


Environmental Policy & Regulation

Content hubs, information flows, and reactions for pesticide-related discussions on Twitter/X

Tiancheng Zhou^{1,*}, Yu Huang², Raghavendhran Avanas³, Richard A. Brain³, Mattia Prosperi⁴ , and Jiang Bian²

¹Department of Health Outcomes & Biomedical Informatics, University of Florida, Gainesville, FL, United States

²Department of Biostatistics and Health Data Science, Indiana University, Bloomington, IN, United States

³Syngenta Crop Protection, LLC, Greensboro, NC, United States

⁴Department of Epidemiology, University of Florida, Gainesville, FL, United States

*Corresponding author: Tiancheng Zhou. Email: tzhou@ufl.edu

Abstract

Pesticides are essential in modern agriculture for controlling pests and enhancing food production. However, concerns about their human and environmental health impacts have broadened discussions on their use, regulation, ethics, and sustainability. Scientific research, media coverage, and input from corporations, governments, and nongovernmental organizations (NGOs) shape public opinions and potentially influence regulatory decisions. This project analyzed pesticide-related discussions on Twitter/X from 2013 to 2022, focusing on information influence and propagation among individuals and organizations, advancing over prior research that looked at topic frequency, trends, and geography. Using a validated snowball sampling method, we collected over 3 million tweets from 1 million users and identified key network influencers, i.e., information hubs, analyzing their content, popularity, and characteristics. Machine learning and a tailored information flow score were used to explore the dynamics of information flow and sentiment across hubs. Our analysis revealed that organizational hubs, particularly NGOs and media, were more active and had higher follower-to-following ratios than individual influencers. Media and NGOs also dominated the pesticide-related discourse, while individual influencers had a lesser role. Information sources were unevenly distributed, with a dominance of retweets, news, and media posts, and a low prevalence of scientific sources. Information flow was high through NGOs, academia, and individuals, but poor from government accounts. Pesticide-focused hubs were more active and targeted in their information dissemination, with public sentiment largely negative. By delving deeper into the dynamics of information dissemination and influence networks, this study provides insights that emphasize (1) the need for better communication strategies to integrate diverse stakeholder perceptions and values, and (2) prioritizing the dissemination of credible scientific information, while also addressing sectoral disparities. Together, they can help policymakers and industry stakeholders build trust, promote transparency, and advance sustainable pesticide regulation.

Keywords: machine learning, pesticide perceptions, information dissemination, sentiment analysis, Twitter/X

Introduction

In the last decade, social media platforms have become a global assembly point for society (Burgess et al., 2019; Jackson et al., 2021; Kapoor et al., 2018; Wang et al., 2019). In 2023, about 83% of U.S. adults used video-based platforms such as YouTube, 68% have used Facebook, 47% have used Instagram, and 22% have used Twitter/X and Reddit (Gottfried, 2024). Social media provides the largest and most dynamic representation of human behaviors, offering profound insights into relationships and communication patterns of individuals and organizations. As a result, it is an invaluable source of often publicly available data for research, enhancing our understanding of digital interactions, and opens new avenues for studying complex social phenomena (Debreceeny et al., 2019). Although not the largest in terms of users, Twitter/X currently stands as one of the most popular and extensively analyzed social media platforms (Walsh, 2024) as a large portion of its data has been public for a long time, and research has been facilitated by application programming interfaces (APIs), catalyzing data

services and analytics (Twitter API Documentation, 2024). The application (app) is widely used by individuals as well as organizational entities to share, discuss, and promote health-related content (Burgess et al., 2019; Debreceeny et al., 2019). As for nearly all content covered on Twitter/X, health-related discussions are cluttered with polarization, conflicts of interest, and misinformation (Bortleson & Davis, 1987; Himelboim & Han, 2014; Sotrender, 2022; Yin et al., 2021). A variety of studies use Twitter/X as the source of data to understand public perception and activism on health risks for consumers' products, e.g., e-cigarettes (Sugawara et al., 2012), dietary supplements (Wood-Doughty et al., 2018), genetically modified organisms, and pesticides (Jun et al., 2023).

Pesticides enhance agricultural productivity by increasing crop yields and improving produce quality, supporting economic growth, community health, and food security. They are also vital in controlling the invasion of foreign pests. Additionally, pesticides are used in various nonagricultural settings, including public health (e.g., control of vector-borne diseases) and civil

Received: December 11, 2024. Revised: January 24, 2025. Accepted: February 17, 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of the Society of Environmental Toxicology and Chemistry.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

engineering infrastructure (e.g., parks, landscaping, infrastructure maintenance). Notwithstanding the tremendous benefits, pesticides, when not used as required by law (i.e., the label), can potentially lead to unintended health risks to humans (Eddleston, 2000; Pimentel et al., 2007); similarly, the environmental impact of such use can be potentially disruptive and lead to biodiversity loss (Aktar et al., 2009), e.g., for aquatic life, and decline in soil fertility due to effects on beneficial microorganisms (Andreu & Picó, 2004; Bortleson & Davis, 1987).

A prior study (Jun et al., 2023) used public Twitter/X data to study public promotion, dissemination, discussions, as well as the general public's perception of safety, regulation, and health risks of pesticide use in the United States. Using natural language processing and machine learning, we identified and classified pesticide-related topics by user types—individuals and organizations, further stratifying by time trends, geographic distributions, and sentiments. Twitter/X discussions on pesticide safety exhibited high volume, with topic ramifications from human to environmental health, governance, and liability (Cha et al., 2012; Jackson et al., 2021; Jun et al., 2023; Moloko et al., 2010; Pascual-Ferrá et al., 2022), involving both social circles of individuals (including laypersons and scientists together) and organizational entities (e.g., media, industry, advocacy groups, academia, and government agencies), with different geographic distribution and peaks in time trends. One limitation of this study was that the analysis did not evaluate how the topics and sentiments spread through the network and who influenced the information flows.

Thus, an important next step after topic and sentiment modeling is to characterize how this structured information flows through and influences the social media circles, i.e., study the topological characteristics of users' social network related to the subject of interest; in fact, a number of studies have applied network analysis methods to study influencers and the virality of health-related discussions. On Twitter/X, Himelboim and Han (2014) investigated cancer-related talks, identifying interconnected users and their most followed hubs under various network metrics. Pascual-Ferrá et al. (2022) examined different network properties to analyze the factors hindering public health message dissemination. Sugawara et al. (2012) analyzed highly influential cancer patient accounts' interactions and behavior patterns, finding a rapidly evolving network of cancer patients who exchange information extensively. As a limitation, these studies assessed several off-the-shelf social network metrics without further analyzing the users in terms of their positions in the information flow or the dynamics of social network users. In terms of information flow analysis, many efforts have been put forward to study the dynamic information flow of Twitter/X by using a wide variety of novel approaches. Some recent network analysis works stratified different user groups or analyzed the effects of certain users on information dissemination. Cha et al. (2012) collected a vast volume of Twitter/X data to investigate the different roles that users play in information flow. Yin et al. (2021) developed an opinion-delay susceptible-forwarding immunized model to examine the influence of opinion involving opinion leaders.

In this study, we aimed to understand how pesticide-related information spreads across the Twitter/X network, and the different roles that user types—individuals, media, government, industry, academia, and nongovernmental organizations (NGOs)—play in the network dynamics.

Accordingly, we defined five research questions apt to characterize network dynamics, information flow, influencers' characteristics, and sentiments, as follows:

1. What are the characteristics of the most influential information hubs of the Twitter network on pesticide discussion?
2. Are information sources (e.g., scientific journals, news media articles, opinion pieces) heterogeneously distributed among information hubs?
3. How does information flow across the network, and how do hubs influence such flow?
4. Do hubs influence sentiment dynamics across the Twitter network?
5. What role do chemical companies play within the network compared to other entities like NGOs, academia, and government, particularly in influencing information flow and sentiment?

To address these research questions, we used pesticide-related tweets broadcasted between 2013 and 2022 using a previously validated snowball sampling procedure developed by our team, extending collection up until the feeds were accessible publicly using API allowed by X. We identified influencers, i.e., information hubs, based on well-accepted network indices, and used an original information flow score, together with machine learning models, to predict the user's sentiments toward pesticide-related hubs.

Materials and methods

Data collection and hub identification

We used keyword-based data collection through snowball sampling according to a previously validated and published procedure. For a detailed description of our methodology and a comprehensive data summary, please refer to the Data Collection section of our prior study (Jun et al., 2023). In summary, we established an initial set of keywords to extract data on pesticides, pesticide safety, and related regulations. These keywords were continually refined and expanded through iterative web searches, categorized into core (directly related to pesticides, i.e., synonyms) and topic-related (frequently used with pesticides) terms, which were used to generate permuted combinations. Our data sampling involved extracting tweets from the full public Twitter/X archive, covering the period from January 1, 2013 to December 30, 2022—an extension from the previous cutoff of October 10, 2021. We used the Twitter API that permits keyword searches within specific dates and tweet retrieval via user handles through their timeline. Through two certified academic research accounts, we accessed up to 20 million tweets monthly, capturing comprehensive tweet metadata including timestamp, geolocation, text, handle, and user description. Since the source data are an extension of a previous study, the dynamic changes in topics within the network show similar trends. From 2013 to 2015, Twitter discussions focused primarily on general pesticide risks, safety, and potential environmental impacts, including concerns about pollinator health. In 2016 and 2017, attention shifted toward potential human health risks, particularly regarding glyphosate and associated litigation, resulting from the Group 2A “probably carcinogenic to humans” reclassification by the International Agency for Research on Cancer (IARC in 2015). This period also showed increased advocacy and campaigning for banning certain pesticides. For example, in 2018, discussions surrounding alleged “ROUNDUP®” (glyphosate) cancer claims and the EU's expanded ban on neonicotinoid pesticides peaked in intensity. During the COVID-19 pandemic (2019–2021), pesticide-related tweet volumes declined as public focus shifted to pandemic-related issues, although discussions on regulatory

changes and health risks continued. By 2021–2022, pesticide-related conversations had begun to resurge, emphasizing environmental sustainability and farming practices.

Based on our previous study and others (El Ali et al., 2018; Moloko et al., 2010), we categorized collected tweets and their corresponding accounts as individual or organizational. Organizational accounts were further divided into seven categories: media, government, industry, academia, and three types of NGOs, including environmental advocacy groups, farmer advocacy groups, and others. We also used an independent list of organizational outlets in addition to the tweet/account selection by keyword-based snowball sampling. Of note, the data capture procedure considered retweets and mentions, which could result in the inclusion of certain accounts that do not post actively on pesticides, yet they are involved in the information flow of pesticide-related discussions by being called out or retweeted in pesticide-related tweets.

Another case is when an account is included because of its clear association with pesticide-related discussions, e.g., an environmental agency, but the API search on the public Twitter/X sample does not retrieve any pesticide-related post. To identify the most influential hubs within our data, we calculated the in-degree of each user, which is defined as the total number of tweets mentioning or tagging a user's username. We then ranked the users by their in-degree values, as users within the Twitter network frequently mention highly influential users in their posts (Mary, 2021). We initially selected the top 100 users from this ranking, but the performance of sentiment prediction was highly biased due to limited data and imbalanced hub group distribution. Eventually, we selected the top 600 users from this ranking as our final list of hubs, which is sufficient to achieve decent performance on sentiment prediction tasks. However, this set included users such as WHO, despite being highly influential, might not be focused on discussing pesticides or associated topics regularly (e.g., the user “@realDonaldTrump” would be a top influencer but may not tweet actively about pesticides).

We refer to such nonpesticide-specific accounts as general hubs. To refine our analysis to hubs more relevant to pesticides, we manually reviewed the account descriptions, posted content, and associated websites of the top influencers. Through this process, we selected only accounts directly related to pesticides, environment, agriculture, farms, and food safety discussions, and we focused on this subset of hubs for deeper analysis.

Hub characteristics

To explore the characteristics of the most influential hubs in the network of pesticides and the distribution of information sources, we conducted a detailed analysis of features associated with each user account (Syngenta Crop Protection, LLC). This analysis was performed separately for all general hubs and the pesticide-focused hubs. First, we assessed the in-degree score distribution for these groups, calculating differences in distribution metrics (average, median, 10th percentile, 90th percentile, etc.). We then analyzed the volume of pesticide-related tweets posted by each hub, employing network metrics to compare how active the general and pesticide-specific hubs were regarding disseminating information within the network. We also examined the tenure of these users by comparing the differences in the years of experience using Twitter/X to determine whether having a longer experience with the app correlated with a more significant influence. Additionally, we collected the number of followers each hub had and the number of users they were currently following. We categorized the number of followers and following for each hub into three groups: fewer than 10,000 (moderate), 10,000–100,000

(large), and over 100,000 (very large). By calculating the distribution of the number of followers and following, we could further understand how pesticide-related hubs would impact other users differently than general hubs in terms of gaining followers and expanding their social networks by following others. Lastly, we verified the presence of associated websites (excluding links to other social media profiles) for each hub by manually reviewing their account pages. This step assessed whether owning a website contributed to a hub's influence.

Information source distribution

To address whether information sources distribute heterogeneously among information hubs, we examined the contents of the 20 most recent tweets posted by each pesticide-related hub. These samples provide a relevant snapshot of current topics and trends, and we found that the selected sample size is manageable, consistent across hubs, and large enough to capture diverse content while identifying key patterns. We classified the sources of information in these tweets into six categories, based on a prior study (Dann, 2010): retweets, official statements, or press releases (Castillo et al., 2011; such as an official announcement of a new product or event), news media articles (Mitchell, 2021), opinion pieces based on personal experiences, interviews and podcasts, and scientific articles typically posted by researchers.

Our analysis focused on assessing the variety of these sources within each hub's tweets. Preliminary findings revealed a high prevalence of retweets as the information source among the content shared by these hubs. We then further examined the sources within these retweets to determine the extent of heterogeneity in information distribution among the hubs. This approach allowed us to understand the diversity and spread of information sources across the network of pesticide-related users.

Information flow

To examine how information flows across the network and how hubs influence such flow, we focused on defining and quantifying the information flow value for each hub. We adapted an established methodology (Cha et al., 2012) to account for the sparseness and asymmetry of follower–following pairs due to sampling. Originally, the method defined user A, who potentially delivered information, as the source of user B if three criteria were met: (1) both users discussed the same topics in their tweets, (2) there was a follower relationship where B followed A, and (3) user A posted a tweet about these topics before B did. In our case, due to the scarcity of detailed follower data, we substituted the following–follower relationship with instances of mentioning or tagging. To determine the topical similarity between two tweets, we initially checked if two tweets contained the same URLs. If no common URLs were found, we analyzed the tweets for overlapping pesticide-relevant keywords. Since most of the hub tweets contain two keywords, we said that a tweet included the same topics as the tweet posted by the hub if at least two keywords matched, such as “pesticide,” “bees,” “farm,” and “health.” For each hub, we calculated the proportion of the total number of tweets that either mentioned or tagged the hub that met the above criteria. This proportion served as a key indicator of a hub's impact on information flow. We subsequently analyzed and compared these metrics across different hub groups, plotting the average values of information flow to assess how hubs from various groups facilitated information dissemination. Furthermore, there are hubs that do not have any pesticide-related tweets; e.g., “@realDonaldTrump” and “@WHO,” which are the general and pesticide-related hubs based on their in-degree ranking, respectively, did not post any pesticide-related

tweets. These hubs can lower the average information flow score, so we created two plots with and without these hubs to illustrate the distribution of the average score on each hub group.

Sentiment analysis

To investigate the influence of pesticide-related hubs on sentiment dynamics within discussions, our study analyzed data at both the hub and individual tweet levels. We applied the Valence Aware Dictionary and Sentiment Reasoner (VADER; [Hutto & Gilbert, 2014](#)), a sentiment analysis tool to assign a compound score to each tweet, categorizing their sentiments as positive, neutral, or negative. We aggregated tweets that mentioned or tagged each hub and calculated the average compound scores for these hubs to determine the overall sentiment associated with them. We compared the sentiment distributions among all hubs and pesticide-related hubs and the sentiment distribution in each group to illustrate how different hub groups can sway users' attitudes and opinions in discussions about pesticide-related topics.

We enhanced our study by employing machine learning to predict sentiment toward hubs—i.e., others' sentiments with respect to a hub—and the sentiment of each tweet, assessing which hub characteristics or features most significantly influenced sentiment dynamics. We categorized sentiments into three groups: positive, neutral, and negative. To simplify the prediction tasks and improve the performance, positive and negative sentiments were combined into a single category named others, and we performed sentiment predictions for neutral vs. others and separately for positive vs. negative. We utilized four machine learning methods: random forest, gradient boosting, XGBoost, and logistic regression.

For hub-based predictions, our feature set included the in-degree score of each hub, the group to which the hub belongs, the number of years a hub has been active on Twitter/X, its follower count, the number of users a hub is following, the number of followers each hub gained per year on average, the number of pesticide-related tweets each hub had posted, the number of total posts a hub currently had, whether the hub had an associated account, and if this hub was related to pesticide. For predictions based on individual tweets, we included features related to their mentioned hubs, such as the user's group affiliation, the tweet's sentiment toward its mentioned hub, and the individual user's

follower count. We conducted experiments 10 times and used SHapley Additive exPlanations (SHAP; [Lundberg & Lee, 2017](#)) values to identify the features most strongly correlated with the prediction tasks, illustrating the significant drivers of sentiment in our models.

Results and discussion

Hub characteristics comparison

Our updated Twitter/X dataset included 3,149,223 tweets from 917,945 unique user accounts, which we ranked by their in-degree score. From the most influencing 600 influencers/hubs, we manually identified 230 accounts directly associated with pesticide-related fields (albeit they might not tweet frequently on pesticides) or specifically discussing pesticides. [Figure 1](#), left image shows the distribution frequency of general and pesticide-related hubs stratified by each group.

Pesticide-related hubs have larger in-degree than general hubs. Specifically, the average in-degree score for all 600 hubs was 617.5, compared to 773.7 for pesticide-related hubs, i.e., 25.3% higher for the latter, while the median in-degree was 268 vs. 274, i.e., 2.3% higher for pesticide-related hubs. When looking at hubs with in-degree scores in the top 90th percentile, i.e., the most popular groups in the network, the most prevalent ones overall were individuals, followed by industry and government. However, for the 230 pesticide-related hubs, government and NGO_Env groups were most represented, followed by individuals. The in-degree distribution comparison is presented in [Table 1](#) (bold text represents the highest values), and [Figure 1](#), right image illustrates the group distribution for hubs in the top 90th percentile.

Regarding the frequency of pesticide-related tweets, we found that 70.7% of all hubs posted at least one tweet related to pesticides, while this figure increased to 87.4% for specific pesticide-associated hubs. The average number of tweets posted by general influencers was 24.5 tweets/year of being active on Twitter, whereas pesticide-associated hubs posted 45.9 tweets/year, i.e., an 87.3% increase. The distribution showed that 80% of general hubs posted between 0.2 and 56.2 tweets/year, and half posted at least 3.7 tweets/year. Of the 230 pesticide-focused hubs, 80% posted between 1.7 and 94.8 tweets/year, with half posted at

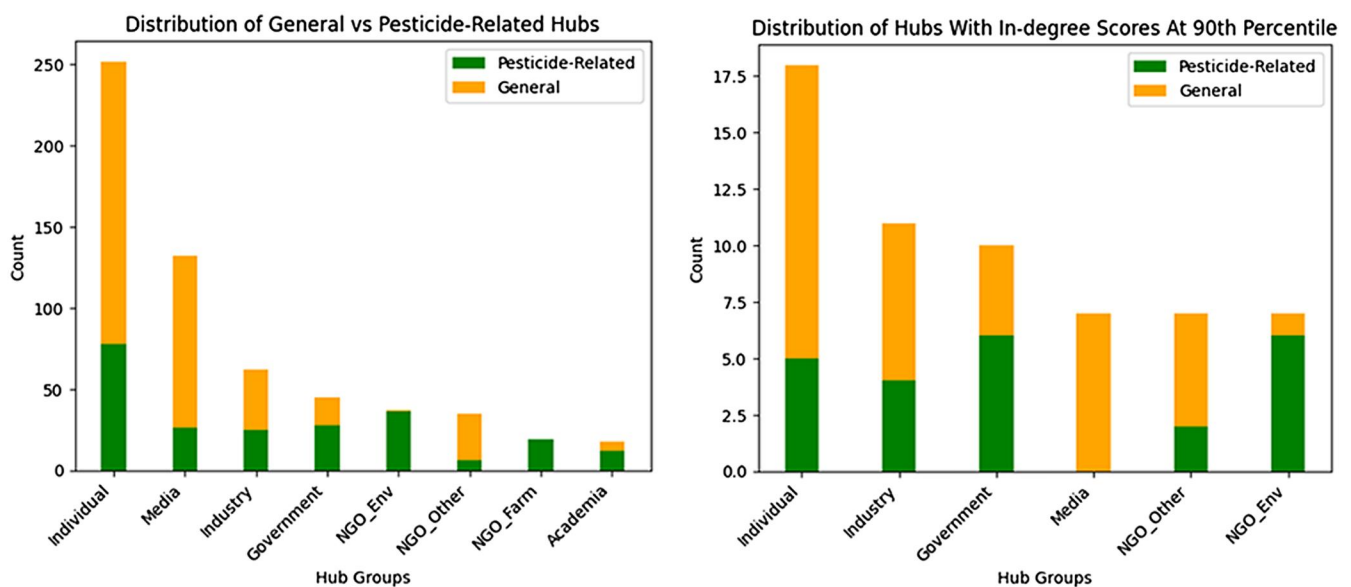


Figure 1. Group distribution of general and pesticide-related hubs. NGO: nongovernmental organization.

Table 1. In-degree summaries for user accounts directly/indirectly involved in pesticide-related discussions (2013–2021).

	Min	Max	Average	Standard deviation	Median	10 th percentile	90 th percentile
All hubs	140	13,304	617.5	1,128.2	268	151	1,300.2
Pesticide-related hubs	140	13,304	773.7	1,491.8	274	151	1,828.8

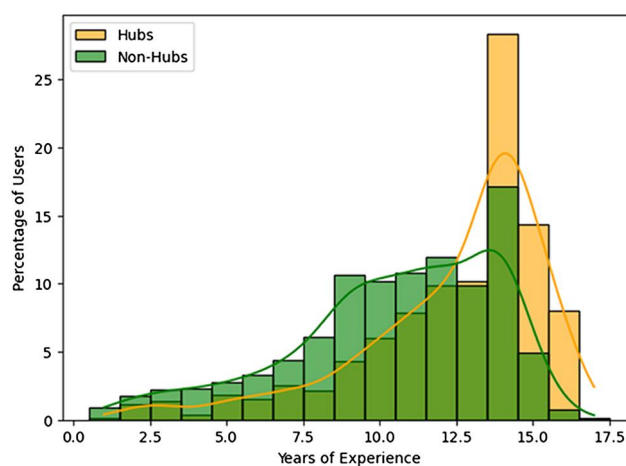
Table 2. Summaries of tweeting activity for user accounts directly/indirectly involved in pesticide-related discussions (2013–2021).

	Percentage of hubs that have at least one post	Min	Max	Average	Standard deviation	Median	10 th percentile	90 th percentile
All hubs	70.7% (424)	0.1	1,402.6	24.5	87.4	3.7	0.2	56.2
Pesticide-related hubs	87.4% (201)	0.1	1,402.6	45.9	122	13.7	1.7	94.8

least 13.7 tweets/year, indicating that these users were much more active in the targeted discussions—in other words, general influencers may post a lot but do not necessarily dive into specific discussions, and especially for pesticide-related discussions. The details of the comparison can be found in [Table 2](#) (bold text represents the highest values). The group distribution for hubs actively posting pesticide-related tweets is shown in [online supplementary material, Figure S1](#). The individual group dominates in terms of activity, being the most active users for all hubs and pesticide-related hubs. Additionally, since there is a significant gap between the median and the average amount of tweets posted by both the whole set and the pesticide-specific hubs, we inferred that some hubs, despite being influential and pesticide-focused, might not be actively tweeting about the subject in depth. In contrast, some hubs are highly active, eager to disseminate information, and frequently post tweets, contributing significantly to the average tweet count. To explore this further, we investigated the group distribution among the hubs whose volume of pesticide-related tweets was in the 90th percentile. We found that the individual group again leads in tweet volume, followed by NGO_Env for both categories of hubs, indicating that the individual users and the organizations that focus on environment-related fields are the most active in creating posts and expressing their opinions about pesticides, as illustrated in [online supplementary material, Figure S2](#).

To investigate how the experience on Twitter/X impacted users' status as a hub, we compared the distribution of the number of years of activity for hubs/influencers vs. other user accounts (see [Figure 2](#)). On average, hubs have 12.4 years of experience on Twitter/X, notably higher than the 10.5 years observed for nonhub users. Looking at the 10-year longevity, 84.67% of hubs met the mark vs. only 65.64% of nonhubs, demonstrating that longer experience with the app is associated with higher, sustained influence.

The number of followers and following are critical metrics for assessing a user's popularity within a network, complementary to the in-degree score. Unlike in-degree influence, which indicates whether a user is a key player in shaping opinions or being relevant in real-time discussions, a user with many followers can have a significant impact by broadcasting messages to a large audience but might not necessarily mean that the content is being actively engaged with or discussed. For general hubs, the average follower count was 3,725,852, with 55% of them having more than 100,000 followers. In contrast, the specified pesticide-related hubs had a significantly lower average of 328,627 followers, and only 31.7% had more than 100,000

**Figure 2.** Distribution of years of experience with Twitter.

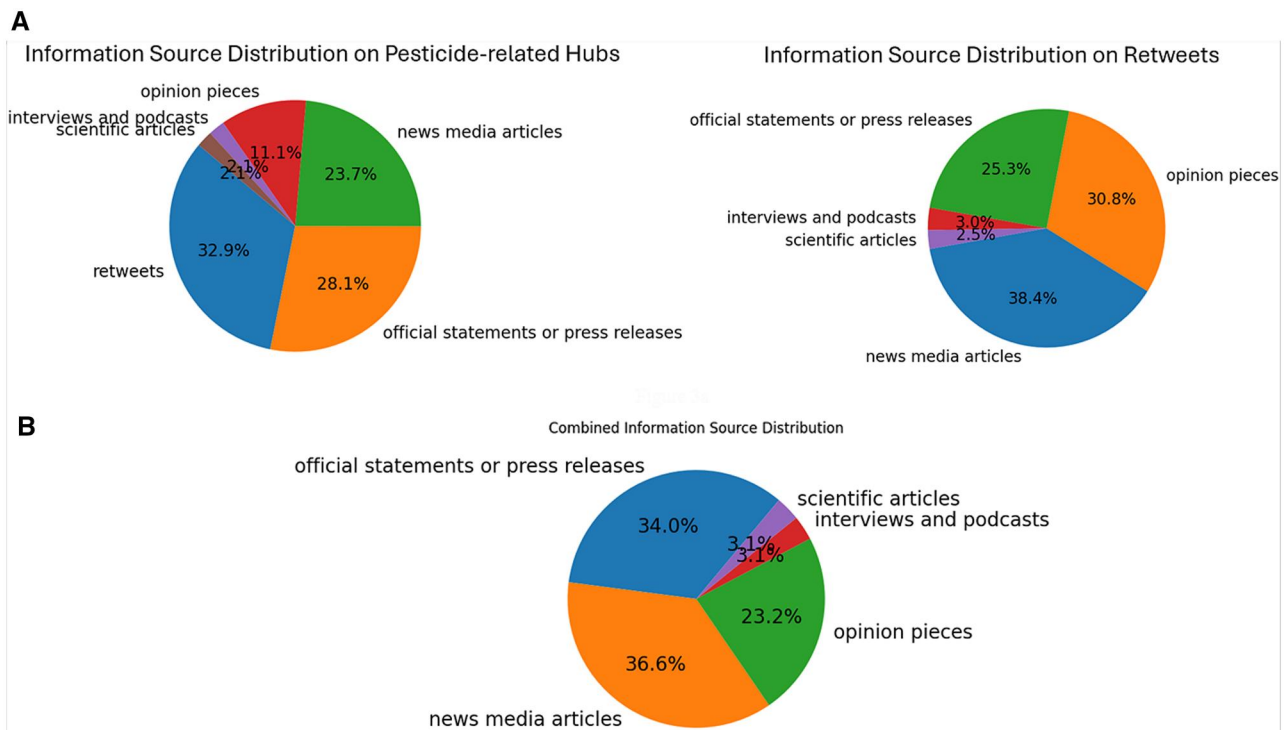
followers (see [Table 3](#)). Regarding group distribution among the hubs that have the highest number of followers, individuals had the greatest number of hubs, followed by media and government organizations with the number of followers in the top 90th percentile, highlighting their substantial presence and popularity within this specific discourse.

Interestingly, our analysis of the number of users that hubs follow presented an inverse relationship compared to their follower counts. Pesticide-related hubs followed an average of 8,643 users, which is notably higher than the average of 7,180 users followed by all hubs in general. Among pesticide-related hubs, 80% follow between 290 and 14,360 users, with a median of 2,154. In contrast, for all hubs, 80% followed between 103 and 11,710 other users, with a median of 1,309, which was 39.2% less than pesticide-related hubs. In terms of commonality between all and pesticide-related hubs, 12% of hubs have followed over 100,000 users for both hub categories. However, a significant difference was observed in the intermediate range: 57% of pesticide-related hubs followed between 10,000 and 100,000 users, compared to only 45% of all hubs, as detailed in [Table 3](#) (bold text represents the highest values).

Furthermore, when looking at the group distribution of hubs at the 90th percentile of the following, the individuals' group led both categories of hubs, highlighting their strong intention on connecting others. However, while the media group was expectedly the second most prevalent among all hubs, the NGO_Env group took the second place in the pesticide-related hubs.

Table 3. Descriptive statistics on the number of followers and following on all hubs and pesticide-focused hubs.

	Average	Standard deviation	Median	10 th percentile	90 th percentile	Percentage of hubs that have less than 10,000 followers	Percentage of hubs that have 10,000–1,00,000 followers	Percentage of hubs that have more than 1,00,000 followers
Number of followers								
All hubs	3,725,852.11	3,085,425.81	56,000.05	536.5	8,972,422.3	14% (85)	31% (184)	55% (331)
Pesticide-related hubs	3,28,627.7	1,473,167.5	41,700.5	3,947.9	4,14,948.6	20% (46)	48.2% (111)	31.7% (73)
Number of following								
All hubs	7,180.7	32,090.4	1,309.5	103.3	11,710	42.3% (254)	45.7% (274)	12% (72)
Pesticide-related hubs	8,642.7	33,308.7	2,154	290.7	14,360	30% (69)	57.4% (132)	12.6% (29)

**Figure 3.** Information source distribution.

Further, we examined the prevalence of website ownership among all hubs compared to pesticide-related hubs. We found that the distribution of hubs with a website was nearly identical between the two categories, showing that both general and pesticide-related hubs maintain comparable outside-app online presence.

Information source distribution

As outlined in the methodology section, we classified the sources of information into six categories: official statements or press releases, news media articles, retweets, opinion pieces, interviews or podcasts, and scientific articles. Our analysis revealed a heterogeneous distribution of these sources across the pesticide-related hubs. Specifically, retweets constituted 32.9% of the information sources in these hubs, followed by official statements and press releases at 28.1%, and news media articles at 23.7%. In contrast, a mere 2.1% of the hubs featured information from interviews and podcasts, and scientific articles, as shown in the left panel of Figure 3A.

Given that retweeting emerges as the predominant method of information dissemination, we conducted an in-depth analysis of the content of these retweets. As the distribution shown in the right panel of Figure 3A, retweeted content primarily consists of news media articles, representing 38.4% of all retweets; 30.8% of

retweets convey other individuals' opinions, thoughts, or personal experiences; and an additional 25.3% of the retweets comprised official statements or press releases. Finally, we combined the first two panels from Figure 3A by replacing the retweets section with the other five information source categories, as shown in Figure 3B. Overall, news media articles accounted for 36.6%, official statements or press releases for 34.0%, opinion pieces for 23.2%, scientific articles for 3.1%, and interviews and podcasts for 3.1%.

Information flow across the network

To assess the impact of pesticide-related hubs on information dissemination within the social network, we calculated the adapted information flow as specified in the methods. On average, hubs exhibited an information flow score (matched ratio) of 61.6%, and half of these hubs had an information flow score of at least 75.2%, with a score of 99.4% as the 90th percentile. The distribution of information flow scores is shown in [online supplementary material, Figure S3](#). Note that 12.6% of pesticide-focused hubs had a zero score, i.e., failed to meet any of the matching criteria. No pesticide-relevant tweets were retrieved by the API (except passive mentions) for 13 out of 28 hubs categorized under government, resulting in the lowest average matched ratio score among the groups. After

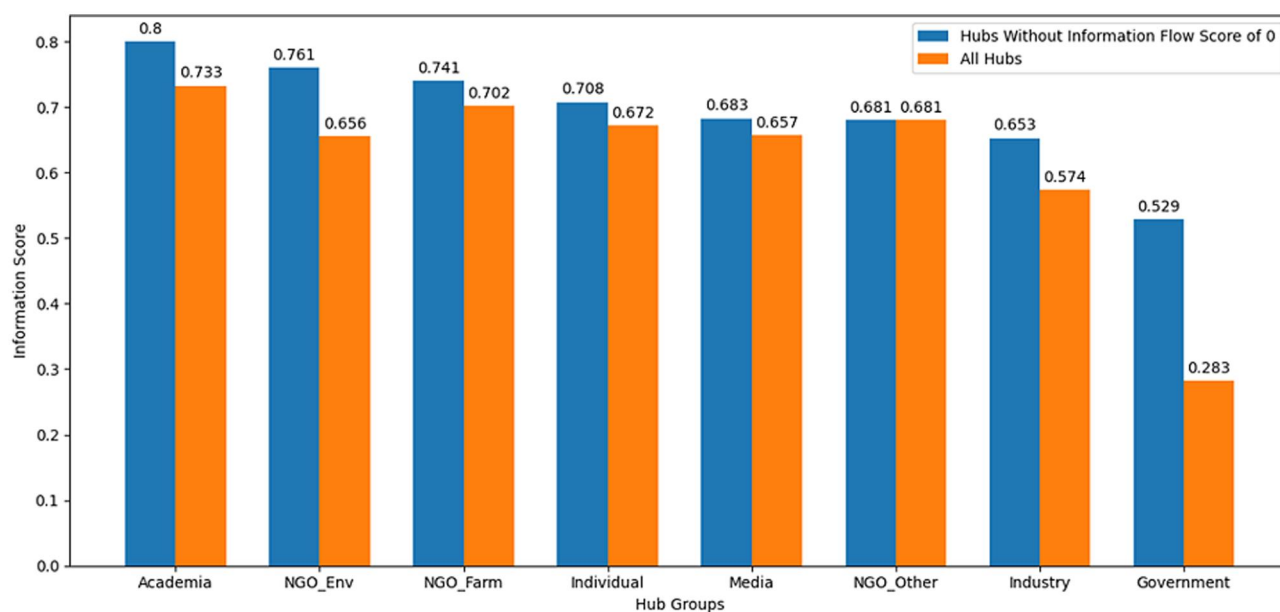


Figure 4. Average information score for pesticide-focused hub groups. NGO: nongovernmental organization.

excluding all hubs with a zero score, the average score for government hubs increased from 0.283 to 0.529. Despite this improvement, it remained the lowest average score compared to other groups, while academia held the highest score, as shown in Figure 4.

Looking at nonzero hubs that achieved scores in the top 90th percentile (see [online supplementary material, Figure S4](#)), we found that the NGO_Env group contains the highest number of hubs within this top tier, followed by the individuals.

Sentiment analysis

We selected all the tweets that mentioned/tagged the generic and pesticide-specific hubs and inferred their sentiments toward the hubs. [Online supplementary material, Figure S5](#) illustrates how the negative sentiment was predominant in both hub categories but was 25.8% higher in the general hubs compared to the pesticide-focused ones (59.2% vs. 43.9%). The same held for neutral sentiment, which was the second most prevalent, while the positive was the least.

[Online supplementary material, Figures S6–S8](#) show the detailed breakdown of each sentiment per account group. There was high heterogeneity among groups and hub types. It is worth noting that, among pesticide-focused hubs, NGO_Farm and industry hubs had the largest share of positive sentiment, while individuals and media had the highest share of negative sentiments; individuals also dominated the neutral sentiments, and NGO_Env had high shares of both negative and positive tweets.

When looking at intragroup sentiment distribution in [Figure 5](#), the media stand out among others for extremely negative leaning, while NGO_Farm had greater positive takes.

To understand the key features of user accounts that would predict one sentiment over another, we applied machine learning methods. Out of the four machine learning methods applied, we excluded gradient boosting due to extensive training time and suboptimal performance.

For the hub-based model fit, the random forest model outperformed others, achieving an average F1 score of 0.651 and an AUROC of 0.718 on the prediction of neutral vs. others. In predicting negative and positive sentiments, the random forest model

again achieved superior performance, with an average F1 score of 0.712 and an AUROC of 0.751 (see [Table 4](#), where bold text represents the highest values).

Additionally, we analyzed the correlation of different features with the prediction outcomes using SHAP values. For the neutral vs. others category, the features with the highest correlations were the media group, followed by the government, NGO_Env, in-degree score, and NGO_Other groups. For the positive vs. negative predictions, the media group also showed the highest correlation, followed by NGO_Env, NGO_Farm, industry, and government (see [online supplementary material, Figure S9](#)).

For the individual tweets-based sentiment prediction, XGBoost was the top-performing model. Specifically, for predicting neutral vs. others sentiments, XGBoost achieved the highest F1 score of 0.726 and AUROC of 0.796. In the classification of positive vs. negative sentiments, XGBoost maintained the highest AUROC at 0.750. However, logistic regression outperformed XGBoost in this scenario with the best F1 score of 0.670. Detailed comparisons of these performances are presented in [Table 5](#) (bold text represents the highest values).

For the neutral vs. others sentiment prediction, the feature with the highest impact on sentiment classification was the number of posts from the mentioned hubs and was closely followed by the hub in-degree score, which reflects the high connectivity of a hub within the network. Additionally, the number of users each hub followed, the number of followers each tweet author had, and the volume of pesticide-related tweets emanating from each hub were also significant contributors to the predictive model's accuracy.

In contrast, for the negative vs. positive sentiment prediction, the sentiments expressed toward the hubs stood out as the most influential feature. This feature's impact was markedly higher than other variables, suggesting that the general sentiment toward hubs is critical in distinguishing between positive and negative sentiments. This feature significance for both scenarios is visually detailed in [online supplementary material, Figure S10](#), clearly representing how different features disproportionately affect sentiment classification outcomes in distinct categories.

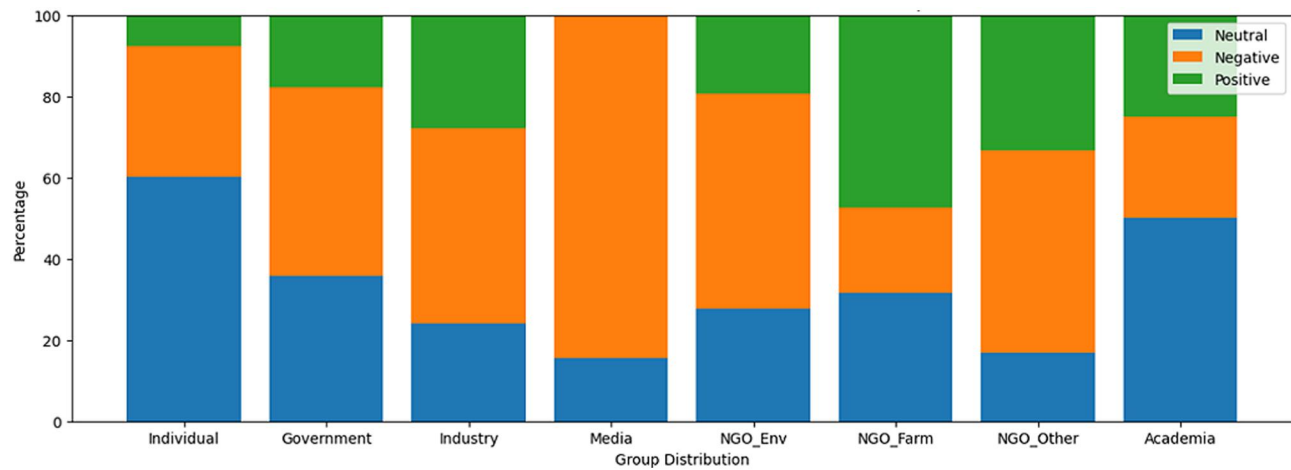


Figure 5. Sentiment distribution for pesticide-focused hub group. NGO: nongovernmental organization.

Table 4. Overview of hub-based sentiment prediction performance.

	Accuracy	Precision	Recall	F1-score	AUROC
Neutral vs. others					
Random forest	0.652	0.651	0.652	0.651	0.718
Gradient boosting	0.614	0.642	0.614	0.616	0.679
XGBoost	0.584	0.603	0.584	0.587	0.640
Logistic regression	0.622	0.633	0.622	0.625	0.643
Negative vs. positive					
Random forest	0.681	0.788	0.681	0.712	0.751
Gradient boosting	0.664	0.789	0.664	0.699	0.738
XGBoost	0.659	0.795	0.661	0.696	0.721
Logistic regression	0.664	0.766	0.664	0.697	0.657

Table 5. Overview of tweets-based sentiment prediction performance.

	Accuracy	Precision	Recall	F1-score	AUROC
Neutral vs. others					
Random forest	0.687	0.824	0.687	0.725	0.795
XGBoost	0.689	0.825	0.689	0.726	0.796
Logistic regression	0.688	0.767	0.681	0.714	0.672
Negative vs. positive					
Random forest	0.662	0.690	0.662	0.667	0.746
XGBoost	0.662	0.695	0.662	0.668	0.750
Logistic regression	0.676	0.6682	0.676	0.670	0.683

Discussion

In this work, we analyzed the influencers' characteristics, information flows, and mutual reactions for pesticide-related discussions on Twitter/X using a keyword- and account-curated sample of over 3 million tweets and almost a million users between 2013 and 2022, extending a prior study that looked at topic content and individual sentiments (Jun et al., 2023). About a third of the most influential general hubs (identified using accounts' in-degree measure) included accounts that were highly specific to pesticide and pesticide-related discussions (230 out of 600, each reviewed manually), e.g., environmental agency. Among the general hubs, individuals were the most prevalent (over 40%), closely

aligned with findings from a previous study (Wood-Doughty et al., 2018). However, the most influential hubs were not individuals—only two were in the top 10, i.e., @realDonaldTrump and @JunckerEU, with the remainder being government, industry, and NGO_Other accounts. For pesticide-focused hubs, government and NGO_Env entities were the most represented, followed by individuals—yet, again, no individuals were in the top 10. This would suggest that users of the app were keener to engage with organizations than other individuals regarding pesticide-related topics. The pesticide-focused accounts were on average more active than general hubs, since general influencers may have posted a lot but did not necessarily dive into specific discussions, especially for pesticide-related discussions. The high activity is in line with prior findings (Wojcik & Hughes, 2019) where the top active users created 80% of tweets in the network, and the tweets from these hubs typically included news sharing and announcing new products or events. In terms of tweet volume, NGO_Env, individuals, and media largely surpassed government and other outlets (tweeting on more generic content). All hubs on average had a longer app presence than nonhubs, and individuals posted the largest volume of tweets per unit time (year), followed by NGO_Env.

Besides the in-degree score, we also looked at hubs from the number of followers, which is an essential metric for gauging an account's influence within the network (Woodward, 2022). The generic hubs contained the top accounts with the most followers on Twitter/X, such as @ElonMusk, @BarackObama, @YouTube, and @CNN, and their average number of followers was almost 10 times more than pesticide-related hubs. This is not unexpected, as pesticide-related discussions are not among the most viral on Twitter/X, although in a prior study, we observed virality spikes corresponding to key events (Jun et al., 2023). On the other hand, when looking at the number of followings, the pesticide-related hubs had a higher average than generic hubs by 20%. According to Cha et al. (2012), grassroots or ordinary users and evangelists, including opinion leaders, local businesses, and organizations, tend to reciprocate (follow back) most of their followers and are more willing to follow others to expand their network activity. In our analysis, apart from individual users, NGO_Env hubs did not only contribute to a high tweet volume but also actively followed others in the network (second place in the pesticide-related hubs in terms of following counts), making them one of the most engaged groups in networking. Finally, over 83% of pesticide-focused hubs had an associated website; although it is not

uncommon for influencers to have websites, these websites could be used as additional sources to scrape pesticide information—in conjunction with the other sources here considered.

The information sources were heterogeneously distributed among hubs. First, retweets were the most common way of posting, accounting for one-third of the total. After looking at the retweet original content, the most common information sourcing were news media articles (over 35%), official statements or press releases (34%), and opinion pieces (23%). Scientific articles had a very low share (3%), similar to interviews and podcasts. This pattern underscored the influence of authoritative and opinion-driven content on Twitter user interactions and information dissemination (Yan, 2024), while the core source of scientific literature is inconsequential relative to opinion, platitudes, and hyperbole.

Further, a prior study showed that about 70% of Twitter/X users purchased products from the accounts they followed, indicating how product placement by influencers directly impacts users' interest, purchase choices, or business endeavors (Jay, 2019).

The information flow through hubs was remarkable—with a median value of over 75%. Academia, NGO_Farm, and individuals exhibited the highest scores, albeit the other groups were close, with the exception of the government that showed the lowest flow.

According to Jain and Sinha (2020), opinion leaders are often highly influential in the Twitter network, and the “academia” group consists of the majority of educational organizations, such as research centers, and “opinion leaders,” such as professors, who are pivotal in spreading scientific knowledge and attracting attention from peers and students. One potential explanation is that the pursuit of truth in science is often undermined by the sensationalism and negative bias prevalent in media, especially in fields like ecotoxicology. Humans are naturally drawn to negative information, a trait exploited by media outlets to capture attention, which often results in exaggerated or skewed representations of scientific findings (not just limited to scientific articles; Brain & Hanson, 2021). Thus, although “academia” held the highest information flow score based on our experimental results, it is essential to always critically evaluate the veracity of any information source. Conversely, the “Government” group exhibited the lowest score on average, with nearly half of its hubs not posting about pesticides despite their relevance to environmental and health concerns. This reflected a relatively low level of engagement from government entities in pesticide-related discussions. When looking at the reception of hubs' content (i.e., sentiments toward an influencer), we found that negative dominated over neutral and positive for both general and pesticide-focused hubs (44%–60%). This is a common finding with several environmental and health-related foci, e.g., trusting organic foods (Singh & Glińska-Neweś, 2022). The media group, in particular, exhibited the highest proportion of negative reception among other groups, in line with prior research (Brain & Hanson, 2021; Sirvydaitė, 2021). Although media outlets are not directly linked to pesticides, the result underscores a general discontent or distrust from app users. Also, from the machine learning analysis, we found that the sentiment from other users toward a hub on average is the most important feature for sentiment prediction at the tweet level, indicating that the hubs' reception further influences the single postings.

While this study did not specifically track the temporal dynamics of information flow, our analysis provided insights into general patterns of engagement and influence within the pesticide-related discussion network. Observed trends suggested it is likely that information flow followed a cyclical pattern

influenced by key events such as policy changes, media reports, and scientific findings (Cha et al., 2012). Previous research has shown that social media activity often spikes around major events and gradually declines over time (Jun et al., 2023), suggesting that pesticide-related information may have experienced similar trends. The dominance of organizational hubs, such as NGOs and media, indicated that structured and targeted communication efforts likely contributed to sustained information flow over extended periods. Additionally, engagement metrics and sentiment analysis results suggested that public interest in pesticide-related discussions is periodically revitalized by news coverage, regulatory updates, and advocacy efforts.

Conclusions

This study provided a comprehensive analysis of pesticide-related discussions on Twitter/X focusing on information hubs and flows, revealing significant insights into the dynamics of information reception, dissemination, and influence.

Building on strong evidence that influencers shape public opinions, behaviors, and consumer decisions, this study introduced valuable metrics and insights into the dynamics of pesticide-related information dissemination and influence. These insights can inform science-driven environmental management by enabling policymakers to design credible evidence-based communication strategies that address public concerns while promoting accurate and balanced narratives about pesticides. By recognizing the dominant role of NGOs and media in shaping public sentiment, regulators and industry stakeholders can integrate diverse stakeholder values more effectively, enhance transparency, and foster trust through tailored messaging. Moreover, this work highlighted the importance of amplifying information obtained from credible scientific sources and addressing sectoral disparities, particularly the limited engagement from governmental accounts and organizations. These findings can also aid in balancing sectoral differences and improving collaboration across academia, industry, and civil society to advance sustainable pesticide management and policy. By leveraging the dynamics of information flow, decision-makers can enhance public engagement, align communication strategies with stakeholder expectations, and foster more informed and equitable environmental governance.

Based on the findings from this study, we offer several key recommendations. First, policymakers should develop clear and transparent communication strategies prioritizing credible, scientifically validated information. Second, critical thinking and informed skepticism among the public should be advocated to counter narrative-based commentary and facilitate informed discourse. Third, the governmental sector should be more engaged to provide perspective and clarity regarding regulatory decisions and positions. Fourth, multipartite collaboration among academia, industry, nongovernmental organizations, and government should be pursued to promote balanced and representative dialogue and debate in the public domain. Lastly, decision-makers should actively monitor public sentiment to dynamically adapt communication strategies to address emerging concerns and misinformation, ensuring that messaging remains accurate, relevant, timely, and impactful.

Supplementary material

Supplementary material is available online at *Integrated Environmental Assessment and Management*.

Data availability

The data that support the findings of this study are available through Twitter Academic Research access, but restrictions apply (i.e., a researcher needs academic research access and uses the same keywords that this study used to extract historical tweet volume within the time window). More details about the Twitter API Academic Research access are listed at <https://developer.x.com/en/products/x-api>

Author contributions

Tiancheng Zhou (Data curation, Formal analysis, Investigation, Methodology, Project administration), Yu Huang (Conceptualization, Investigation, Methodology, Project administration, Supervision), Raghavendhran Avanasi (Conceptualization, Investigation, Methodology, Project administration, Supervision), Richard Brain (Conceptualization, Investigation, Methodology, Project administration, Supervision), Mattia Prosperi (Conceptualization, Investigation, Methodology, Project administration, Supervision), and Jiang Bian (Conceptualization, Investigation, Methodology, Project administration, Supervision)

Funding

This research was funded by Syngenta Crop Protection, LLC under the project number TK0251998.

Conflicts of interest

None declared.

Acknowledgments

R.A. and R.B. are employees of Syngenta Crop Protection. The authors thank Syngenta Crop Protection, LLC for resource assistance as well as research sponsorship. The authors thank the University of Florida for computing resources.

References

- Aktar, M. W., Sengupta, D., & Chowdhury, A. (2009). Impact of pesticides use in agriculture: Their benefits and hazards. *Interdisciplinary Toxicology*, 2, 1–12.
- Andreu, V., & Picó, Y. (2004). Determination of pesticides and their degradation products in soil: Critical review and comparison of methods. *TrAC Trends in Analytical Chemistry*, 23, 772–789.
- Bortleson, G., & Davis, D. (1987). *Puget sound basin*. U.S. Geological Survey & Washington State Department of Ecology.
- Brain, R. A., & Hanson, M. L. (2021). The press sells newspapers, we should not sell ecotoxicology. *Environmental Toxicology and Chemistry*, 40, 1239–1240.
- Burgess, J., Marwick, A., & Poell, T. (2019). *The SAGE handbook of social media*. SAGE Publications.
- Castillo, C., Mendoza, M., & Poblete, B. (2011). Information credibility on Twitter. In *Proceedings of the 20th International Conference on World Wide Web* (pp. 675–684). ACM. <https://doi.org/10.1145/1963405.1963500>
- Cha, M., Benevenuto, F., Haddadi, H., & Gummadi, K. P. (2012). The world of connections and information flow in Twitter. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 42, 991–998.
- Dann, S. (2010). Twitter content classification. *First Monday*, 15. <https://doi.org/10.5210/fm.v15i12.2745>
- Debreceeny, R. S., Wang, T., & Zhou, M. (2019). Research in social media: Data sources and methodologies. *Journal of Information Systems*, 33, 1–28.
- Eddleston, M. (2000). Patterns and problems of deliberate self-poisoning in the developing world. *QJM: Monthly Journal of the Association of Physicians*, 93, 715–731. <https://doi.org/10.1093/qjmed/93.11.715>
- El Ali, A., Stratmann, T. C., Park, S., Schöning, J., Heuten, W., & Boll, S. (2018). Measuring, understanding, and classifying news media sympathy on Twitter after crisis events. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM. <https://doi.org/10.1145/3173574.3174130>
- Gottfried, J. (2024). *Americans' social media use*. Pew Research Center. <https://www.pewresearch.org/internet/2024/01/31/americans-social-media-use/>. Date accessed May 13, 2024.
- Himmelboim, I., & Han, J. Y. (2014). Cancer talk on Twitter: Community structure and information sources in breast and prostate cancer social networks. *Journal of Health Communication*, 19, 210–225.
- Hutto, C., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In Y.-R. Lin, Y. Mejova, & M. Cha (Eds.), *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 8, No. 1, pp. 216–225).
- Jackson, M., Brennan, L., & Parker, L. (2021). The public health community's use of social media for policy advocacy: A scoping review and suggestions to advance the field. *Public Health*, 198, 146–155.
- Jain, S., & Sinha, A. (2020). Identification of influential users on Twitter: A novel weighted correlated influence measure for Covid-19. *Chaos, Solitons, and Fractals*, 139, 110037.
- Jay, A. (2019). 85 Twitter statistics you must know: 2024 market share analysis & data. *FinancesOnline*. <https://financesonline.com/twitter-statistics/>. Date accessed May 28, 2024.
- Jun, I., Feng, Z., Avanasi, R., Brain, R. A., Prosperi, M., & Bian, J. (2023). Evaluating the perceptions of pesticide use, safety, and regulation and identifying common pesticide-related topics on Twitter. *Integrated Environmental Assessment and Management*, 19, 1581–1599.
- Kapoor, K. K., Tamilmani, K., Rana, N. P., Patil, P., Dwivedi, Y. K., & Nerur, S. (2018). Advances in social media research: Past, present, and future. *Information Systems Frontiers*, 20, 531–558.
- Lundberg, S., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. arXiv, <http://arxiv.org/abs/1705.07874>, preprint: not peer reviewed.
- Mary. (2011). *Indegree, retweets, and mentions: What does it take to be influential on Twitter?* DiploFoundation. <https://www.diplomacy.edu/blog/indegree-retweets-and-mentions-what-does-it-take-be-influential-twitter/>. Date accessed June 15, 2024.
- Mitchell, A. (2021). *News on Twitter: Consumed by most users and trusted by many*. Pew Research Center. <https://www.pewresearch.org/journalism/2021/11/15/news-on-twitter-consumed-by-most-users-and-trusted-by-many/>. Date accessed May 15, 2024.
- Moloko, M. G., Petersen, B., et al.; E Squared Investments. (2010). *The 100 most influential news media Twitter accounts*. Memeburn. <https://memeburn.com/2010/09/the-100-most-influential-news-media-twitter-accounts/>. Date accessed May 15, 2024.
- Pascual-Ferrá, P., Alperstein, N., & Barnett, D. J. (2022). Social network analysis of COVID-19 public discourse on Twitter: Implications for risk communication. *Disaster Medicine and Public Health Preparedness*, 16, 561–569.
- Pimentel, D., Cooperstein, S., Randell, H., Filiberto, D., Sorrentino, S., Kaye, B., Nicklin, C., Yagi, J., Brian, J., O'Hern, J., Habas, A., & Weinstein, C. (2007). Ecology of increasing diseases: Population growth and environmental degradation. *Human Ecology: An Interdisciplinary Journal*, 35, 653–668.

- Singh, A., & Glińska-Noweś, A. (2022). Modeling the public attitude towards organic foods: A big data and text mining approach. *Journal of Big Data*, 9, 2.
- Sirvydaitė, E. (2021). Digital media interactions: Sentiment analysis of users' comments of "BBC news" and "CNN" YouTube videos. <https://doi.org/10.13140/RG.2.2.31928.06403>
- Sotrender. (2022). What are mentions and how do I use them? Sotrender Blog. https://www.sotrender.com/blog/knowledge_base/what-are-mentions-and-how-do-i-use-them/. Date accessed May 27, 2024.
- Sugawara, Y., Narimatsu, H., Hozawa, A., Shao, L., Otani, K., & Fukao, A. (2012). Cancer patients on Twitter: A novel patient community on social media. *BMC Research Notes*, 5, 699.
- Twitter API Documentation. (2024). <https://developer.x.com/en/docs/twitter-api>. Date accessed June 25, 2024.
- Walsh, S. (2024). The top 10 social media sites & platforms. *Search Engine Journal*. <https://www.searchenginejournal.com/social-media/social-media-platforms/>
- Wang, Y., McKee, M., Torbica, A., & Stuckler, D. (2019). Systematic literature review on the spread of health-related misinformation on social media. *Social Science & Medicine* (1982), 240, 112552.
- Wojcik, S., & Hughes, A. (2019). Sizing up Twitter users. Pew Research Center. <https://www.pewresearch.org>. Date accessed May 15, 2024.
- Wood-Doughty, Z., Mahajan, P., & Dredze, M. (2018). Johns Hopkins or Johnny-Hopkins: Classifying individuals versus organizations on Twitter. In *Proceedings of the Second Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media*. <https://doi.org/10.18653/v1/w18-1108>
- Woodward, M. (2022). Twitter user statistics 2024: What happened after "X" rebranding? SearchLogistics. <https://www.searchlogistics.com/learn/statistics/twitter-user-statistics/>
- Yan, E. (2024). *Prompting fundamentals and how to apply them effectively*. eugeneyan.com. <https://eugeneyan.com/writing/prompting/>. Date accessed July 1, 2024.
- Yin, F., Zhu, X., Shao, X., Xia, X., Pan, Y., & Wu, J. (2021). Modeling and quantifying the influence of opinion involving opinion leaders on delayed information propagation dynamics. *Applied Mathematics Letters*, 121, 107356.