

Demo: The Accountability Project search tool



Presenter: Jacob Fenton
email us: accountability@irworkshop.org !

Outline

- Outline (2 MIN)
- Talk about [the site](#) (5-10 MIN)
- New data announcement (3 MIN)
- Demo it quickly (10-15 MIN)
- Stuff that doesn't fit (5-15 MIN)
- Questions (REST OF THE CONFERENCE)
- Your searches, there's too much to demo

Privacy, passwords, etc

- Voters and property records are password protected. This is a hard balance to strike.
- There's a link to request a login on the site, it just tells us so we can send you a code. Here's the code: [Removed from tipsheet, click on 'request an account']

Insight: find people / companies / places across data

- Most search sites we know and love focus on a single subject. This makes sense, as custom search should fit the data.
- A “[name](#)” search looks across all datasets; a “[dataset](#)” search only searches that one. Search progression is usually general to specific.
- But we take the opposite tack: we want to be able to surface people, companies and addresses *across* different data sets.
- Simple idea: lots of data searchable in one place
- Standardize on place
- This isn't a replacement for those sites, either, we need them to exist
- We want to add whatever's useful, rather than trying to be “complete”

What's on TAP now? (dumb pun intended)

- [See listings](#) with “browse datasets,” can also filter by state.
- A billion campaign finance records, most states and federal government, but lags. We capture the street addresses in FEC records.
- 100 million voter records. Must be signed in to see these.
- 41 million nonprofit records: directors, grants, contractors. Perhaps best nonprofit data freely available, thanks Scott!
- 39 million property records, general parcel ownership, not sales. Hit or miss: this is mostly east coast states.
- 30 million public employee salary records (some repeat per year). New: Federal Employees.
- Lots more: licenses, planes, doctors, nursing homes. Next slide: thing we're announcing today

NEW: OFAC licenses list

- Little known fact: US sanctions have holes. They are called “licenses”.
- Without sidetracking this demo too much
- Release [story](#) - more [here](#) including download of file and explanatory letter



U.S. Department of the Treasury Office of Public Affairs

Press Release: FOR IMMEDIATE RELEASE
March 2, 2022

Contact: Treasury Public Affairs; Press@Treasury.gov

U.S. Treasury Takes Additional Steps to Strengthen Compliance with Russia-Related Sanctions

Simple example

- Eric Trump [name search](#). Search across all datasets. Filter by collection. Each result is a unique combination of name and address.
- If you click on any, you get the matching results from that dataset. Refine by making minor edits if you want.
- Use a “transaction model” of the data, what is searched are “parties” in a transaction.
- Search results show where *1 party in a transaction matched the search criteria* ; if half of the terms appear in one party and half appear in the other, that’s not a match.

Workflow / Use Case

- Scrambling on deadline for contacts, neighbors, anything
- Research for investigative pieces
- This is better at finding wealthy folks, and people who aren't hiding from the law

But also, we can do things beyond the site's functionality

- "Repeated searches" We can script lookups.
- Other.

Caveats

- To date we strongly favor including vast quantities of data, so we're willing to sacrifice timeliness. Many of the "high value" data sets require considerably more work than just downloading a file (e.g. FOIA-ing a government body).
- We have an administrative process for uploading data
- Although I hate to say it, there are limits on how much stuff we can fit. We haven't hit them yet.

More examples

- [3389 Sable Chase Ln](#)
- [Patrick Sookhdeo](#) (British)
- [Steven Latulippe](#) (Oregon Doctor)
- [Mordechai Korf](#) (Alleged launderer, see [Forward story](#))
- [N419K](#) (Surveillance plane; no “N”)
- [Legacy Emanuel](#) (nonprofit hospital)
- [Multnomah Athletic Club](#) (nonprofit athletic center)

Sex abuser Patrick Sookhdeo still has key role at Barnabas Fund



WILLAMETTE WEEK NEWS • RESTAURANTS • BARS • MUSIC • ARTS • POTLANDER • GUIDES • GET BUSY

Three Homeland Security Flights Circled Portland Protesters. Here Are Their Flight Paths.

“The Trump administration has failed to answer basic questions from June about federal law enforcement surveillance flights over peaceful protests in Portland.”

Is search the right tool?

- Datasette lets us do things differently
- Some published, like this one [on COVID data](#)
- Some [semi-secret](#), it's efficient for us to spin these up. Also see [this](#).



But the data we need is what *you* think is missing (and that we can ethically source)

- We've basically built an infrastructure for processing this sorta thing
- We can help on some preparation stuff, we've seen it all
- I've done the PDF extraction at NICAR talk a number of times, if it's in a electronic PDF it *is* extractable.
- We may need your expertise about state records law, it's awesome if you know the gory details, because chances are I don't