Revista Latina de Comunicación Social
ISSN 1138-5820 / No. 83, 1-14.
https://doi.org/10.4185/rlcs-2025-2418

X Data-Based Scientific Research: A Review of Trends and Challenges

Lucía Rivadeneira
Technical University of Manabí. Ecuador.
lucia.rivadeneira@utm.edu.ec

Ignacio Loor
Technical University of Manabí. Ecuador.
ignacio.loor@utm.edu.ec

How to cite this article / Standard reference:

Rivadeneira, Lucía & Loor, Ignacio (2025). X Data-Based Scientific Research: A Review of Trends and Challenges [*Los Datos de X en la Investigación Científica: Tendencias y Desafíos*]. *Revista Latina de Comunicación Social*, 83, 1-14. https://www.doi.org/10.4185/RLCS-2025-2418

Date Received: 03/10/2024
Acceptance Date: 18/12/2024
Date of Publication: 20/03/2025

ABSTRACT

Introduction: The growth of social media, especially X (formerly Twitter), has become a key resource for scientific research. This literature review identifies the factors driving its use, forecasts trends, and addresses challenges faced by researchers. **Methodology:** The review, based on a systematic search in Scopus, employed thematic mapping to identify interdisciplinary applications, methodological innovations, and the impact of global events. Key among these innovations was natural language processing (NLP) for data analysis, which grew 268% from 2019 to 2023. **Results:** NLP has established itself as a vital tool. However, publications based on X data showed a slowdown between 2021 and 2023, while Instagram and TikTok-based publications accelerated, signaling increased interest in these platforms. X remains the most used platform, followed by Facebook. **Conclusions:** The review highlights the need for more advanced analysis methods, stronger ethical standards concerning privacy and consent, and interdisciplinary approaches in social media research.

Keywords: social media; Twitter; X; research; trends.

1. INTRODUCTION

The incorporation of data from X, formerly known as Twitter, into scientific research represents a fundamental intersection between social network analysis and various academic disciplines. This article, based on a literature review, examines the use of data generated by Twitter between 2019 and 2023, the period during which its rebranding as X took place. The study identifies the factors that have driven the use of data from this platform in scientific research, anticipates future trends, and addresses the main barriers faced by researchers, such as data access restrictions, issues of representativeness, and the emergence of new platforms.

At least since Asur and Huberman (2010), Twitter has consolidated itself as an essential platform for collecting data for the analysis of social, discursive and communicative phenomena, thanks to its ability to provide real-time data on public opinion, social mobilization and information dissemination. However, the rise of platforms such as TikTok and Instagram, together with changes in the demographic structure of X users, pose new challenges for its use as a primary data source in interdisciplinary research.

Over the course of Twitter's trajectory in Scopus-indexed manuscripts, 377 literature review articles were included in the content analysis of this platform. With such a large volume of reviews, it can be difficult to identify aspects that have not yet been explored in depth. This review distinguishes itself by addressing the methodological and ethical difficulties identified in previous studies and by adopting a forward-looking perspective, highlighting both opportunities and challenges for social network research in an ever-changing digital environment. Its importance lies in providing an up-to-date overview of the state of the art, highlighting current trends, emerging challenges and practical recommendations to strengthen research in social networks.

2. OBJECTIVES

Three objectives guided the scope of this research. First, we sought to unravel the factors that have contributed to the rise of Twitter data-driven research, focusing on analyzing the use of social media data in academia and its impact on knowledge generation. Secondly, we aim to anticipate future trends in the field by identifying emerging themes and potential areas of growth in research using X data. Finally, we aim to outline the main difficulties that researchers have reported when working with Twitter data, covering concerns related to access, analysis and ethical issues.

3. METHODOLOGY

The study consisted of a systematic literature review focusing on the use of Twitter data in scientific research. The process began with a search of the Scopus database, selected for its broad coverage of international and multidisciplinary academic publications. The search strategy included the terms Twitter *AND* data *OR* content *AND* analysis, limiting the results to the years 2019-2023 and to document types such as articles, conference proceedings and book chapters. This initial search yielded a total of 11,058 publications.

In order to refine the results, additional filters were applied to include only publications in English and Spanish and to restrict the search to thematic areas related to the social sciences. These criteria reduced the number of publications to 3,227. The titles and abstracts of these publications were then exported into an Excel file for further analysis. At this stage, a systematic keyword-based process was implemented to align the results with the objectives of the study. Keywords for this filtering process were associated with each objective: to identify factors contributing to the increase in Twitter-based research, terms such as growth, adoption, popularity, increase, impact and use of Twitter were used; to anticipate future trends, terms such as future, emerging issues, opportunities, development and progress were included; while to examine the main difficulties, terms such as challenges, limitations, barriers, access, privacy, ethics and bias were used.

Using Excel formulae, each abstract was automatically scored for the presence of these key terms and those that addressed one or more objectives were flagged for inclusion. The abstracts were then ranked according to the number of objectives they met, with priority given to those that addressed at least two of the objectives of this study. This step allowed us to refine the set to 80 abstracts. Finally, a manual review of these abstracts was performed to verify their relevance and alignment with the study objectives, resulting in the final selection of 43 manuscripts. Table 1 below details the steps involved in the search and selection of articles.

Table 1. *Steps for searching and selecting relevant publications*

	Details	Code / formula
Search in Scopus	Initial search strategy in Scopus focusing on Twitter and data analysis between 2019 and 2023.	Twitter AND data OR content AND analysis
Filter by language	Publications limited to English and Spanish.	Applied directly in Scopus
Filter by subject area	Restricted to areas related to social sciences.	Applied directly in Scopus
Filter by keyword (Excel)	Identification of relevant summaries according to objectives: Factors, Trends and Difficulties.	=IF(OR(ISNUMBER(SEARCH("growth"; C2)); ISNUMBER(SEARCH("adoption"; C2)); ...), "Yes"; "No") (for each objective)
Rating	Prioritizing abstracts that address multiple objectives.	=COUNTIF(D2:F2; "Yes")
Manual review	Review of the 80 most relevant abstracts to ensure alignment with research objectives.	Non-code subjective assessment of abstracts flagged during filtering.

Source: Own elaboration (2024).

The 43 selected manuscripts were examined using thematic analysis following the methodology of Soaita et al. (2020), which involves systematic coding of data and identification of recurring patterns and themes. To ensure the validity of the thematic analysis, the process was collaborative between the study authors, who discussed and revised the initial codings in several iterative rounds. This allowed emerging categories to be identified and refined, and grouped into broader themes in a consensual manner, thereby reducing the potential for individual bias. In addition, decisions made during the analysis were assessed for their consistency with the study objectives, so that the final themes were directly informed by the manuscript data.

4. RESULTS

Research based on social media data has experienced significant growth since 2010, especially in the case of Twitter, which has established itself as a key platform for the analysis of social, political and economic phenomena. This boom responds both to its ability to capture social interactions in real time and to technological advances that have made it possible to handle large volumes of unstructured data with greater precision. However, the increasing reliance on these platforms also poses critical challenges, such as the need to ensure sample representativeness, overcome data access limitations, and address emerging ethical issues. The findings presented below examine the factors behind the growth of research using Twitter data, the trends and opportunities that are shaping the future of the field, and the challenges researchers face in accessing and analyzing it, shaping a landscape that is constantly evolving.

4.1 Factors behind the growth of the research using Twitter data

Over the past five years, research based on Twitter's data has grown rapidly, driven by a combination of technological, social and academic factors. The number of publications in Scopus analyzing Twitter content increased from 1,827 in 2019 to 2,723 in 2023, reflecting a 49% increase that evidences the growing relevance of this platform in academic research. This phenomenon has consolidated Twitter as a key tool for addressing contemporary research questions, offering unprecedented methodological depth in diverse disciplines. Among the most notable factors are the expansion of the platform's reach, defined by its number of users, technological advances, emerging preferences in academia, and the integration of big data analysis in the social sciences.

The global reach and ability to reflect social and global dynamics in real time have been essential to its widespread adoption in scientific research. Nellore et al. (2023) note that the accessibility of this platform has enabled the study of complex and diverse social phenomena, making it an indispensable resource in disciplines ranging from public health to communication studies. Examples of these applications include the analysis of social responses to critical events such as pandemics, political unrest, and natural disasters (Cvetojevic & Hochmair, 2018; Raja et al., 2016). These studies have demonstrated how Twitter data capture the pulse of social interactions at key moments, providing a unique window to study how societies respond to high-impact events. In addition, Henry et al. (2018) highlight that this platform also serves as a channel for the transnational dissemination of information, reinforcing its global relevance in research.

However, Twitter's ability to reflect social dynamics in real time poses significant challenges. The ephemeral and fragmented nature of interactions can introduce bias into analyses, prioritizing emergent narratives while neglecting deeper structural aspects. For example, while Cvetojevic and Hochmair (2018) and Raja et al. (2016) have demonstrated the potential of the platform to analyze social reactions, the risk of overrepresenting the perspectives of active or digitally connected users may exclude less visible communities.

Twitter has also become a space where voices can be amplified and public opinion can be mobilized, challenging traditional forms of information dissemination. Mao et al. (2023), in the context of the COVID-19 pandemic, explore how Twitter's interaction with traditional media helps shape news flows, highlighting the platform's role in disseminating critical information in real time. Yuan (2017), in the context of the U.S. presidential election, emphasizes its role in the analysis of communication patterns, showing how Twitter can be used to identify dynamics in the spread of political messages and predict future trends. In the same vein, Dersan Orhan (2020), by analyzing the reactions to Donald Trump's tweets during his presidency regarding Iran, demonstrates how the platform can act as a real-time barometer to inform strategic decisions, providing a detailed insight into public perceptions on foreign policy issues. In this regard, Efanova (2023), in a study of U.S. digital diplomacy, highlights the use of Twitter as a diplomatic tool, evidencing its role in narrative construction and symbolic negotiation in complex international contexts.

However, these apparent priorities are not without limitations. The increasing use of bots (Ghosh et al., 2023) and automated accounts poses additional challenges, raising doubts about the authenticity of interactions and thus the validity of analyses based on Twitter data. This highlights the need for more robust methods to distinguish between human behavior and automated activity to ensure the reliability of results.

On the other hand, technological advances, from devices with greater storage and processing capacity to more sophisticated algorithms, have strengthened the ability of researchers to analyze the data generated on Twitter. Advanced tools such as machine learning (ML) and natural language processing (NLP) have been instrumental in handling large amounts of unstructured data (Yu & Muñoz-Justicia, 2022). Unstructured data refers to information that does not follow a predefined model, such as large amounts of text. In particular, NLP has emerged as a key method for interpreting textual content on the platform, providing a more nuanced understanding of digital discourse and its context. These innovations have broadened the scope of scholarly analysis, allowing us to explore social dynamics and patterns of communication in greater depth.

At the same time, in their efforts to advance analytical methods, data scientists have relied heavily on Twitter data to test and validate models and algorithms, which has significantly boosted the growth of research based on this platform. For example, Budiharto and Meiliana (2018) used Twitter data related to presidential elections in Indonesia to evaluate the predictive and explanatory power of their models. Similarly, Rangel et al. (2020) used these data to develop models focused on the study of misinformation and fake news. These applications reflect not only a growing interest in digital phenomena, but also a shift in the research paradigm toward more interdisciplinary approaches adapted to the digital environment.

Finally, the integration of big data analytics has redefined the possibilities for social science research. Schroeder and Cowls (2019) and Steinert-Threlkeld (2018) highlight how the use of large Twitter datasets has enabled a more detailed understanding of human behavior and fostered the development of new theoretical models. Applications in financial analysis, such as those of Pagolu et al. (2017) and Zou and Herremans (2023), illustrate the ability of these data to integrate economic theories with advanced computational methods. Similarly, Wang et al. (2019) highlight how Twitter data have linked social and environmental responses in disaster management, reinforcing the value of interdisciplinary approaches.

In sum, the factors behind the growth of Twitter data-driven research reflect a dynamic academic environment in which the platform not only facilitates the exploration of complex social phenomena but also redefines research practices. While technological advances and scholarly interest have solidified its relevance, challenges related to data representativeness and authenticity remain. This invites the scholarly community to continue to innovate in both methodological development and a critical approach to the use of these data to ensure that the opportunities offered by Twitter are used responsibly and effectively.

4.2. Trends and opportunities in research with data from X

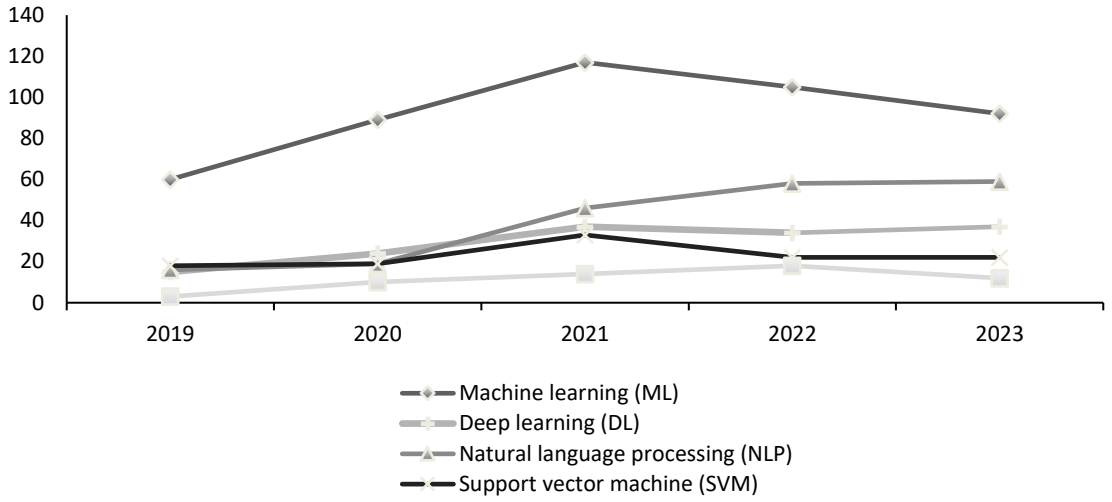
Research with Twitter data has experienced a multi-dimensional expansion in recent years, encompassing new application areas and methodological advances. On the one hand, these data have been progressively integrated into a growing number of disciplines, finding applications in fields such as public health, disaster management, environmental studies, and predictive analytics. On the other hand, data mining and analysis methods have become increasingly sophisticated, enabling more accurate predictions and more transparent representations. Machine learning techniques have emerged as essential tools for analyzing large volumes of structured and unstructured data with remarkable accuracy. At the same time, their conceptual scope has evolved, leaving behind purely descriptive approaches and giving way to the construction of more complex theoretical models that offer innovative perspectives on social and global phenomena.

A clear example of this multidisciplinarity is public discourse analysis. In this area, Fütterer et al. (2023) and Hu et al. (2021) have demonstrated how computational methods facilitate the exploration of collective attitudes and perceptions on topics of common interest. These applications reflect not only the versatility of the platform, but also the growing recognition of its potential to address complex problems from multiple disciplines in an integrated manner.

The rise of interdisciplinary applications of X data highlights its potential to address complex problems from multiple perspectives. However, it also raises the challenge of integrating findings from different methodologies and priorities. For example, disaster management may use X data to monitor immediate responses and coordinate actions, while public health studies may focus on long-term impacts on collective well-being. Far from being an obstacle, this methodological diversity underscores the need to develop analytical frameworks that link seemingly disparate findings to build more complete narratives that reflect the inherent complexity of social and global phenomena.

In parallel, technological developments have redefined the methods for analyzing X data, prioritizing advanced techniques such as ML and NLP. Ioannides et al. (2023) highlight that NLP has seen a steady increase in adoption due to its ability to extract patterns from the dynamic and often informal language of X. This method has proven essential for sentiment analysis, prediction, and understanding complex social dynamics. Figure 1 shows how NLP has outperformed other methods, such as support vector machines (SVM) and decision trees, both of which have seen significant declines in recent years. Nguyen et al. (2024) note that deep learning (DL) peaked in adoption in 2022, although its use declined slightly in 2023, suggesting a possible maturation of these methods.

Figure 1. *Trends in the use of analytical methods in Twitter data (2019-2023).*

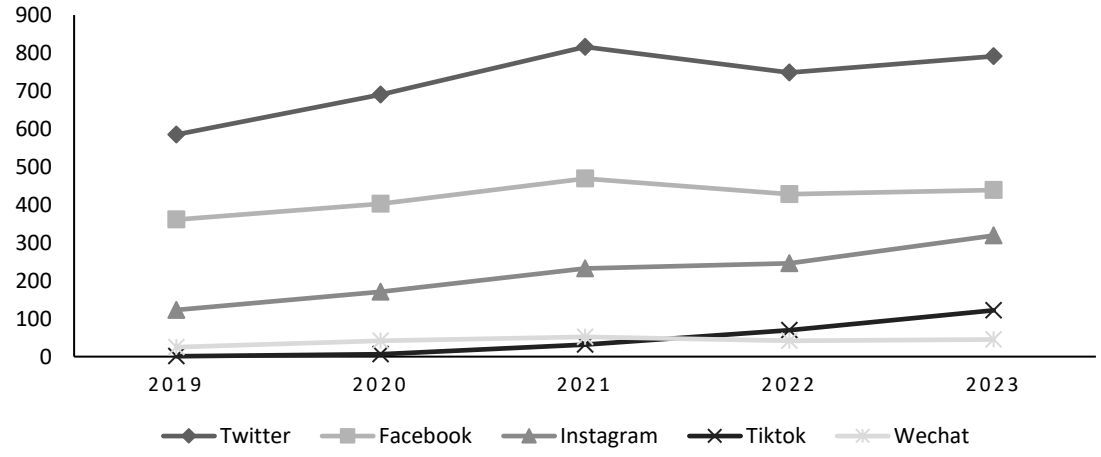


Source: Scopus' dababase.

Although advanced tools such as NLP and DL have enabled more efficient handling of X-generated data, their increasing reliance in research poses challenges related to the opacity of these methods. The "black box" of algorithms used in ML and DL can make it difficult to interpret and replicate results, thus compromising scientific transparency (Mazhar & Dwivedi, 2024). Therefore, when choosing a tool for X content analysis, it is important for researchers to consider not only the efficacy of these tools, but also how to communicate their limitations and methodological choices.

Methodological advances aside, the trend toward cross-platform analysis is shaping a more diverse landscape for academic research. Figure 2 compares the frequency of publications based on data from X, Instagram, TikTok, and WeChat between 2019 and 2023. While X peaked in 2021, 2022 shows a significant slowdown in its use in research, which may signal a maturity phase in its adoption. In contrast, TikTok and Instagram show sustained growth, with TikTok standing out for its ability to capture the social dynamics of younger audiences through visual and participatory formats. Sinnenberg et al. (2017) suggest that while X remains irreplaceable for certain types of research, diversification to these platforms offers new opportunities to explore social phenomena from a comparative perspective.

Figure 2. *Comparison of publications based on social media data (2019-2023)*



Source: Scopus' database.

Looking to the future, the integration of advanced artificial intelligence platforms such as OpenAI's GPT, IBM Watson, and Julius.ai promises to revolutionize X data research. These tools not only automate complex processes, but also enable more sophisticated analysis of large volumes of textual data. According to Nguyen et al. (2024), combining NLP and DL with advanced AI models could transform sentiment analysis and social trend prediction. However, this transition to artificial intelligence raises ethical and epistemological questions that require critical reflection on the transparency and impact of these technologies on social research.

Finally, future opportunities for research with X data will depend on the ability of researchers to develop more sophisticated and ethically responsible methodologies. The dynamic and ephemeral nature of X data presents unique challenges for collection and analysis, especially given the limitations imposed by the platform's API and the demographic biases of its user base. Dixon (2023) warns that X's declining popularity among young users may limit its representativeness in studies focused on youth culture, underscoring the need to diversify data sources and adapt methodologies to new platforms.

In sum, the trends and opportunities in X data-driven research reflect a constantly evolving field characterized by a balance between methodological innovation and adaptation to a diversified digital ecosystem. While X remains a valuable source of knowledge generation, the growth of alternative platforms and advances in artificial intelligence promise to expand the horizons of academic research into more inclusive and interdisciplinary areas. Figures 1 and 2 show how these dynamics are reshaping the landscape, underscoring the importance of a critical and flexible approach to maximizing the potential of social media data in a global context.

4.3. Challenges in accessing and analyzing data from X

Despite the many opportunities offered by X data analysis for scientific research, this field faces challenges related to data access, sample representativeness, and methodological and ethical complexities. These challenges have not only conditioned the evolution of research, but have also required critical reflection on the methodologies employed and the epistemological implications of working with social media data. These implications relate to how the intrinsic characteristics of social media —such as the unstructured nature of the data, the inherent biases in the representation of certain social groups, and the volatility of digital interactions— can influence the research questions, the interpretation of the results, and the construction of knowledge, raising questions about the validity, generalizability, and theoretical scope of the findings derived from these data.

One of the main obstacles is limited access to data. Since the early stages of Twitter-based research, restrictions imposed by its API have made it difficult to obtain complete and deep datasets. Weller (2014) highlights that these restrictions force researchers to rely on third-party tools, such as CrowdTangle or NodeXL, which can introduce biases and compromise the integrity of the results. De Vreese and Tromble (2023) point out that while these external tools circumvent some limitations, they present their own challenges, such as inconsistencies in the quality of data collected, additional costs, and processing difficulties. These barriers not only complicate data collection, but also create inconsistencies between studies using different approaches and tools, reducing the comparability of results (Maci et al., 2024). In addition, the increasingly restrictive access policies of X not only limit the ability to obtain data in real time and in a comprehensive manner, but also discourage the academic community from considering it as a viable source for future research. This trend threatens to erode X's relevance as an analytical tool in favor of more accessible and transparent platforms.

Access to data is not the only issue. The representativeness of samples also raises important concerns. Twitter's user base, which traditionally has higher incomes and education levels than the general population, introduces a significant demographic bias (Blank, 2017; McLachlan, 2024). This bias affects the validity of studies that attempt to extrapolate findings to larger or more representative populations. Furthermore, recent

demographic changes in the composition of Twitter users exacerbate this problem, limiting the applicability of research that relies on diverse representation. Dixon (2023) warns that the decline in Twitter use among younger audiences may affect its usefulness in studies focused on youth culture, requiring a reevaluation of the platform as a data source.

The decline of young users on Twitter raises not only practical challenges, but also theoretical questions about how to interpret the data generated by a platform whose demographic profile is rapidly changing. While young users tend to be a key group for studies related to culture, trends, and digital consumption, their migration to other platforms such as TikTok and Instagram suggests that Twitter may be losing relevance as a space for social interaction for these segments. This phenomenon not only limits the representativeness of the data, but also calls for a rethinking of what types of social dynamics can be effectively captured on the platform, and which may be underrepresented or absent in future studies.

These issues are compounded by ethical challenges related to privacy and consent. Nellore et al. (2023) highlights that the use of social media data raises complex dilemmas, particularly with regard to the protection of personal data and compliance with evolving regulations. This ethical landscape has become more complicated due to growing public awareness of the risks associated with handling personal data and the need for stricter standards. Bluemke et al. (2023) suggest that regulatory pressure and public scrutiny will force researchers to implement more rigorous strategies to ensure ethical compliance and transparency in their methods.

In addition, the methodologies used to collect and analyze data present additional limitations. Cleaning and organizing large amounts of unstructured Twitter data requires sophisticated and often costly processes that consume significant time and resources. Kim et al. (2013) highlight that these steps, if not handled carefully, can introduce additional biases that affect the accuracy and reliability of analyses. The growing presence of bots and automated accounts on the platform further complicates this scenario, as these actors generate content that does not reflect real human behavior, skewing the conclusions of studies.

Finally, challenges related to data integrity and sampling strategies are also significant. The abundance of data on Twitter does not necessarily guarantee its quality or relevance. Researchers must be critical of the inherent limitations of available data and methodologies. Rivadeneira (2023) argues that concerns about representativeness and ethics require continuous revision of research practices to adapt to changes in the digital ecosystem.

Taken together, these challenges highlight the need for more sophisticated and ethically responsible methods to maximize the potential of Twitter data in scientific research. While the barriers are significant, they also present opportunities for innovation in study design and the development of analytical tools to overcome these limitations. In an ever-changing digital context, the ability of researchers to adapt and anticipate new challenges will be critical to ensuring the relevance and impact of their scholarly contributions.

5. DISCUSSION AND CONCLUSIONS

This literature review comprehensively addresses the factors that have driven the rise of Twitter data-driven research, the emerging trends in this field, and the main difficulties faced by the academic community when working with this type of data. Based on the stated objectives, the findings reflect how the platform has transformed the generation of academic knowledge, characterized by its ability to capture social dynamics in real time and provide unprecedented access to public discourses, not without limitations.

The analysis of the factors that have contributed to the growth of research using Twitter data shows that the combination of accessibility, technological advances, and academic interest in understanding contemporary social phenomena has strengthened its adoption. However, the results also show that limited access to the data,

limitations imposed by the API, and reliance on third-party tools have limited the scope of this research. These barriers affect not only the quality and consistency of the studies, but also their ability to capture complete and unbiased representations of social reality. This scenario raises the need to develop approaches that reduce these limitations and ensure more robust and comparable analyses.

In terms of future trends, the results point to a diversification of data sources, with a growing interest in platforms such as TikTok and Instagram. This phenomenon responds in part to the demographic shifts in Twitter's user base and the different interaction dynamics offered by other platforms. Exploring a cross-platform approach, identified as an emerging trend, promises to broaden the understanding of social dynamics and allows us to compare patterns of public participation in different digital environments. This perspective offers an opportunity to enrich the analysis, but also requires consistent methodological tools that allow the integration of data from heterogeneous sources.

Finally, in terms of difficulties reported by researchers, the findings highlight that, in addition to challenges in accessing data, there are ethical concerns related to privacy and consent. The increasing use of bots and automated accounts complicates the authenticity of the data collected and raises questions about its validity in representing real human behavior. These difficulties, along with the need to ensure transparency and interpretability in the use of advanced tools such as machine learning, underscore the importance of establishing clear standards for the ethical use of social media data.

In this context, the design of specific methodologies that address the demographic biases identified in the Twitter data is proposed. These methodologies could include sampling strategies that explicitly consider the demographic and behavioral characteristics of users, as well as analytical techniques that compensate for the underrepresentation of certain groups. It is also recommended that clear and widely accepted ethical standards be promoted to guide research based on social media data. This includes the implementation of protocols that ensure responsible use of privacy, informed consent, and minimization of potential harm to the communities being studied.

In summary, the objectives of this review are reflected in the findings, which highlight both the opportunities and the limitations of using X-data for academic research. This exercise not only demonstrates the relevance of the platform for interdisciplinary knowledge production, but also invites us to reflect on the need to adapt research practices to an ever-changing digital landscape. Connecting the results with future trends, reported difficulties, and practical recommendations allows us to outline a path that prioritizes functional methodological approaches, ethical considerations, and a critical understanding of social dynamics in an ever-changing digital environment.

6. REFERENCES

Asur, S., & Huberman, B. A. (2010). Predicting the future with social media. HP Laboratories Technical Report, 53. Scopus. https://acortar.link/295pNj

Blank, G. (2017). The Digital Divide Among Twitter Users and Its Implications for Social Research. *Social Science Computer Review*, *35*(6), 679-697. https://doi.org/10.1177/0894439316671698

Bluemke, E., Collins, T., Garfinkel, B., & Trask, A. (2023). Exploring the relevance of data Privacy-Enhancing technologies for AI governance use cases. *arXiv preprint arXiv:2303.08956*. https://doi.org/10.48550/ARXIV.2303.08956

Budiharto, W., & Meiliana, M. (2018). Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis. *Journal of Big Data*, *5*(1). https://doi.org/10.1186/s40537-018-0164-1

Cvetojevic, S., & Hochmair, H. H. (2018). Analyzing the spread of tweets in response to Paris attacks. Computers, Environment and Urban Systems, 71, 14-26. https://doi.org/10.1016/j.compenvurbsys.2018.03.010

de Vreese, C., & Tromble, R. (2023). The Data Abyss: How Lack of Data Access Leaves Research and Society in the Dark. *Political Communication*, *40*(3), 356-360. https://doi.org/10.1080/10584609.2023.2207488

Dixon, S. J. (13 de septiembre de 2023). *U.S. annual X/Twitter growth 2025*. Statista. https://www.statista.com/statistics/238729/twitters-annual-growth-rate-in-the-us/

Efanova, E. V. (2023). Twitter diplomacy in shaping the foreign policy agenda of the United States of America during the presidency of D. Trump. *Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 4. Istoriya. Regionovedenie. Mezhdunarodnye otnosheniya*, *28*(3), 97-103. https://doi.org/10.15688/jvolsu4.2023.3.9

Fütterer, T., Fischer, C., Alekseeva, A., Chen, X., Tate, T., Warschauer, M., & Gerjets, P. (2023). ChatGPT in education: Global reactions to AI innovations. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-42227-6

Ghosh, S., Fernandez, J. M. Z., González, I. Z., Calle, A. M., & Shaghaghi, N. (2023). Detecting Fake News Spreaders on Twitter Through Follower Networks. En R. Hou, H. Huang, D. Zeng, G. Xia, K. K. A. Ghany., & H. M. Zawbaa (Eds.), *Big Data Technologies and Applications. BDTA BDTA 2022 2021. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering* (Vol. 480, pp. 181-195). Springer, Cham. https://doi.org/10.1007/978-3-031-33614-0_13

Henry, D., Stattner, E., & Collard, M. (2018). Information Propagation Routes between Countries in Social Media. En Companion Proceedings of the The Web Conference 2018 (WWW '18) (pp. 1295-1298). International World Wide Web Conferences Steering Committee. https://doi.org/10.1145/3184558.3191569

Hu, T., Wang, S., Luo, W., Zhang, M., Huang, X., Yan, Y., Liu, R., Ly, K., Kacker, V., She, B., & Li, Z. (2021). Revealing public opinion towards covid-19 vaccines with twitter data in the united states: Spatiotemporal perspective. *Journal of Medical Internet Research*, *23*(9). https://doi.org/10.2196/30854

Ioannides, G., Jadhav, A., Sharma, A., Navali, S., & Black, A. W. (2023). Compressed models for co-reference resolution: Enhancing efficiency with debiased word embeddings. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-45677-0

Kim, A. E., Hansen, H. M., Murphy, J., Richards, A. K., Duke, J., & Allen, J. A. (2013). Methodological considerations in analyzing twitter data. *Journal of the National Cancer Institute - Monographs*, 47, 140-146. https://doi.org/10.1093/jncimonographs/lgt026

Maci, S., Demata, M., McGlashan, M., & Seargeant, P. (Eds.). (2024). *The Routledge handbook of discourse and disinformation*. Routledge.

Mao, Y., Menchen-Trevino, E., & Cronin, J. (2023). Communicating environmental issues across media: An exploration of international news flows between twitter and traditional media. *Journal of International Communication*, *29*(1), 39-61. https://doi.org/10.1080/13216597.2022.2149605

Mazhar, K., & Dwivedi, P. (2024). Decoding the black box: LIME-assisted understanding of Convolutional Neural Network (CNN) in classification of social media tweets. *Social Network Analysis and Mining*, *14*(1), 133. https://doi.org/10.1007/s13278-024-01297-8

McLachlan, S. (7 de mayo de 2024). 19 X (Twitter) Demographics for Marketers in 2024. *Hootsuite*. https://blog.hootsuite.com/twitter-demographics/?utm_source=chatgpt.com

Nellore, N., Zimmer, M., & Apostol, B. (2023). Updating the Topology of Twitter Research: Towards A Systematic Review of Twitter Research from 2013-2022. In Companion Proceedings of the 2023 ACM International Conference on Supporting Group Work (GROUP '23) (pp. 4-5). Association for Computing Machinery. https://doi.org/10.1145/3565967.3570972

Nguyen, H.-D., Nguyen, D. Q., Nguyen, C.-D., To, P. T., Nguyen, D. H., Nguyen-Gia, H., Tran, L. H., Tran, A. Q., Dang-Hieu, A., Nguyen-Duc, A., & Quan, T. (2024). Supervised learning models for social bot detection: Literature review and benchmark. *Expert Systems with Applications*, 238. https://doi.org/10.1016/j.eswa.2023.122217

Dersan Orhan, D. (2021). Making Foreign Policy Through Twitter: An Analysis of Trump's Tweets on Iran. In E. Esiyok (Ed.), *Handbook of Research on New Media Applications in Public Relations and Advertising* (pp. 380-394). IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-7998-3201-0.ch022

Pagolu, V. S., Reddy, K. N., Panda, G., & Majhi, B. (2017). Sentiment analysis of Twitter data for predicting stock market movements. In *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)* (pp. 1345-1350). IEEE. https://doi.org/10.1109/SCOPES.2016.7955659

Raja, H., Ilyas, M. U., Saleh, S., Liu, A. X., & Radha, H. (2016). Detecting national political unrest on Twitter. In *2016 IEEE International Conference on Communications (ICC)* (pp. 1-7). https://doi.org/10.1109/ICC.2016.7511393

Rangel, F., Giachanou, A., Ghanem, B., & Rosso, P. (2020). Overview of the 8th Author Profiling Task at PAN 2020: Profiling Fake News Spreaders on Twitter. In *CEUR workshop proceedings* (Vol. 2696, pp. 1-18). Sun SITE Central Europe. https://acortar.link/RAgq5H

Rivadeneira, L. (2023). Análisis del comportamiento de decisión usando datos de Twitter: Una revisión de la literatura. *Revista Ibérica de Sistemas e Tecnologias de Informação*, E61, 307-321. https://acortar.link/kzucCg

Schroeder, R., & Cowls, J. (2019). Big Data Approaches to the Study of Digital Media. En J. Hunsinger, M. Allen y L. Klastrup (Eds.), *Second International Handbook of Internet Research* (pp. 957-977). https://doi.org/10.1007/978-94-024-1555-1_13

Sinnenberg, L., Buttenheim, A. M., Padrez, K., Mancheno, C., Ungar, L., & Merchant, R. M. (2017). Twitter as a tool for health research: A systematic review. *American Journal of Public Health*, *107*(1), e1-e8. https://doi.org/10.2105/AJPH.2016.303512

Soaita, A. M., Serin, B., & Preece, J. (2020). A methodological quest for systematic literature mapping. *International Journal of Housing Policy*, *20*(3), 320-343. https://doi.org/10.1080/19491247.2019.1649040

Steinert-Threlkeld, Z. C. (2018). *Twitter as data*. Cambridge University Press. https://doi.org/10.1017/9781108529327

Wang, Z., Lam, N. S. N., Obradovich, N., & Ye, X. (2019). Are vulnerable communities digitally left behind in social responses to natural disasters? An evidence from Hurricane Sandy with Twitter data. *Applied Geography*, 108, 1-8. https://doi.org/10.1016/j.apgeog.2019.05.001

Weller, K. (2014). What do we get from twitter- and what not? A close look at twitter research in the social sciencest. *Knowledge Organization*, *41*(3), 238-248. https://doi.org/10.5771/0943-7444-2014-3-238

Yu, J., & Muñoz-Justicia, J. (2022). Free and Low-Cost Twitter Research Software Tools for Social Science. *Social Science Computer Review*, *40*(1), 124-149. https://doi.org/10.1177/0894439320904318

Yuan, Y. (2017). Modeling inter-country connection from geotagged news reports: A time-series analysis. In Y. Tan, H. Takagi y Y. Shi (Eds.) *Data Mining and Big Data. DMBD 2017. Lecture Notes in Computer Science* (Vol. 10387, pp.183-190). Springer, Cham. https://doi.org/10.1007/978-3-319-61845-6_19

Zou, Y., & Herremans, D. (2023). PreBit — A multimodal model with Twitter FinBERT embeddings for extreme price movement prediction of Bitcoin. *Expert Systems with Applications*, 233. https://doi.org/10.1016/j.eswa.2023.120838

AUTHOR CONTRIBUTIONS, FUNDING AND ACKNOWLEDGEMENTS

Authors' contributions:

Conceptualization: Rivadeneira, Lucía. **Software:** Rivadeneira, Lucía. **Validation:** Loor, Ignacio. **Formal analysis:** Rivadeneira, Lucía y Loor, Ignacio. **Data curation:** Rivadeneira, Lucía y Loor, Ignacio. **Writing-Preparation of the original draft:** Rivadeneira, Lucía. **Writing-Revision and Editing:** Loor, Ignacio. **Visualization:** Rivadeneira, Lucía. **Supervision:** Rivadeneira, Lucía y Loor, Ignacio. **Project management:** Rivadeneira, Lucía y Loor, Ignacio. **All authors have read and accepted the published version of the manuscript:** Rivadeneira, Lucía y Loor, Ignacio.

Funding: No external funding was received for this research.

Interest conflict: None.

AUTHORS:

Lucía Rivadeneira
Technical University of Manabí.
Lucía Rivadeneira holds a Ph.D. in Business and Management from the University of Manchester, United Kingdom; a Master's degree in Information Systems from Nanyang Technological University, Singapore; and a degree in Systems Engineering from the Technical University of Manabí, Ecuador. She is currently a lecturer and researcher at the Faculty of Computer Science, Technical University of Manabí, where she specializes in the analysis of social media data for modeling social phenomena. Her areas of interest include artificial intelligence, machine learning, unstructured data analysis, sentiment analysis, social network analysis, and classification and optimization models.
lucia.rivadeneira@utm.edu.ec
Orcid ID: https://orcid.org/0000-0001-5989-7703
Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=57221283412

Ignacio Loor
Technical University of Manabí.
Ignacio Loor holds a Ph.D. in Human Geography from the University of Manchester, United Kingdom; a Master's degree in International Business from Nova Southeastern University, United States; and an Economics degree from the Universidad Católica Santiago de Guayaquil, Ecuador. He currently works as a researcher in urbanism and sustainable development and as the Vice Dean of Research at the Faculty of Humanistic and Social Sciences at the Technical University of Manabí, Ecuador. His areas of interest include informal settlement infrastructure, green infrastructure, social organization of informal communities, informal practices, and the transition to net zero carbon emissions.
ignacio.loor@utm.edu.ec
Índice H: 6
Orcid ID: https://orcid.org/0000-0003-4806-1032
Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=57224464879
ResearchGate: https://www.researchgate.net/profile/Ignacio-Loor

Artículos relacionados:

Arce Garcia, S., Cano Garcinuño, M. I., Quiles Cano, C., & Cano Pérez, J. (2023). Vacunas anticovid y trombosis: el miedo en las redes sociales. *Revista de Comunicación y Salud*, 14, 1-19. https://doi.org/10.35669/rcys.2024.14.e307

Demuner Flores, M. del R. (2021). Uso de redes sociales en microempresas ante efectos COVID-19. *Revista de Comunicación de la SEECI*, 54, 97-118. https://doi.org/10.15198/seeci.2021.54.e660

Martínez-Fresneda Osorio, H., & Sánchez Rodríguez, G. (2022). La influencia de Twitter en la agenda setting de los medios de comunicación. *Revista de Ciencias de la Comunicación e Información*, 27, 1-21. https://doi.org/10.35742/rcci.2022.27.e136

Moreno Cabanillas, A., & Castillero Ostio, E. (2023). Comunicación política y redes sociales: análisis de la comunicación en Instagram de la campaña electoral del 13-F. *Vivat Academia,* 156, 199-222. https://doi.org/10.15178/va.2023.156.e1461

Pérez Altable, L., & Serrano-Tellería, A. (2021). Communications patterns and power dynamics in the digital public sphere: A case study of the conversation about Minimum Living Income on Twitter. *European Public & Social Innovation Review*, *6*(1), 1-15. https://epsir.net/index.php/epsir/article/view/148