P-R-R & Y-O-U

An inquiry into public records data

Kevin McCraney, Open Data Literacy Project
August 10, 2018

Scope

Can we analyze **public records** data across western Washington to support **proactive disclosure** of datasets*?

*and possibly provide a pipeline for expediting this process

Two questions:

What are public records requests (PRRs)?
What is proactive disclosure?

What are PRRs?

What are PRRs?

A last-ditch effort to get information.
An opportunity to provide value.

What is preemptive disclosure?

What is preemptive disclosure?

A panacea.

What is preemptive disclosure?

Releasing information proactively so stakeholders can get what they need *before* they ask.

What is preemptive disclosure?

<u>The Three R's of Preemptive Disclosure</u>

1. **Research -** Pull data out of centralized platform and look for broad trends/domains of interest

2. **Release -** Disclose new data, foregrounded by research

3. **Relax -** Enjoy accolades from your constituency and favorable reportage in the news media (c.f. SFD releasing a super-popular dataset)

What is preemptive disclosure?

A ~~panacea~~.
pain

What is preemptive disclosure?

<u>The Three C's of Preemptive Disclosure</u>

1. **Collect -** Aggregate data

2. **Cower -** Realize that the data we have is unstructured, missing tons of information and metadata, and that we cannot easily derive real insight; get sort of down about the scope of the project and feel like there is no way we can address such a large problem in the span of eight weeks

3. **Capitulate -** Try some complicated stuff to figure out what's happening.

Companies use NLP to find trends in unstructured text and deliver value.

Why can't the government?

Procedures

pdr_table.csv

Port Orchard PRR Data.xls

StudentPRRLog_06212018_2.xlsx

pdr_table.csv

Port Orchard PRR Data.xls

StudentPRRLog_06212018_2.xlsx

Not many features in common…😩

mmon…

Not many features in common…

…but lots of raw text to process!

What does a request look like?

5/15/2018 Burglary ADDRESS ! Burglary

Auto Theft 06/07/2018 : ADDRESS : PERSON

Police report on a domestic violence allegation filed against me on 6/1/2018

Any criminal history dating back to 2005. Also what is currently on my record for employment purposes

Traffic Collision Police Report

Police - Incident Reports/Calls For Service Reports

Transaction/F REFERENCE# I Insurance Company: STATE FARM CLAIMS COMPASS Report Type: Auto Accident Report/Case Number: CASE NUMBER I Date of Occurr
THURSTON State: WA Insured Driver: PERSON

Transaction/F REFERENCE# I Insurance Company: LIBERTY MUTUAL FTP Report Type: Auto Accident Report/Case Number: CASE NUMBER Date of Occurrence: 6/10
Driver: PERSON

What was the date that the "Transportation 2030 Maps" were adopted into the Comprehensive Plan via Ordinance 7104 and approved by the City Council?

REQUESTING ACCIDENT REPORT UNK VEHICLE CRASHED INTO PARKED CARS AND PUSSHED PARKED CARS INTO UTILITY EASEMENT DAMAGING CENTURYLINK PROPERTY. CE

Medical and billing records for I PERSON an incident on 6/14/2017.

Any and All permits for ADDRESS Olympia tax parcel REFERENCE# note there are two little houses (both built in 1901) on this tax parcel, so please pull p

Police - Incident Reports/Calls For Service Reports

06/06/2018; 12:18 am; ADDRESS ; Vehicle was recovered theft from Pierce Co. Case #: CASE NUMBER

CASE NUMBER

Please provide me with the public works sewer maintenance logs for ADDRESS . Please provide me with the City's entire file on my property I ADDRE
payments to me, City communications, emails, and meeting minutes regarding my property since 1991. I will accept electronic documentation and files. I will also pick

Police report and any associated documentation related to a welfare check on PERSON who had reported suicidal ideation on Thursday 6/14/18.

All plans for our recently purchased property in Olympia, parcel REFERENCE# ADDRESS . Specifically I would like the lot measurements so I can appro
REFERENCE# , and REFERENCE#

5/15/2018 Burglary ADDRESS ! Burglary

Auto Theft 06/07/2018 : ADDRESS : PERSON

Police report on a domestic violence allegation filed against me on 6/1/2018

Any criminal history dating back to 2005. Also what is currently on my record for employment purposes

Traffic Collision Police Report

Police - Incident Reports/Calls For Service Reports

Transaction/F REFERENCE# I Insurance Company: STATE FARM CLAIMS COMPASS Report Type: Auto Accident Report/Case Number: CASE NUMBER I Date of Occurr
THURSTON State: WA Insured Driver: PERSON

Transaction/F REFERENCE# I Insurance Company: LIBERTY MUTUAL FTP Report Type: Auto Accident Report/Case Number: CASE NUMBER Date of Occurrence: 6/10
Driver: PERSON

What was the date that the "Transportation 2030 Maps" were adopted into the Comprehensive Plan via Ordinance 7104 and approved by the City Council?

REQUESTING ACCIDENT REPORT UNK VEHICLE CRASHED INTO PARKED CARS AND PUSSHED PARKED CARS INTO UTILITY EASEMENT DAMAGING CENTURYLINK PROPERTY. CE

Medical and billing records for I PERSON an incident on 6/14/2017.

Any and All permits for ADDRESS Olympia tax parcel REFERENCE# note there are two little houses (both built in 1901) on this tax parcel, so please pull p

Police - Incident Reports/Calls For Service Reports

06/06/2018; 12:18 am; ADDRESS ; Vehicle was recovered theft from Pierce Co. Case #: CASE NUMBER

CASE NUMBER

Please provide me with the public works sewer maintenance logs for ADDRESS . Please provide me with the City's entire file on my property I ADDRE
payments to me, City communications, emails, and meeting minutes regarding my property since 1991. I will accept electronic documentation and files. I will also pick

Redacted, for safety.

Police report and any associated documentation related to a welfare check on PERSON who had reported suicidal ideation on Thursday 6/14/18.

All plans for our recently purchased property in Olympia, parcel REFERENCE# ADDRESS . Specifically I would like the lot measurements so I can appro
REFERENCE# , and REFERENCE#

PRRs span a continuum. From terse…

CASE NUMBER

Traffic Collision Police Report

5/15/2018 Burglary ADDRESS : Burglary

Any and All permits for ADDRESS : Olympia tax parcel REFERENCE# : note there are two little houses (both built in 1901) on this tax parcel, so please pull permits for both thanks in advance for your help

PRRs span a continuum. From terse…

CASE NUMBER

Traffic Collision Police Report

5/15/2018 Burglary ADDRESS : Burglary

Any and All permits for ADDRESS : Olympia tax parcel REFERENCE# : note there are two little houses (both built in 1901) on this tax parcel, so please pull permits for both thanks in advance for your help

…to verbose.

March 19, 2018 VIA ELECTRONIC REQUEST City of Olympia City Clerk's Office City Hall (Address . Re: Washington Public Records Act Request To the Public Records Officer of the City of Olympia: Please consider this a request under the Washington Public Records Act, RCW 42.56.001, et seq., for public records identifying the following: Documents showing the number of employees in the City of Olympia â€¢ Documents showing the locations and dimensions of all time limited and/or metered public on-street parking in Olympia. â€¢ Documents showing the total count of all time limited and/or metered public on-street parking in Olympia. â€¢ Documents showing the locations and addresses of all City owned parking lots with public parking. â€¢ At each City owned parking lot with public parking please provide the following information: o Total number of public parking spaces. o Number of public â€œregularâ€ accessible parking spaces. o Number of public â€œvanâ€ accessible parking spaces. o Original construction date of the parking lot. Please provide a copy of the original construction drawings used for construction. o Significant alteration dates of the parking lot. Specifically, any renovations and alterations that occurred between July 26, 1992 to September 15, 2010; September 15, 2010 to March 14, 2012; or March 15, 2012 to present. ? If so, any alterations occurred between these date, please provide a detailed description of the scope of work for these renovations and alterations including their corresponding dates for submission to the authority having jurisdiction for construction permits or a construction start date where a construction permit was not required or obtained, and a date for completion of the work, as well as the overall construction cost, excluding fees for professional services performed by design professionals, for each and every individual renovation or alteration that has occurred, within the dates identified. ? Please produce all documents which pertain to your answer of the previous bullet. o For each parking lot which has been altered, please provide a detailed description of the scope of work for the alterations. ? As part of the detailed description of the scope of work, please include the corresponding dates for submission to the authority having jurisdiction for construction permits or a construction start date where a construction permit was not required or obtained, and the date for completion of the work, as well as the overall construction cost, excluding fees for professional services performed by design professionals. ? Please produce all documents which pertain to your answer of the previous bullet. o Please provide the completed AS BUILT drawings for any renovations and alterations which occurred between July 26, 1992 to September 15, 2010; September 15, 2010 to March 14, 2012; or March 15, 2012 to present. The dates of the renovations and alterations should be clearly visible on every sheet of the drawings provided. o Please produce a copy all documents, including permits, in the City of Olympiaâ€™s possession, custody or control that pertain to, relate to, or discuss ADA accessibility for these parking lots. o Please produce a copy of all documents that pertain to inspection of the renovations of these parking lots including inspections prior to construction and post construction. o Please produce a copy of the Cityâ€™s Americans with Disabilities Act (ADA) Title II Self-Evaluation and Transition Plan as required by 28 CFR 35.105 and 28 CFR 35.150(d). If these records are not in your custody or control, then I ask that you promptly notify us and also state in detail to the best of your knowledge and belief the reason for the absence of the records from your custody or control, the location of the records, and identify the person(s) now havi

Unstructured Text

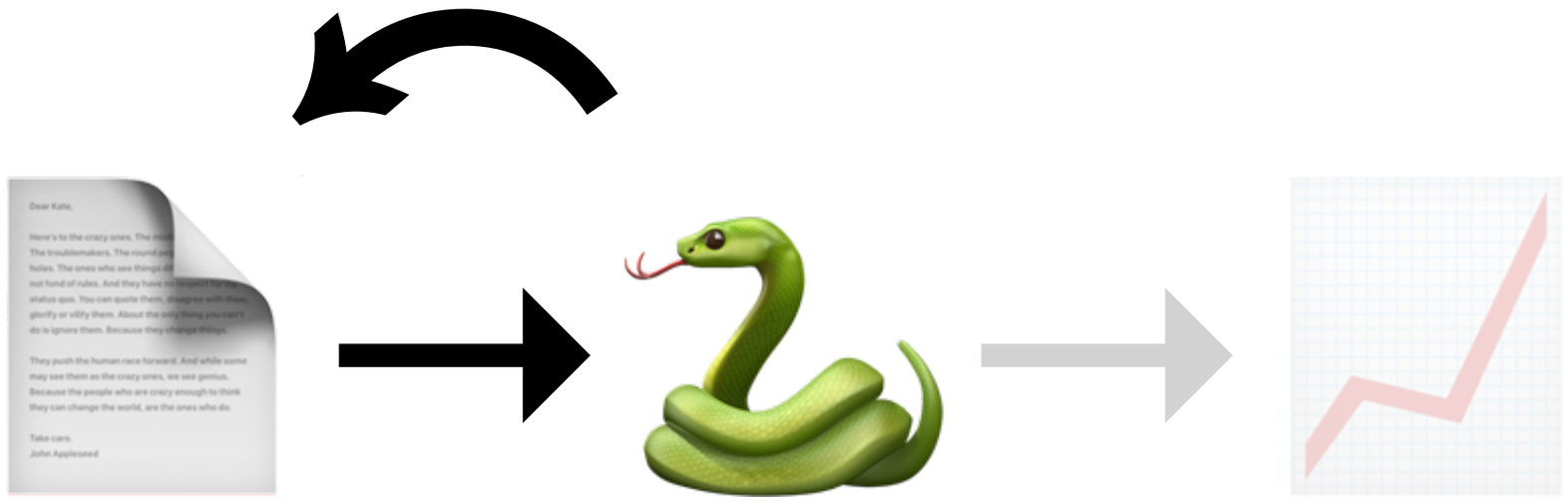
Python Pipeline

Interactive Visuals

Unstructured Text

Python Pipeline

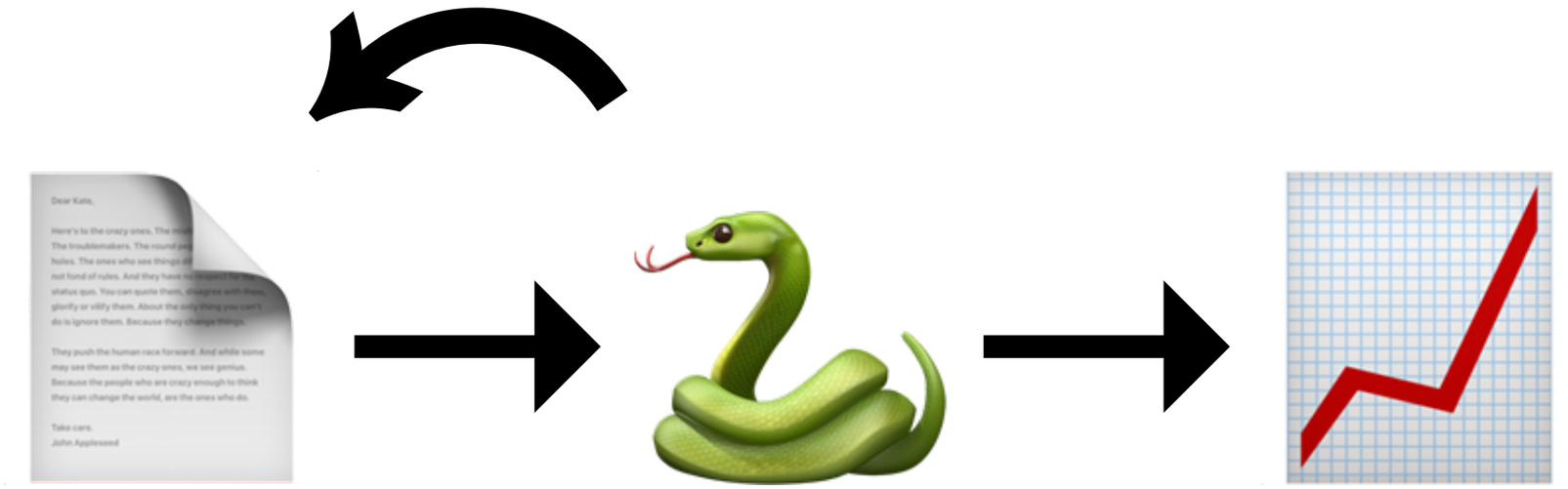
Interactive Visuals



Unstructured Text

Python Pipeline

Interactive Visuals

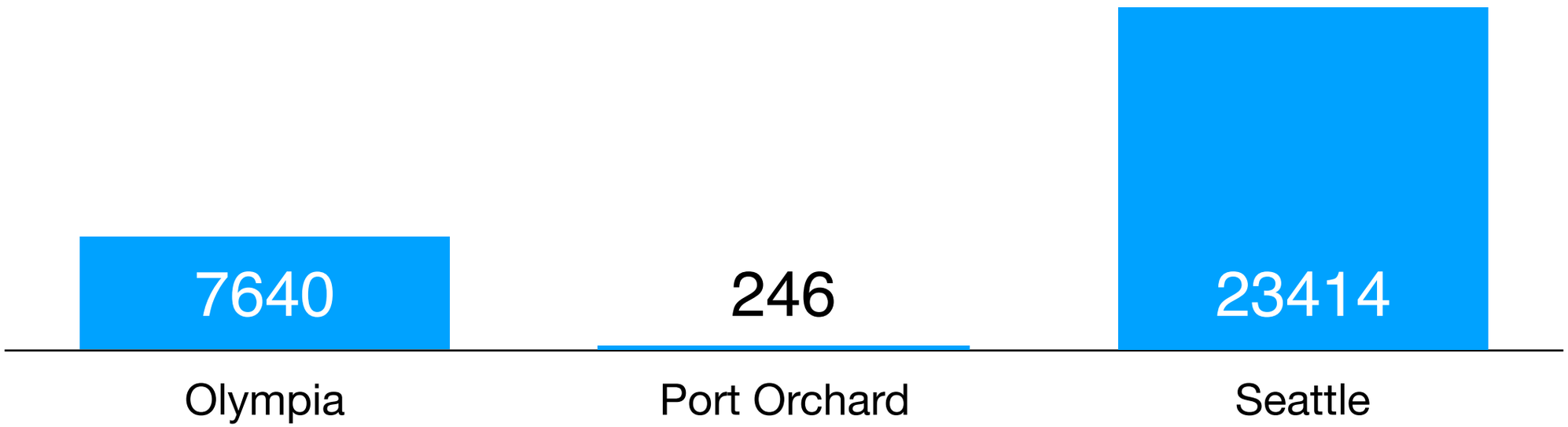


Unstructured Text

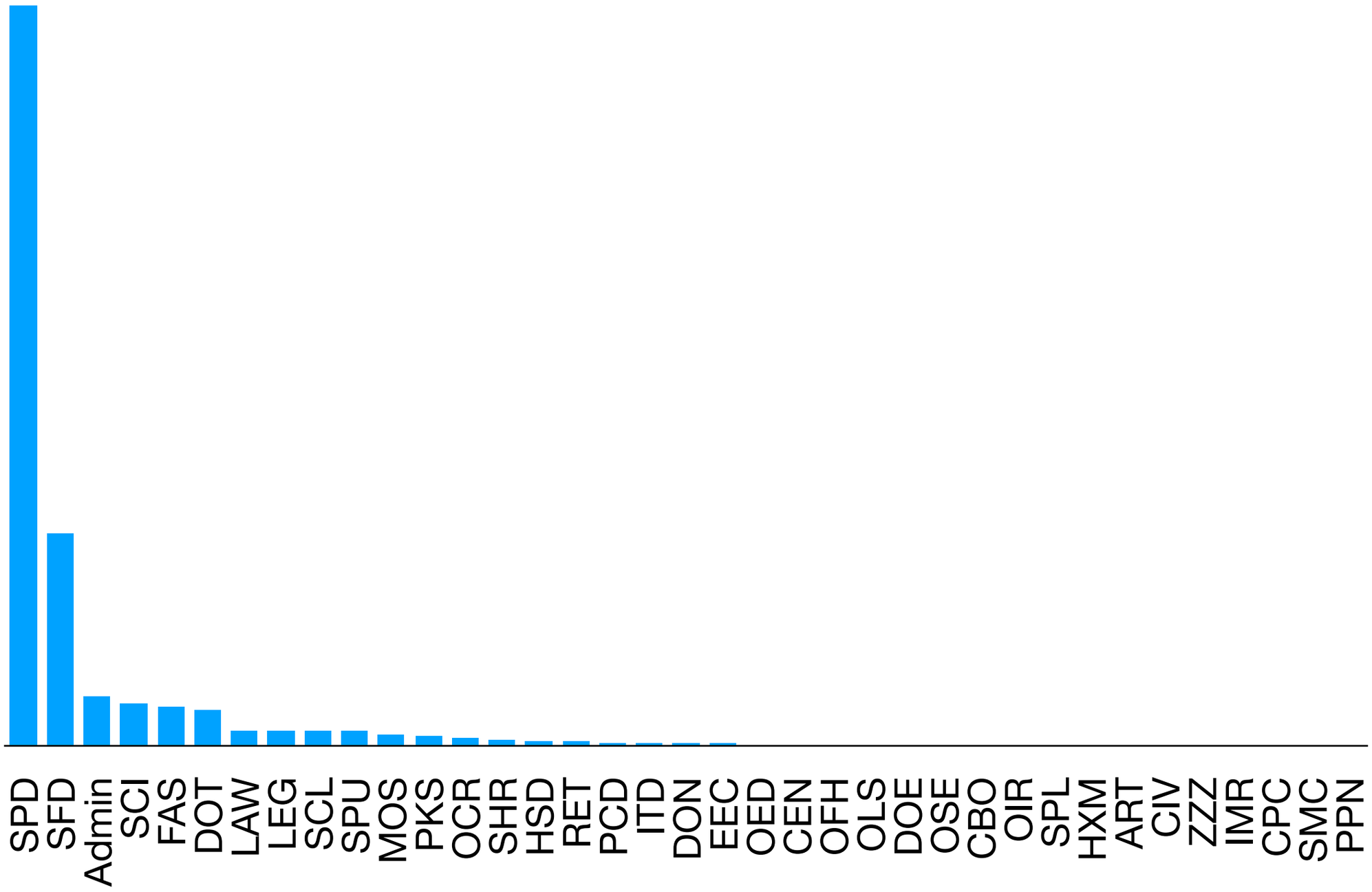
Python Pipeline

Interactive Visuals

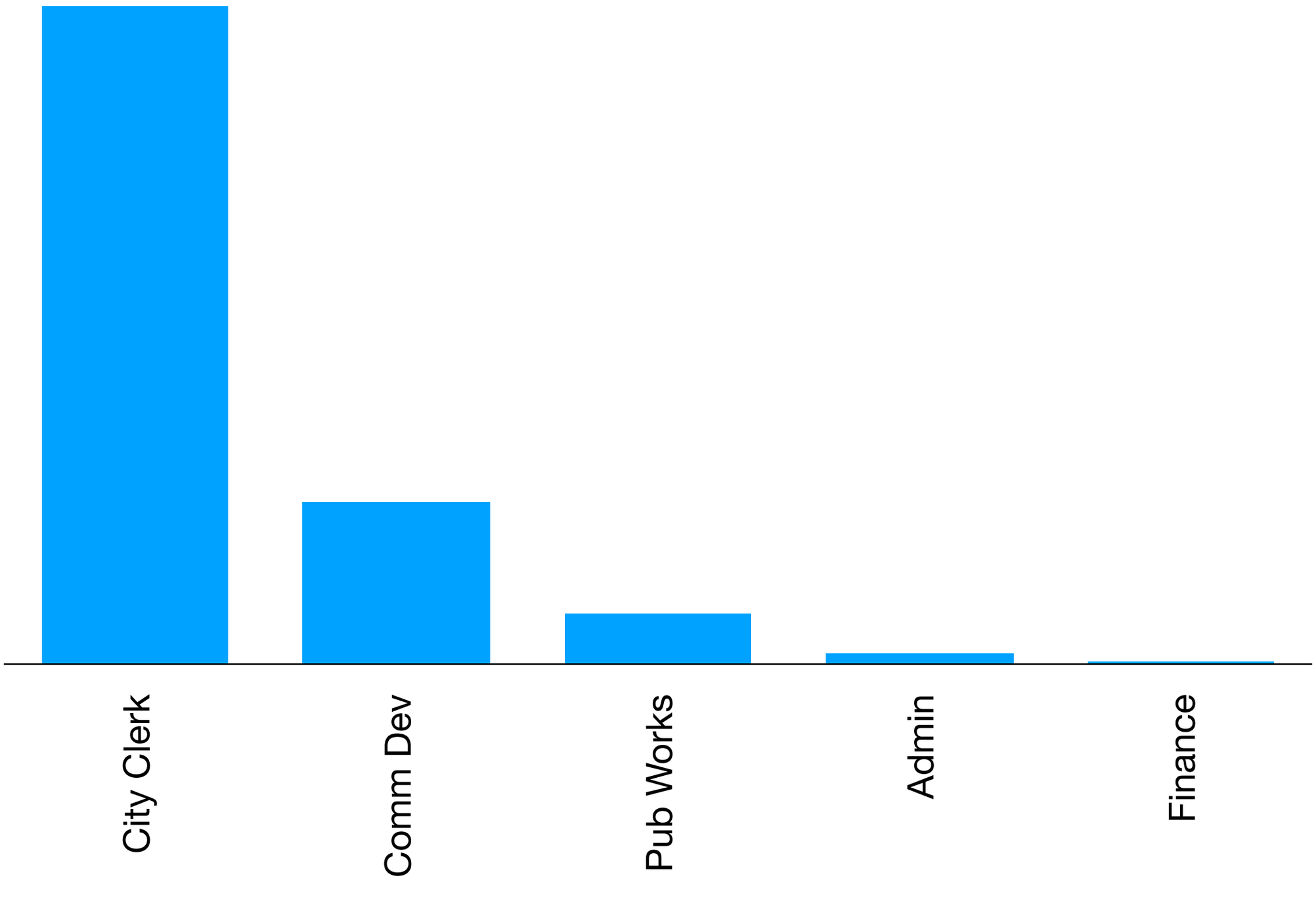
Number of Requests



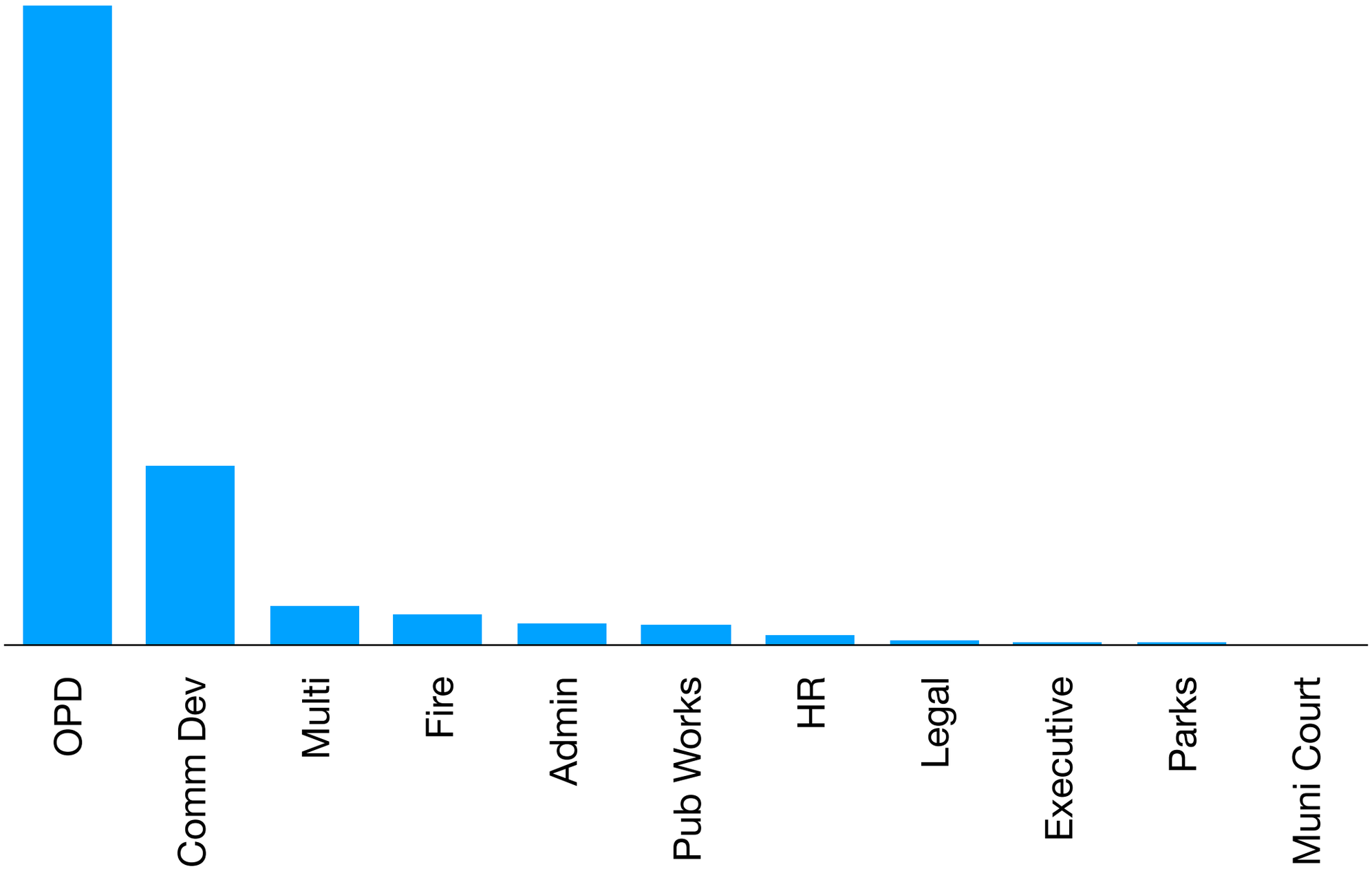
Number of Requests by Department - SEA



Number of Requests by Department - POR



Number of Requests by Department - OLY

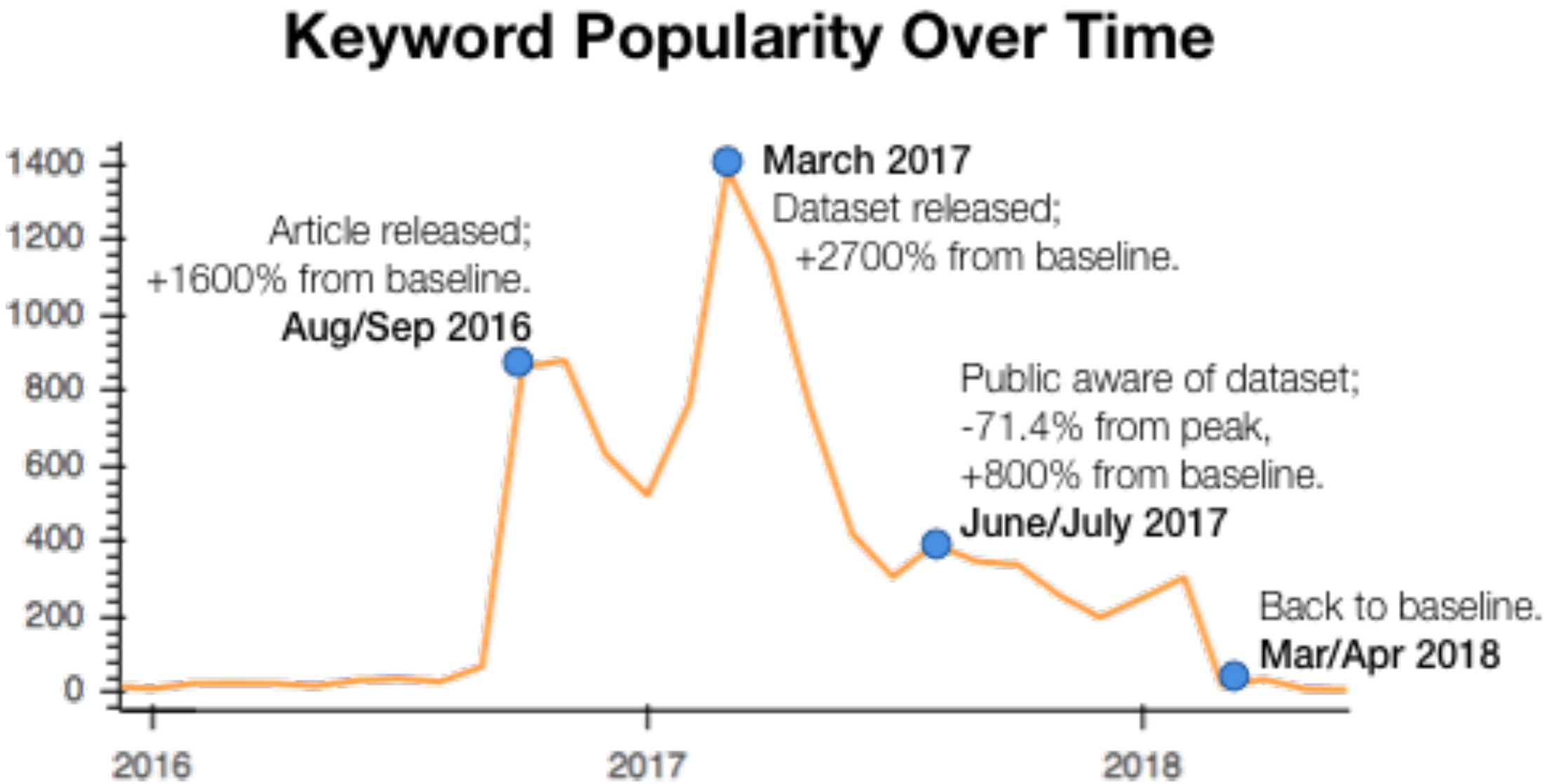


Project Outcomes

Findings

Text takes *many* forms.

We can see history!



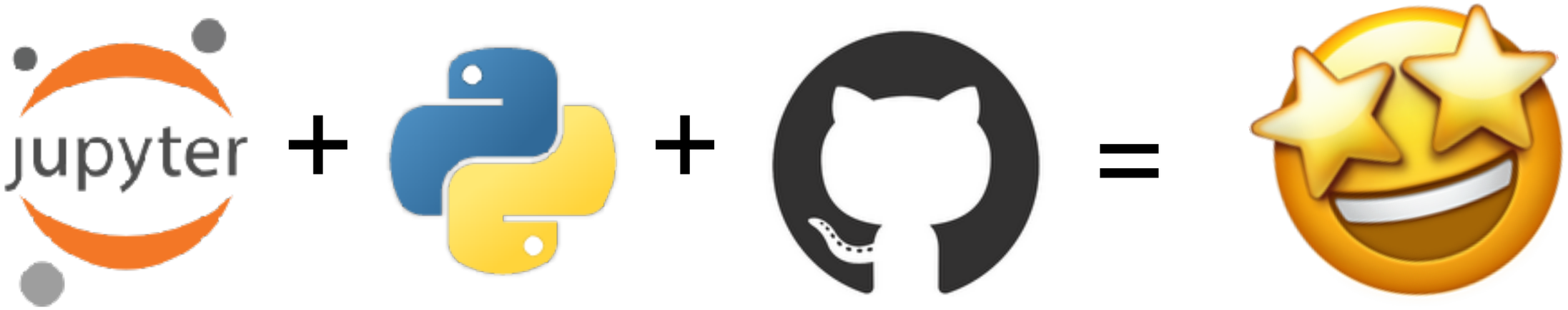
Fruits

Data Analysis Pipeline

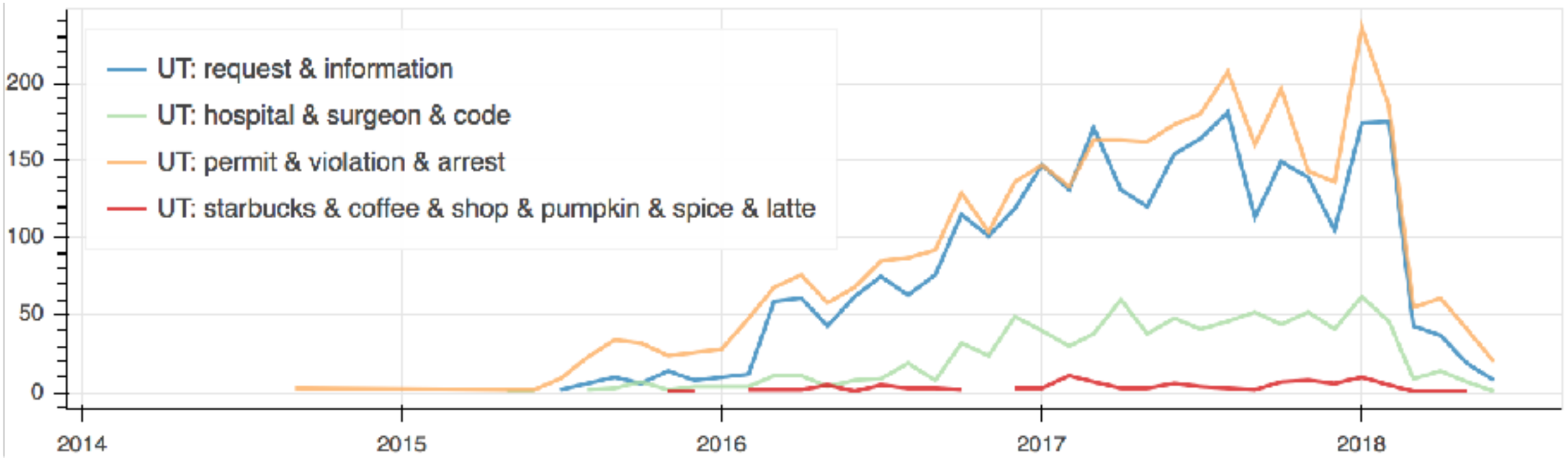
NLP on PRR: Data Processing Pipeline & Theoretical Models

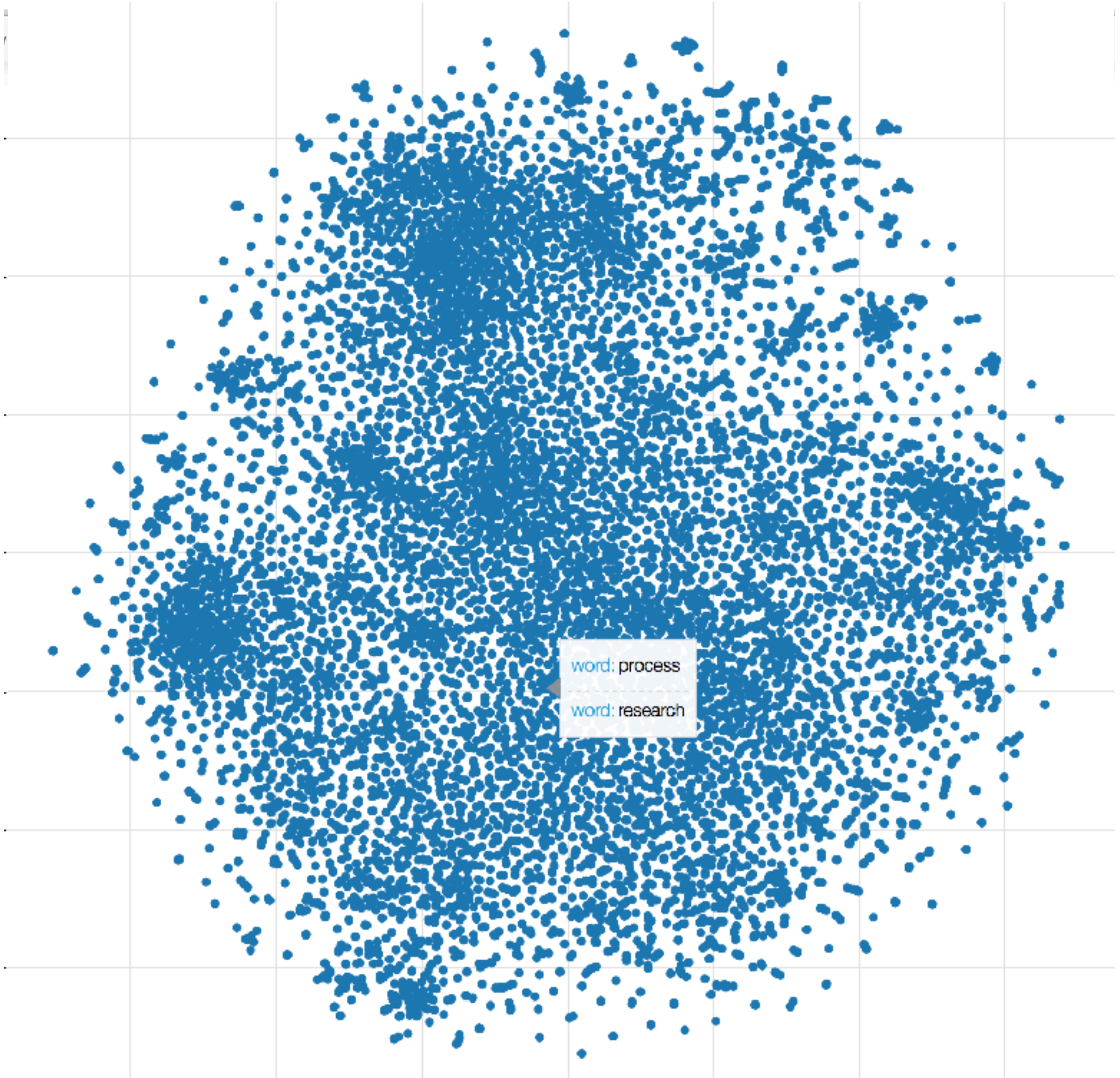
In [1]:
```
import spacy
import pandas as pd
import xlrd
#from sklearn.manifold import TSNE
from spacy.tokenizer import Tokenizer
from gensim.models import Word2Vec, ldamodel
from gensim.models.phrases import Phrases, Phraser
from gensim.corpora import Dictionary
from collections import Counter
import itertools
from spacy.lang.en.stop_words import STOP_WORDS
import re
import numpy as np
import matplotlib.pyplot as plt
import pprint
import time
import os
from sklearn.manifold import TSNE

# this isn't strictly an import, but it's used globally
pp = pprint.PrettyPrinter()

# change the model for different word vectors
nlp = spacy.load('en_core_web_lg')

# and finally, suppress an annoying warning that pandas throws
# when using straightforward indexing methods
pd.options.mode.chained_assignment = None
```

Interactive Visualizations



video
send
record
relate
mirro
underground
protester
dean
incident
corning
lemerande
name
public
wazid
information
christa
jama
decommission
seattle
vigilant
include
ase
tanks
facial
regard
would
department
call
unattended
officer
email
slip
mauricio
make
follow
atkyn
storage
beyond
police
number
datum
multi
reportna


UT: request & information
UT: hospital & surgeon & code
UT: permit & violation & arrest
UT: starbucks & coffee & shop & pumpkin & spice & latte
200
150
100
50
0
2014
2015
2016
2017
2018


word: process
word: research

In [217]:
```
# if you want to try out your own analogy, type it here...
word_analogy('surveillance', 'audio', 'video')
```

Out[217]: 'surveillance is to recordings as audio is to video.'

Why this platform?

Too easy.

Builds enduring skills.

Try it out on GitHub.

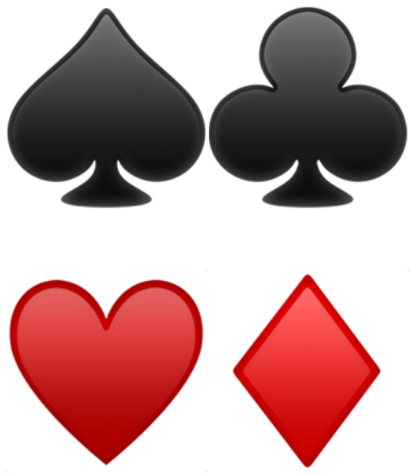
launch binder

Future

Stronger data standards.

Better topic isolation.

!=



Predictive modeling.

?

Any questions?

Thanks for listening!