Who Talks to Whom: Quantifying Echo Chamber Effects in Emerging Social Media Platforms

Mao Li[1], Xinyi Chen[1]
[1]University of Michigan - Ann Arbor
maolee@umich.edu, chloexc@umich.edu

Abstract

This study investigates interaction patterns on three emerging social media platforms in the "post-Twitter era"—BlueSky, Mastodon, and Truth Social—focusing on reply networks as highly interactive, opinion-expressive forms of engagement within a polarization context. We employ a multi-method approach to quantify the echo chamber effect, analyzing (1) the role of influential users in shaping narratives (using PageRank and LexRank), (2) ideological distribution within reply network communities (leveraging the Leiden algorithm for community detection and LLMs to assess ideological leanings), and (3) patterns of connectivity favoring ideologically similar users (via biased random walks). To ensure rigorous cross-platform comparability, we apply strict controls on topic and timeframe. Our findings reveal a relatively weaker echo chamber effect on Mastodon and BlueSky, though the platforms display subtle differences: Mastodon fosters a diverse ideological landscape with frequent cross-ideological exchanges, while discussions on BlueSky are more centered around moderate voices. In contrast, Truth Social exhibits a pronounced echo chamber effect, with the majority of users supporting right-leaning or pro-Trump narratives, in alignment with the platform's founding intent as a dedicated space for Trump supporters. Despite this homogeneity, we observe occasional instances of critical engagement and fact-checking on Truth Social, suggesting that even within ideologically driven spaces, counter-narratives can emerge, though their influence remains uncertain. These findings underscore the need for further research into the unique dynamics of emerging social media platforms, particularly those shaped by ideological leanings. Understanding the roles echo chambers play in shaping public opinion is essential for gauging their impact on (mis)information diffusion and political polarization in an increasingly fragmented digital landscape.

Code — https://github.com/LiMaoUM/echo-chamber
Datasets — https://doi.org/10.5281/zenodo.17410680

Introduction

The rise of social media has transformed how individuals consume information and interact with one another, shifting users' roles from passive consumers to active content creators and opinion shapers (Kwak et al. 2010; Bode 2016; Del Vicario et al. 2016). Social platforms facilitate broad interactions, yet their structures often foster environments where users engage primarily with like-minded individuals, reinforcing existing beliefs and forming what are commonly called "echo chambers" (Garimella et al. 2018a,b; Garrett 2009; Jamieson and Cappella 2010). Within these virtual spaces, users may not realize the extent of homogeneity in their interactions, leading to an insular exchange of ideologically aligned information—a phenomenon that selective exposure and confirmation bias theories help to explain (Klapper 1960; Nickerson 1998). This echo chamber effect is further amplified by the growing polarization online, particularly around political topics (Persily and Tucker 2020). As political landscapes have polarized, especially in democratic contexts, social media has become a fertile ground for reinforcing group narratives, reducing exposure to diverse perspectives, and potentially exacerbating political divides.

The situation is complicated by recent shifts within Twitter, once viewed as a potential "digital public sphere" (Kuncoro et al. 2024). In early 2022, former President Donald Trump launched Truth Social following his bans from mainstream platforms, prompting many of his followers to leave Twitter (THE ASSOCIATED PRESS 2024). This migration marked an ideological shift in Twitter's user base, which was further affected by Elon Musk's acquisition and rebranding of the platform as X. Musk's relaxed content moderation policies, controversial statements, and platform changes have pushed more centrist and left-leaning users toward alternatives like Mastodon and Bluesky (Grant 2022). Although Twitter continues to report a highly active daily user base, independent estimates indicate a marked decline in users following the platform's acquisition (Hern 2024). This decline suggests a migration of users to alternative social media platforms, potentially leading to a more ideologically segmented online environment.

These events have led to the emergence of ideologically distinct social media spaces. Truth Social, for instance, has drawn a primarily conservative audience and can be broadly categorized as an alt-tech platform—a niche but increasingly influential type of social media often focused on "free speech" while hosting far-right content (Dehghan and Nagappa 2022). In contrast, platforms like Mastodon and BlueSky emphasize decentralization and community-driven governance, attracting users who seek alternatives to corporate-controlled platforms. For some, these alternatives offer appealing new features or moderation policies, while others simply prefer a space outside Twitter's shifting landscape under Musk's ownership (Faverio and Anderson 2025; Odabaş 2023; Roose 2024; Cava, Aiello, and Tagarelli 2023).

In light of these ideology-driven trends, it is crucial to study the interaction patterns on these emerging social platforms, with particular focus on the presence and extent of echo chambers. This focus stems from the fact that echo chambers and political polarization reinforce each other, creating a self-perpetuating feedback loop (Sunstein 1999; Iandoli, Primario, and Zollo 2021). Investigating whether this feedback loop is mitigated or amplified within evolving social media landscapes is a crucial area of research.

Prior research indicates that even when exposed to ideologically diverse content, users preferentially engage with like-minded information, reinforcing echo chambers through selective exposure(Bakshy, Messing, and Adamic 2015). This suggests that user behavior, rather than platform algorithms alone, is a critical driver of ideological clustering.

Most existing research on political interactions in social media has focused on repost networks, such as retweets on Twitter, where sharing content typically signifies endorsement (Boyd, Golder, and Lotan 2010; Metaxas et al. 2015). However, reposts do not always reflect active engagement with ideas, nor do they necessarily reveal ideological contestation. In contrast, reply interactions represent a richer form of discourse: users may agree, disagree, challenge, or amplify opinions.

Therefore, in this study, we conducted a comparative study of Truth Social, BlueSky, and Mastodon, which measured the existence and strength of echo chambers across these emerging platforms. Specifically, we quantified the echo chamber effect by analyzing users' most engaging interactions - replies, which explicitly show users' opinions via textual content. Our findings shed light on how emerging platforms like Truth Social, BlueSky, and Mastodon foster unique patterns of ideological clustering and echo chambers. By comparing these platforms, this study provides valuable insights into the role of social media structures in shaping political discourse and user interactions.

Related Work

Echo Chamber and Polarization on Social Media

Echo chambers on social media have become a critical issue - users often gravitate toward information that aligns with their preexisting beliefs and join groups that reinforce these shared narratives (Garrett 2009; Hershey 2009; Garimella et al. 2017). Broadly, an echo chamber can be described as an environment where individuals' opinions, political leanings, or beliefs are reinforced through repeated interactions with ideologically aligned users or content sources (Garimella et al. 2017; Cinelli et al. 2021; Alatawi, Sheth, and Liu 2024). Such environments limit exposure to diverse perspectives, creating insular communities that reinforce biases and reduce opportunities for balanced discourse.

According to group polarization theory (Sunstein 1999), echo chambers and political polarization are closely linked, forming a bidirectional, reinforcing feedback loop. On the one hand, echo chambers intensify political polarization by isolating individuals from opposing viewpoints, reinforcing their existing beliefs, and driving them toward more extreme positions over time. On the other hand, heightened political polarization increases individuals' tendency to seek like-minded communities, further fostering echo chambers. This dynamic is especially evident on social media platforms (Iandoli, Primario, and Zollo 2021), where factors such as influential users, media outlets, and algorithmic content delivery encourage ideological clustering (Alatawi, Sheth, and Liu 2024).

For example, Flamino et al. (2023) analyzed how political influencers shaped discourse during the 2016 and 2020 U.S. elections, reinforcing divides between left- and right-wing ideologies through selective content sharing. Similarly, Cota et al. (2019) quantified information spread within politically homogeneous networks, showing how echo chambers limit exposure to diverse perspectives, thereby amplifying polarization. Additionally, Dehghan and Nagappa (2022) found that the alt-tech platform Gab fosters environments conducive to fringe ideologies and radical discourse, often intensified by echo chamber dynamics. In a comparative study, Cinelli et al. (2021) observed that homophilic clustering—where users gravitate toward like-minded communities—is most pronounced on Facebook and Twitter, while less so on Gab and Reddit.

Furthermore, polarized individuals tend to seek out like-minded communities and information sources, reinforcing their preexisting views through a process known as selective exposure (Garrett 2013). This tendency strengthens echo chambers by drawing individuals toward ideologically consistent groups and media, further limiting exposure to opposing perspectives.

The reinforcing cycle between polarization and echo chambers makes bridging partisan divides increasingly challenging, weakening the public sphere and democratic dialogue. Collectively, these studies underscore the persistence of echo chambers within social media and highlight the urgent need to understand and address this phenomenon amid rising political polarization.

New Social Media Platforms

In this study, we focus on three emerging social media platforms—Truth Social, Mastodon, and BlueSky. BlueSky, with its single-instance setup, closely resembles Twitter in terms of user habits. Early research suggests that BlueSky's network topology mirrors those of larger platforms, showing similar patterns in user connectivity and engagement (Failla and Rossetti 2024; Quelle and Bovet 2025). Its user base leans left-center, likely attracting those dissatisfied with Twitter's recent content moderation. Mastodon, however, boasts over 7,500 instances, including the main server Mastodon Social and numerous smaller communities focused on specific topics. This decentralized structure allows for greater community autonomy and quality engagement over sheer user numbers (Zulli, Liu, and Gehl 2020). For example, a study on Mastodon's structure highlights its unique approach to community management and content dissemination (Zignani, Gaito, and Rossi 2018).

Conversely, Mastodon offers a highly decentralized experience, fostering niche communities through thousands of independent instances that differ substantially from Twitter's centralized structure (Zignani, Gaito, and Rossi 2018). Truth Social, based on Gerard, Botzer, and Weninger (2023)'s preliminary analysis, aligns with alt-tech platforms, initially drawing a predominantly right-leaning audience.

Existing research has extensively examined Twitter and Reddit due to their popularity and data accessibility. These studies have shown that traditional social media platforms often exhibit strong echo chamber effects and, to some extent, polarized structures (Iandoli, Primario, and Zollo 2021). However, our understanding of newer platforms like Truth Social, BlueSky, and Mastodon remains limited, despite their growing user bases. It is also increasingly evident that the future of social media will feature a greater diversity of platforms and structures (Chen 2023). Thus, it is crucial to investigate the founding motivations behind these emerging platforms, to examine whether users' interaction patterns align with or diverge from those on traditional platforms, and to assess whether they are fostering a more diverse or increasingly polarized online environment.

Quantifying Echo Chambers Effect

Before asserting the presence of echo chambers on social media, we must first establish a principled approach to quantify them. As no standardized methodology exists, prior studies have explored this measurement from various angles. For instance, Cinelli et al. (2021) analyzed echo chambers by labeling users' ideological leanings and comparing these leanings within user networks to detect clustering. They further examined community structures and modeled opinion spread through susceptible-infected-recovered (SIR) dynamics. Similarly, Cota et al. (2019) employed epidemic models (SIS and SIR) to study information dissemination during Brazil's impeachment debate, quantifying polarization within politically aligned networks.

Other approaches have focused on content sources and ideological drift. Flamino et al. (2023), for example, categorized news outlets and influencers by political bias and tracked behavioral shifts across two U.S. elections, revealing increased polarization over time. Across these efforts, a common goal emerges: determining whether social interactions are structured by ideological affinity, thus reinforcing homogeneity and limiting exposure to diverse perspectives.

Notably, many of these studies rely on repost networks — retweets, shares, or reblogs — as proxies for endorsement or ideological alignment. For example, Garimella et al. (2018a) leveraged retweet patterns to quantify polarized discourse on Twitter. However, repost networks predominantly capture behaviors of agreement or amplification (Boyd, Golder, and Lotan 2010; Metaxas et al. 2015), potentially missing critical interactions such as disagreement, debate, or ideological contestation.

Furthermore, studies like the one by Bakshy, Messing, and Adamic (2015) examined ideological filtering through exposure to and engagement with news links on Facebook. While crucial for understanding selective exposure, such analyses capture content-less behavioral signals — user actions (e.g., clicking) that indicate preference but do not reveal the substance of ideological exchanges. In contrast, reply networks represent content-bearing interactions: every link in the network reflects a direct communicative act, embedding both relational ties and the textual substance of discussions. Replies offer a broader view of how users affirm, contest, or negotiate ideological positions, thereby providing a richer lens for studying the dynamics of echo chambers.

Building on this distinction, our study introduces a multifaceted model of echo chambers that integrates both network structure and narrative influence. While existing research has largely emphasized structural features — such as ideological clustering or information diffusion — it has underexplored how influential users actively steer discussions and frame dominant narratives within clusters. We propose that echo chambers are not solely defined by who connects with whom, but also by whose voices shape the content that circulates within these communities.

Specifically, we operationalize the echo chamber effect through three complementary dimensions: (1) narrative dominance by central users, measured by alignment between user centrality (PageRank) and content representativeness (LexRank); (2) community-level ideological homogeneity, assessed through stance analysis; and (3) connectivity bias toward ideologically similar users, modeled through biased random walks. By combining structural and textual analysis within reply networks, our approach offers a more comprehensive and dynamic view of echo chambers, capturing both the formation of ideological silos and the narratives that sustain them.

Data

We collected posts and replies from BlueSky, Mastodon, and Truth Social to compare opinions across these platforms. For Mastodon, we focused on the main server with the largest user base, "mastodon.social." Previous research revealed a lack of comprehensive comparative studies on the echo chamber effect across social media platforms, with inconsistencies in data collection time frames and topics that hinder comparability. For example, Cinelli et al. (2021) analyzed echo chambers on Facebook, Twitter, Reddit, and Gab, but the data collection periods varied widely: Facebook data spanned from 2010, Twitter from 2016, and Reddit and Gab from 2017, covering durations from as short as 14 days to as long as 7 years. Additionally, the topics analyzed differed by platform (e.g., "abortion," "gun control," and "ObamaCare" on Twitter, versus "vaccines," "Science/Conspiracy," and "News" on Facebook).

To address these inconsistencies, we standardized our data collection to a fixed period, from May 30 to June 30, 2024. For BlueSky, Mastodon, and Truth Social, we used each platform's "search posts" endpoint in their public API to retrieve posts containing the keywords "Biden" and "Trump." These keywords were chosen to capture a wide range of discussions about these presidential candidates, with Biden still representing the Democratic Party during this period.

During this timeframe, two significant public events occurred: Donald Trump became the first former U.S. president convicted of felony crimes (Michael R. Sisak et al. 2024), and Hunter Biden, son of President Joe Biden, was convicted on three federal felony charges in a gun trial (Randall Chase et al. 2024). These high-profile events likely generated strong public reactions, presenting a unique opportunity to explore whether discussions about these incidents align with our hypothesis on the echo chamber effect. We anticipate that these events catalyzed extensive discourse, providing a rich dataset to examine ideological clustering across these platforms.

The Bluesky API provided endpoints for creating posts, managing sessions, and fetching post threads. The thread endpoint was particularly useful for retrieving conversation threads around specific posts, enabling a comprehensive view of discussions. Detailed documentation is available at the Bluesky API documentation site[1].

Similarly, the Mastodon API offered extensive functionality, including the context endpoint, which allowed us to retrieve entire conversation threads surrounding a given post. This helped in understanding the full scope of discussions related to specific topics. Detailed information can be found in the Mastodon API documentation[2].

Truth Social, although lacking publicly available API documentation, shares an API structure similar to Mastodon's. This similarity allowed us to use the context endpoint in a manner akin to Mastodon's, facilitating the collection of surrounding conversations for target posts and enabling seamless integration into our data collection methodology.

We used the search endpoint for each platform to conduct keyword searches, retrieving all posts containing our target keywords. Additionally, we employed supplementary endpoints to ensure comprehensive coverage and a deeper understanding of the discussions. For Mastodon and Truth Social, we used the context endpoint, which retrieves surrounding posts related to a specific post. This approach helps us capture the broader conversation and understand the flow of discussions around our keywords. On BlueSky, we used the thread endpoint, which captures entire threads or connected series of posts, offering insights into how conversations evolve over time.

The context search is a recursive process: we searched both the ancestors and descendants of each post retrieved through keyword searching. For each ancestor and descendant, we recursively searched their ancestors and descendants until reaching either the end or the top of the reply tree. By leveraging these various endpoints, we gathered a rich dataset that includes not only individual posts but also the surrounding context and extended discussions.

We constructed reply networks for Mastodon, BlueSky, Truth Social, and Reddit, with particular emphasis on interaction homophily as measured through textual content. In the context of our user interaction network, each user is represented as a node, and replies form the directed edges between them as reply interactions—an edge from user A to user B indicates that A replied to B. Basic statistics of our networks are shown in Table 1. As we can see, Mastodon's network comprises 18,635 nodes and 48,438 edges, with an average degree (in-degree and out-degree) of 2.60. The relatively high proportion of isolated nodes (10.73%) indicates that a considerable portion of users within this topic space did not participate in any reply-based interactions. BlueSky's network has 138,027 nodes and 827,337 edges, and a higher average degree of 5.99. However, the proportion of isolated nodes (24.90%) suggests that nearly a quarter of users did not engage in discussions related to the selected keywords. Truth Social shows a higher degree of interaction, with 206,097 nodes, 1,570,746 edges, and an average degree of 7.62. Its isolated node proportion (0.80%) is notably low, indicating that most users within this subset were active in reply exchanges.

[1]https://docs.bsky.app/docs/category/http-reference

[2]https://docs.joinmastodon.org/api/

Method

In existing studies, repost networks (such as retweets on Twitter) are commonly used to model social interactions (Galuba et al. 2010; Barberá et al. 2015; Garimella et al. 2018b; Flamino et al. 2023), while few incorporate reply behaviors (Cota et al. 2019; Alatawi, Sheth, and Liu 2024). However, we argue that echo chambers are not solely reinforced through reposts; they are primarily sustained by direct interactions that reaffirm user biases within the network. Additionally, repost behaviors do not explicitly convey opinions and do not necessarily constitute an endorsement (Garimella et al. 2018b), whereas reply content often contains direct expressions of viewpoints and stances. On the other hand, the limited studies that consider reply networks have neglected interaction dynamics, a critical feature of the echo chamber effect. Instead, they focus solely on user spreading capacity (Cota et al. 2019) or rely solely on text embedding methods (Alatawi, Sheth, and Liu 2024), independent of the network structure. Therefore, we constructed reply networks for Mastodon, BlueSky, Truth Social, and Reddit. Each network consists of nodes representing users, with directed edges capturing reply relationships, where an edge from one user to another indicates a reply. As such, the network statistics primarily describe the interaction patterns within these topic-focused communities.

Building on the framework outlined in the related work section, we propose three methods to quantify echo chambers, focusing on influential users, community ideological alignment, and connectivity patterns in reply-based interactions.

Influence-Representativeness Model

Traditional studies of echo chambers have largely focused on structural clustering or exposure patterns. However, at the content level, "echo" also manifests when influential figures propagate or even introduce narratives that are rapidly amplified and repeated by others (Starbird 2017; Asatani et al. 2021). In this sense, echo chambers involve not only who interacts with whom but also what language and frames become dominant. Therefore, our first approach aims to vi-

Platform	Nodes (Users)	Edges	Average Degree	Isolated Nodes (%)	Number of Posts
Mastodon	18,635	48,438	2.60	10.73	111,183
BlueSky	138,027	827,337	5.99	24.90	2,392,413
Truth Social	206,097	1,570,746	7.62	0.80	3,556,587

Table 1: Network Statistics of Social Media Platforms

sualize the echo chamber effect by leveraging two algorithms: PageRank and LexRank, which respectively capture user centrality and textual representativeness. Each algorithm offers complementary insights—structural influence within the network and semantic centrality within the content—allowing us to assess how user prominence corresponds with message representativeness across the entire discourse.

PageRank (Page et al. 1999) is a graph-based algorithm originally developed to rank web pages by importance. It operates by iteratively assigning a score to each node (in our case, a user) based on the centrality of the nodes pointing to it. The algorithm assumes that connections from important nodes contribute more to the rank of a given node, thereby capturing not only direct but also indirect influence. Users with high PageRank scores are those who receive significant engagement from other central users, marking them as influential in the social structure.

LexRank (Erkan and Radev 2004) is another graph-based algorithm, but it operates over a corpus of texts rather than users. LexRank constructs a similarity graph where each sentence or post is a node, and edges between nodes are weighted by the cosine similarity between their content. The algorithm assigns importance scores to each text based on how well it connects with other central texts, following principles similar to PageRank. A post with a high LexRank score is thus highly representative of the broader discourse, as it is semantically aligned with a large portion of the corpus. In this paper, we handle a corpus of millions of posts by implementing FastLexRank (Li, Conrad, and Gagnon-Bartsch 2025), an optimized version of LexRank, which enables efficient processing of large datasets.

The combination of PageRank and LexRank provides a novel way to quantify the echo chamber effect. PageRank identifies which users are structurally influential in the interaction network, while LexRank determines how representative (central) their posts are in relation to the overall corpus. If a significant overlap exists between the two—i.e., if users with high PageRank scores also produce posts with high LexRank scores—it indicates that influential users are shaping the narrative in a way that aligns with the overall corpus. This suggests that the views of prominent users dominate both the interaction structure and the content, leading to high network homogeneity.

This model further captures this dynamic by analyzing the overlap between user centrality (PageRank) and content centrality (LexRank). A strong alignment indicates that a few influential users are steering the conversation and that their narratives are widely echoed across the network. For instance, the adoption of terms like "Sleepy Joe," initially coined by Donald Trump and quickly repeated by his supporters, exemplifies how narrative frames can spread from central actors into the broader discourse, reinforcing ideological homogeneity. By quantifying this narrative alignment, our method offers a new dimension to understanding how echo chambers operate at both the structural and semantic levels.

Community-Stance Model

Our second method examines echo chambers by combining community detection with stance analysis. This approach aims to identify tightly connected communities and assess the ideological alignment within them. Clusters based on community detection represent groups of users with intensive interactions. Identifying such tightly connected subgroups helps uncover the structural layout of social interactions. To further assess the internal homophily of these interaction communities, we extend our analysis by performing stance detection on the textual content within these communities. This additional step allows us to measure whether users within a community tend to align ideologically, reinforcing the presence of echo chambers.

For community detection, we applied the Leiden algorithm (Traag, Waltman, and van Eck 2019) to identify communities within the interaction network. The Leiden algorithm is an improved version of the Louvain algorithm (Blondel et al. 2008), which maximizes modularity to detect groups of nodes with dense intra-group connections. While the Louvain algorithm efficiently identifies hierarchically nested communities through iterative modularity optimization, it can sometimes yield fragmented clusters. To address this, the Leiden algorithm introduces refinements that ensure connectedness within communities and produce more stable partitions, enhancing the reliability of community detection. Therefore, we selected Leiden over Louvain for its improved accuracy and stability in detecting cohesive communities.

We then employ the `Llama-3.1-8B-Instruct` model to perform stance detection on the posts, categorizing each one into one of five stances: left, lean-left, center, lean-right, or right. Given the dialogic nature of the reply-based network, where individual responses often rely on preceding context, the model considers conversational flow to enhance the accuracy of stance annotations. Each stance is then assigned a numerical value: -1 for left, -0.5 for lean-left, 0 for center, 0.5 for lean-right, and 1 for right.

Two researchers evaluated the model's output on a sample of 200 instances. Inter-rater reliability was assessed using Cohen's κ coefficient. During the coding process, our annotators encountered cases where labeling was subjective and nuanced, making it difficult to fully harmonize their judgments. We acknowledge this subjectivity as an inherent aspect of the task. Therefore, instead of enforcing a single "correct" label, we used Cohen's κ to measure inter-rater reliability, as it accounts for agreement beyond chance and treats both annotators symmetrically. We consider the model's agreement with each annotator acceptable as long as the κ values are not too low, reflecting that the model aligns reasonably well with human interpretations despite some ambiguity in the labeling task. The agreement between the two human annotators was substantial, with $\kappa = 0.78$. The agreement between the LLM and Human 1 was $\kappa = 0.64$, while the agreement between the LLM and Human 2 was $\kappa = 0.73$. We also examined the cases where the annotators disagreed with the model. In most of these instances, the model struggled to determine the individual's political ideology due to insufficient information. While human coders typically labeled such cases as neutral, LLM often attempted to infer an ideological alignment. Occasionally, these guesses were accurate, but more often, we believe the ambiguity made it genuinely difficult to draw a definitive conclusion. Overall, we find that the annotation accuracy of the LLM is reasonably strong. Moreover, since post-level political ideology labels are aggregated to the user level, individual annotation errors are less likely to impact the overall outcome significantly. The detailed prompt can be seen in the Appendix.

To aggregate these stance scores at the user level, we calculate a political ideology leaning score based on each user's posts. For a given user i with n_i posts, denoted $P_i = \{p_1, p_2, \ldots, p_{n_i}\}$, the user's political ideology score S_i is calculated as follows:

$$S_i = \frac{1}{n_i} \sum_{j=1}^{n_i} c_j$$

where c_j is the stance score for post p_j in the set P_i.

Once the user's political ideology scores have been inferred, we looked at the distribution of users' ideology scores across the communities that were detected by the Leiden algorithm for each social media platform. Since the community detection algorithm normally returns a large number of small communities with a small number of large communities, we picked the ten largest communities for each social media platform.

Based on the above process, we quantified the degree of homogeneity at the community level and assessed the extent to which ideological uniformity reinforces echo chamber effects. A high degree of alignment between structural and ideological homogeneity suggests stronger echo chambers, while more diverse distributions indicate weaker or non-existent echo chambers.

Biased Random Walk Model

We propose a biased random walk model for simulating influence spread within a network. This model captures the tendency of individuals to interact preferentially with others who share similar ideological leanings, quantified by a political ideology score. The objective is to further examine the extent to which an initial node (seed) can propagate its influence based on a bias toward ideological similarity, without imposing a strict distance threshold.

The interaction network $G = (V, E)$ consists of a set of nodes V and edges E, where each node represents an individual and edges represent interactions between them. Each node $v \in V$ has an associated political ideology score p_v, where $p_v \in [-1, 1]$. Scores near -1 represent left-leaning ideologies, scores near 0 indicate centrism, and scores close to 1 represent right-leaning ideologies.

For each random walk, an initial node $v_0 \in V$ is selected randomly. At each step, a transition probability $P(v \to u)$ determines the likelihood of moving from the current node v to a neighboring node u, weighted by the similarity in their political ideology scores. This probability is defined by:

$$P(v \to u) = \begin{cases} \alpha & \text{if } |p_v - p_u| \leq \delta \\ 1 - \alpha & \text{if } |p_v - p_u| > \delta \end{cases}$$

where α represents the bias strength, this value should reflect real-world levels of ideological bias in interaction. Studies estimate 60–80% of user interactions on politically charged topics are with ideologically similar others; therefore, we set $\alpha = 0.7$ to reflect a moderate preference for similar than dissimilar neighboring nodes (Barberá et al. 2015; Bakshy, Messing, and Adamic 2015). The maximum political ideology similarity threshold, δ, is 1, permitting the random walk to explore nodes of diverse ideological distances.

The random walk continues for a maximum of $s = 25$ steps, or until no further connections are available. To obtain a comprehensive view of influence propagation across the network, this process is iterated over 1000 walks with randomly selected starting nodes. For each walk, we record the influence set—a collection of nodes reached from the seed node under the bias constraint. Recognizing that parameter settings could influence the results, we conducted a sensitivity analysis for both the bias strength (alpha) and the number of walk steps (s), as detailed in the Appendix.

Under this approach, we assess the strength of the echo chamber effect by examining the ideological composition of the influence sets reached by the biased random walk. When influence sets for both left- and right-leaning seed nodes encompass a diverse range of ideological positions, this suggests a weaker echo chamber effect with greater cross-ideological exposure. Conversely, a strong echo chamber effect is indicated if the influence sets of seed nodes are predominantly composed of individuals with similar ideological leanings, as the influence remains largely confined to ideologically homogeneous groups, reinforcing limited exposure to opposing views.

Results

Influence-Representativeness Alignment

The results of our Influence-Representativeness comparison are displayed in Figure 1. To address potential processing errors, we mapped the relationship between the top-rated users identified by the PageRank algorithm and their posts (specifically, the 20 most representative posts) in Figure 1. Each curve connects a user to their posts ranked by the FastLexRank algorithm, with color coding applied based on

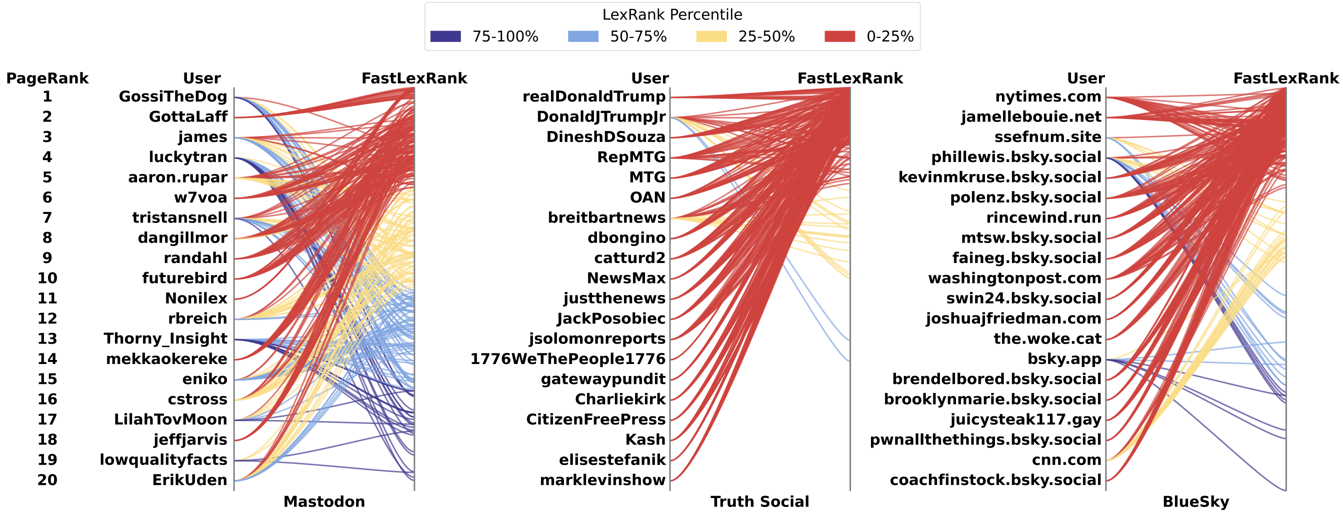

Figure 1: Comparison of User Influence (PageRank) and Content Representativeness (FastLexRank) on Mastodon, Truth Social, and BlueSky

quartile position: red for the top 25%, light red for the second quartile, yellow for the third, and blue for the fourth quartile. While PageRank and FastLexRank offer useful proxies for user influence and content representativeness, we note that they capture network centrality and semantic centrality, respectively, rather than direct measures of real-world influence or narrative control. To further support the validity of these measures, we conducted a manual inspection and a human evaluation study, summarized below and detailed in the Appendix. The user-content centrality mapping on Truth Social is particularly striking. Nearly all of the top 20 influential users who attract significant engagement from other central users fall within the top 25% quartile for LexRank scores. This suggests that these influential users are highly aligned with the platform's dominant narratives. Notably, a large proportion of these central users are associated with Donald Trump and Republican-related content (e.g., RealDonaldTrump, DonaldJTrumpJr, RepMTG), indicating that discussions on Truth Social are heavily concentrated around a specific ideological framework, reinforcing a strong echo chamber effect.

In contrast, the interaction patterns on BlueSky and Mastodon show notable differences. The LexRank scores of the top 20 influential users on these platforms are more evenly distributed across quartiles, suggesting that the central users on these platforms do not as effectively represent the views of the broader user base. This distribution reflects a more diverse exchange of opinions, allowing for more ideological variation and reducing the concentration of discussion around a single narrative.

Additionally, on BlueSky, many of the most influential users are official accounts, including prominent media outlets (e.g., NYTimes), which support the platform's role in disseminating diverse information. On Mastodon, however, the most influential users are primarily private accounts, aligning with the platform's decentralized structure and emphasis on user autonomy. These variations underscore how differences in platform architecture and user demographics shape interaction patterns and narrative formation, influencing the extent to which echo chambers and ideological concentration emerge.

Community-Based Ideological Distribution

As shown in Figure 2, the results of our community-stance analysis are consistent with earlier findings on user interactions and narrative dynamics across platforms. The modularity scores of the reply networks were 0.2515 for Mastodon, 0.0639 for Truth Social, and 0.2125 for BlueSky. Mastodon and BlueSky both exhibit moderate modularity, suggesting the presence of structurally distinct interaction clusters—an observation that aligns with the ideological diversity and decentralized engagement patterns seen on these platforms. In contrast, the very low modularity on Truth Social does not imply an absence of ideological clustering. Rather, it reflects the dominance of a single, densely connected community centered around high-profile figures—particularly Donald Trump—whose posts attract widespread attention and replies. As demonstrated in our influence-representativeness analysis, a small number of central users not only concentrate user interactions but also shape the prevailing narratives, producing a tightly interconnected echo chamber rather than multiple isolated ideological groups. To evaluate ideological consistency, we focused on the ten largest communities on each platform, which together encompass 37% of the total nodes on Mastodon, 91% on Truth Social, and 74% on BlueSky—capturing the majority of meaningful interaction within each network.

On Truth Social, the 6 largest communities dominate in size, and within each of these communities, users are predominantly labeled as "right" or "lean-right." There is minimal representation of opposing viewpoints, indicating a significant degree of ideological clustering, with little cross-

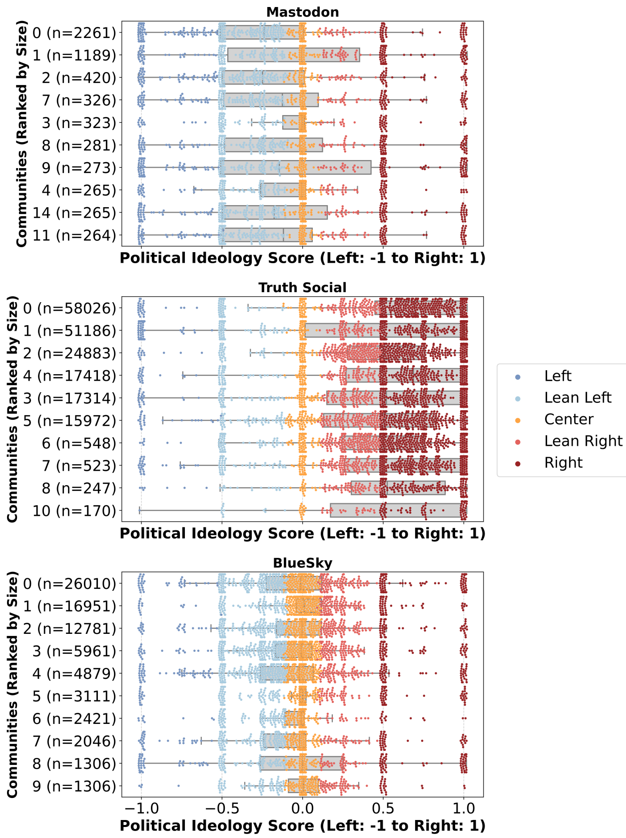

Figure 2: Distribution of Political Ideology Scores within the Largest Communities across Mastodon, Truth Social, and BlueSky

ideological discussion. This pattern suggests the presence of homogeneous communities, consistent with previous insights about ideological concentration on this platform.

In contrast, the ideological distribution among users in Mastodon's communities is more dispersed, although it leans slightly to the right overall. The communities on Mastodon are smaller and more evenly distributed, reflecting the platform's decentralized structure and the diversity of its user interactions. This pattern indicates a broader range of perspectives, though ideological preferences still skew rightward in aggregate.

On BlueSky, the communities show a more centrist ideological tendency, with user ideologies distributed symmetrically around the mean. Community sizes on BlueSky are larger than those on Mastodon but smaller than those on Truth Social, with the three largest communities accounting for a significant portion of the total nodes. This pattern reflects a balance between ideological diversity and structure, suggesting that BlueSky facilitates discussions with more varied viewpoints compared to the other two platforms.

These findings illustrate how different platform architectures and user behaviors influence the development of communities and the extent of ideological clustering. They further highlight the varying degrees to which echo chambers may form across platforms, contributing to different dynamics of political discourse and engagement.

Results and Analysis of Biased Random Walks

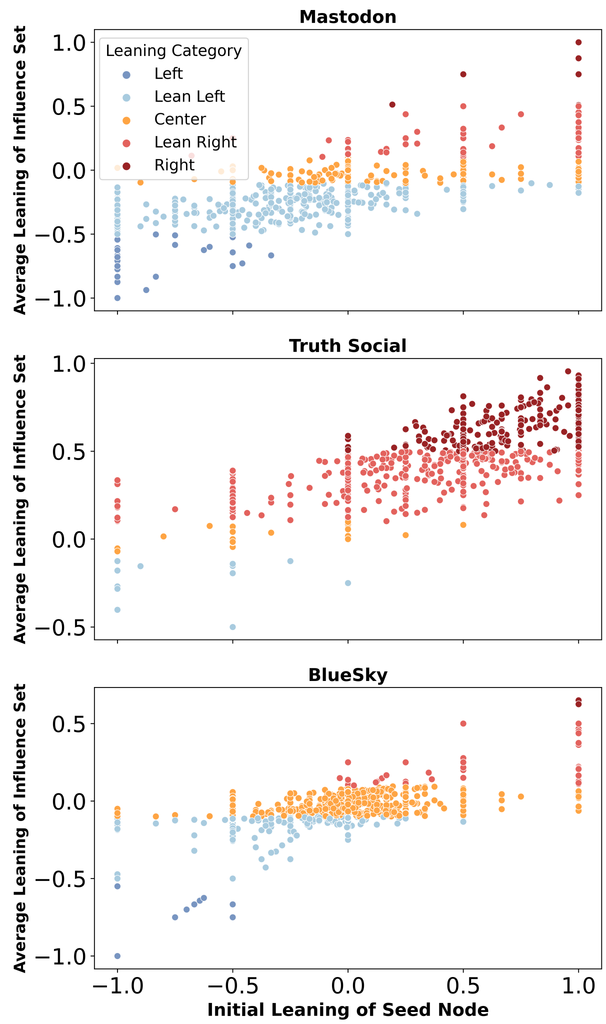

Figure 3: Ideological Influence Spread Across Social Networks via Biased Random Walks

To clearly illustrate the ideology of influence sets resulting from the biased random walk, we present the results in a scatter plot. This plot represents the relationship between the initial political ideology score of each seed node and the average political ideology score of the nodes within its influence set. Each point on the scatter plot corresponds to one random walk, with the x-axis representing the initial leaning of the seed node and the y-axis showing the average leaning of its influence set.

Each platform demonstrates unique ideological clustering patterns. On Truth Social, a significant clustering effect appears on the right, with right-leaning seed nodes reaching influence sets that are also predominantly right-leaning. This pattern suggests a strong echo chamber effect, where users primarily interact with ideologically similar peers, reinforcing in-group biases and limiting exposure to diverse viewpoints.

In contrast, Mastodon displays a more balanced ideological spread. Influence sets for both left- and right-leaning seed nodes cover a broader range of ideologies, indicating weaker echo chambers and greater cross-ideological exposure. The clustering around the center line implies a platform structure that fosters a variety of viewpoints, reducing ideological homogeneity.

BlueSky exhibits a central clustering pattern, where most influence sets converge around the center regardless of the initial seed's leaning. This suggests that BlueSky's network structure promotes cross-ideological interactions, resulting in a more balanced exchange of perspectives and a reduced echo chamber effect. Users with extreme leanings are more likely to encounter diverse viewpoints, which can moderate ideological clustering.

The initial ideological leaning of seed nodes heavily impacts the composition of influence sets, particularly on Truth Social, where right-leaning nodes predominantly influence right-leaning sets. This reinforcement of ideological consistency suggests a mechanism that may contribute to polarization. Mastodon shows more ideological diversity in influence sets, implying a structure that encourages broader interaction. BlueSky's central clustering further reflects a dilution of extreme ideologies, potentially reducing bias in influence patterns.

These findings imply that network structure influences ideological exposure and polarization. Truth Social's strong echo chambers may deepen ideological divides by limiting cross-ideological interactions. In contrast, Mastodon and BlueSky show more diverse interactions, which may counteract polarization by fostering a balanced exchange of ideas. The centrist clustering on BlueSky, in particular, points to a platform environment where extreme ideologies are less likely to dominate, possibly due to platform design or user behavior that encourages broader connectivity.

Conclusion and Discussion

In this study, we focused on the most engaging and opinion-expressive form of interaction—the reply network—to analyze interaction patterns on three major emerging social media platforms in the "post-Twitter era": BlueSky, Mastodon, and Truth Social, within a polarization context. To quantify the echo chamber effect, we applied three methods that examine (1) the role of influential users in shaping narratives, (2) the ideological alignment within reply network communities, and (3) the connectivity patterns favoring ideologically similar users. Our analysis used strict controls on topic and time frame in data collection to ensure rigorous cross-platform comparisons.

Our findings reveal a relatively weaker echo chamber effect on both Mastodon and BlueSky, where discussions span a broader ideological spectrum. However, there are subtle differences in the patterns observed on these two platforms: (1) On Mastodon, the top influential accounts are predominantly private users, while on BlueSky, several top accounts are official entities, including neutral news media. Mastodon's most influential users display the most dispersed representativeness in their content, suggesting a highly diverse range of opinions within the interaction network. (2) In the communities detected on Mastodon, participant ideologies are widely distributed, covering nearly the entire ideological spectrum. By contrast, interactions on BlueSky tend to cluster around users with centrist ideologies. (3) In our biased random walk analysis, seed nodes on Mastodon reach users across a broad ideological range, while on BlueSky, they are largely clustered around centrist users. We, therefore, categorize these interaction patterns as follows: Mastodon fosters rich ideological diversity with cross-ideological exchanges, whereas BlueSky's discussions are more centered around moderate voices.

By contrast, Truth Social exhibits a pronounced echo chamber effect, with the majority of its user base leaning right or supporting Trump. This strong ideological alignment aligns with the platform's founding purpose, which was to provide Trump supporters with a dedicated space for digital engagement following his Twitter ban. The high degree of ideological homogeneity on Truth Social suggests that discussions are more insular, reinforcing shared perspectives and potentially limiting exposure to opposing viewpoints.

Beyond quantifying echo chamber effects, our content analysis of Truth Social offers additional qualitative insights into user dynamics within this platform. On the one hand, the ideological clustering observed within Truth Social raises concerns about the potential for echo chambers to amplify misinformation or reinforce collective biases. For instance, influential figures, such as Trump himself, play a prominent role in shaping narratives on the platform. A post by Trump on May 31, 2024, about his Manhattan trial drew 2,142 replies, 5,969 reblogs, and 22,400 favorites, illustrating the considerable reach and influence of his commentary among followers. This environment, where influential voices dominate the discourse, could increase the likelihood of recirculating and entrenching specific narratives within the community.

On the other hand, it is encouraging to observe that Truth Social's ideological alignment does not entirely preclude critical engagement or fact-checking. Even within this predominantly right-leaning space, we identified instances where users engage in debate, challenging each other's narratives with counter-evidence. For example, in discussions around Trump's alleged remarks about veterans, while some users defended Trump by citing figures like former Chief of Staff Gen. John Kelly, who allegedly denied hearing disparaging comments, others directly challenged these defenses. One user asserted, "Critics claim Trump disparaged veterans, but key figures deny hearing him call anyone 'losers.'" This statement was countered by another user who responded, “John Kelly has, in fact, confirmed Trump’s disparaging remarks, a fact easily verified with a simple search.” Such exchanges suggest that fact-checking and critical discourse do occur, even within a platform skewed toward a single ideological orientation.

In summary, while echo chambers on platforms like Truth Social can potentially intensify ideological conformity and misinformation, they may also host instances of fact-checking and debate, adding complexity to our understanding of these spaces. Future research could delve further into the dynamics within echo chambers, investigating whether they solely promote polarization and misinformation or also provide opportunities for critical engagement and truth-checking. This nuanced perspective could help clarify the role of echo chambers in shaping public opinion and the quality of discourse in online communities.

Limitations

Our analysis focuses on a specific one-month period (May 30–June 30, 2024) and centers on political discussions involving keywords related to Biden and Trump. This tight topical and temporal framing allows for rigorous cross-platform comparability during a period of high political salience. However, it also constrains the generalizability of our findings. Interaction patterns and echo chamber effects may differ across other topics (e.g., health, entertainment) or in non-election periods. Future research should expand this approach to different time frames and broader topic domains to examine whether similar patterns of ideological clustering and narrative convergence emerge under varied conditions. Nevertheless, by focusing on a politically charged context, our study captures a ”stress test” scenario where echo chamber dynamics are likely to be most pronounced.

References

Alatawi, F.; Sheth, P.; and Liu, H. 2024. Quantifying the Echo Chamber Effect: An Embedding Distance-based Approach. In *Proceedings of the 2023 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, ASONAM ’23, 38–45. New York, NY, USA: Association for Computing Machinery.

Asatani, K.; Yamano, H.; Sakaki, T.; and Sakata, I. 2021. Dense and influential core promotion of daily viral information spread in political echo chambers. *Scientific Reports*, 11(1): 7491.

Bakshy, E.; Messing, S.; and Adamic, L. A. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239): 1130–1132.

Barberá, P.; Jost, J. T.; Nagler, J.; Tucker, J. A.; and Bonneau, R. 2015. Tweeting from left to right: is online political communication more than an echo chamber? *Psychological Science*, 26(10): 1531–1542.

Blondel, V. D.; Guillaume, J.-L.; Lambiotte, R.; and Lefebvre, E. 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10): P10008.

Bode, L. 2016. Political News in the News Feed: Learning Politics from Social Media. *Mass Communication and Society*, 19(1): 24–48.

Boyd, D.; Golder, S.; and Lotan, G. 2010. Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter. In *2010 43rd Hawaii International Conference on System Sciences*, 1–10. Piscataway, NJ.: IEEE.

Cava, L. L.; Aiello, L. M.; and Tagarelli, A. 2023. Drivers of social influence in the Twitter migration to Mastodon. *Scientific Reports*, 13(1): 21626.

Chen, B. X. 2023. The Future of Social Media Is a Lot Less Social. https://www.nytimes.com/2023/04/19/technology/personaltech/tiktok-twitter-facebook-social.html. Accessed: 2025-04-22.

Cinelli, M.; De Francisci Morales, G.; Galeazzi, A.; Quattrociocchi, W.; and Starnini, M. 2021. The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9): e2023301118.

Cota, W.; Ferreira, S. C.; Pastor-Satorras, R.; and Starnini, M. 2019. Quantifying echo chamber effects in information spreading over political communication networks. *EPJ Data Science*, 8(1): 1–13.

Dehghan, E.; and Nagappa, A. 2022. Politicization and Radicalization of Discourses in the Alt-Tech Ecosystem: A Case Study on Gab Social. *Social Media + Society*, 8(3): 20563051221113075.

Del Vicario, M.; Bessi, A.; Zollo, F.; Petroni, F.; Scala, A.; Caldarelli, G.; Stanley, H. E.; and Quattrociocchi, W. 2016. The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3): 554–559.

Erkan, G.; and Radev, D. R. 2004. LexRank: Graph-based Lexical Centrality as Salience in Text Summarization. *Journal of Artificial Intelligence Research*, 22: 457–479.

Failla, A.; and Rossetti, G. 2024. “I’m in the Bluesky Tonight”: Insights from a year worth of social data. *PLoS One*, 19(11): e0310330.

Faverio, M.; and Anderson, M. 2025. X users’ views of politics, experiences on the site differ by party. https://www.pewresearch.org/short-reads/2025/06/05/republicans-and-democrats-on-x-differ-over-the-sites-politics-and-their-experiences/. Accessed: 2025-03-16.

Flamino, J.; Galeazzi, A.; Feldman, S.; Macy, M. W.; Cross, B.; Zhou, Z.; Serafino, M.; Bovet, A.; Makse, H. A.; and Szymanski, B. K. 2023. Political polarization of news media and influencers on Twitter in the 2016 and 2020 US presidential elections. *Nature Human Behaviour*, 7(6): 904–916.

Galuba, W.; Aberer, K.; Chakraborty, D.; Despotovic, Z.; and Kellerer, W. 2010. Outtweeting the twitterers - predicting information cascades in microblogs. In *Proceedings of the 3rd Wonference on Online social networks*, WOSN’10, 3. USA: USENIX Association.

Garimella, K.; De Francisci Morales, G.; Gionis, A.; and Mathioudakis, M. 2017. The Effect of Collective Attention on Controversial Debates on Social Media. In *Proceedings of the 2017 ACM on Web Science Conference*, WebSci ’17, 43–52. New York, NY, USA: Association for Computing Machinery.

Garimella, K.; De Francisci Morales, G.; Gionis, A.; and Mathioudakis, M. 2018a. Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18*, 913–922. Lyon, France: ACM Press. ISBN 978-1-4503-5639-8.

Garimella, K.; Morales, G. D. F.; Gionis, A.; and Mathioudakis, M. 2018b. Quantifying Controversy on Social Media. *Trans. Soc. Comput.*, 1(1): 3:1–3:27.

Garrett, R. K. 2009. Echo chambers online?: Politically motivated selective exposure among Internet news users1. *Journal of Computer-Mediated Communication*, 14(2): 265–285.

Garrett, R. K. 2013. Selective Exposure: New Methods and New Directions. *Communication Methods and Measures*, 7(3-4): 247–256.

Gerard, P.; Botzer, N.; and Weninger, T. 2023. Truth Social Dataset. In *Proceedings of the International AAAI Conference on Web and Social Media*, 1034–1040. Washington, DC: AAAI Press.

Grant, N. 2022. Security Training Group Asks Musk to Rid Twitter of Antisemitism. https://www.nytimes.com/2022/11/01/technology/musk-twitter-antisemitism-security-group.html. Accessed: 2025-04-22.

Hern, A. 2024. Twitter usage in US 'fallen by a fifth' since Elon Musk's takeover. https://www.theguardian.com/technology/2024/mar/26/twitter-usage-in-us-fallen-by-a-fifth-since-elon-musks-takeover. Accessed: 2025-04-22.

Hershey, M. R. 2009. Echo Chamber: Rush Limbaugh and the Conservative Media Establishment. *Perspectives on Politics*, 7(2): 415–417.

Iandoli, L.; Primario, S.; and Zollo, G. 2021. The impact of group polarization on the quality of online debate in social media: A systematic literature review. *Technological Forecasting and Social Change*, 170: 120924.

Jamieson, K. H.; and Cappella, J. N. 2010. *Echo Chamber: Rush Limbaugh and the Conservative Media Establishment*. New York, NY: Oxford University Press.

Klapper, J. T. 1960. *The effects of mass communication*. New York, NY: Free Press.

Kuncoro, H. R.; Hasanah, K.; Sari, D. L.; and Kurniawati, E. 2024. Algorithmic Influence: Twitter's Role in Shaping Public Discourse and Amplifying Radical Content. In *Proceedings of the 2nd International Conference on Advance Research in Social and Economic Science (ICARSE 2023)*, 549–553. Paris, France: Atlantis Press.

Kwak, H.; Lee, C.; Park, H.; and Moon, S. 2010. What is Twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web*, 591–600. Raleigh, NC: ACM.

Li, M.; Conrad, F.; and Gagnon-Bartsch, J. 2025. FastLexRank: Bring Order into Social Media Posts Using Lexical Ranking Algorithm. *Workshop Proceedings of the 19th International AAAI Conference on Web and Social Media*, 2025: 48.

Metaxas, P.; Mustafaraj, E.; Wong, K.; Zeng, L.; O'Keefe, M.; and Finn, S. 2015. What Do Retweets Indicate? Results from User Survey and Meta-Review of Research. In *Proceedings of the International AAAI Conference on Web and Social Media*, 658–661. Washington, DC: AAAI Press.

Michael R. Sisak; Jennifer Peltz; Eric Tucker; Michelle L. Price; and Jill Colvin. 2024. Guilty: Trump becomes first former US president convicted of felony crimes. https://apnews.com/article/trump-trial-deliberations-jury-testimony-verdict-85558c6d08efb434d05b694364470aa0. Accessed: 2025-04-22.

Nickerson, R. S. 1998. Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology*, 2(2): 175–220.

Odabaş, S. D. a. M. 2023. 8 facts about americans and twitter as it rebrands to X. https://www.pewresearch.org/short-reads/2023/07/26/8-facts-about-americans-and-twitter-as-it-rebrands-to-x/. Accessed: 2025-04-29.

Page, L.; Brin, S.; Motwani, R.; and Winograd, T. 1999. The PageRank citation ranking: Bringing order to the web. Technical report, Stanford infolab.

Persily, N.; and Tucker, J. A. 2020. *Social Media and Democracy*. Cambridge, UK: Cambridge University Press.

Quelle, D.; and Bovet, A. 2025. Bluesky: Network topology, polarization, and algorithmic curation. *PLOS ONE*, 20(2): e0318034.

Randall Chase; Claudia Lauer; Michael Kunzelman; Alanna Durkin Richer; and Colleen Long. 2024. President Joe Biden's son, Hunter Biden, is convicted of all 3 felonies in federal gun trial. https://apnews.com/article/hunter-biden-gun-trial-federal-charges-delaware-5dd8a9380235c6360a1ddb691ef24a06. Accessed: 2025-04-02.

Roose, K. 2024. Bluesky, smiling at me. https://www.nytimes.com/2024/11/22/technology/bluesky-x-alternative.html. Accessed: 2025-04-29.

Starbird, K. 2017. Examining the alternative media ecosystem through the production of alternative narratives of mass shooting events on twitter. In *Proceedings of the International AAAI Conference on Web and Social Media*, 230–239. Washington, DC: AAAI Press.

Sunstein, C. R. 1999. The Law of Group Polarization.

THE ASSOCIATED PRESS. 2024. What we know about Truth Social, Donald Trump's social media platform. https://apnews.com/article/truth-social-donald-trump-djt-ipo-digital-world-7437d5dcc491a1459a078195ae547987. Accessed: 2024-06-19.

Traag, V. A.; Waltman, L.; and van Eck, N. J. 2019. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific Reports*, 9(1): 5233.

Zignani, M.; Gaito, S.; and Rossi, G. P. 2018. Follow the "Mastodon": Structure and Evolution of a Decentralized Online Social Network. In *Proceedings of the International AAAI Conference on Web and Social Media*, 1, 541–550. Washington, DC: AAAI Press.

Zulli, D.; Liu, M.; and Gehl, R. 2020. Rethinking the "social" in "social media": Insights into topology, abstraction, and scale on the Mastodon social network. *New Media & Society*, 22(7): 1188–1205.

Paper Checklist

1. For most authors...
   (a) Would answering this research question advance science without violating social contracts, such as violating privacy norms, perpetuating unfair profiling, exacerbating the socio-economic divide, or implying disrespect to societies or cultures? Yes
   (b) Do your main claims in the abstract and introduction accurately reflect the paper's contributions and scope? Yes
   (c) Do you clarify how the proposed methodological approach is appropriate for the claims made? Yes
   (d) Do you clarify what are possible artifacts in the data used, given population-specific distributions? Yes
   (e) Did you describe the limitations of your work? Yes
   (f) Did you discuss any potential negative societal impacts of your work? No, because we did not foresee major negative societal impacts
   (g) Did you discuss any potential misuse of your work? No, it is not possible to misuse our work
   (h) Did you describe steps taken to prevent or mitigate potential negative outcomes of the research, such as data and model documentation, data anonymization, responsible release, access control, and the reproducibility of findings? Yes
   (i) Have you read the ethics review guidelines and ensured that your paper conforms to them? Yes
2. Additionally, if your study involves hypotheses testing...
   (a) Did you clearly state the assumptions underlying all theoretical results? Yes
   (b) Have you provided justifications for all theoretical results? Yes
   (c) Did you discuss competing hypotheses or theories that might challenge or complement your theoretical results? Yes
   (d) Have you considered alternative mechanisms or explanations that might account for the same outcomes observed in your study? Yes
   (e) Did you address potential biases or limitations in your theoretical framework? Yes
   (f) Have you related your theoretical results to the existing literature in social science? Yes
   (g) Did you discuss the implications of your theoretical results for policy, practice, or further research in the social science domain? Yes
3. Additionally, if you are including theoretical proofs...
   (a) Did you state the full set of assumptions of all theoretical results? Yes
   (b) Did you include complete proofs of all theoretical results? Yes
4. Additionally, if you ran machine learning experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? No, but we will do so upon the acceptance
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? No, we did not train any models
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? No, we did not train any models
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? No, but we will disclose the details upon the acceptance
   (e) Do you justify how the proposed evaluation is sufficient and appropriate to the claims made? Yes
   (f) Do you discuss what is "the cost" of misclassification and fault (in)tolerance? Yes
5. Additionally, if you are using existing assets (e.g., code, data, models) or curating/releasing new assets, **without compromising anonymity**...
   (a) If your work uses existing assets, did you cite the creators? Yes
   (b) Did you mention the license of the assets? Yes
   (c) Did you include any new assets in the supplemental material or as a URL? No, but we will do so upon the acceptance
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? Yes
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? No, because we used publicly available data and no PII involved
   (f) If you are curating or releasing new datasets, did you discuss how you intend to make your datasets FAIR? Yes
   (g) If you are curating or releasing new datasets, did you create a Datasheet for the Dataset? Yes
6. Additionally, if you used crowdsourcing or conducted research with human subjects, **without compromising anonymity**...
   (a) Did you include the full text of instructions given to participants and screenshots? NA
   (b) Did you describe any potential participant risks, with mentions of Institutional Review Board (IRB) approvals? NA
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? NA
   (d) Did you discuss how data is stored, shared, and deidentified? NA

Large Language Models Evaluation

Prompt

Replies Analyze the following post for the user's political leaning, based on these detailed guidelines:

1. **Left**: - The post praises or supports Joe Biden, Democrats, liberal values, or progressive policies. - The post criticizes or opposes Donald Trump, conservatives, or right-leaning positions.

2. **Lean Left**: - The post expresses moderate or qualified support for left-leaning positions (e.g., support for Biden or Democratic policies but with some reservations). - The post mildly criticizes right-leaning positions or Trump but does not express strong or radical views.

3. **Center**: - The post takes a neutral or balanced stance without strong support or opposition for either side. - The post avoids taking a clear left or right stance and may present both sides of an issue equally.

4. **Lean Right**: - The post expresses moderate or qualified support for right-leaning positions (e.g., support for Trump or conservative values but with some reservations). - The post mildly criticizes left-leaning positions or Biden without expressing extreme or radical views.

5. **Right**: - The post praises or supports Donald Trump, conservatives, right-wing ideologies, or conservative policies. - The post criticizes or opposes Joe Biden, Democrats, or liberal values.

Consider that this post is a response to the following context:

Replying to: "{reply_to_post}"

Post: "{post}"

Please classify the political leaning using one of the following options: - Left - Lean Left - Center - Lean Right - Right

ONLY provide the output in the following JSON format without any additional text: {{ "Political_Leaning": "Left-/Lean Left/Center/Lean Right/Right" }}

Original Post Analyze the following post for the user's political leaning, based on these detailed guidelines:

1. **Left**: - The post praises or supports Joe Biden, Democrats, liberal values, or progressive policies. - The post criticizes or opposes Donald Trump, conservatives, or right-leaning positions.

2. **Lean Left**: - The post expresses moderate or qualified support for left-leaning positions (e.g., support for Biden or Democratic policies but with some reservations). - The post mildly criticizes right-leaning positions or Trump but does not express strong or radical views.

3. **Center**: - The post takes a neutral or balanced stance without strong support or opposition for either side. - The post avoids taking a clear left or right stance and may present both sides of an issue equally.

4. **Lean Right**: - The post expresses moderate or qualified support for right-leaning positions (e.g., support for Trump or conservative values but with some reservations). - The post mildly criticizes left-leaning positions or Biden without expressing extreme or radical views.

5. **Right**: - The post praises or supports Donald Trump, conservatives, right-wing ideologies, or conservative policies. - The post criticizes or opposes Joe Biden, Democrats, or liberal values.

Post: "{post}"

Please classify the political leaning using one of the following options: - Left - Lean Left - Center - Lean Right - Right

ONLY provide the output in the following JSON format without any additional text: {{ "Political_Leaning": "Left-/Lean Left/Center/Lean Right/Right" }}

Sensitivity Analysis

To assess the robustness of our results, we conducted a sensitivity analysis on two key parameters of the biased random walk process: the number of walk steps s, and the bias strength b. For each setting, we report the average influence leaning across multiple simulation runs, along with 95% confidence intervals (CI).

Sensitivity to Number of Walk Steps

We varied the number of walk steps s from 5 to 30, holding the bias strength fixed at $b = 0.7$. For each step length, we performed 100 simulation iterations. As shown in Figure 4, the average influence leaning remains stable across different walk lengths, suggesting that relatively short random walks (e.g., $s \leq 25$) are sufficient to capture the local structure across all three platforms (Mastodon, TruthSocial, and Bluesky).

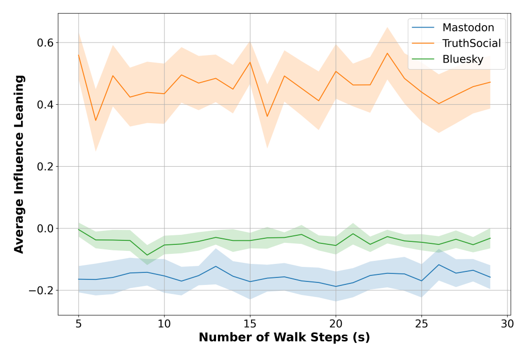

Figure 4: Sensitivity of average influence leaning to the number of walk steps.

Sensitivity to Bias Strength

We next varied the bias strength α between 0.5 (nearly unbiased) and 0.95 (strongly biased), using 10 evenly spaced values, while fixing the number of walk steps at $s = 25$. For each bias level, we conducted 100 simulation iterations. As shown in Figure 5, the influence leaning is largely stable across different bias strengths, especially for Mastodon and Bluesky. Truth Social exhibits slightly more variability but maintains a consistent overall trend. This suggests that our findings are not overly sensitive to the bias parameter.

Overall, the sensitivity analysis indicates that the results are robust to reasonable variations in the number of walk

Platform	Accuracy	Centrality Level		
		High (P/R/F1)	Medium (P/R/F1)	Low (P/R/F1)
BlueSky	40%	0.80 / 0.20 / 0.32	0.43 / 0.15 / 0.22	0.35 / 0.85 / 0.50
Mastodon	58%	0.67 / 0.40 / 0.50	0.39 / 0.35 / 0.37	0.67 / 1.00 / 0.80
Truth Social	67%	0.80 / 0.80 / 0.80	0.50 / 0.50 / 0.50	0.70 / 0.70 / 0.70

Table 2: Human evaluation results for FastLexRank validation across platforms.

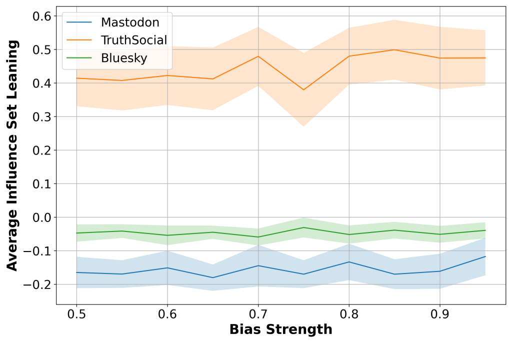

Figure 5: Sensitivity of average influence leaning to bias strength.

steps and bias strength. In other words, locally biased random walks effectively capture influence patterns without substantial parameter sensitivity.

Validation of PageRank and FastLexRank

PageRank Validation

To verify that PageRank identifies genuinely influential users, we manually inspected the top 20 users by PageRank score on each platform (Mastodon, Truth Social, and BlueSky). The top-ranked accounts included well-known figures such as realDonaldTrump, DonaldJTrumpJr, NewsMax, and the New York Times, consistent with expectations of influence on these platforms. No anomalous or irrelevant accounts appeared among the top 20 users. Thus, PageRank appears to capture user centrality and influence reliably in the constructed reply networks.

FastLexRank Validation

We conducted a human evaluation study to assess whether FastLexRank scores correspond to human perceptions of content representativeness. For each platform (BlueSky, Mastodon, and Truth Social), we sampled 60 posts (20 high, 20 medium, 20 low) based on FastLexRank scores.

A single expert annotator (from the research group) was asked to classify posts into high, medium, or low representativeness categories without knowing the true labels. The annotator was selected based on their extensive prior exposure to the dataset and familiarity with platform-specific discourse, ensuring a high level of domain expertise. Given the annotator's in-depth knowledge, this evaluation can be considered an expert review rather than a crowd-sourced annotation. While we acknowledge that a single annotator may introduce subjective bias, the goal of this validation was to assess the general recognizability of representativeness rather than produce an exhaustive labeling effort.

The human evaluation results are summarized in Table 2.

Overall, the results demonstrate that annotators could distinguish posts of different representativeness levels substantially better than random guessing (random baseline = 33% accuracy). We recognize that this classification task is inherently challenging, particularly when posts are short, topically nuanced, or when "representativeness" itself can be subjective. The relatively lower accuracy observed for BlueSky (40%) may reflect platform-specific characteristics, such as more decentralized or less polarized discourse compared to Mastodon or Truth Social, making it harder for annotators to differentiate post centrality purely based on textual content.

These limitations notwithstanding, the results provide supporting evidence that FastLexRank captures aspects of semantic centrality in platform discourse that are perceptible to human evaluators.