



Selected Papers of #AoIR2025:
The 26th Annual Conference of the
Association of Internet Researchers
Niterói, Brazil / 15 – 18 Oct 2025

MISSING THE BIG PICTURE: PLATFORM OPACITY WEAPONIZED FOR DISINFORMATION IN THE TWITTER FILES BRAZIL CASE

Carlos Eduardo Barros
Universidade Federal do Rio de Janeiro - UFRJ

Rose Marie Santini
Universidade Federal do Rio de Janeiro - UFRJ

Thiago Ciodaro Xavier
Universidade Federal do Rio de Janeiro - UFRJ

Felipe Grael
Universidade Federal do Rio de Janeiro - UFRJ

Fernando Ferreira
Universidade Federal do Rio de Janeiro - UFRJ

Introduction

Under Musk's administration, X discontinued transparency reports on government requests, published by Twitter since 2012, and imposed barriers to data access for research via API, hindering disinformation research on the platform (Hickey et al., 2023). However, in April 2024, Musk released to one American and two Brazilian journalists some of these requests from Brazilian authorities, along with internal records, accusing them of violating users' privacy and free speech rights (Oxford Analytica, 2024). Relying on legal differences between Brazil and the U.S, the accusations followed refusals by X to cooperate on national security efforts against disinformation, hate speech, and various criminal investigations. Mirroring the U.S. 'Twitter Files' case (Bond, 2022), the Twitter Files Brazil (TFB) intensified the debate on state censorship, escalating further, with Musk demanding the resignation or impeachment of Supreme Court Minister Alexandre de Moraes (Vick, 2024). Far-right leaders and communities from Brazil and the U.S. amplified government persecution narratives, framing requests related to accounts investigated for crimes such as drug cartels as if they were targeted by posting conservative opinions (Caetano, 2024).

Against this backdrop, the sequence of data restrictions and claims against state intervention fits as part of a broader trend on platforms building politically induced

Suggested Citation (APA): Barros, C. E., Santini, R. M., Xavier, T. C., Grael, F., & Ferrreira, F. (2025, October). *Missing The Big Picture: Platform Opacity Weaponized For Disinformation In The Twitter Files Brazil Case*. Paper presented at AoIR2025: The 26th Annual Conference of the Association of Internet Researchers. Niterói, Brazil: AoIR. Retrieved from <http://spir.aoir.org>.

ignorance to avoid accountability (Napoli, 2024a; 2024b). Furthermore, the geopolitical implications surrounding the TFB position it as a key case for analyzing social media manipulation (Bradshaw & Howard, 2018; DiResta, 2018) and also through the lens of digital colonialism critique in the Global South (Faustino & Lippold, 2023; Pinheiro-Machado & Vargas-Maia, 2023). Within this framework this paper examines the debate surrounding the TFB on X to investigate the spread of computational propaganda and disinformation targeting Brazilian government and institutions. How has the TFB controversy contributed to the political manipulation of platform governance and accountability?

Methods and Results

To contextualize, we first reviewed platform communications regarding changes to transparency reports and data access for Brazil following Musk's acquisition, such as every document released on the TFB. To capture the international debate after his call for Moraes' removal from office, we then collected and analyzed posts on X from April 8–9, 2024, using a query in both English and Portuguese. We developed network analysis to group accounts by sharing interactions, as indicators of influence and a key factor to map disinformation. The accounts were compared across creation dates and other aspects. Since the API restrictions rendered established detection models for coordinated inauthentic behavior (CIB) unfeasible (Hickey et al., 2023), we relied on historical data to map accounts flagged for CIB in earlier research about the Brazilian political debate (Santini et al., 2023). The 10 most shared posts on each group were analysed for disinformation.

Our collection found 90,434 unique publications matching, of which 45.8% were shares. The sharing network analysis revealed a clear polarization, leading to a division of two distinct communities. Musk is situated in Community 1, which comprises 25,988 profiles (88.2%). Top shared posts on this community came from far-right U.S. figures such as Simon Goddek, Rob Schneider and Alex Jones, besides Musk himself. Community 2, with 3,390 profiles (11.8%), was led by Brazilian progressivist influencers holding the top shared posts, including the country's first lady, Janja Lula. Only 15,517 sharing accounts (52.82%) matched available data about CIB analysis from previous research. This is due to a large number of U.S.-based accounts that had not previously engaged with the Brazilian political debate and accounts created after Musk's acquisition (October, 2022). However, a notable share of 4,709 accounts (30.34% of those with available data) were flagged for CIB, highlighting the platform's limited effectiveness in curbing this issue (Hickey et al., 2023). Figure 1 shows the ratio of recent created accounts by community.

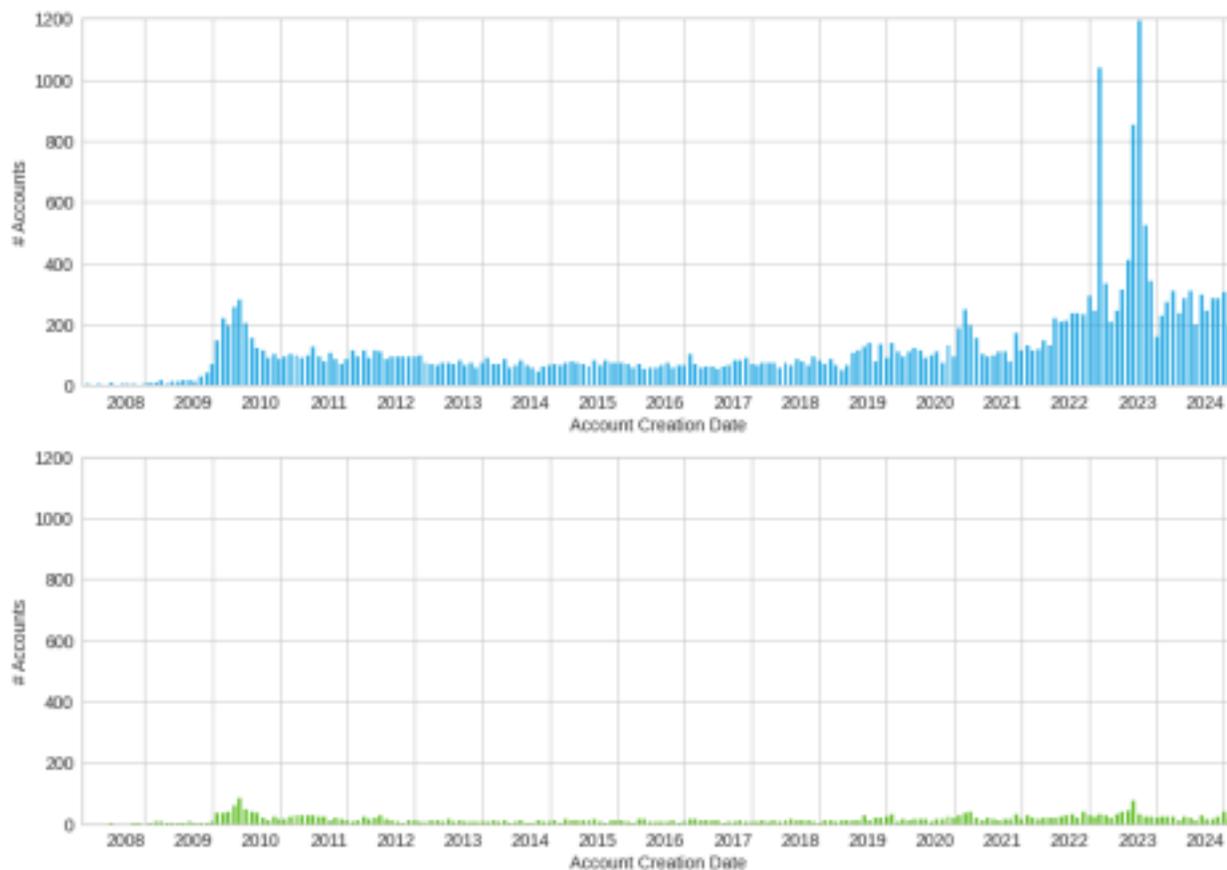


Figure 1: Account Distribution by Creation Date.
 Source: the authors.

Despite the concentration of recently created accounts on Community 1, the proportion of CIB historical accounts was higher in this community (31.51%) compared to Community 2 (23.94%). Although U.S. accounts were the most shared in Community 1, the community is largely composed of Brazilian far-right groups and supporters of former President Jair Bolsonaro, as the federal deputy Nikolas Ferreira, among the national hubs. Both communities featured posts in English, with many Brazilian accounts seeking international attention, and in Portuguese, focusing on the domestic debate over the legitimacy of the moderation requests. Figure 2 shows a graph of the sharing network, where profiles are grouped into two communities: Community 1 (blue) and Community 2 (green). Profiles without historical CIB information are gray, while CIB-marked profiles are red.

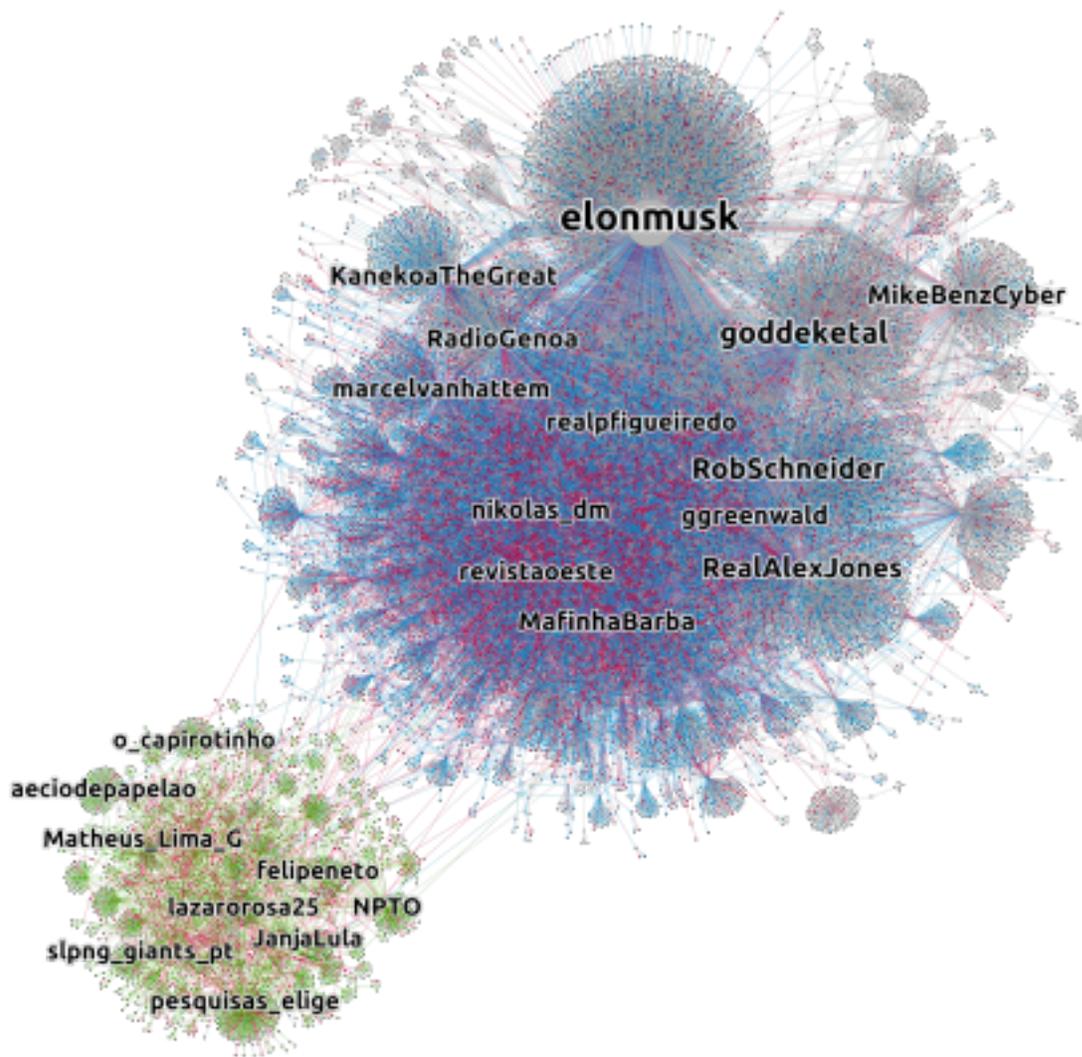


Figure 2: Network of accounts connected by share relations.
Source: the authors.

As the major hub of the network, Musk's account is also the most shared by early flagged CIB accounts, with his main post mocking that Alexandre de Moraes had become the dictator of Brazil, with Lula "on a leash". Nikolas Ferreira responded to this post asking for more information, to which Musk replied that he would do a "full data dump" after X employees were safe from prison threats. However, the alleged prison threat did not involve Moraes or the Supreme Court, but rather a local investigation into a known cartel, as later admitted by the three journalists responsible for publishing the TFB (Caetano, 2024).

In the most shared post on the network, Simon Goddek screenshots the interaction between the CEO of X and the Brazilian congressman, claiming that Brazil would be a "quasi-dictatorship". By focusing on the legality of Brazil's requests, X deflects responsibility for its own moderation powers (Napoli, 2024a), continuously arbitrating removals without accountability and enabling falsehoods driven by political propaganda (Hickey et al., 2023; DiResta, 2018).

Musk dismissing Brazilian requests by releasing accounts engaged in disinformation, hate speech, and other crimes highlights the alignment between the far-right in Brazil and the U.S. in discrediting the government. The most shared posts in Community 2, however, denounce inconsistencies in the TFB, but were also amplified by historical CIB accounts, albeit to a lesser extent. This indicates that this technique is not exclusive, even though some groups benefit more from it than others due to the asymmetries of the information industry.

Scientific and Societal Contribution

Amidst concerns over social media manipulation (Bradshaw & Howard, 2018; DiResta, 2018), this study highlights how the TFB controversy has been strategically used by far-right leaders to delegitimize Brazilian institutions, amplifying false claims about Brazil's judiciary and reinforcing narratives of illegal censorship and persecution. This aligns with efforts to challenge democratic structures and attract international attention to movements undermining governance.

Our analysis reveals continued challenges, with historical CIB accounts amplifying false claims against Brazilian authorities, highlighting platform policy shortcomings. These findings addressed not only domestic manipulation in Brazilian politics but also the exportation of disinformation to international audiences, addressing the debate on digital platforms and colonialism (Faustino & Lippold, 2023).

Furthermore, the lack of transparency under Musk's administration worsened conditions for the spread of false claims, as a "theater of transparency" (Bouko; van Ostaeyen; Voué, 2021), aiming to direct public attention to selected topics and sources while hindering broader and independent analyses. It reinforces the argument that platform opacity weaponizes opacity, jeopardizing democratic governance (Napoli, 2024b).

References

- Bradshaw, S., & Howard, P. N. (2018). Challenging truth and trust: A global inventory of organized social media manipulation. *The computational propaganda project*, 1, 1-26. <https://philhoward.org/wp-content/uploads/2022/08/ct2018.pdf>
- Bond, S. (2022, December 14). Elon Musk is using the Twitter Files to discredit foes and push conspiracy theories. *NPR*. <https://www.npr.org/2022/12/14/1142666067/elon-musk-is-using-the-twitter-files-to-discredit-foes-and-push-conspiracy-theor>
- Bouko, C., Van Ostaeyen, P., & Voué, P. (2021). Facebook's policies against extremism: Ten years of struggle for more transparency. *First Monday*. <https://firstmonday.org/ojs/index.php/fm/article/view/11705>
- Caetano, F. (2024, April 10). Advogado não foi ameaçado pelo STF, como afirmou autor dos Twitter Files, mas pelo MP-SP em investigação sobre o PCC. *O Globo*. <https://oglobo.globo.com/politica/noticia/2024/04/10/advogado-nao-foi-ameacado-pelo-stf-como-afirmou-autor-dos-twitter-files-mas-pelo-mp-sp-em-investigacao-sobre-o-pcc.ghtml>

DiResta, R. (2018). Computational propaganda: If you make it trend, you make it true. *The Yale Review*, 106(4), 12-29.

<https://muse.jhu.edu/pub/1/article/791857/summary>

Faustino, D., & Lippold, W. (2023). *Colonialismo digital: por uma crítica hacker-fanoniana*. Boitempo Editorial.

Hickey, D., Schmitz, M., Fessler, D., Smaldino, P. E., Muric, G., & Burghardt, K. (2023, June). Auditing Elon Musk's impact on hate speech and bots. In Proceedings of the international AAAI conference on web and social media (Vol. 17, pp. 1133-1137). <https://ojs.aaai.org/index.php/ICWSM/article/view/22222>

Oxford Analytica. Brazil social media legal vacuum poses controversies. Emerald Expert Briefings, n. oxan-db, 2024. <https://doi.org/10.1108/OXAN-DB286367>

Pinheiro-Machado, R., & Vargas-Maia, T. (Eds.). (2023). *The rise of the radical right in the Global South*. Abingdon: Routledge.

Perriam, J., Birkbak, A., & Freeman, A. (2020). Digital methods in a post-API environment. *International Journal of Social Research Methodology*, 23(3), 277-290. <https://www.tandfonline.com/doi/abs/10.1080/13645579.2019.1682840>

Santini, R. M., Salles, D., Ferreira, F., & Grael, F. (2023). Gotcha Bot Detection: Context, Time And Place Matters. In SciELO Preprints. <https://doi.org/10.1590/SciELOPreprints.5974>

Vick, M. (2024, April 11). Disinformation in the Twitter Files. And the Supreme Court's action in the X case. *Nexo Jornal*. <https://www.nexojournal.com.br/expresso/2024/04/11/x-antigo-twitter-transparencia-fake-news>

Yang, K. C., Varol, O., Hui, P. M., & Menczer, F. (2020, April). Scalable and generalizable social bot detection through data selection. In Proceedings of the AAAI conference on artificial intelligence (Vol. 34, No. 01, pp. 1096-1103). <https://ojs.aaai.org/index.php/AAAI/article/view/5460>

Napoli, P. M. (2024a). Agnotology, Free Speech, and the Precarious Politics of Media Research (SSRN Scholarly Paper No. 4695907). Social Science Research Network. <https://doi.org/10.2139/ssrn.4695907>

Napoli, P. M. (2024b). In pursuit of ignorance: The institutional assault on disinformation and hate speech research. *The Information Society*, 41(1), 1–17. <https://doi.org/10.1080/01972243.2024.2419105>