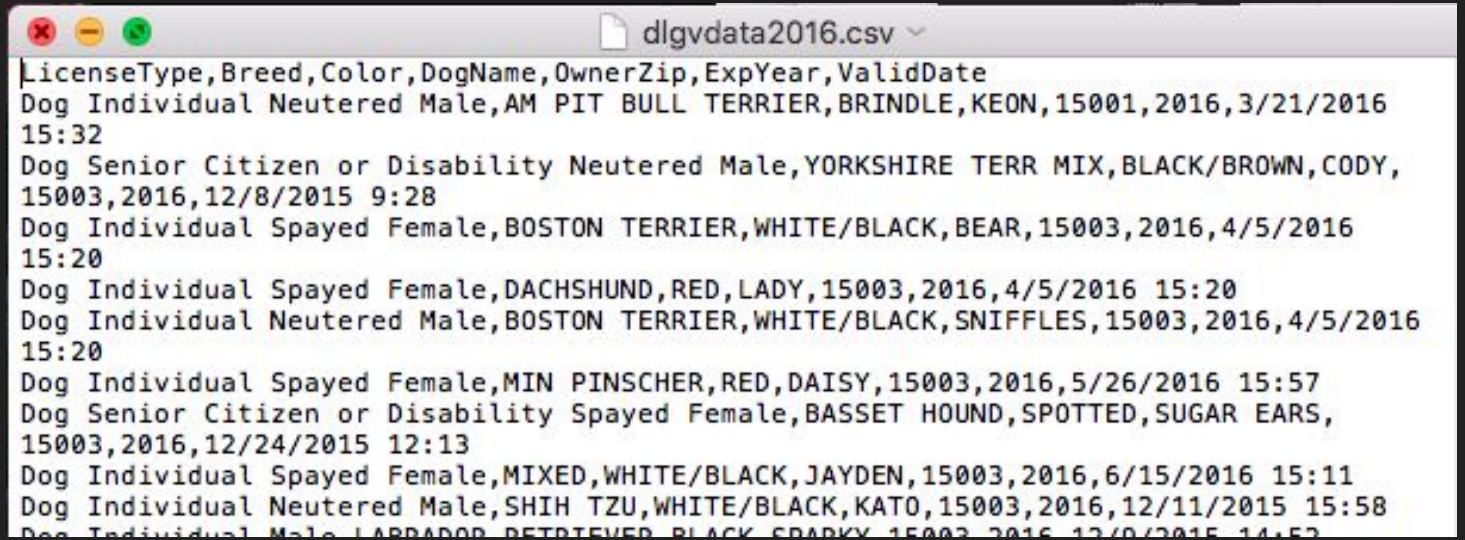
How do I get that delicious data?

Option 1: Download CSV (comma-separated values) files



How do I get that delicious data?

Option 1: Download CSV (comma-separated values) files

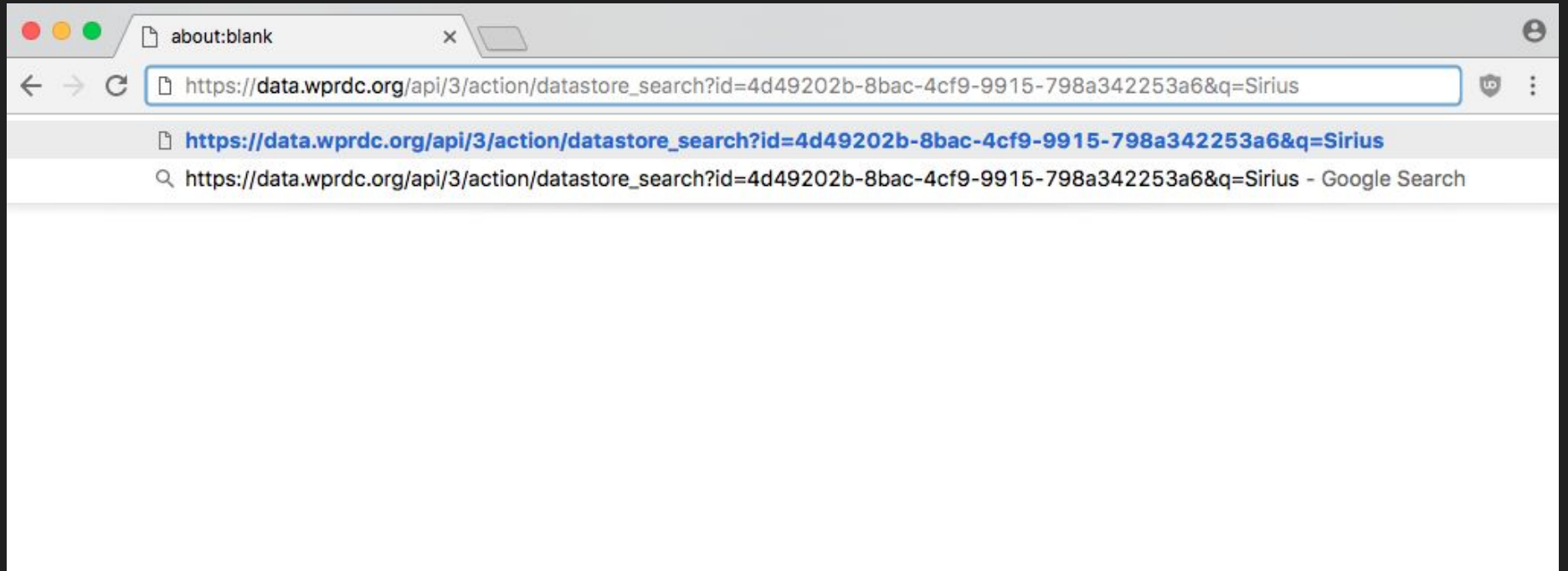


How do I get that delicious data?

Option 2: Use our beautiful API (Application Programming Interface)

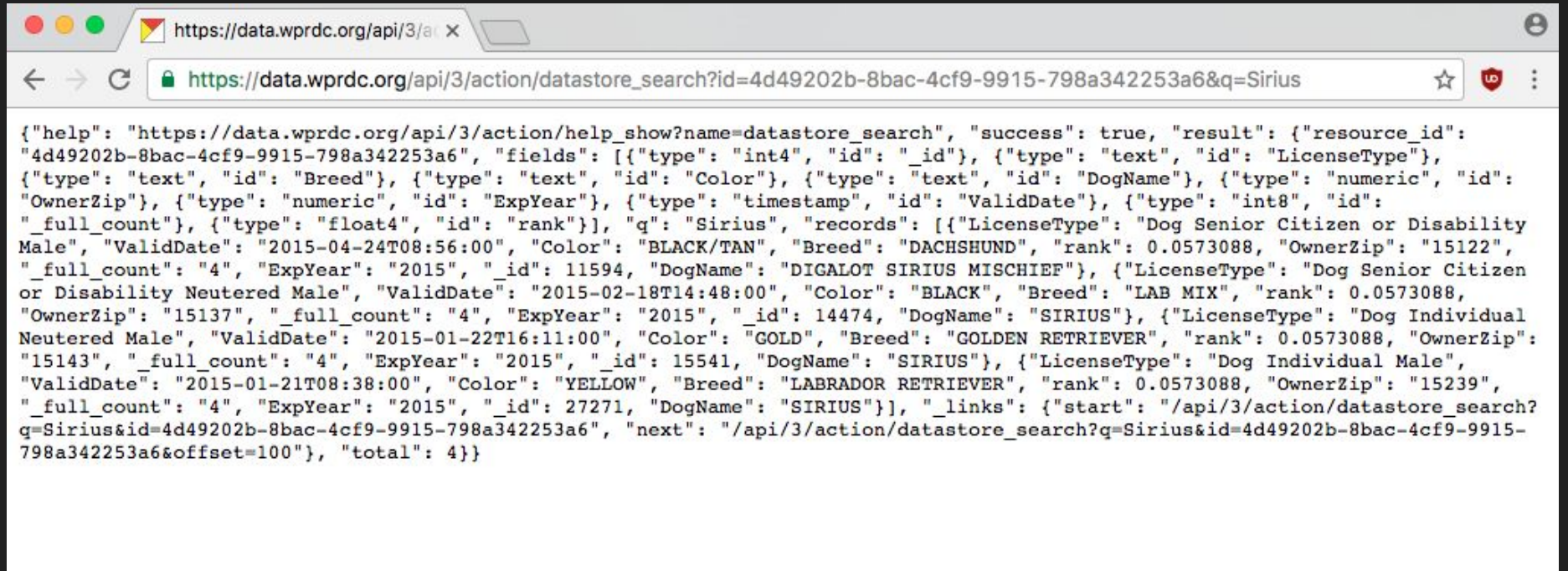
How do I get that delicious data?

Option 2: Use our beautiful API (Application Programming Interface)



How do I get that delicious data?

Option 2: Use our beautiful API (Application Programming Interface)



Mad Libs with Open Data?

Can you explain an example API call?

https://data.wprdc.org/api/3/action/datastore_search?id=1797ead8-8262-41cc-9099-cbc8a161924b&limit=5

Root URL
API URL - URL where all API endpoints are located
API endpoint - the action we want to perform
Question mark - indicates that text afterwards represent parameters
Parameter/value - key-value pair of endpoint-specific parameter mapped to a value we send
Ampersand - separates parameters

Tiny examples, anyone?

Questions?

What can I do with the API?

1) Full-text search:

https://data.wprdc.org/api/action/datastore_search?id=4d49202b-8bac-4cf9-9915-798a342253a6&q=Sirius

What can I do with the API?

2) More powerful queries:

https://data.wprdc.org/api/action/datastore_search_sql?sql=SELECT * FROM '4d49202b-8bac-4cf9-9915-798a342253a6' WHERE 'DogName' LIKE '%SIRIUS%'

These are SQL queries (where SQL is a special language for dealing with databases).

What can I do with the API?

2) More powerful queries:

https://data.wprdc.org/api/action/datastore_search_sql?sql=SELECT * FROM '4d49202b-8bac-4cf9-9915-798a342253a6' WHERE 'DogName' LIKE '%SIRIUS%' AND "Breed" LIKE 'DACHSHUND'

These are SQL queries (where SQL is a special language for dealing with databases).

What can I do with the API?

2) More powerful queries:

https://data.wprdc.org/api/action/datastore_search_sql?sql=SELECT%20*%20FROM%20%224d49202b-8bac-4cf9-9915-798a342253a6%22%20WHERE%20%22DogName%22%20LIKE%20%27%SIRIUS%%27%20AND%20%22Breed%22%20LIKE%20%27DACHSHUND%27

These are SQL queries (where SQL is a special language for dealing with databases).

How about that Twitterbot demo?

When should I download CSV files? When should I use the API?

Download CSVs when:

- The data is not going to change.
- You want to manipulate the whole dataset.

Use the API when:

- The data is being regularly updated.
- You don’t want to download a 500,000-row dataset just to find the sales prices for one house.
- You want to get the types of the data fields.

What should I do next?

- Some ideas:
  - Open-Data Mad Libs Poetry Slam!
  - Play with the city's Burgh's Eye View (where you can download data)
  - Explore the Tiny Examples notebook
  - Find an interesting dataset and make your own API calls
  - Make your own Twitterbot
  - Build "Even Madder Libs"
  - Create an automatically updating map of the potholes reported in your neighborhood

Is this the final slide?

Yes.

Go to https://github.com/WPRDC/api-workshop

for the Tiny Examples notebook, the Twitterbot demo, links to where you can run these notebooks online, and lots of documentation.

If you make something cool, send us a link to it! We'll be blogging about our API talks and would like examples to include in our blog post.

WPRDC site: http://wprdc.org / E-mail: wprdc@pitt.edu / Twitter: @wprdc