Weekly-aggregated records for Facebook. Filename: fb_week_##_YYYY.csv. These records were collected from the official Meta CrowdTangle API. The coverage of the data includes all posts by state legislators' accounts during the period January 2020 to December 2021. As of this version, the DAPR database does not include any data on Facebook posts beyond December 2021.

Column	Description
official_id	Unique identifier for each official in the dataset.
urls	List of the URLs to each post
calendar_week	ISO week and year (YYYY-WW)
total_posts	The total number of posts made by the official in week XX of a given year YYYY.
avg_total_interactions	Average number of interactions on the posts
avg_likes	Average number of likes on the posts
avg_comments	Average number of comments on the posts
avg_love	Average number of loves on the posts
avg_wow	Average number of wows on the posts
avg_haha	Average number of hahas on the posts
avg_sad	Average number of sads on the posts
avg_angry	Average number of angrys on the posts
avg_care	Average number of cares on the posts
avg_shares	Average number of shares of the posts
Terms Columns	Frequency of each word, hashtag, and account mention within the official's tweets. Represents the vocabulary used across tweets and serves as features in a TFIDF model.

Weekly-aggregated records for X/Twitter. Filename: X_week_##_YYYY.csv. From January 2020 through the end of May 2023, every tweet posted by a U.S. state legislator was collected using the Academic Twitter API. In June 2023 our access to the Academic Twitter API was shut down as X moved to a paid model for API access. In Fall 2023 we received a grant from the National Science Foundation (awards 2318460 and 231846) to support the expansion of the scope of officials we were collecting, and funds to purchase X API access. In June 2024 we resumed collection of X posts. In the data beginning in 2024 we collect data from a random sample of 1,000 elected officials each month. The fields represented in the data are the same. The two key differences in the 2024 data are that (1) we cover all elected officials included in the database, not just state legislators, and (2) we collect posts from a random sample of officials, not all of them, each month.

Column	Description
official_id	Unique identifier for each official in the dataset.

urls	List of the URLs to each post.
calendar_week	ISO week and year (YYYY-WW)
total_posts	The total number of posts made by the official in week XX of a given year YYYY.
avg_retweets	Average number of retweets among the posts
avg_likes	Average number of likes among the posts
avg_replies	Average number of replies among the posts
avg_quotes	Average number of quote tweets among the posts
Terms Columns	Frequency of each word, hashtag, and account mention within the official's posts. Represents the vocabulary used across posts and serves as features in a TFIDF model.

Metadata on officials. Filename: official_data.csv. This dataset is current in its coverage of elected officials in the U.S. as of December 2024.

Column	Description
official_id	Unique identifier for each official
name	Name of the official
party	Party affiliation (e.g, Democratic, Republican)
race	Racial coding of the official, available for select officials
yr_elected	Year in which the official was elected
bp_url	Ballotpedia URL
vote_pct	Vote percentage in the latest general election
yr_vote	Year in which the latest election took place
state_fips	State Federal Information Processing Standards (FIPS) codes
lastname	Last name of the official
firstname	First name of the official
gender	Gender of the official
OCDID	Open Civic Data Division Identifiers
state	State abbreviation
office_name	Ballotpedia office name, available for select officials
office_level	Ballotpedia office level, available for select officials
office_branch	Ballotpedia office branch, available for select officials
district_name	Ballotpedia district name, available for select officials
district_type	District type (e.g., State Legislative, Congress, State)
first_post_fb	Date of first Facebook post in the database for the official
first_post_X	Date of first twitter/X post in the database for the official
num_follower	Number of followers on X, available for select officials
num_following	Number following on X, available for select officials