

OPINION

Rethinking data sources for crisis management: How restricted social media access is reshaping current practices

Roberto Interdonato ^{1,2*}, Rémy Decoupes^{1,3}, Mathieu Roche ^{1,2}, Mehtab Alam Syed^{1,2}, Maguelonne Teisseire^{1,3}, Sarah Valentin ^{1,2}

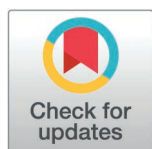
1 TETIS, Univ. Montpellier, AgroParisTech, CIRAD, CNRS, INRAE, Maison de la Télédetection, Montpellier, France, **2** CIRAD, Montpellier, France, **3** INRAE, Montpellier, France

* roberto.interdonato@cirad.fr

1 From early adoption to current practices

Since their massive diffusion during the first decade of the 2000s, it became clear that social media could serve as a valuable source of real-time information about any topic, acting as crowdsourced newsfeeds [1], and being able to support situational awareness and decision-makers during mass emergencies [2]. While almost any available social media has been used in literature as an information source, Twitter quickly emerged as the most suitable option due to several factors, including its micro-blogging format based on frequent short posts, and the availability of user-friendly APIs [3]. Most importantly, even if a subscription fee was required for full access, the free version of the APIs allowed enough data to be crawled to support research work in different domains. Other popular sources struggled to gain the same success in the scientific community, either for difficulties in accessing the data (Facebook) or for the challenges of processing multi-modal data with the methodologies available at that time (YouTube, Instagram, Flickr). In fact, early research primarily relied on network analysis to explore community structures and information diffusion, alongside Natural Language Processing (NLP) to identify crisis events and assess public perception, among other tasks. Along with the use of social media data, researchers have characterised their limitations. These include various inherent biases in users' selection (e.g., demographic), as well as technical challenges in data analysis due to the volume, unstructured nature and multilingualism of their content. Additionally, the overflow of unverified content also facilitates the dissemination of misinformation [4]. More recently, the integration of transformer-based architectures and Large Language Models (LLMs) has marked a significant shift in crisis monitoring. BERT-based models have dominated event detection and classification tasks, while decoder-only models such as GPT-4 and LLaMA-2 have gained traction for their reasoning and few-shot capabilities [6,5]. Despite their promising performance, these models still struggle with the informal language and domain-specific jargon that are characteristic of crisis communication.

However, the popularity of using social media data for crisis management (i.e., prevention, preparedness, response, adaptation) has decreased over the years, facing substantial disruption following Elon Musk's acquisition of Twitter in 2022 and



OPEN ACCESS

Citation: Interdonato R, Decoupes R, Roche M, Syed MA, Teisseire M, Valentin S (2026) Rethinking data sources for crisis management: How restricted social media access is reshaping current practices. PLOS Complex Syst 3(6): e0000112. <https://doi.org/10.1371/journal.pcsy.0000112>

Editor: Roger Cremades, University of Leeds, UNITED KINGDOM OF GREAT BRITAIN AND NORTHERN IRELAND

Published: June 30, 2026

Copyright: © 2026 Interdonato et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

the subsequent operational changes, including the platform's rebranding to X, altered content moderation policies, and changes in the API access (i.e., end of free API in February 2023) [7]. These developments have prompted many scientists and professionals to reconsider their reliance on the platform for crisis communication, marking a potential end to the era when Twitter served as a primary conduit for scientific discourse during public emergencies. Following these events, alternative micro-blogging platforms are emerging (e.g., Bluesky, Mastodon, Threads) and have been the subject of massive user migration from Twitter/X (Fig 1). However, there is still uncertainty about when (and if) some of these networks will reach a critical mass to make them worth considering in the context of crisis management [8].

Many other challenges in this context are related to the General Data Protection Regulation (GDPR) implemented by the EU since May 2018. The GDPR is probably the first (and only) example of a unified law about the privacy and use of personal data at the continental scale. While this law is nowadays a fundamental component of European Union (EU) privacy and human rights laws (also regulating the extent to which data can be used outside the EU), it also represents a significant limitation on how social media data can be used in the context of research projects (and especially EU-funded ones such as Horizon or ERC projects). In particular, it may hinder reproducibility, as researchers are often unable to archive and share datasets containing information that could directly or indirectly identify individuals. Moreover, researchers must ensure that when users withdraw their consent from a platform, all corresponding records are removed from the datasets. This implies the development of monitoring systems capable of detecting consent withdrawal events (e.g., account deletions) and updating datasets accordingly. Over time, as more users delete their accounts or revoke consent, datasets must be continuously updated and may consequently shrink, further complicating longitudinal analyses and reproducibility. Some further concerns have emerged about the new Twitter/X policies no longer fully fulfilling key obligations required by the GDPR. While the existing literature on the application of the GDPR to the humanitarian aid sector is rather limited [9], some studies suggest that people may agree to share their personal data during emergency situations [10].

2 Future directions for crisis monitoring

The recent increase in the cost of access to Twitter/X APIs, combined with regulatory constraints, strongly questions the sustainability of relying on a single dominant platform for crisis monitoring. A natural alternative to Twitter/X is leveraging other social media platforms such as Mastodon, Reddit, Telegram, or TikTok. However, existing studies show that none of these platforms can fully substitute Twitter/X's role in disaster management due to lower and more uneven coverage and strong access limitations. To name a few, TikTok lacks a real-time public stream, Reddit and Telegram rely on fragmented access (content is dispersed across independent subreddits and channels that must be individually identified and monitored, rather than aggregated into a single public stream), Bluesky and Mastodon have extremely limited adoption. These limitations position these platforms as valuable complements rather

Crisis Management Articles Mentioning a Social Network in the Title by Year

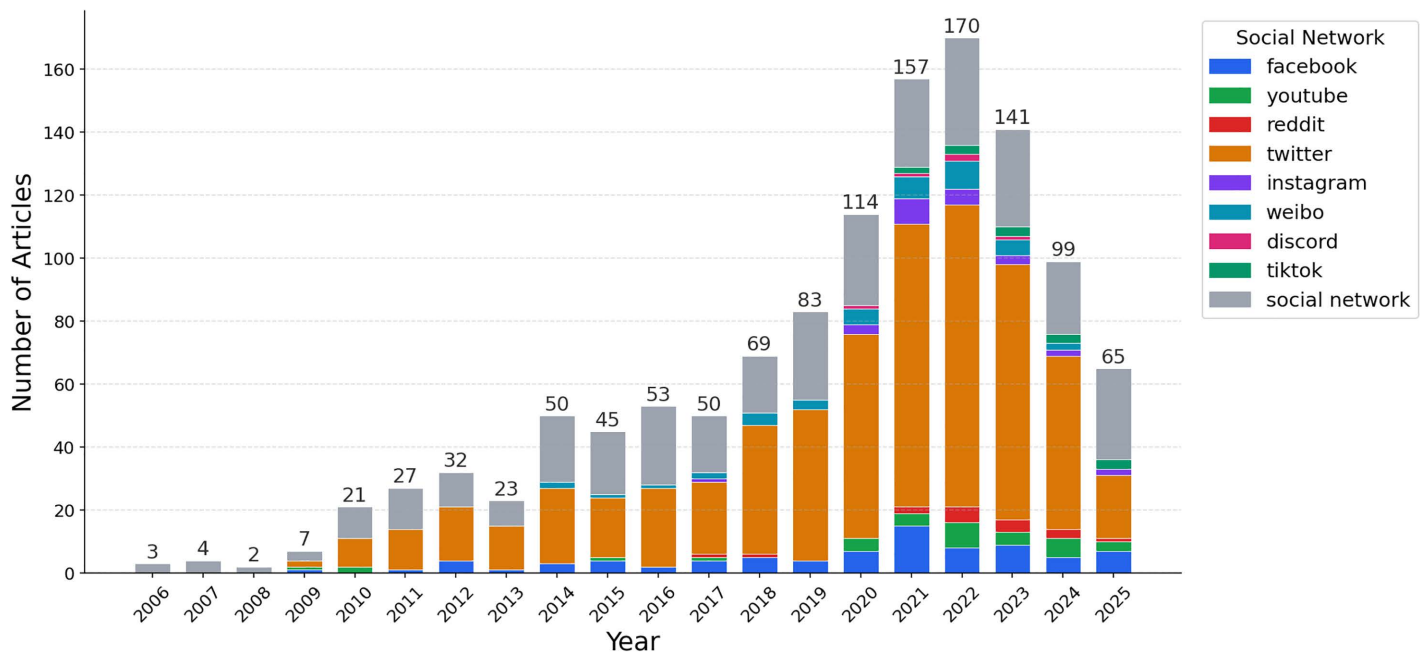


Fig 1. Number of articles per year, including the name of an Online Social Network in the title, alongside a crisis management keyword. Articles including the generic “Social Network” mention in the title are also included. Titles were retrieved via the Scopus API in February 2026, limiting the search to four subject areas involving data oriented/computational approaches: Computer Science, Mathematics, Engineering, Decision Sciences. Code and list of keywords at <https://github.com/interdonatos/ScopusCrawl/>.

<https://doi.org/10.1371/journal.pcsy.0000112.g001>

than standalone replacements [8]. Moreover, each social media platform must be used according to its audience and user behaviors, to match regional contexts and analytical objectives.

From platform dependency to multi-source crisis monitoring

We argue for a shift towards multi-source crisis monitoring frameworks, combining online news media [11] and radio broadcasts [12]. Beyond improving coverage, such diversification reduces the selection biases inherent to any single source by capturing heterogeneous communities and communication practices. That being said, collective efforts are needed to establish governance frameworks that allow controlled access to social media data during crises, enabling their integration in support of operational decision-making.

Benefits and limitations of media-based sources

While press and radio data are also subject to editorial and political biases, they offer complementary advantages for crisis monitoring. Local radio stations, in particular, remain key information channels in low-connectivity contexts and provide access to under-represented community-level perceptions that are largely invisible on global online platforms [12]. When combined with national and international press, these sources partially compensate for selection biases and contribute to more geographically equitable situational awareness.

Compared to user-generated social media content, media-based sources enable better control of misinformation through source filtering and credibility assessment, and their use raises fewer legal and ethical challenges with respect to data protection regulations such as the GDPR.

Open challenges: Multilingualism and information fusion

Moving towards multi-source crisis monitoring infrastructures introduces new technical challenges. First, the same events are often reported across multiple channels, requiring robust deduplication and information fusion methods to avoid redundancy and contradictory signals. Second, multilingual processing remains a major bottleneck, particularly for under-resourced languages that are prevalent in many crisis-prone regions. This is an important issue to consider for Africa, which has more than 2000 indigenous languages, with the production of dedicated multilingual models [13] useful to deal with local languages [13]. Without dedicated efforts on low-resource NLP, multi-source systems risk reproducing existing geographical and linguistic inequities in crisis response.

Conclusion

In order to tackle crisis management, future work should integrate a large panel of data (multimodal data, multilanguage texts, etc.) that provides complementary information. But in this context, we need to tackle two important challenges: data fusion/ingestion of heterogeneous data and bias mitigation because bias can reproduce and amplify long-standing inequalities with intrinsic features of generative AI [14].

References

1. Kwak H, Lee C, Park H, Moon S. What is Twitter, a social network or a news media?. In: Proceedings of the 19th international conference on World wide web, 2010. 591–600. <https://doi.org/10.1145/1772690.1772751>
2. Imran M, Castillo C, Diaz F, Vieweg S. Processing social media messages in mass emergency. *ACM Comput Surv.* 2015;47(4):1–38. <https://doi.org/10.1145/2771588>
3. Pilipiec P, Samsten I, Bota A. Surveillance of communicable diseases using social media: a systematic review. *PLoS One.* 2023;18(2):e0282101. <https://doi.org/10.1371/journal.pone.0282101> PMID: 36827297
4. Cinelli M, Quattrocioni W, Galeazzi A, Valensise CM, Brugnoli E, Schmidt AL, et al. The COVID-19 social media infodemic. *Sci Rep.* 2020;10(1):16598.
5. Xu F, Ma J, Li N, Cheng JCP. Large language model applications in disaster management: an interdisciplinary review. *Int J Disaster Risk Reduc.* 2025;127:105642. <https://doi.org/10.1016/j.ijdr.2025.105642>
6. Cantini R, Cosentino C, Marozzo F, Talia D, Trunfio P. Harnessing prompt-based large language models for disaster monitoring and automated reporting from social media feedback. *Online Social Networks and Media.* 2025;45:100295. <https://doi.org/10.1016/j.osnem.2024.100295>
7. Murtfeldt R, Paik S, Alterman N, Kahveci I, West JD. Rip twitter api: a eulogy to its vast research contributions. 2025. <https://doi.org/10.26434/chemrxiv-2025-2404>
8. Vračević N, Schmidt S, Keskin M, Hanny D, Resch B. More than just Tweets: the potential of alternative geo-social media data for disaster management. *Soc Netw Anal Min.* 2025;15(1). <https://doi.org/10.1007/s13278-025-01494-z>
9. Gazi T, Gazis A. Humanitarian aid in the age of COVID-19: a review of big data crisis analytics and the General Data Protection Regulation. *Int rev Red Cross.* 2020;102(913):75–94. <https://doi.org/10.1017/s1816383121000084>
10. Aldehoff L, Dankenbring MD, Reuter C. Renouncing privacy in crisis management? People's view on social media monitoring and surveillance. In: Proceedings of the 16th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2019), Valencia, Spain, 2019. p. u 1184–97.
11. Valentin S, Arsevska E, Rabatel J, Falala S, Mercier A, Lancelot R, et al. PADI-web 3.0: a new framework for extracting and disseminating fine-grained information from the news for animal disease surveillance. *One Health.* 2021;13:100357. <https://doi.org/10.1016/j.onehlt.2021.100357> PMID: 34950760
12. Nakatumba-Nabende J, Mukiibi J, Bateesa TS, Murindanyi S, Katumba A, Mutebi C. Machine learning analysis of radio data to uncover community perceptions on the ebola outbreak in Uganda. *ACM J Comput Sustain Soc.* 2024;2(3):1–28. <https://doi.org/10.1145/3685697>
13. Adebara I, Toyin HO, Ghebremichael NT, Elmadany AA, Abdul-Mageed M. Where are we? Evaluating LLM performance on African languages. In: Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2025. 32704–31. <https://doi.org/10.18653/v1/2025.acl-long.1572>
14. Kerche FW, Zook M, Graham M. The silicon gaze: A typology of biases and inequality in LLMs through the lens of place. *Platforms Society.* 2026;3. <https://doi.org/10.1177/29768624251408919>