

IOWA

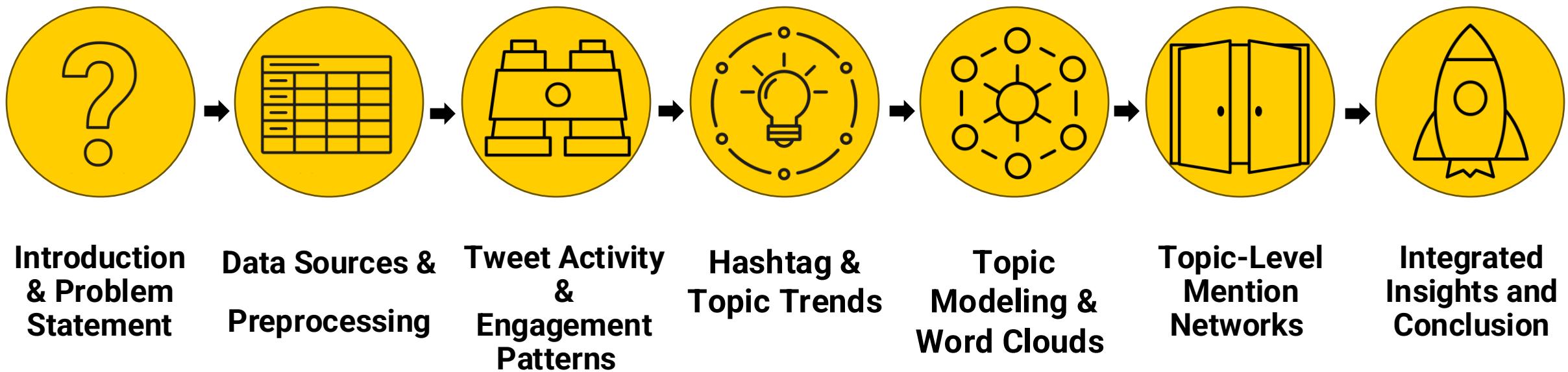
Group 6

Analyzing Political Discourse on Twitter

Lucas Assis, Kaleb Sailer, Bereket Ware

December 3rd, 2025

Agenda



Research Questions & Motivation

To guide our analysis, we begin by clarifying our focus and why this time period matters.

Overview

- We analyze how political conversations emerged on Twitter between **Oct 2012–Dec 2014**.
- Focus on **Congressional accounts** and their public communication patterns.
- Combine **text analysis**, **temporal analysis**, and **social network analysis**.
- Goal: Understand **what topics were discussed**, **how engagement evolved**, and **who influenced information flow**.

Problem Statement / Aim

- Investigate how members of Congress used Twitter during 2012–2014.
- Identify **issue salience, changes in engagement, and party-level differences.**
- Measure how information spreads through **mention networks.**
- Understand which actors functioned as **central communicators.**

Why This Analysis Matters

- Captures political communication during Twitter's early rise in government use.
- Provides historical baseline before the major shifts of 2016+.
- Rich dataset (**245K tweets**, 548 user accounts).
- Combines **network and text analytics**, the two pillars of modern social analytics.
- Insights applicable to issues like **polarization, agenda setting, and public engagement**.

Implications of Results

- Reveals early digital influence patterns inside Congress.
- Highlights major topics and issue framing strategies.
- Identifies how engagement differed across parties over time.
- Shows which political figures acted as bridges in conversation networks.
- Helps explain how political narratives spread online.

Data Sources & Preprocessing

Before analyzing political behavior on Twitter, we first built a clean dataset that captures congressional activity from 2012–2014.

Data Sources

- Monthly Twitter archives from **Oct 2012 – Dec 2014** (27 months).
- **245,858 tweets** after removing retweets.
- **548 Congressional accounts** with party labels (D, R, I).
- Joined tweet-level and user-level datasets.

Sample of Data

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
1	user_id	created_at	screen_name	name	description	location	lang		followers_count	friends_count	listed_count	statuses_count	verified	protected	geo_enabled	time_zone	utc_offset	url	party
2	2.915E+09	#####	AkGovBillW	Governor B	Official Twit	Alaska	en		5870	550	190	115	1642	TRUE	FALSE	TRUE			http://t.co/I
3	33537967	#####	amyklobouti	Amy Klobouti	U.S. Senator from Minn	Minneapolis	en		170717	91751	2493	24	5567	TRUE	FALSE	FALSE	Central Time	-18000	http://t.co/D
4	1.378E+09	#####	AnthonyBrc	Anthony G. Congressm	Rep. Mitchellville	Arkansas	en		6067	369	170	90	1878	TRUE	FALSE	TRUE	Eastern Time	-14400	https://t.co/D
5	269992801	#####	AsaHutchin	Gov. Asa Hutchinson	Rep. Husband of Arkansas	Arkansas	en		25747	598	420	440	3778	TRUE	FALSE	TRUE	Central Time	-18000	https://t.co/R
6	234797704	#####	AustinScott	Rep. Austin Scott	I am proud	Ashburn, Georgia	en		13048	313	735	12	2022	TRUE	FALSE	TRUE			http://t.co/R
7	82453460	#####	BennieGThompson	Rep. Thompson	Bolton, Mis	Mississippi	en		6707	428	484	27	756	TRUE	FALSE	FALSE	Eastern Time	-14400	http://t.co/D
8	55677432	#####	BillCassidy	Bill Cassidy	U.S. Senato	Baton Rouge	en		20716	1187	1257	80	4526	TRUE	FALSE	TRUE	Central Time	-18000	http://t.co/R
9	26103389	#####	BillHaslam	Gov. Bill Haslam	Governor o Tennessee	Tennessee	en		70622	7851	1209	25	1649	TRUE	FALSE	FALSE	Central Time	-18000	http://t.co/R
10	15394954	#####	boblatta	U.S. Rep. Bob Latta	Rep. Bob Latta	Ohio	en		29159	5854	1267	468	5785	TRUE	FALSE	TRUE	Eastern Time	-14400	http://t.co/R
11	30216513	#####	BradSherman	Rep. Brad Sherman	Rep. Brad Sherman	California	en		23126	512	912	0	1117	TRUE	FALSE	FALSE	Pacific Time	-25200	https://t.co/D
12	47747074	#####	brianschatz	Brian Schatz	United States Senator	Washington	en		33198	1923	892	889	4508	TRUE	FALSE	TRUE	Hawaii	-36000	https://t.co/D
13	36396752	#####	ButchOtter	Idaho Gov	If you are in Boise, Idaho	Idaho	en		16878	1662	555	15	1450	FALSE	FALSE	TRUE	Mountain Time	-21600	http://t.co/R
14	305620929	#####	Call_Me_Di	Dutch Rupp	U.S. Congre	Washington	en		11206	1350	616	244	2431	TRUE	FALSE	TRUE	Quito	-18000	http://t.co/D
15	3.298E+09	#####	CaptClayHiggins	Captain Clay Higgins	Louisiana, Len	Louisiana, Louisiana	en		4442	112	63	771	716	FALSE	FALSE	FALSE	Pacific Time	-25200	https://t.co/R

Data Preprocessing

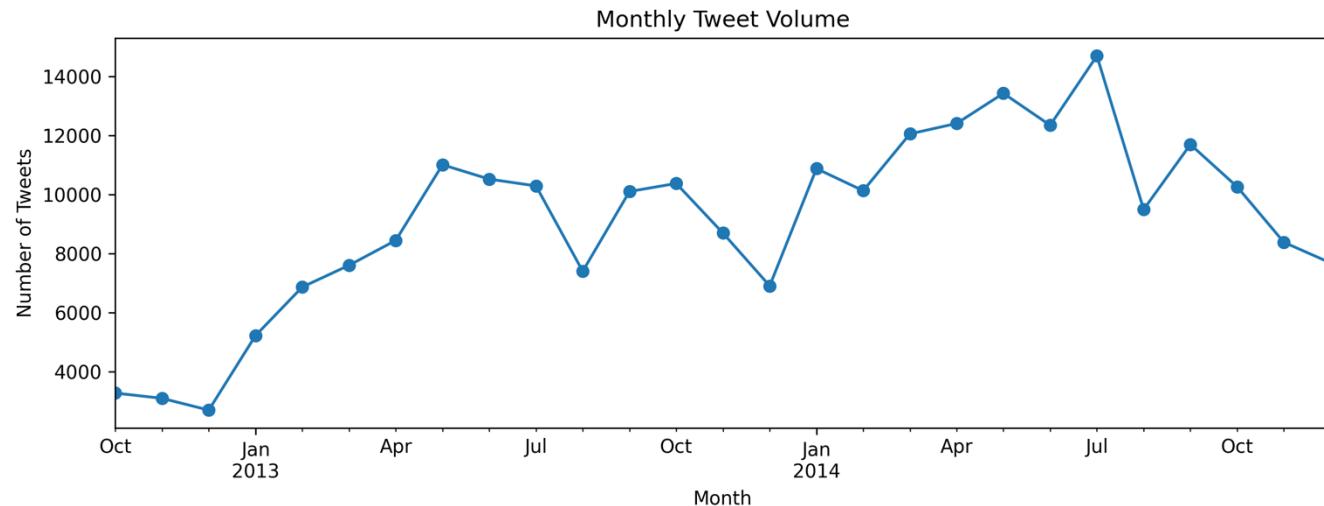
- Converted timestamps → monthly periods for temporal analysis.
- Removed retweets to focus on **original messaging**.
- Extracted hashtags and mentions.
- Created engagement metric: **likes + retweets**.
- Prepared text for topic modeling (stopwords, tokenization, filtering).

Tweet Activity & Engagement Patterns

With our data prepared, we begin by analyzing how active Congressional accounts were and how audiences engaged with them over time.

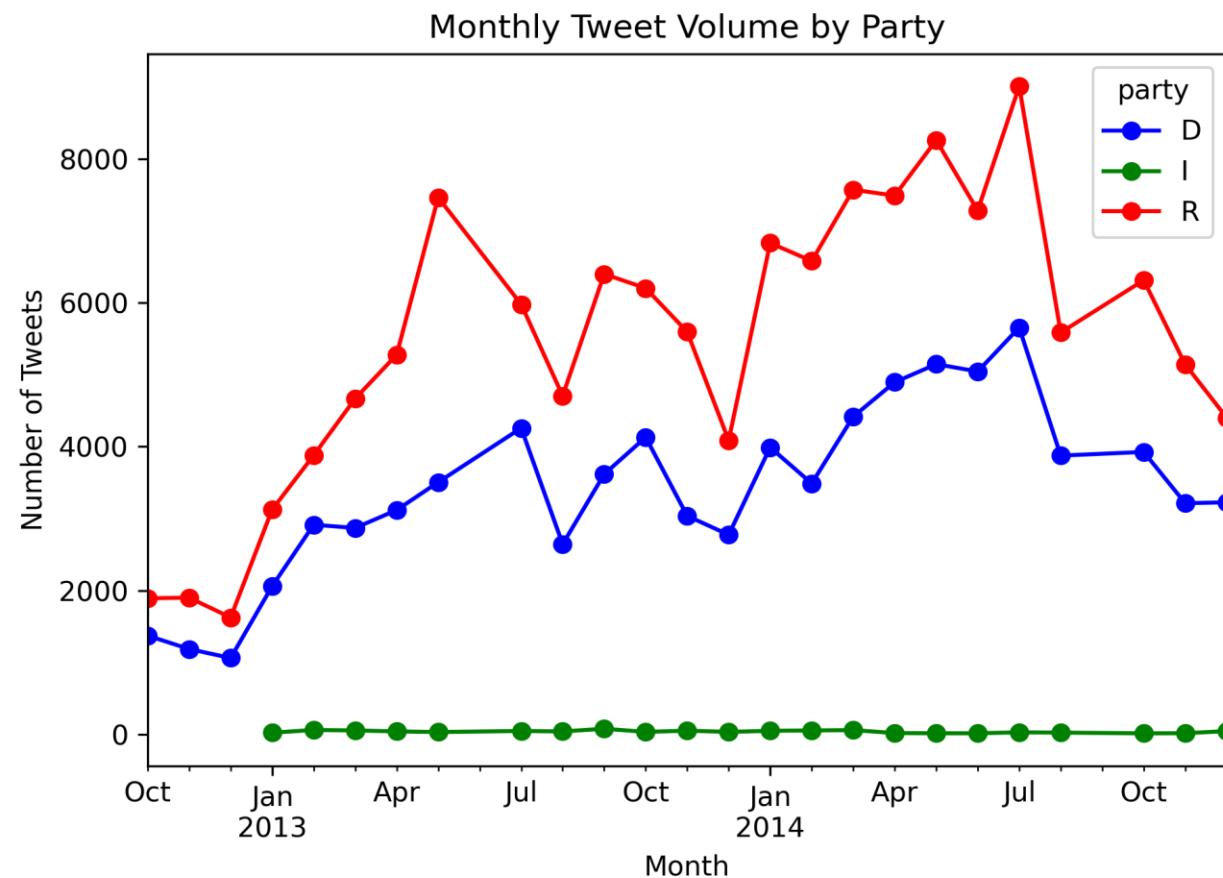
Tweet Volume Over Time

- Tweet activity rises sharply through 2013–2014.
- Peaks correspond to major legislative periods.
- Consistent activity provides strong foundation for longitudinal analysis.



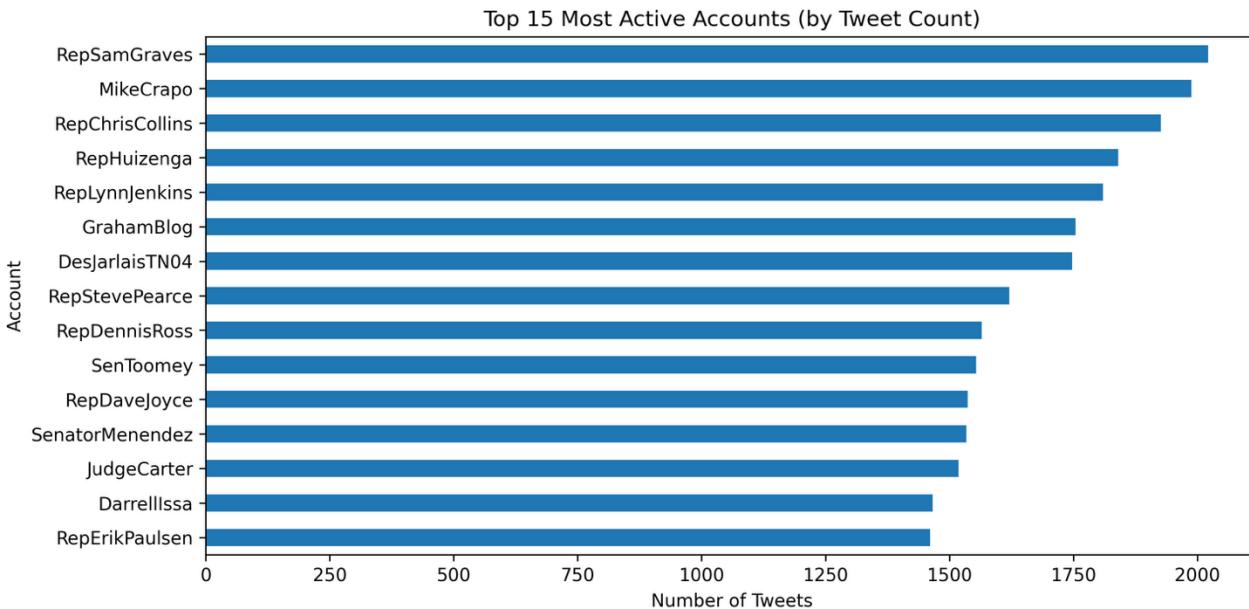
Tweet Volume by Party

- Republicans consistently tweeted more than Democrats.
- Independents show minimal volume but still included.
- Reflects different communication strategies.



Top 15 Most Active Members

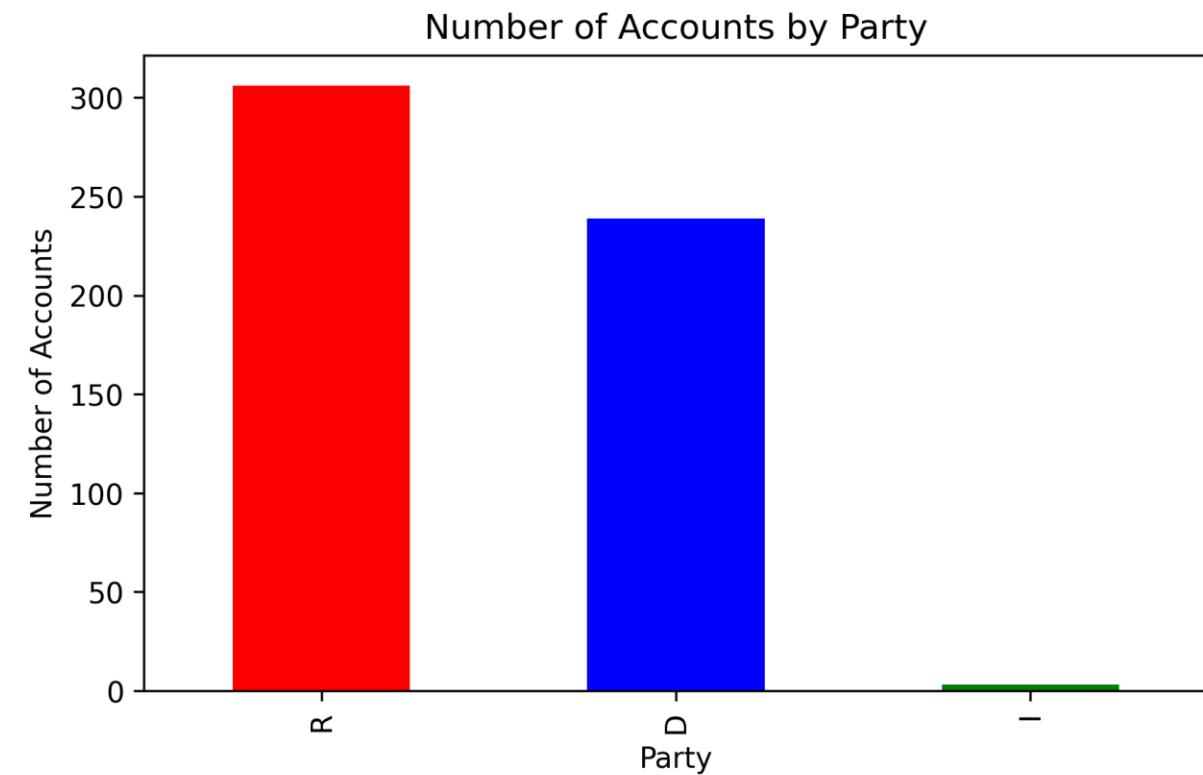
- Many top accounts belong to House members.
- Activity concentrated among a small subset of users.
- Important for interpreting influence and network structure.



Accounts per Party

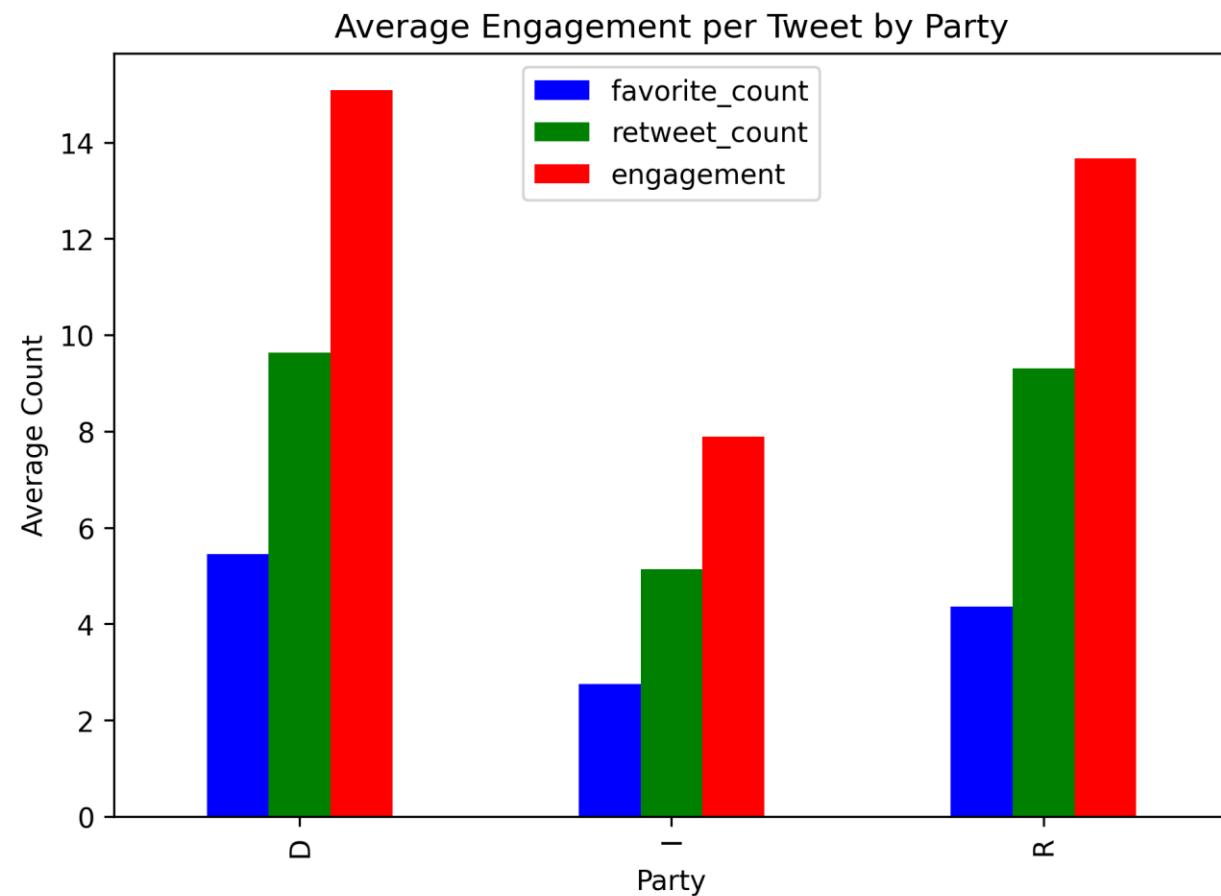
Number of Congressional Accounts by Party

- Republican accounts: 306
- Democratic accounts: 239
- Independent accounts: 3
- Reflects seat distribution and social media adoption.



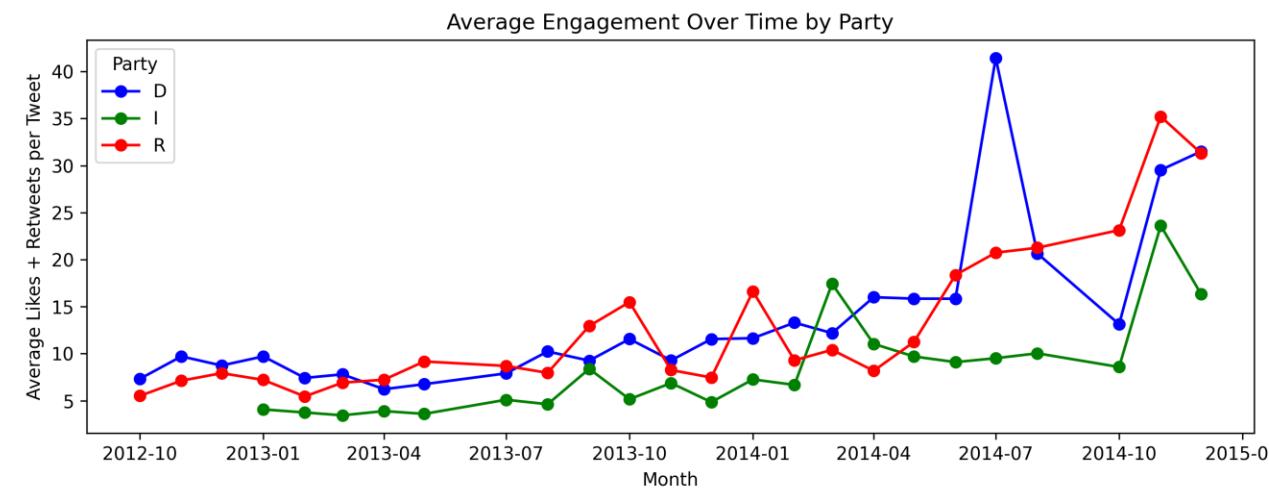
Engagement by Party

- Democrats receive the highest engagement per tweet.
- Republicans also receive strong engagement despite higher tweet volume.
- Independents show lower levels due to smaller audience base.



Engagement Over Time

- Engagement rises significantly in 2014.
- Spikes correlate with policy debates and national events.
- Suggests growing use of Twitter to mobilize public sentiment.

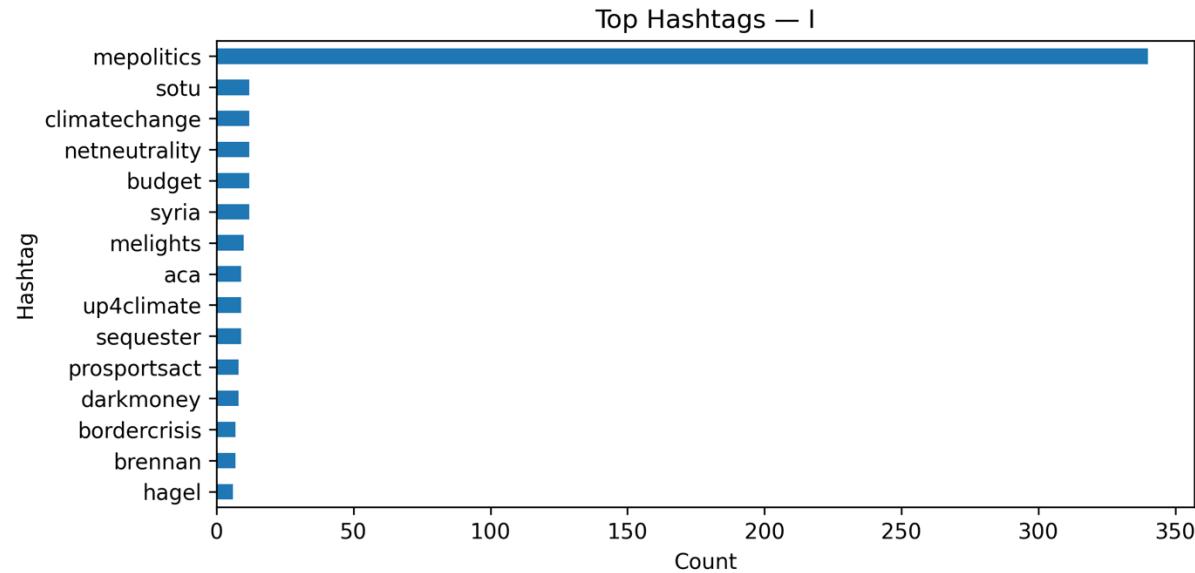


Hashtag Trends

Next, we explore the key issues that dominated congressional communication through hashtag analysis.

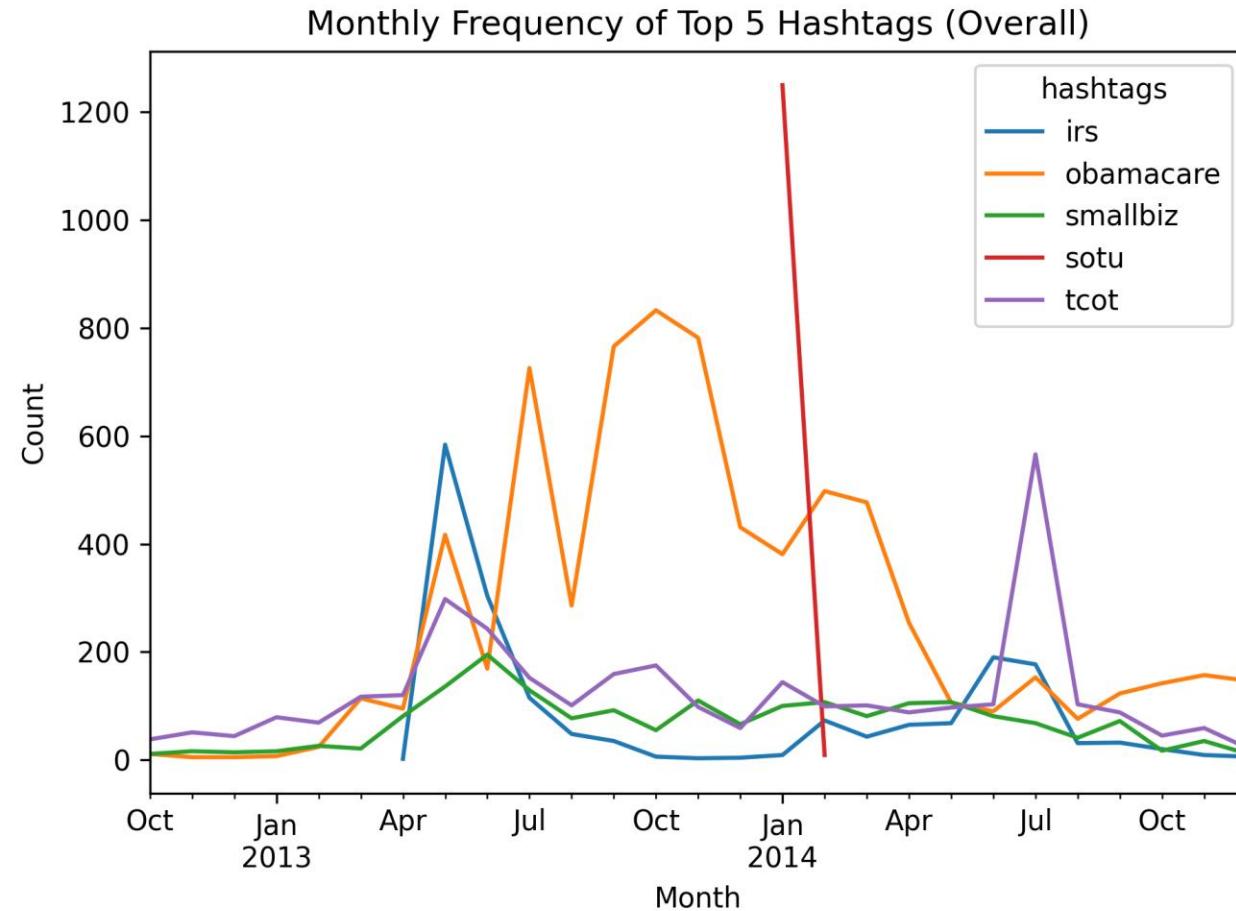
Most Frequent Hashtags

- #obamacare dominates discussion.
- Other major topics: #tcot, #smallbiz, #sotu, #irs.
- Represents policy priorities and partisan messaging.



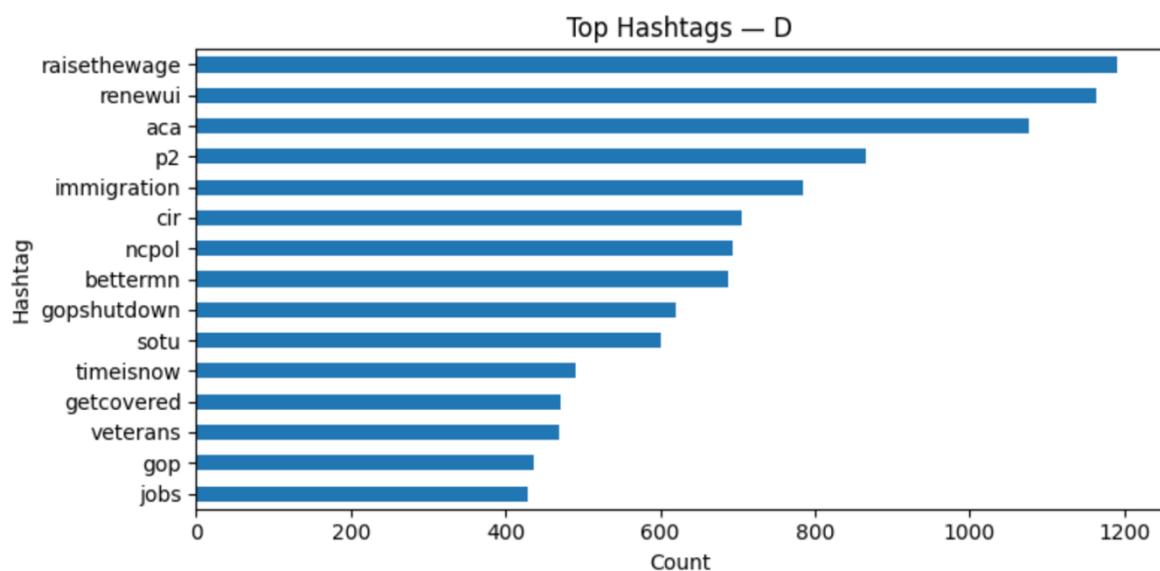
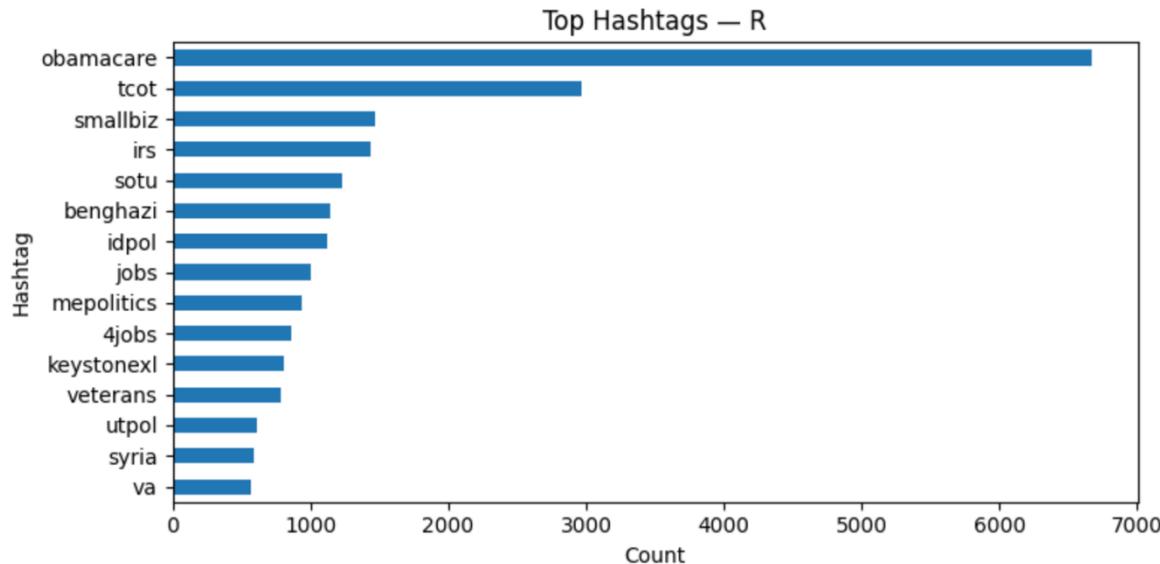
Issue Attention Over Time

- #obamacare shows repeated spikes during legislative debates.
- #irs peaks during early 2013 hearings.
- Issue salience shifts month to month.



Hashtags by Party

- Republicans emphasize #obamacare, #irs, #benghazi.
- Democrats highlight #raisethewage, #renewui, #veterans.
- Indicates clear partisan framing.

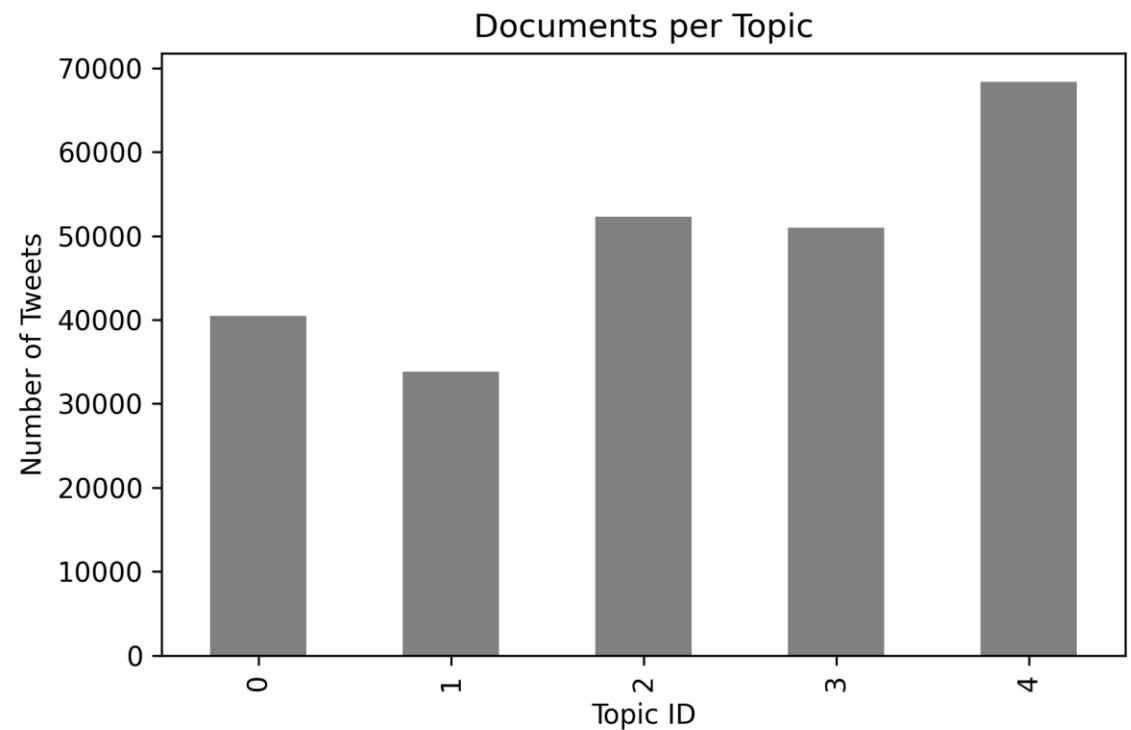


Topic Modeling, Word Clouds & Networks

To synthesize themes across hundreds of thousands of tweets, we apply topic modeling to identify the major issues being discussed.

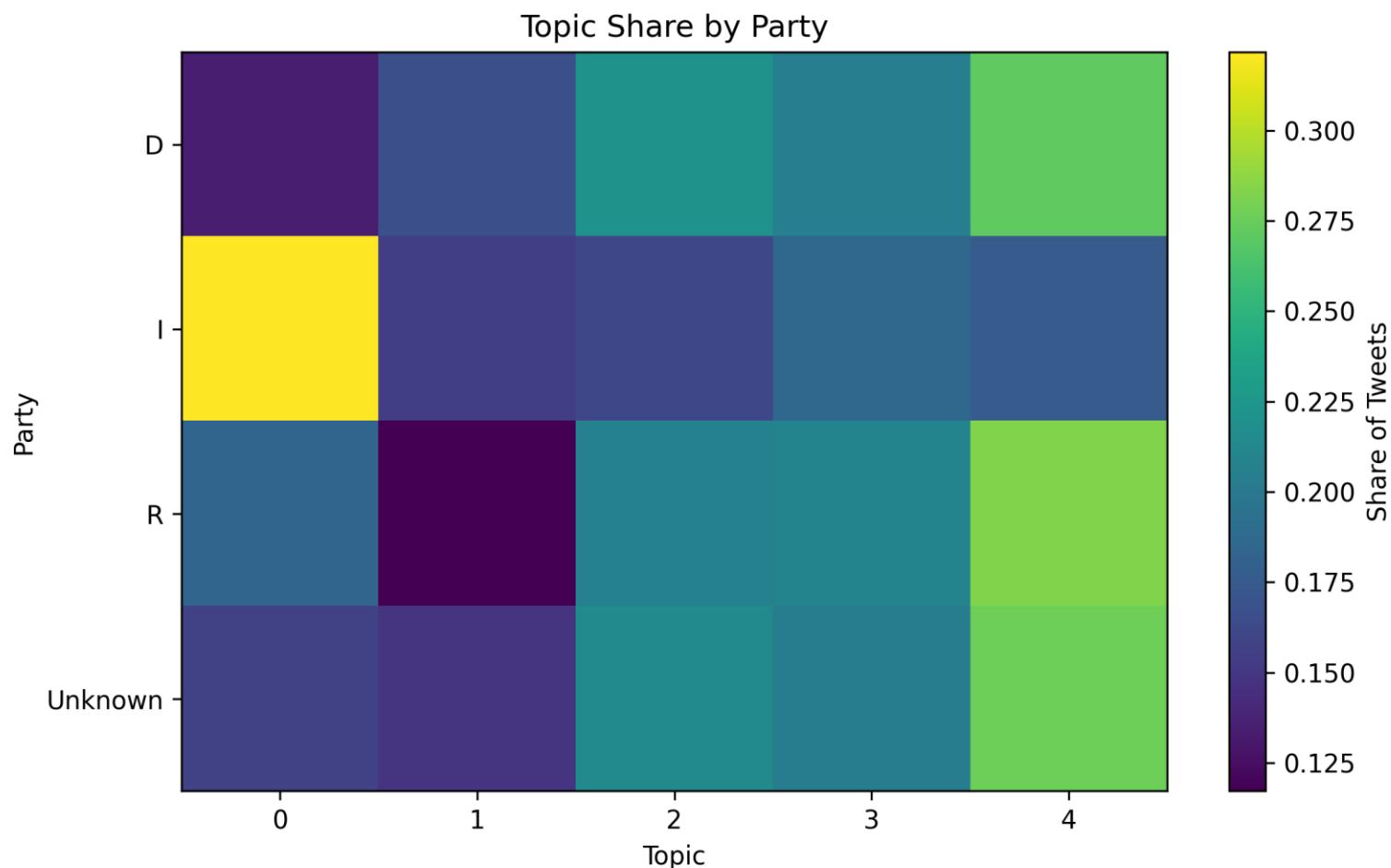
LDA Topic Overview

- Extracted **5 major topics** representing broad issue clusters.
- Topics include: hearings/live updates, jobs/economy, military/service, school/work, and Obamacare/policy debate.
- Large sample ensures stable topic structure.



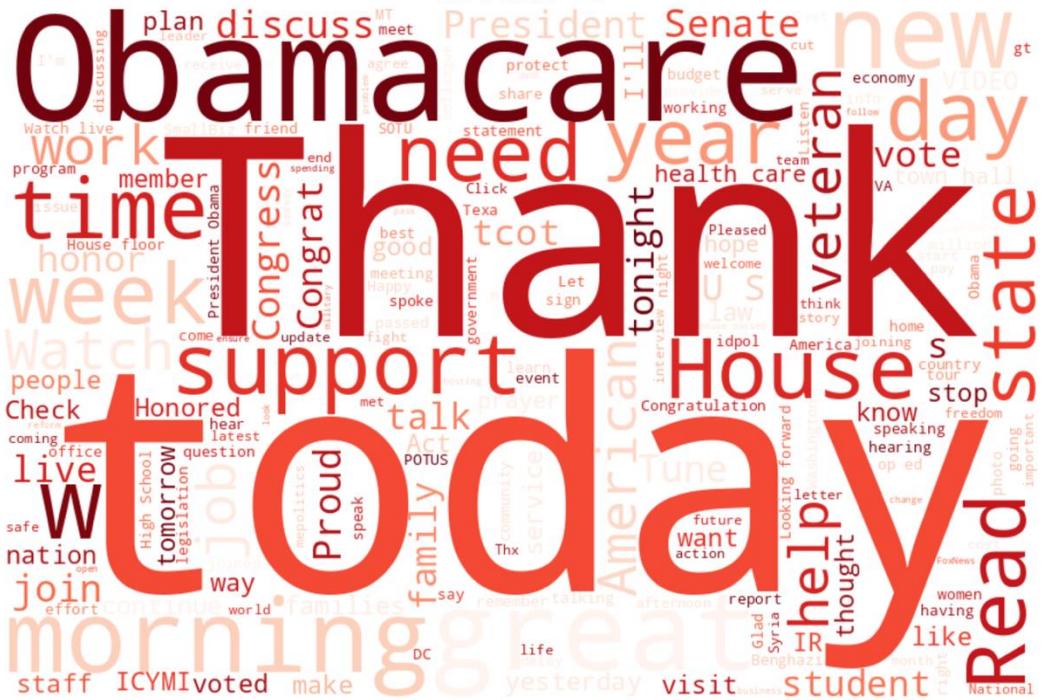
Topic Share by Party

- Both parties contribute to all topics but with varying emphasis.
- Republicans strongly associated with **Topic 4** (Obamacare).
- Democrats more present in **Topic 1** (jobs, economy).
- Shows partisan issue prioritization.



Word Clouds by Party

Word Cloud — R



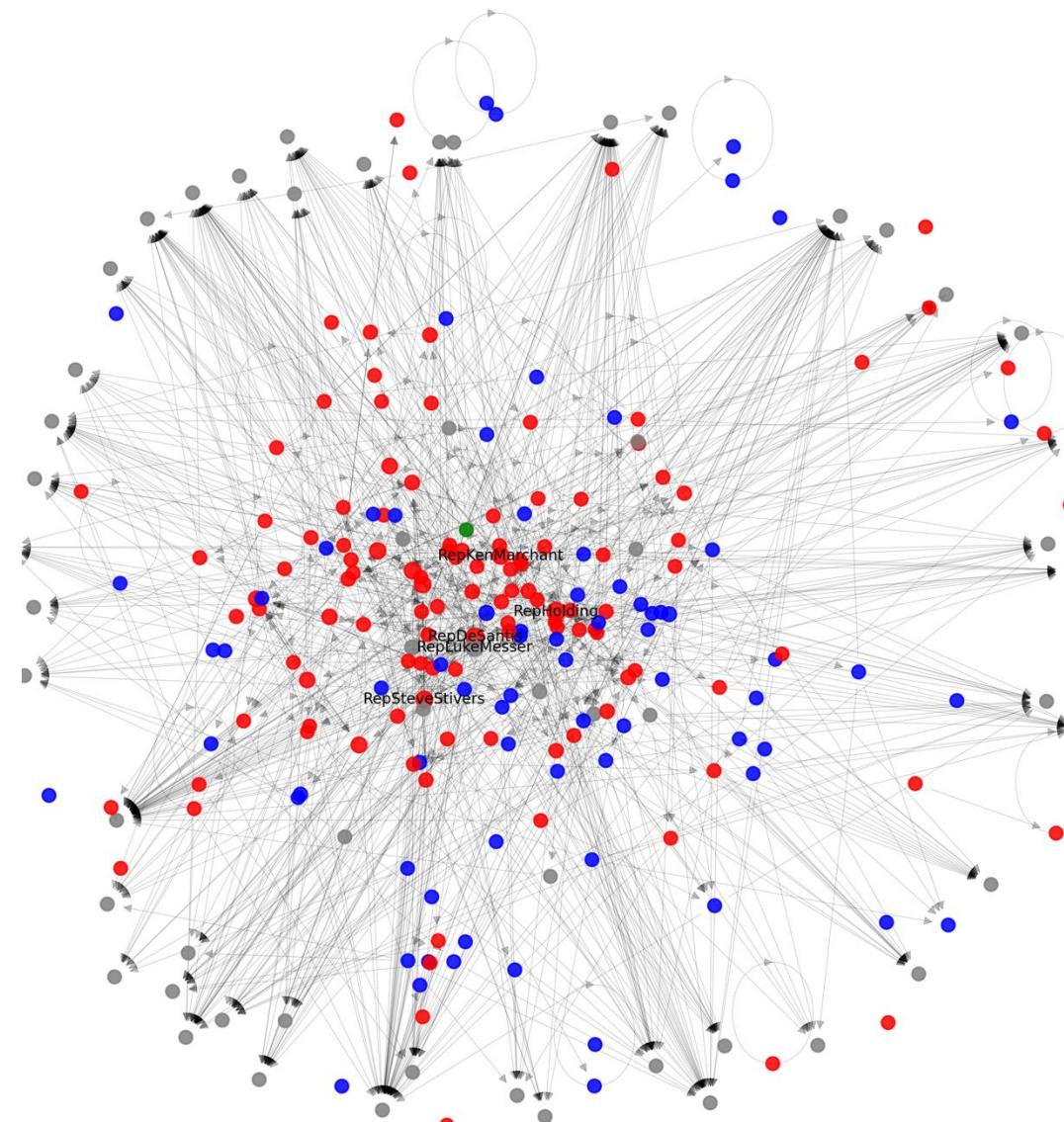
Word Cloud – D



IOWA

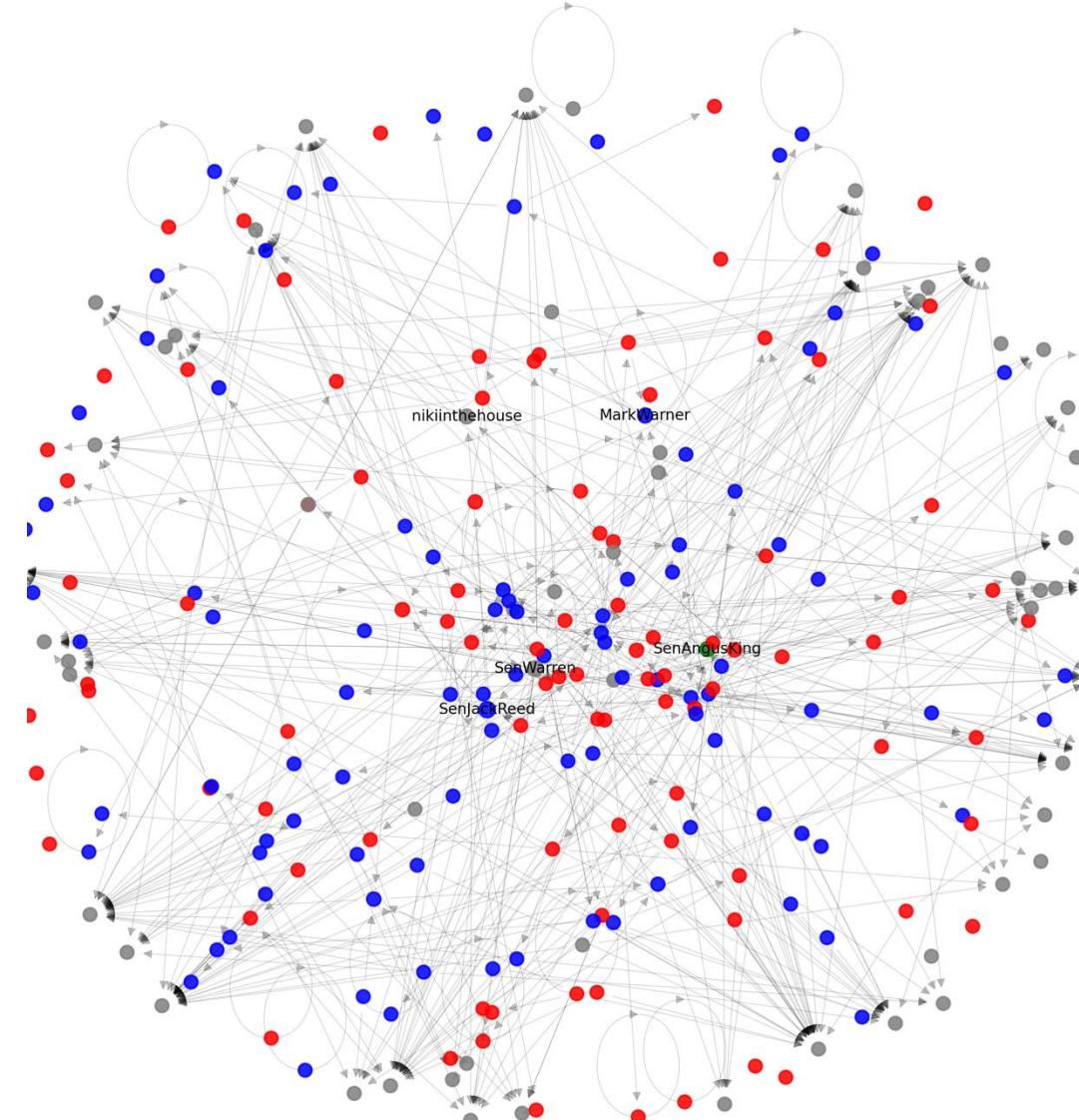
Topic 0 – Hearing/Event Updates

- Denser network with many cross-party mentions.
- Key bridges: RepHolding, RepDeSantis.
- Reflects procedural updates & announcements.



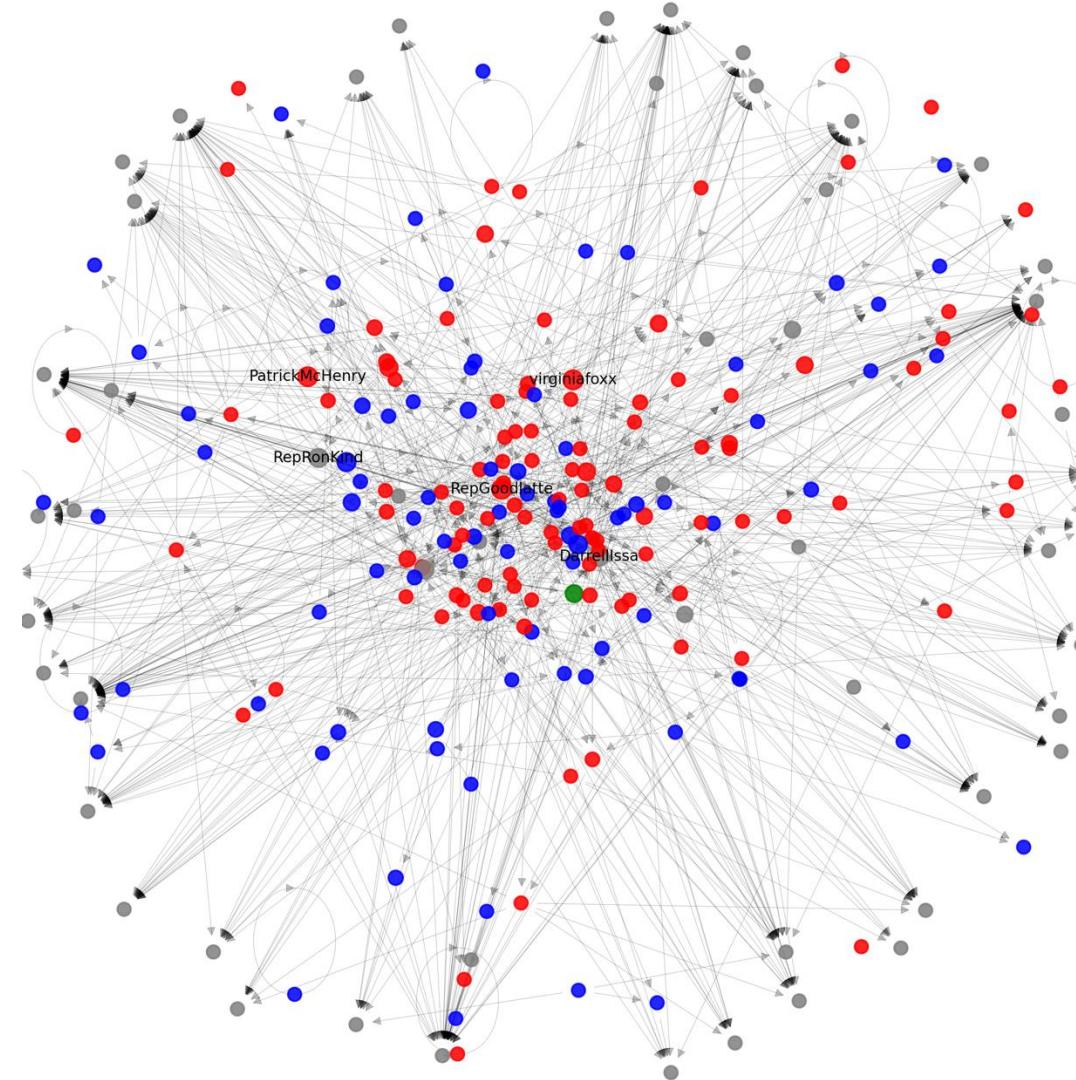
Topic 1 – Jobs & Economy

- More distributed structure.
- Key connectors: SenJackReed, SenWarren.
- Reflects bipartisan focus on economic messaging.



Topic 4 – Obamacare Debate

- Highly polarized with strong clustering.
- Top influencers: RepGoodlatte, virginiafoxx, DarrellIssa.
- Demonstrates partisan amplification.



Limitations & Conclusion

What our analysis cannot fully capture, and final takeaways on congressional political communication

Limitations

- No visibility into private interactions.
- Topic modeling abstracts nuanced political language.
- Mention networks based on explicit mentions only.

Lessons Learned

- Doing network analysis on this large of a scale is difficult and time consuming.
- Understanding of the value network analysis can bring in uncovering insights.

Key Takeaways

- Republicans tweet more, but Democrats receive higher engagement.
- Issue cycles heavily influence engagement spikes.
 - State of the Union, Election Seasons
- Mention networks identify influential communicators & bridge nodes.
- Congressional Twitter use became more structured and strategic over time.

IOWA

Group 6

Questions?

Thank You!

IOWA