🌻 Animated social map of US news

Data source and processing

Data source

We used **Google News RSS search** (site-filtered by source domains).

- Data route: Google News RSS query endpoint (`news.google.com/rss/search`)
- Scope: US-English (`hl=en-US`, `gl=US`)
- Source control: fixed source-site allowlist in the collection script
- Proportional sample sizing: January 2020 to February 2026

Query-term list (US politics focus):

`trump`, `covid`, `jobs`, `employment`, `economy`, `biden`, `vance`, `musk`, `kamala harris`, `maga`, `republican`, `democrat`, `white-house`, `president`, `election`, `supreme court`, `voter`, `votes`, `voting`, `campaign`, `senate`, `house of representatives`, `congress`, `epstein`, `tariffs`, `doge`, `ice`, `immigration`, `deportation`, `border`, `politics`, `party`.

A row is only kept if the assigned `query_term` appears in that row's title or snippet text.

Processing pipeline

1. Build month-by-month RSS queries by term + source site + `after/before` date bounds.
2. Fetch RSS results in small parallel batches with retry/backoff and pacing.
3. Parse headline, link, date, source label, and short snippet from each item.
4. Keep only rows where the `query_term` is present in title/snippet text.

5. Deduplicate by URL and story key.

6. Convert publication date to UTC-like timestamp string.

7. Draw a month-balanced, source-balanced sample to `top_n`.

8. Save monthly checkpoints to output CSV during the run.

9. If rerun with an existing output CSV, resume from the month after the latest completed month in that file.

Target row count is proportional to the selected timespan (anchored to 12,000 rows for Jan 2020 to Feb 2026).

Link identification for social map

From sampled text, links are coded using the same social-link logic described here !:

- Nodes represent social actors (people, groups, institutions).
- Directed links represent actor-to-actor relationships (not abstract causal mechanisms).
- Relationship labels/tags are attached to each link.
- Sentiment and time labels are added where possible.

The resulting table is converted to app-compatible link columns (`Cause`, `Effect`, `Tags`, plus metadata) for mapping.

Animation method

Animation follows the Causal Map app's **Animate filter** behavior documented in `causal-map-extension/webapp/README.md`:

- Select one link field to define frames (for example month or another time field).
- Optional cumulative mode includes all earlier frames up to the current frame.
- Frame construction respects pipeline order, so filter ordering affects what appears in each animation step.
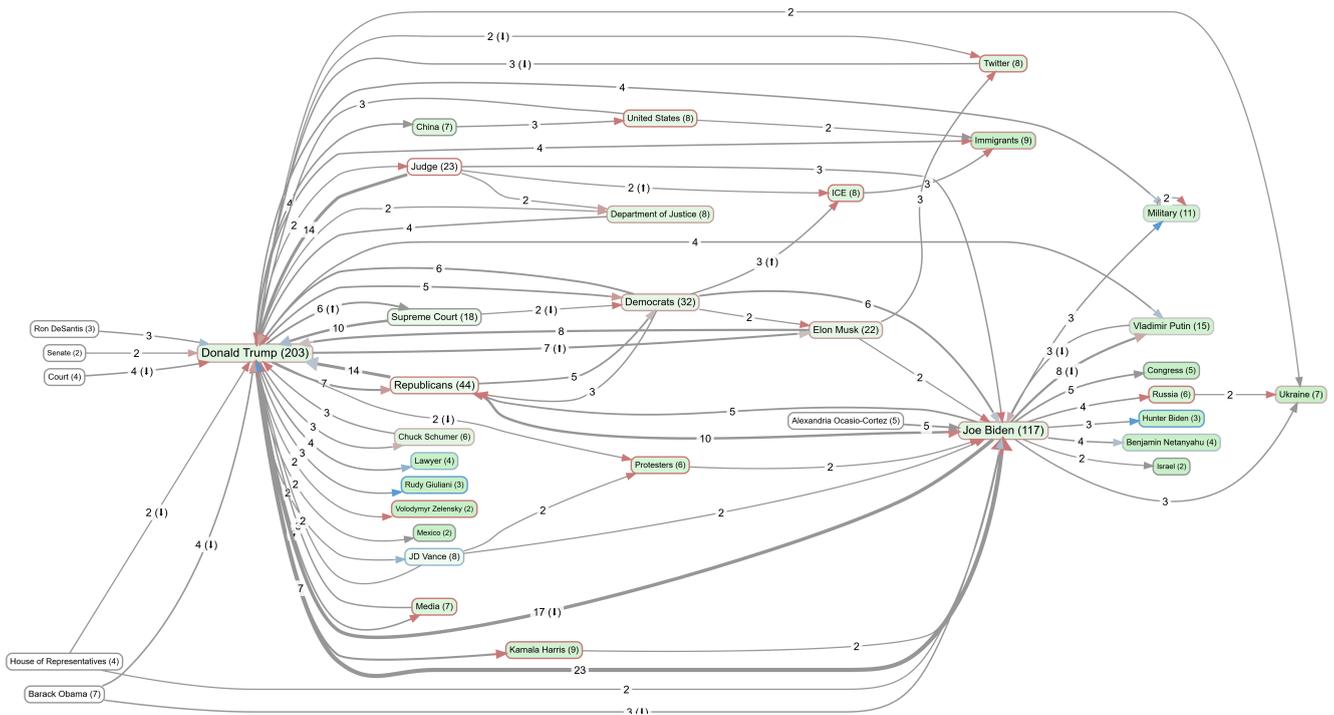
Cumulative animation

Filename: us-news2. Citation coverage 16% of all sources: 316 citations shown out of 1920. Factors — size: citation count; numbers: citation count; colour: outcomeness; border: avg incoming sentiment (blue=positive, grey=neutral, red=negative). Links — width: citation count; labels: custom links label; arrowheads: effect sentiment (blue=positive, grey=neutral, red=negative). Filters applied: Sources included: All sources. Factor freq: top 38 by citation count; Link freq: minimum 2 citations; Animate: s_month. Bookmark #1209 2026-02-24 09:03

Attribution note

Google News RSS is an aggregated discovery feed, not a publisher ranking API. Terms such as "top" in this workflow refer to **sampling rules in processing** (filters and balancing), not editorial prominence.

Significant changes over time

Filename: us-news2. Citation coverage 16% of all sources: 316 citations shown out of 1920. Factors — size: citation count; numbers: citation count; colour: outcomeness; border: avg incoming sentiment (blue=positive, grey=neutral, red=negative). Links — width: citation count; labels: custom links label; arrowheads: effect sentiment (blue=positive, grey=neutral, red=negative); ↑ increasing over numeric groups; ↓ decreasing. Filters applied: Sources included: All sources. Factor freq: top 38 by citation count; Link freq: minimum 2 citations; Animate: s_month; Custom link label: s_month, mode=chisq-counts, counts=citations, ordinal. Bookmark #1208 2026-02-24 08:59