

Civic Data: Local Reps Call (2) Re-cap **Apr 30, 2025**

Decisions:

- Create working groups on specific topics (See conversation in Unified:Civic Data chat)

Next Steps:

- Propose working groups in Civic Data (In progress. See above.)
 - Each working group should have at least (2) volunteer leads to share the workload of coordinating.
- Create new sub-channels for each working group (admin)
- Note: If you are interested in joining a working group, please join Unified 'Civic Data' and let us know!

Agenda/Working doc:

- See:
https://docs.google.com/document/d/12CINqYshMfjvE9kT72dIJsbp3gsZVMsL/edit?usp=drive_link&ouid=105992325138979778362&rtpof=true&sd=tr

Discussion

- Provided a quick overview of prior calls for newcomers
- Began by asking those who have projects to define what their project can offer, what it needs and worked to pair the needs with those who can contribute. See: Project Needs Overview in Agenda/working doc.
<https://docs.google.com/document/d/12CINqYshMfjvE9kT72dIJsbp3gsZVMsL/edit>
- 1. Data Acquisition
 - Discussed scope
 - Emil/Kas: Need for local data beyond top 100 cities
 - Need for list of official/reliable source links for elected officials in local offices.
 - Need is true both for AI approaches and traditional scraper approaches.
 - Discussed potential to use AI/LLM to source sources to augment/speed research of sourcing
 - Some folks have already developed sourcing documents
 - Files shared with group:
 - OpenStates: Local Places - Places - Local Places.csv:
 - https://docs.google.com/spreadsheets/d/1e0BF3ppRKKLM0PXbFP0ebUFOBE5CiLMNtFyzh8sndXE/edit?usp=drive_link
 - Stump: Cities - UScities:
https://docs.google.com/spreadsheets/d/11A_Ok6TTsiRY2I7t6kgmnHpVX-Whmj6dq7gRiP9ktlg/edit?usp=drive_link
 - Sources shared with group:
 - Michelle: LA
<https://voterportal.sos.la.gov/ELECTEDOFFICIALS>
 - Discussed definition of schema for sourcing objects.

- Make suggestions to proposed schema in .csv shared by OpenStates
- Jeff?: Place identifier should be `county-place` to avoid need for disambiguation for two places named the same thing.
- KAS: We should include `last\_updated`, `last\_updated\_by` to schemas in general
- We may want to provide a data model that defines the types in addition to the shared .csv
- Tim: openstates schema allows for 'alias' ids from outside sources.
- Is there anything we can get from existing Google Civic Information team
 - Emil: Shared outreach to Center for Tech and Civic Life (see below)
 - Kas: Reached out again by email... tbd.
- Discussed AI output
 - Jeff: Shared a link to AI output:
 - https://docs.google.com/spreadsheets/d/1nxEHFkMQfFtjI0_XbRNM9tHpqRFSaQYDvCsq4vSmzIY/edit?usp=sharing
- Cost sharing.
 - Est cost:
 - Jeff/Kas: between \$2-6 per record (paid human researchers)
 - Michelle: 0.15 per record (AI)
 - Can cost share by having folks run code using their own AI tokens.
- 2. Data delivery
 - We did not cover this agenda item. TBD
- 3. Entry points for collaboration
 - Decided to start working groups focused on specific goals
 - Working groups to be decided in Civic Data chat on Unified
 - Some potential working groups include:
 - AI Scrapers Working Group
 - Crowdsourcing User Interface Working Group
 - Local Jurisdictions List Consolidation

open civic data local reps -- chat notes.

Jesse Mortenson

1:02 PM

I missed it, who we talking about?

Jeff Pierce

1:04 PM

Hi ! Jeff, co-founded VoteLight.

Jesse Mortenson

1:04 PM

Jesse Mortenson

Plural - Open States

Engineer: Python, JS, etc.

Dheeraj Chand

1:04 PM

Dheeraj Chand, Siege Analytics, Scala, Python, PostGIS

You

1:04 PM

Agenda:

https://docs.google.com/document/d/1TVctb4DAbRAR9WTQOE1aBi1DgyzBz_7DEh4nm9JBXiU/edit?tab=t.0#heading=h.xd64l3ogwhzd

keep

Pinned

Marjorie Roswell

1:04 PM

Margie. Python. JS, GIS.

Brian Hedden

1:04 PM

Brian Hedden, Founder of <https://stump.vote> - built on bubble.io

Michelle Truong

1:04 PM

Michelle - developer (ruby/js) - volunteer/no org @ civicpatch.org

Tim Showers

1:05 PM

Tim Showers - CTO GovHawk / Openstates Contributor -- Python, PHP, Scraping.

Henry Lai

1:05 PM

Hi, I'm the cofounder of Populist - we aggregate election, politician, and legislation data and help newsrooms create content from that data (we helped Minnesota Public Radio create these candidate guides for a recent special election

<https://www.mprnews.org/story/2025/04/22/minnesota-senate-district-6-special-election-heintzeman-slipy>)

Virginia Dooley

1:05 PM

Virginia Dooley, python - most recently at Democracy Club

Amy Farrell - Q2AADDA

1:05 PM

Amy Farrell, GSA, PHP, js, whatever.

I'm just here to see what's up; probably won't have time to contribute and will just be tracking the project.

Emil Guareno

1:05 PM

Emil - Founder @ Uproot

Javascript, Node, MongoDB

You

1:07 PM

We're focused on office holders on this call, correct? There is a geo call coming up...

You

1:09 PM

For me, the scope is whatever my user needs... which most of the time falls out of the top 100 cities as most users live in peri-urban areas.

Emil Guareno

1:10 PM

same ^ it would be great to let anyone know regardless of where they live that they can find information about their elected official

Jeff Pierce

1:13 PM

@Emil @Kas - Ok, consensus seems to be all local places with a local government (city, town, CDP with govt, etc.)

Henry Lai

1:15 PM

Populist is open to contributing our data but we have our own custom schema

You

1:17 PM

@henry lai we can help write a mapper for your data if you like....

Henry Lai

1:18 PM

we do store IDs from other APIs in our database, like VoteSmart

You

1:19 PM

I like the id of essentially assigning maintainers to focus on specific sets of data within the full data set.

Tim Showers

1:19 PM

@Henry – the openstates schema allows for 'alias' ids from outside sources.

Frank Gould

1:20 PM

Is there anything we can obtain from Google that they used for the existing Civic Data?

Emil Guareno

1:22 PM

Hi Emil,

Thank you again for expressing your interest in our Governance Project data. I'm following up to update that after careful consideration, we've decided to not stand up our Governance Project data in an API.

In light of the Google Civic Information API sunsetting its coverage of representative data, we'd like to share BallotReady, Ballotpedia, and Cicero as alternative data vendors that offer officeholder data and candidate data in their respective APIs.

If you have any questions, p

You

1:23 PM

Yeah... what Emil said:) I did send a formal email.

Emil Guareno

1:23 PM

They never responded after that, they sent me this email back in September 13th 2024

Michelle Truong

1:23 PM

Civic Patch is WA/OR (with plans to expand outward)

You

1:24 PM

Unified will like purchase data and work with you all to scale in the following way:

for each users get: city council, school board, prosecutors, judges.

But not scaling horizontally. Working to get more depth for users where we are missing information.

Dheeraj Chand

1:26 PM

Henry, talk to me off line re: campaign finance data.

Henry Lai

1:26 PM



Michelle Truong

1:27 PM

Civic Patch - needs to scale -

- * AI/Scraper approach

- * figure out state-level source for city officials data (optionally) - potentially maintaining a maximum of 50 scrapers

- * \$\$ for calling 3rd party APIs OR volunteers (if we spread the load between each person it can look like \$5-10 per person)

- * Does NOT include CDP/government

Henry Lai

1:27 PM

@dheeraj what's your email?

Dheeraj Chand

1:29 PM

Dheeraj@siegeanalytics.com

You

1:30 PM

Henry... former journalist. Have a direct connection to a large community.

Dheeraj Chand

1:30 PM

Telephone / text is faster +12025507110

Henry Lai

1:31 PM

@Kas awesome, i love meeting and networking with journalists :)

You

1:32 PM

On unified we have tooling to create "Actions" --- so a user could verify their reps on Unified.

You

1:33 PM

Michelle... I think we could programmatically generate that list of sources and then vet it.

You

1:34 PM

Some of the best sources for this data (Texas directory, NYC green book, etc.) are paid.

Tim Showers

1:38 PM

<https://www.nlc.org/membershipold/state-municipal-leagues/>

You

1:39 PM

Tim: Get with Creyton Loud.... he is collecting this I believe.

Tim Showers

1:39 PM

thumbs up

You

1:39 PM

I think we should pair sourcing by ocd-ids

You

1:41 PM

Candidate data and current incumbent office holder data are very, very different.

Frank Gould

1:41 PM

Does Populist also include appointed officials details?

You

1:42 PM

Different sources, the candidate that won, may or may not be the office holder 1-2 years into their term...

Dheeraj Chand

1:42 PM

Guys I have to jump unexpected, but please do be in touch, I have some hours to give and a lot of familiarity with the problem

Henry Lai

1:48 PM

@Kas yes candidate data and office holders are different, but we also do live election results in MN, so we track who wins the elections and update our database accordingly

You

1:49 PM

@Henry, yes.. that works for sure. The issue is when they resign. (happens more than you think)

Henry Lai

1:50 PM

very true, but typically there is then a new special election so then we track those - we actually have a bunch of special elections coming up in MN in the next couple weeks

but yeah, maintaining our database is also a challenge :)

You

1:53 PM

Jesse/Tim: I did review your code/repo... we use a similar stack at Unified.

Jesse Mortenson

1:53 PM

+1

Jeff Pierce

1:53 PM

@Henry Lai: Let's talk, we work with news organizations. Use jeff@votelight.com. Thx!

Henry Lai

1:54 PM

@Jeff Pierce that's great, I will definitely reach out, thanks!

Jeff Pierce

1:54 PM

Ok!

You

1:55 PM

@henry.... sql or no sql db?

Jesse Mortenson

1:56 PM

we can work that piece out, for sure

You

1:58 PM

Crowd sourcing could be no code, low-code solution.. i.e. Airtable?

Henry Lai

1:58 PM

@Kas we use GraphQL

You

1:58 PM

(has an api. so can be integrated into apps)

Jesse Mortenson

1:59 PM

agree my guess is we're pretty sparse on muni

You

2:03 PM

Tim: ... maybe define a schema? Pydantic model, data classe or some such for source objects.

Jesse Mortenson

2:04 PM

I think Michelle found that cities can exist in more than one county :)

Michelle Truong

2:04 PM

Imao

It didn't fit neatly into folders! gnis is what I'm using.

Michelle Truong

2:05 PM

++1 I forgot to mention grabbing city websites is needed for the state-level source lists as well.

You

2:06 PM

Jeff: Noted we can ask both Ballotpedia and Ballotready to join our next call.

Tim Showers

2:06 PM

https://docs.google.com/spreadsheets/d/1r66wDQd_jWFxVfwHvrNJA3CdI0gOZn5hEEZYUyD08LA/edit?usp=sharing

You

2:08 PM

Tim: Can we add last_updated, updated_by...

Jesse Mortenson

2:08 PM

+1

You

2:11 PM

Michelle: I ask my friend Ray to join the Civic Data group on unified. Ton of experience doing LLM evaluation (also self-hosting open source LLMs)

I want to try to connect you with him. Will try to get him on the demo call to start.

Brian Hedden

2:11 PM

This is the list of cities I have

<https://docs.google.com/spreadsheets/d/1GVRSPn3TCW8DgoHI35WAWUoestsLcxzcYGZJjk3AnEQ/edit?usp=sharing>

You

2:12 PM

Brian -- ok if i download that and share it with the group?

Michelle Truong

2:12 PM

adding a self-hosted LLM is something I'd like to look into for the project, but I don't have any experience of that. Would be great to connect, thank you.

Brian Hedden

2:12 PM

Re: Kas - yes, go for it

You

2:13 PM

@brian Can you give me permission?

Jesse Mortenson

2:14 PM

could be used as a test/validation dataset - does your new app generate stuff that matches it great minds :)

Brian Hedden

2:15 PM

@Kas - should work now

You

2:16 PM

Brian – tks!

You

2:18 PM

Yeah it was \$6 per prosecutor district ... they were tough.

Virginia Dooley

2:18 PM

Need to jump off, it's late here. Looking forward to the next meeting/steps. Thanks for organising Kas, Jesse.

You

2:18 PM

Tks, Virginia!

I'll send out notes.

Henry Lai

2:18 PM

are there email notifications for the Unified group? like when there is activity in it?

Michelle Truong

2:19 PM

I can't tell you the costs currently since I'm using a more targeted approach and making use of ChatGPT free tiers, but at its worst the AI+scraper cost was \$0.15.

You

2:19 PM

No, not yet, but there are mobile notifications if you have the mobile app.

Your feature suggestion has been noted!

Michelle Truong

2:20 PM

Louisiana has one!

<https://voterportal.sos.la.gov/ELECTEDOFFICIALS>

Tim Showers

2:20 PM

I need to run too, thanks all!

You

2:20 PM

Thanks Tim... any next steps you want to suggest?

You

2:22 PM

Because our group has enough people in it... we could consider calling state SOS offices and doing a public records request.

FOIA can be slow but they work.

We can even automate the FOIA request.

Michelle Truong

2:22 PM

Are they obligated to do this? They pointed me to a county last time I asked.

Over email though. And not very assertively.

You

2:23 PM

Yes, they are required to respond to FOIA requests.

Henry Lai

2:23 PM

State SoS offices are not obligated to do this

You

2:23 PM

States have different laws w/r/t to FOIA requests, but many do honor them.

Henry Lai

2:24 PM

oh yeah sorry they are required to respond, i thought you were asking about something else

You

2:24 PM

Guys... call is set to end in 6m.. should we start to wrap up?

Henry Lai

2:24 PM

i mean they aren't obligated to have records that contain that data

You

2:24 PM

@henry... yes you are right on that...

1. use NJ as test case for source URLs

You

2:26 PM

2. Use NH or another state for checking veracity of AI

Jeff Pierce

2:29 PM

AI OUTPUT LINK:

https://docs.google.com/spreadsheets/d/1nxEHFkMQfTjI0_XbRNM9tHpqRFSaQYDvCsq4vSmzIY/e
dit?usp=sharing