# Rationale for Analysing Bluesky Posts on the Migration from X

## Introduction

Bluesky is a decentralised social media platform that started as a project under X (formerly Twitter) but later became independent. On November 5, 2024, the day of the US election, Bluesky saw a rise in new users, many of whom shared why they left X. This shift reflects growing dissatisfaction with X, driven by political and management divisions. While there has been research on social media migration, Bluesky has not been widely studied as an emerging platform.

The phenomenon observed on the US election day is a new development, but existing research on social media platform migration provides valuable context. Ryan Kor-Sins explored the political dynamics of alternative social media, focusing on Gab through the lens of heterogeneous engineering. His analysis argues that Gab has become a haven for alt-right ideologies (Kor-Sins, 2021). A team from the University of Notre Dame studied reputation transfer by tracking user IDs migrating from X to Threads, Mastodon, and Truth Social (Radivojevic et al., 2024). Additionally, researchers from Arizona State University established foundational work in 2011 by defining social media migration patterns (Kumar et al., 2011). They recently expanded their research to examine migration from X to Threads, Bluesky, and Mastodon, employing a methodology that combined account data analysis via APIs, post text analysis to measure activity levels, and sentiment word clouds to evaluate user loyalty (Jeong et al., 2023). These studies have inspired the approach and methodology of this project.

This project investigates the discourse surrounding this migration, focusing on how users perceive their move to Bluesky and the factors driving their decision.

## Research Question

How do Bluesky users perceive this phenomenon? What underlying reasons or patterns can be identified in this migration?

**Data Collection**

Bluesky's open-source nature and accessible HTML architecture made it an excellent choice for data collection, especially in contrast to X's restrictive API policies.

For this project, posts containing the keyword 'left X' were gathered using the 'top' sorting option, which prioritises highly engaged content. Only English-language posts were included to maintain consistency during analysis.

The scrapping process utilized the BlueSkyScraper GitHub project by F.C. Gouveia (Gouveia, 2024), with adjustments made to collect timestamps alongside the post content. Selenium WebDriver was used to scrape the data, while Pandas was employed for management and procession.

A total of 8,265 posts were collected between December 9, 23:12 and December 10, 00:44 (UK time), providing an adequate dataset for analysis.

**Data Analysis**

1. **Timeline Analysis**:

   The frequency of posts over time was visualized using Matplotlib. To facilitate the analysis, raw timestamp strings were converted into Unix timestamps, using only the date for statistical purposes. These timestamps were then plotted and converted back into readable date formats in the graph.

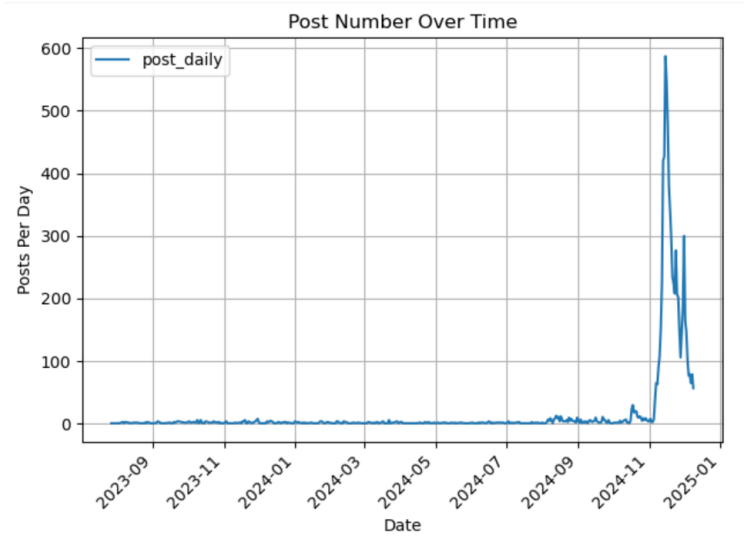


Figure 1

After observing a noticeable surge in activity in the latter half of 2024, the graph was zoomed into the range from August to December 2024, providing a clearer view of the trends. A red vertical line was added to highlight the election day.

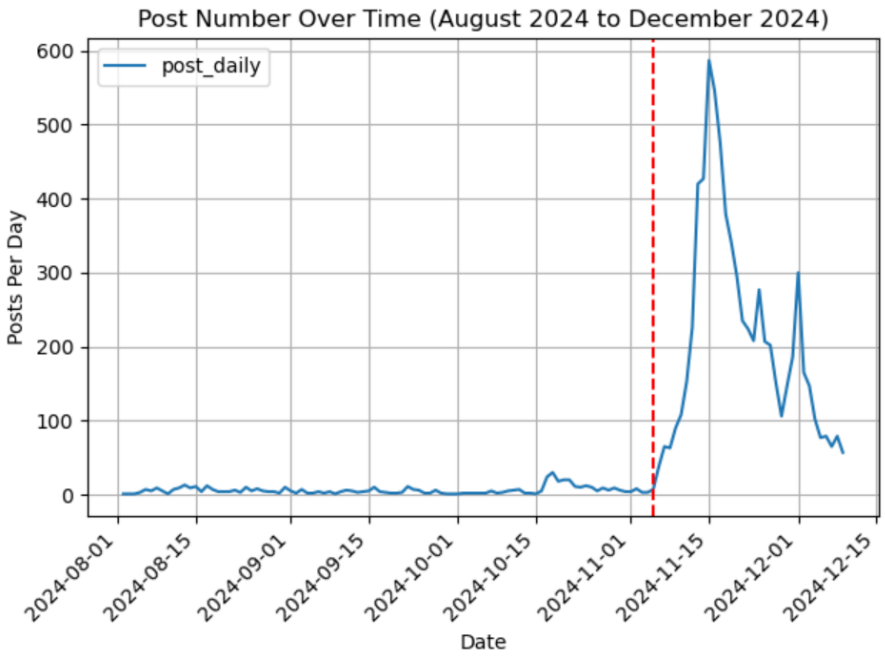


Figure 2

2. **Textual Analysis**:
    - **Stopwords Removal**: NLTK's stopword corpus was used to eliminate unessential words, allowing for a cleaner dataset and more meaningful analysis of frequent terms.
    - **Word Frequency**: The collections.Counter class was used to count word frequencies in the dataset. To refine the analysis, I filtered out words that appeared less than 50 times and excluded the topic-specific words: “left”, “X” and “Bluesky”.
    - **POS Tagging**: Given SpaCy’s text length limitations, the analysis was adjusted by first performing word frequency analysis, then using SpaCy to tag the most frequent terms.
    - **Word Cloud**: A word cloud was created to highlight the most frequent terms, with larger words showing higher frequencies. I focused on adjectives and nouns, colouring adjectives red and nouns orange to make them easier to distinguish. The

word cloud was generated using the WordCloud class from the wordcloud library, and Matplotlib was used to display it clearly.


Figure 3

**Findings and Reflections**

1. The volume of posts mentioning “left X” increased significantly after election day, peaking on November 15.
2. Users expressed predominantly positive emotions about leaving X, using terms such as “good”, “glad” and “happy”. While X was frequently described as “toxic”, highlighting dissatisfaction with the platform's environment.
3. Elon Musk was the most mentioned individual, far surpassing mentions of Donald Trump, despite the election being the triggering event.

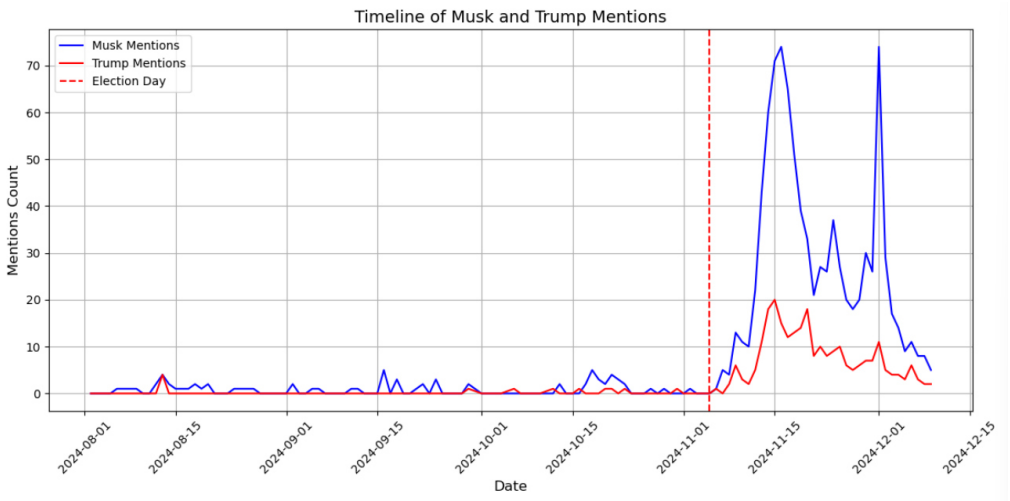


Figure 4

## Challenges and Limitations

1. The ‘latest’ sorting option was unavailable, raising questions about whether the data fully represents the migration discourse.
2. Reliance on the keyword “left X” may have excluded other relevant posts.
3. The length of post text exceeded SpaCy’s parsing capacity, limiting further text analysis.
4. Emojis and images were not analysed, which excluded non-textual elements that might have provided additional insights into user sentiment.
5. Data was collected using the UK time zone, which may have introduced slight discrepancies given the US-centric nature of the event.

## Conclusion

This study confirms the observation that the 2024 US election day prompted a significant migration of English-speaking users from X to Bluesky. Interestingly, while the triggering event was political, the data revealed that dissatisfaction with Elon Musk’s leadership at X was a more dominant factor in user sentiment.

**References**


Kor-Sins, R., 2021. The alt-right digital migration: A heterogeneous engineering approach to social media platform branding. *New Media & Society*, 25, pp. 2321 - 2338. doi:10.1177/14614448211038810.

Radivojevic, K., Adams, D., Laszlo, G., Kery, F., & Weninger, T., 2024. Reputation Transfer in the Twitter Diaspora. *ArXiv*, abs/2405.12040. doi:10.48550/arXiv.2405.12040.

Kumar, S., Zafarani, R., & Liu, H., 2011. Understanding User Migration Patterns in Social Media. *Proceedings of the AAAI Conference on Artificial Intelligence*. doi:10.1609/aaai.v25i1.8089.

Jeong, U., Nirmal, A., Jha, K., Tang, X., Bernard, H., & Liu, H., 2023. User Migration across Multiple Social Media Platforms. , pp. 436-444. doi:10.48550/arXiv.2309.12613.

Fcgouveia/blueskyscraper: A simple bluesky scraper in python using selenium Library, GitHub. Available at: https://github.com/fcgouveia/BlueSkyScraper (Accessed: 04 December 2024).

Balduf, L. et al. (2024) 'Looking at the blue skies of Bluesky', Proceedings of the 2024 ACM on Internet Measurement Conference, pp. 76–91. doi:10.1145/3646547.3688407.

Brown, J., Onik, A.R. and Baggili, I. (2024) 'Blue Skies from (x?s) pain: A digital forensic analysis of threads and Bluesky', Proceedings of the 19th International Conference on Availability, Reliability and Security, pp. 1–12. doi:10.1145/3664476.3670904.

Failla, A. and Rossetti, G. (2024) '"I'm in the bluesky Tonight": Insights from a year worth of social data', PLOS ONE, 19(11). doi:10.1371/journal.pone.0310330.

Mulyadi, D. et al. (2024) 'The negative impact of social media on generation Z in election activities in 2024', Sinergi International Journal of Law, 2(3), pp. 183–197. doi:10.61194/law.v2i3.151.

Puspita, T. et al. (2024) 'Political identity in the 2024 elections through Social Media Instagram', International Journal of Social Service and Research, 4(9). doi:10.46799/ijssr.v4i9.867.

Sinaga, N. et al. (2024) ‘Analysis of political communication on the role of social media towards the 2024 general election’, Proceedings of International Conference on Social Science, Political Science, and Humanities (ICoSPOLHUM), 4, p. 00039. doi:10.29103/icospolhum.v4i.468.