

Analyzing News Engagement on Facebook: Tracking Ideological Segregation and News Quality in the Facebook URL Dataset

Keywords: User engagement, Ideological Segregation, Information Quality, Social media, Facebook

Extended Abstract

Understanding how individuals engage with ideologically aligned content is necessary for assessing the influence of digital platforms on news consumption. Prior research shows that individuals tend to consume ideologically aligned content, with extreme views linked to lower-quality sources and higher engagement [1, 2]. Conservative users, in particular, exhibit greater ideological segregation [3], and Facebook has been found to foster stronger ideological echo chambers than other platforms like Reddit [4]. Additionally, user activity correlates with increased selective exposure to partisan news [5], reinforcing content biases and shaping engagement patterns [6]. These patterns raise concerns about the role of platform-specific mechanisms in shaping news consumption. Beyond user preferences and engagement patterns, algorithmic features play a crucial but often ambiguous role in shaping engagement, polarization and misinformation exposure, as seen in research on the 2020 U.S. Election Study [7]. Despite advances in understanding these dynamics, much research relies on short time frames or small treatment groups that do not account for broader platform changes [3]. Algorithmic modifications, UX design updates, and moderation strategies are rarely disclosed, making it difficult to assess their impact. For example, Facebook's 2020 news feed changes, coinciding with prior research, illustrate how such undisclosed adjustments can potentially affect the conclusions of otherwise well-designed research [8]. Thus, a deeper understanding of the interplay between user behavior and platform characteristics over time is essential.

To address these issues, we conduct a comprehensive longitudinal analysis of engagement with popular news domains using the Facebook Privacy-Protected Full URLs Dataset [9], which provides engagement metrics (views, clicks, likes, shares, comments, and emoji reactions) at the URL-action level. Our study spans four years (January 2017–December 2020) and examines engagement metrics related to news URLs in the U.S. We analyze how engagement with news URLs compares to other content and how it varies by user ideology (**RQ1**) as well as how ideological preferences influence interactions with partisan and low-quality sources (**RQ2**). In a longitudinal analysis, we furthermore investigate how engagement with news URLs evolves over time and identify key change points (**RQ3**), as well as how these trends compare to engagement with other URLs (**RQ4**). Additionally, we examine long-term trends in ideological segregation (**RQ5a**) and the consumption of low-quality news (**RQ5b**).

We incorporate ideological alignment and news quality, along with users' political preferences, to construct weighted averages of ideology and quality of news consumption for liberal, conservative, and moderate audiences. To identify key shifts in engagement patterns, we apply segmented linear regression, which models engagement trajectories as piecewise linear trends and detects significant change points in time. Users' ideological preferences are obtained from Facebook's Political Page Affinity (PPA) scores, which classify users into five ideological categories ranging from most liberal (-2) to most conservative (+2). News quality is assessed using

external domain-level ratings [10], while ideological orientation is determined using audience-based inferred ideology scores [11]. We separate news-related popular domains from non-news domains by intersecting three criteria: high engagement, the presence of ideology scores, and the availability of quality ratings. To quantify ideological segregation in news consumption, we compute the absolute difference in the weighted average ideology of URLs engaged by conservative and liberal users. We also assess exposure to low-quality news by measuring the proportion of consumed URLs with a quality score below a predefined threshold. Given the privacy protections in the dataset, we account for noise introduced through differential privacy mechanisms, ensuring the robustness of our findings through signal-to-noise ratio analyses and uncertainty interval assessments.

Our findings confirm that more extreme users engage more actively, especially through comments and reactions, consistent with prior research [12] (RQ1), as shown in Figure 1A. Additionally, we find that news diets of users, and engagement with such news, are biased toward their ideological leaning (RQ2). However, these biases are not symmetric. Conservative users exhibit a bimodal engagement pattern, splitting interaction between moderate and extreme sources, while liberals follow a more unimodal trend. Additionally, conservative users predominantly consume lower-quality sources, with this trend becoming more pronounced among more extreme users, aligning with previous findings [13].

In terms of the longitudinal analysis, we find two major shifts in engagement trends (Figure 1B): an increase in 2018 and a decrease in 2020 (RQ3). These trends are mirrored in engagement to URLs not in our popular news domain list, which suggests this is a platform-wide pattern (RQ4). Focusing on ideological segregation (Figure 1C) and low-quality news consumption (Figure 1D), the 2018 engagement increase coincides with a moderate rise in both, which is followed by a gradual decline until early 2020. Then a second change in the engagement trend occurs which is accompanied by a sizable increase in the ideological gap and prevalence of lower-quality sources (RQ5a and RQ5b). These trends suggest that the relationship between engagement, misinformation, and ideological segregation is complex. Finally, our analysis also identifies two distinct and abrupt decreases in the proportion of engagement with low-quality sources. Each of these declines is associated with the sudden disappearance of a different set of low-quality domains and is closely aligned with the detected change points (RQ5b).

Connecting these observations to known News Feed algorithm changes, our results suggest that Facebook's 2018 News Feed update, which prioritized engagement-based metrics [14, 15, 16], likely contributed to increased ideological segregation and lower-quality information exposure by favoring highly engaging content. Conversely, the less well-documented 2020 update [15, 16], coincided with reduced engagement but also led to a stronger ideological gap and more prevalence of lower-quality news. While its exact goal remains unclear, it appears to have reduced the boost given to reshares while maintaining high visibility for comments [15, 16].

A key question that arises is whether the observed shifts in engagement trends are directly linked to these algorithmic changes and, if so, what mechanisms underlie these effects. Understanding how different engagement signals—such as comments, shares, and reactions—shape ideological segregation and the spread of low-quality information would require access to more detailed platform data. Our findings highlight the need for greater transparency and further research into platform-wide changes to assess their implications for news consumption and information quality.

References

- [1] Eytan Bakshy, Solomon Messing, and Lada A Adamic. Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132, 2015.
- [2] Ronald E Robertson, Jon Green, Damian J Ruck, Katherine Ognyanova, Christo Wilson, and David Lazer. Users choose to engage with more partisan news than they are exposed to on google search. *Nature*, pages 1–7, 2023.
- [3] Sandra González-Bailón, David Lazer, Pablo Barberá, Meiqing Zhang, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio, Deen Freelon, Matthew Gentzkow, Andrew M Guess, et al. Asymmetric ideological segregation in exposure to political news on facebook. *Science*, 381(6656):392–398, 2023.
- [4] Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini. The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9):e2023301118, 2021.
- [5] Matteo Cinelli, Emanuele Brugnoli, Ana Lucia Schmidt, Fabiana Zollo, Walter Quattrociocchi, and Antonio Scala. Selective exposure shapes the facebook news diet. *PloS one*, 15(3):e0229129, 2020.
- [6] Michela Del Vicario, Gianna Vivaldo, Alessandro Bessi, Fabiana Zollo, Antonio Scala, Guido Caldarelli, and Walter Quattrociocchi. Echo chambers: Emotional contagion and group polarization on facebook. *Scientific reports*, 6(1):37825, 2016.
- [7] David Garcia. Influence of facebook algorithms on political polarization tested. *Nature*, 620(7972):39–41, 2023.
- [8] Chhandak Bagchi, Filippo Menczer, Jennifer Lundquist, Monideepa Tarafdar, Anthony Paik, and Przemyslaw A Grabowicz. Social media algorithms can curb misinformation, but do they? *arXiv preprint arXiv:2409.18393*, 2024.
- [9] Solomon Messing, Christina DeGregorio, Bennett Hillenbrand, Gary King, Saurav Mahanti, Zagreb Mukerjee, Chaya Nayak, Nate Persily, Bogdan State, and Arjun Wilkins. Facebook Privacy-Protected Full URLs Data Set, 2020.
- [10] Hause Lin, Jana Lasser, Stephan Lewandowsky, Rocky Cole, Andrew Gully, David G Rand, and Gordon Pennycook. High level of correspondence across different news domain quality rating sets. *PNAS nexus*, 2(9):pgad286, 2023.
- [11] Ronald E Robertson, Shan Jiang, Kenneth Joseph, Lisa Friedland, David Lazer, and Christo Wilson. Auditing partisan audience bias within google search. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1–22, 2018.
- [12] Antonis Kalogeropoulos, Samuel Negredo, Ike Picone, and Rasmus Kleis Nielsen. Who shares and comments on news?: A cross-national comparative analysis of online and social media participation. *Social media+ society*, 3(4):2056305117735754, 2017.
- [13] Andy Guess, Kevin Aslett, Joshua Tucker, Richard Bonneau, and Jonathan Nagler. Cracking open the news feed: Exploring what us facebook users see and share with large-scale platform data. *Journal of Quantitative Description: Digital Media*, 1, 2021.

- [14] Adam Mosseri. Bringing people closer together. *Meta*, 2018.
- [15] Kearch Hagey and Jeff Horwitz. Facebook tried to make its platform a healthier place. it got angrier instead. *The Wall Street Journal*, 2021.
- [16] Arvind Narayanan. Understanding social media recommendation algorithms, 2023.

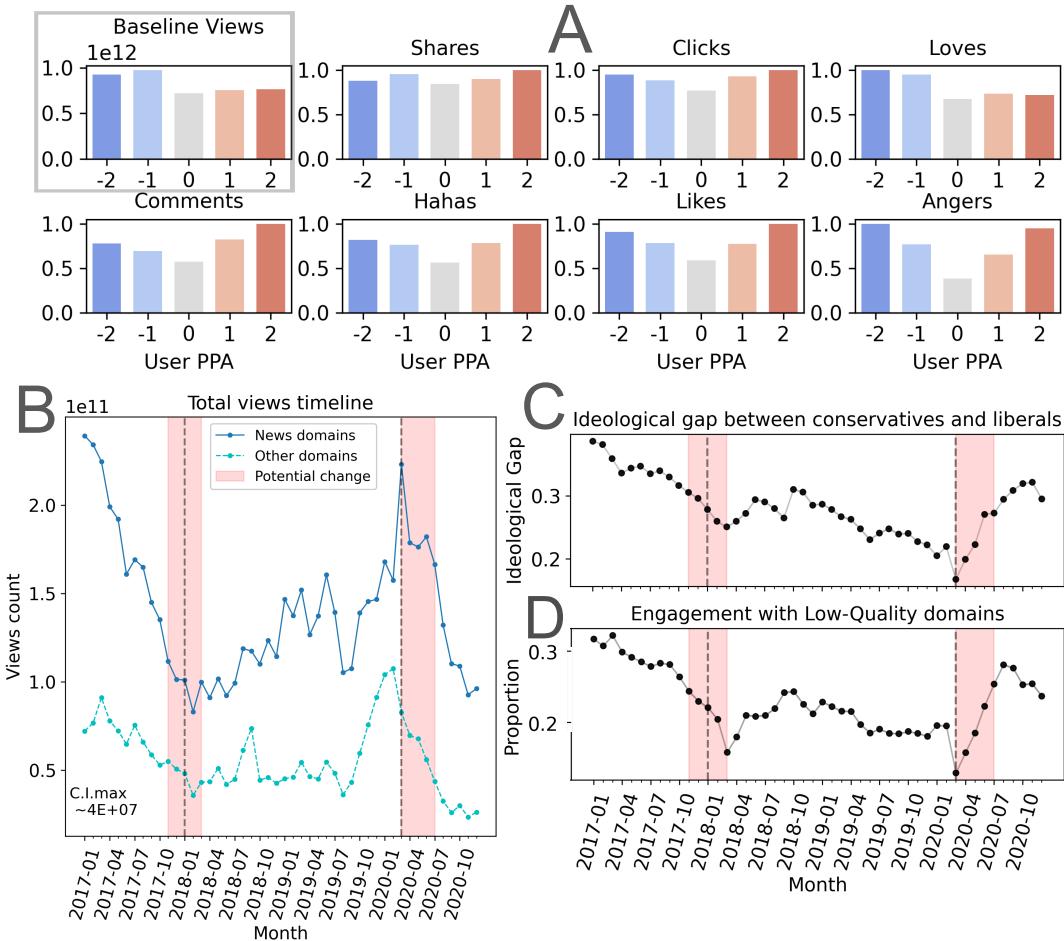


Figure 1: (A) Distribution of different engagement types to news-related content across user political page affinity (PPA), normalized relative to view counts (top left). For engagement types other than views, values are further normalized relative to their maximum across the five PPA categories. Engagement is unevenly distributed, with liberals having significantly more views than conservatives. The U-shaped pattern of engagement (greater engagement by extreme users) becomes more pronounced from left to right and from top to bottom. (B) Total monthly counts of views for news and other domains. Text labels indicate the maximum values of the noise uncertainty confidence intervals (CI). Shaded regions represent potential change points, identified through piecewise linear regression. Dashed lines mark the dates of two major documented algorithmic changes. (C) Ideological gap between the conservative and liberal averages, which can serve as a proxy for ideological segregation of news consumption. (D) Proportion of clicks directed towards low-quality domains, with quality scores lower than 0.6. Noise uncertainty intervals are too small to be visually detected in any of the subfigures.