Weaponization of Social Media Resulting in Unrest in the Society

Vyankatesh Vilasrao Kahale¹*, Dr. Vikas K Jambhulkar²

1 Research Scholar, Department of Political Science, Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur, Maharashtra, India

vanky612@gmail.com

2 Guide and Head of Department of Political Science, Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur, Maharashtra, India

Abstract: The fast growth in social media use has led to a shift in social discourse, political communications, and civic incorporation. Nevertheless, the manipulative aspect of it as a political instrument has caused the weaponization against people through platforms like digital media, which has also been a significant cause of social instability. The paper will focus on the role of algorithmic amplification, misinformation, coordinated disinformation campaigns, and echo chambers in enhancing polarization and fuelling real-life instability. It examines the causality of the connection between propaganda online and offline outcomes, which involve communal strains, protest escalation, mob violence, and institutional warning loss. It applies a doctrinal and analytical methodology to measure structural weaknesses in social media systems and constraints of current regulatory methods. The article posits that the unregulated digital manipulation endangers social unity and democracy. It concludes that dealing with weapons of social media necessitates enhanced digital literacy, open platform management, proportional regulation, and multi-televangel stakeholder cooperation to ensure that digital narratives are not reflected in the unrest of society without infringing on the freedom of expression.

Keywords: Social Media, Weaponization, Misinformation, Digital Propaganda, Social Unrest, Cyber Law, Public Order.

INTRODUCTION

The rapid expansion of social media platforms over the past decade has fundamentally reshaped patterns of communication, political engagement, and collective mobilization. Platforms such as Facebook, X (formerly Twitter), Instagram, and WhatsApp have evolved from networking tools into powerful infrastructures that mediate public discourse, influence

electoral processes, and shape societal narratives. While these platforms have enhanced participatory democracy and enabled real-time information exchange, they have also created new vulnerabilities that can be strategically exploited. The phenomenon commonly described as the "weaponization" of social media refers to the deliberate use of digital platforms to manipulate information, intensify divisions, and incite social, political, or communal unrest.

Unlike traditional media ecosystems, social media operates within decentralized and user-driven environments where content dissemination is rapid, borderless, and often weakly regulated. The convergence of anonymity, virality, and algorithmic personalization allows malicious actors—including political propagandists, extremist groups, and foreign influence networks—to engineer narratives capable of shaping public perception at scale. Allcott and Gentzkow (2017)[1] observe that the low cost of producing and distributing online misinformation has significantly altered the information marketplace, making it easier for false or misleading content to compete with verified news. Similarly, Tucker et al. (2018)[2] note that digital platforms can facilitate both democratic participation and large-scale manipulation, depending on how they are used.

The weaponization of social media is particularly concerning in socially diverse and politically sensitive societies where identity-based mobilization can quickly escalate into conflict. In pluralistic democracies like India, where linguistic, religious, and cultural heterogeneity coexist with high digital penetration, the circulation of inflammatory content has, in several instances, intensified communal tensions and public disorder. Aral (2020)[3] argues that the architecture of social networks enhances the speed and reach of emotionally charged information, thereby amplifying its social consequences. When digital narratives are strategically crafted to exploit existing fault lines, they can catalyze real-world unrest, erode institutional legitimacy, and undermine democratic stability.

Another critical dimension of this issue lies in the transformation of information consumption patterns. More and more, people are getting their news via social media, and they don't always check the veracity or validity of the sources. The blending of opinion, propaganda, satire, and factual reporting creates an ambiguous informational environment in which manipulation can thrive. Persily and Tucker (2020)[4] highlight that digital campaigns can exploit data analytics and targeted messaging to influence voter attitudes and public opinion in subtle yet profound ways. Such developments blur the line between persuasion and manipulation, raising complex ethical and legal concerns.

The consequences of weaponized digital communication extend beyond episodic unrest. Sustained exposure to divisive narratives can foster distrust in democratic institutions, weaken social cohesion, and normalize hostility within public discourse. As digital platforms become deeply embedded in governance, commerce, and everyday life, safeguarding them from strategic misuse has emerged as a pressing policy and security challenge.

In light of these concerns, this paper examines the mechanisms through which social media is weaponized and analyzes its impact on social stability and public order. By situating the discussion within contemporary political, technological, and constitutional-regulatory frameworks, the study seeks to contribute to a nuanced understanding of how digital communication infrastructures can both empower societies and destabilize them when exploited for malicious ends.

REVIEW OF LITERATURE

Recent scholarship conceptualizes the contemporary digital environment as an "information pandemic," where disinformation spreads rapidly across platforms, undermining public trust and state resilience. Surjatmodjo, Unde, Cangara, and Sonni (2024)[5] argue that the scale and velocity of online disinformation resemble a public health crisis, as false narratives weaken institutional credibility and disrupt social stability. In addition to distorting public knowledge, their work shows how digitally transmitted misinformation undermines governance systems that depend on informed citizen engagement.

The structural relationship between social media networks and political polarization has been examined extensively. Through the reinforcement of homophilic relationships, in which users tend to communicate within like-minded clusters, Azzimonti and Fernandes (2023)[6] show that social media networks amplify bogus news and increase ideological conflicts. Their findings suggest that misinformation embedded within polarized networks contributes to sustained political fragmentation and diminished consensus-building.

The concept of social media weaponization has gained prominence in conflict studies. Brezatis (2023)[7] emphasizes the high level of coordination between actors engaged in physical warfare and those operating in digital information spaces. This integration of battlefield operations with strategic online propaganda reflects a hybrid conflict model in which narratives, perception management, and psychological operations are deployed alongside conventional

tactics. Such coordination enhances the capacity of digital platforms to influence public opinion and destabilize societies beyond traditional geographic boundaries.

Automated manipulation further complicates the digital landscape. Rodic (2025)[8] in a comprehensive review of bot detection research, notes the increasing sophistication of social bots that mimic human behavior and evade detection systems. The persistence of automated influence operations makes it difficult to distinguish authentic public discourse from orchestrated campaigns. Marigliano, Ng, and Carley (2024) Marigliano, R., Ng, L. H. X., & Carley, K. M. (2024)[9]. Analyzing digital propaganda and conflict rhetoric: a study on Russia's bot-driven campaigns and counter-narratives during the Ukraine crisis. *Social Network Analysis and Mining*, *14*(1), 170. provide empirical evidence of bot-driven propaganda during the Ukraine crisis, demonstrating how coordinated digital networks shape conflict narratives and counter-narratives in real time.

Algorithmic personalization and artificial intelligence systems also contribute to ideological reinforcement. Rodilosso (2024)[10] argues that AI-driven recommendation systems can unintentionally foster extremism by curating emotionally engaging and confirmatory content. This technological filtering mechanism narrows users' informational exposure, reinforcing polarized worldviews. Complementing this perspective, Fahad and Mustafa (2025)[11] analyze echo chamber dynamics in Delhi, illustrating how online communities can become locked into self-reinforcing cycles of radicalization and identity-based hostility.

Looking at the field of disinformation research via a more holistic lens, Xu, Qian, and Meng (2025)[12] highlight important themes such as polarization, algorithmic amplification, digital propaganda, and social impact, and they trace its history over the last decade. According to their meta-analysis, more and more people are starting to see disinformation as a systemic phenomenon with deep roots in platform infrastructures and social and political contexts, rather than just an isolated incident of poor communication.

Collectively, these studies demonstrate that social media weaponization operates through interconnected mechanisms: network polarization, coordinated propaganda, algorithmic filtering, and automated amplification. While substantial research has examined these dimensions individually, there remains a need for integrated analyses that directly link these processes to societal unrest and weakened public order. This study seeks to contribute to that emerging discourse by synthesizing recent findings and examining the broader implications of weaponized digital communication for social stability.

OBJECTIVES OF THE STUDY

The primary objectives of this research are to examine the concept and mechanisms of social media weaponization, to analyze its impact on social stability and public order, to assess the legal and regulatory challenges in controlling digital unrest, and to propose measures for mitigating the harmful effects of weaponized social media.

RESEARCH METHODOLOGY

In order to investigate the matter in a wide-ranging and methodical fashion, this study makes use of a research methodology that is both qualitative and analytical. Articles from reputable journals, books with a lot of scholarly clout, official government publications, rulings from courts, policy papers, and news items all constitute secondary sources of data used in the research. Using these resources, the researcher hopes to build a solid foundational knowledge of the subject's philosophical, constitutional-regulatory frameworks, and social aspects.

The research is doctrinal in nature, focusing on the critical examination and interpretation of existing constitutional-regulatory frameworks, policy frameworks, and scholarly debates. It emphasizes conceptual clarity and analytical evaluation of how legal and social developments have shaped the discourse surrounding the weaponization of social media. Relevant case studies from India, along with selected international examples, are incorporated to provide contextual depth and comparative insight.

The study does not include empirical surveys, interviews, or statistical modeling. Instead, it relies on descriptive and analytical methods to interpret existing literature and documented instances, thereby enabling a structured and critical assessment of the issue under consideration.

MECHANISMS OF SOCIAL MEDIA WEAPONIZATION

Social media platforms are becoming increasingly weaponized as a result of coordinated digital operations that aim to manipulate information flows, magnify narratives that divide people, and impact public opinion. Making up stories and spreading them quickly is one of the main ways it works. In order to increase the possibility of false or inaccurate material being circulated extensively, it is often designed to elicit emotional reactions. As Soroush Vosoughi (2018)[13] showed, the novelty and emotional effect of false news cause it to spread more quickly and reach more individuals than true information. In a similar vein, Shu et al. (2017)[14]

detailed how social media platforms' inherent characteristics, like network connectivity and peer-to-peer sharing, make it easier for false information to spread rapidly before it can be adequately dispelled.

Algorithmic amplification further intensifies this process. Social media algorithms prioritize engagement-driven content, thereby promoting posts that generate strong reactions. Gillespie (2014)[15] noted that algorithms function as gatekeepers of visibility, shaping public discourse through opaque ranking systems. As a result, sensational or polarizing content often receives disproportionate exposure, reinforcing ideological divides.

The use of social bots and automated accounts is another significant mechanism. Shao et al. (2018)[16] found that bots play a central role in spreading low-credibility content by artificially amplifying its visibility. Frame and Brachotte (2018)[17] also observed that social bots influenced political communication during electoral campaigns by engineering narratives of victory or defeat. Bradshaw and Howard (2019)[18] documented the global rise of organized social media manipulation, highlighting how coordinated digital campaigns are employed to interfere in democratic processes and destabilize public trust.

Echo chambers and filter bubbles further strengthen polarization. Quattrociocchi, Scala, and Sunstein (2016)[19] demonstrated that online communities tend to cluster around shared beliefs, limiting exposure to opposing viewpoints. Cinelli et al. (2021)[20] confirmed that such echo chamber effects are measurable across major platforms and contribute to ideological segregation. Sunstein (2017) as discussed by Aulisio (2018)[21] argued that these fragmented digital environments undermine deliberative democracy by reinforcing group-based divisions.

Collectively, the spread of misinformation, algorithmic prioritization, bot-driven amplification, coordinated disinformation campaigns, and echo chamber dynamics transform social media into a powerful instrument capable of influencing collective behavior and generating social unrest.

IMPACT ON SOCIAL STABILITY AND PUBLIC ORDER

The weaponization of social media has produced measurable disruptions in social stability and public order by reshaping how information influences collective behavior. One of the most serious consequences is the translation of online misinformation into offline conflict. Digitally circulated rumors and manipulated narratives can rapidly mobilize large groups, particularly in emotionally sensitive contexts involving religion, ethnicity, or national identity. Allcott and

Gentzkow (2017)[22] argue that exposure to false political information significantly affects public perceptions, especially when users lack mechanisms for verification. Such distortions in perception can lead to reactive crowd behavior, protest escalation, and, in extreme cases, violence.

The strategic use of computational propaganda further intensifies political instability. Woolley and Howard (2016)[23] describe how automated political communication tools are employed to shape narratives, suppress opposition voices, and artificially magnify ideological support. These coordinated campaigns distort democratic discourse by creating false signals of consensus or crisis. Tucker et al. (2018)[24] emphasize that social media platforms can both facilitate civic engagement and simultaneously undermine democratic accountability when manipulated for partisan objectives. During electoral periods, the circulation of misleading content reduces the quality of public deliberation and increases distrust in electoral outcomes.

Another significant impact is the erosion of institutional trust. Repeated exposure to conspiracy narratives and anti-institutional propaganda contributes to declining confidence in governance, media, and law enforcement agencies. Lewandowsky, Ecker, and Cook (2017)[25] explain that misinformation, once internalized, becomes resistant to correction and can shape long-term belief systems. This persistence of false beliefs fuels skepticism toward official clarifications, making crisis management more difficult for authorities.

The social fragmentation resulting from polarized digital interactions also weakens communal harmony. Bail et al. (2018)[26] found that exposure to opposing political views on social media can sometimes intensify polarization rather than reduce it, particularly when interactions are adversarial. Such dynamics foster identity-based divisions and normalize hostile discourse. Over time, this environment increases the probability that online disputes will spill into offline confrontations.

In addition to political effects, the psychological impact of sustained exposure to inflammatory digital content contributes to collective anxiety and social tension. Recurrent engagement with alarming or sensational material heightens perceptions of threat and insecurity. This atmosphere of perceived instability can trigger reactive mobilization, crowd unrest, and moral panic.

Overall, the weaponization of social media disrupts public order not merely through isolated incidents but through sustained structural influence on political trust, group relations, and

collective emotions. The cumulative effect is a weakened social fabric in which misinformation, polarization, and distrust reinforce one another, creating fertile conditions for unrest.

GOVERNANCE DILEMMAS IN CONTROLLING DIGITALLY ENGINEERED UNREST

The regulation of weaponized social media is no longer confined to questions of censorship or free speech; it has evolved into a broader governance challenge involving platform architecture, state capacity, cross-border coordination, and institutional accountability. Unlike traditional media, digital platforms operate through privately designed algorithmic systems that determine visibility, engagement, and narrative amplification. This creates a regulatory paradox: while the harms manifest publicly, the mechanisms that produce them remain largely proprietary and opaque.

Structural Limits of State-Centric Regulation

Conventional regulatory approaches assume that harmful speech can be addressed through penal provisions, blocking orders, or post-facto prosecution. However, digital propaganda campaigns are decentralized, rapid, and often anonymous. By the time authorities intervene, misinformation may already have produced offline consequences. This reactive model of enforcement exposes a structural weakness: legal remedies operate slower than algorithmic virality.

Further, state-led control mechanisms risk over-dependence on executive discretion. When regulatory authority is concentrated within administrative agencies, concerns arise regarding neutrality, selective enforcement, and political misuse. In highly polarized environments, regulatory interventions may themselves become contested, thereby intensifying rather than resolving public distrust.

Platform Governance and Private Power

A distinctive feature of digital ecosystems is that governance is partially privatized. Social media companies design community standards, moderation rules, and automated detection systems. This shifts significant normative power to private corporations that are neither democratically elected nor fully transparent. Decisions regarding what constitutes "harmful,"

"misleading," or "coordinated" content are frequently shaped by internal policies rather than publicly debated standards.

This hybrid model—where state law coexists with corporate rule-making—creates accountability gaps. When content is removed, users may perceive censorship. When harmful content remains, platforms are accused of negligence. The absence of clear harmonization between public law principles and platform governance frameworks complicates consistent enforcement.

Jurisdictional Fragmentation and Cross-Border Operations

Digital influence operations frequently transcend national boundaries. Coordinated disinformation campaigns may originate outside domestic territory, utilizing proxy accounts, virtual private networks, or bot infrastructures distributed across multiple countries. Traditional territorial jurisdiction struggles to respond effectively to such distributed architectures.

International cooperation mechanisms exist, but they remain procedurally slow and diplomatically sensitive. Moreover, regulatory standards vary significantly between jurisdictions. A post considered unlawful in one country may be protected expression in another. This divergence weakens enforcement consistency and allows malicious actors to exploit regulatory asymmetries.

Encryption, Traceability, and Democratic Safeguards

Encrypted communication systems introduce an additional layer of complexity. Encryption protects journalists, activists, and ordinary citizens from unlawful surveillance. Simultaneously, it can shield coordinated misinformation networks and incitement campaigns from detection. The regulatory question is therefore not simply whether traceability should exist, but how it can be structured without dismantling privacy protections for millions of lawful users.

Broad traceability mandates risk creating systemic vulnerabilities, including potential misuse of surveillance powers. Conversely, absolute anonymity can weaken deterrence. The governance dilemma lies in designing narrowly tailored investigative access mechanisms that operate under judicial supervision rather than generalized monitoring.

Technological Acceleration and Regulatory Lag

Weaponization techniques evolve rapidly. Deepfakes, synthetic media, micro-targeted political advertising, and automated engagement farming represent dynamic tools that outpace legislative cycles. Static legal drafting often fails to anticipate emergent harms, resulting in periodic regulatory patchwork rather than systemic reform.

This regulatory lag suggests the need for adaptable oversight models, including periodic statutory review, technology-neutral drafting, and multi-stakeholder advisory bodies capable of assessing emerging risks without over-criminalizing innovation.

The Risk of Overcorrection

While insufficient regulation enables digital manipulation, excessive intervention can suppress legitimate dissent, satire, and critical journalism. Overbroad definitions of "anti-national," "inflammatory," or "misleading" content risk chilling democratic debate. In polarized political climates, such provisions may be invoked unevenly, thereby undermining institutional credibility.

The sustainability of regulatory frameworks depends not only on their effectiveness but also on public trust. Transparent procedures, independent review mechanisms, and reasoned decision-making processes are essential to ensure that regulatory tools are not perceived as instruments of political control.

Toward a Balanced Regulatory Architecture

An effective response to weaponized social media requires moving beyond purely punitive frameworks. A balanced architecture should include:

- Clear statutory standards aligned with constitutional principles

- Transparent content moderation processes

- Independent appellate or oversight bodies

- Structured cooperation between platforms and law enforcement

- Strengthened digital literacy initiatives

- International collaboration on cross-border influence operations

The objective is not to eliminate harmful digital expression entirely—an unrealistic goal—but to reduce systemic vulnerabilities while preserving democratic freedoms.

In conclusion, regulating weaponized social media is not a binary choice between liberty and security. It is an institutional design challenge requiring proportional safeguards, accountable enforcement, and adaptive governance. Democracies must construct regulatory systems capable of addressing digital destabilization without compromising the very freedoms that define them.

FINDINGS AND SUGGESTIONS

Findings

The paper discloses that despite social media sites being initially meant to improve interconnectedness and democratic processes, the sites have structural elements that render them vulnerable to manipulation. Each engagement-focused content view is the priority of the algorithm-driven systems, and the content with a high emotional charge or polarization can be enhanced. This is structurally designed to allow misinformation and divisive content to circulate rapidly, hence putting the possibility of social instability at a greater risk.

The other notable discovery is that false information is more widely propagated compared to information that has been verified. The fake stories are usually prepared to elicit high emotional interest, e.g., fear, anger, or outrage, which makes them more shareable. Consequently, deceptive information often spreads to the masses before any correction is implemented, thus causing chaos and, in some cases, giving rise to actual effects in the real world.

Another important role in the research is realized by the coordinated digital manipulation. Organized disinformation campaigns, troll farms, as well as automated bot networks can be used to give certain narratives artificial amplification and the sense that the rest are in agreement with them. These well-acted activities misrepresent the discussion of the populace, affect political views, and enhance communal or intellectual strain.

Polarization is also supported by echo chambers and filter bubbles. Automatic personalization doesn't show a user different opinions and reinforces existing ideas. These ideological clusters over the years have led to mistrust, social segregations, and less constructive conversations

between communities. Hostile and adversarial discourse is normalized, and hence it undermines the principles of deliberative democracy.

The paper also concludes that misinformation that occurs on the internet is often converted to offline unrest. Digitally spread rumors in socially sensitive settings have been useful in fanning a mob-fueled violence, protest-filled situation, and communal riots. Rumor-spreading and mobilization have become much faster and occur on a larger scale, due to the speed and volume of digital communication. The other important conclusion has been on the undermining of institutional trust. The continued spread of conspiracy theories and anti-institutional propaganda destroys the trust of people in the systems of government, media, and police departments. This lack of trust complicates the management of the crisis and diminishes whatever is said or can be done to rectify the situation.

Lastly, it is noted that the current legal and regulatory mechanisms are mostly reactive and have a high enforcement problem. The law enforcement issues of jurisdiction, encryption applications, and the transnational nature of social media networks inhibit prompt intervention. Concurrently, the lack of digital literacy in users also leads to susceptibility to manipulation, enhancing the overall effects of weaponized information on society.

Suggestions

Based on these findings, the research proposal that would prevent these might be, as a priority measure, strengthening digital literacy. The schools and the government must encourage critical thinking, fact-checking abilities, and responsible internet usage so as to become less vulnerable to fake news. Citizenry that is digitally informed can become the initial line of protection against the manipulative content. It is also necessary to have an increased level of transparency and accountability in the processes of algorithms. The process of content ranking, recommendation, and amplification should be better explained to social media companies. Algorithms can be audited independently, and this could identify biases and systemic risks that have a role in the polarization and misinformation.

The research also suggests the empowerment of content moderation. The hybrid usage of artificial intelligence instruments with human supervision should appear on platforms to identify orchestrated disinformation, hate speech, and calls to violence. The community guidelines can be used to curb abuse by means of clear and consistent regulations that will not severely constrain the right to expression. Creating stable and constitution-acceptable

regulatory supervision institutions could improve accountability and protect freedom of speech. The regulatory mechanisms will be created to be transparent, proportionate, and reviewable in court to avoid Power abuse. Considering the nature of digital platforms as being international, closer collaboration among countries is essential. Governments ought to come up with coordinated networks of information sharing, border investigation, and synchronized response to widespread digital manipulation exercises.

The research also implies the development of quick response procedures at the time of increased tension, like election or riots. Posting harmful content can be restricted to a viral level if the system detects it early and the government acts in unison with the platform to stop it before it becomes unrest. Lastly, the paper underlines the need to have a multi-stakeholder collaboration. Governments, technologies, civil society organizations, academic institutions, and fact-checking organizations have to come together to ensure digital ecosystems are resilient. Coordinated and balanced actions can help societies to reduce the destructive impact of using social media as a weapon without interfering with democratic freedoms.

CONCLUSION

The weaponization of social media has become a burning issue in the modern digital age that has a significant impact on social peace, democratic procedures, and order in the streets. As much as social media has empowered turnout and information accessibility, the algorithm-based architecture of these sites, combined with personalization and fast content spread, has extended to facilitate misinformation, organization propaganda, and polarizing discourses. The paper will prove that the manipulated content through digital tools may create extremities in the political sphere, undermine the institution of trust, and, in some delicate situations, may lead to an actual state of turmoil and violence. The legal and regulatory mechanisms, despite continued developments, are still struggling with jurisdictional, technological, and constitutional factors. To deal with this problem, a multi-dimensional strategy of harmonizing digital literacy, open-air platform management, balanced regulation, and international collaboration will have to be taken into consideration. A robust digital ecosystem can only be created through the joint efforts of governments, technology organizations, civil society, and people to protect democratic values without encouraging the possibilities of weaponizing social media.

References

1. Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of economic perspectives*, *31*(2), 211-236.DOI: 10.1257/jep.31.2.211

2. Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., ... & Nyhan, B. (2018). Social media, political polarization, and political disinformation: A review of the scientific literature. *Political polarization, and political disinformation: a review of the scientific literature (March 19, 2018)*. http://dx.doi.org/10.2139/ssrn.3144139

3. Aral, S. (2020). *The hype machine: How social media disrupts our elections, our economy, and our health—and how we must adapt*. Currency.

4. Persily, N., & Tucker, J. A. (Eds.). (2020). *Social Media and Democracy*. Cambridge: Cambridge University Press.

5. Surjatmodjo, D., Unde, A. A., Cangara, H., & Sonni, A. F. (2024). Information pandemic: a critical review of disinformation spread on social media and its implications for state resilience. *Social Sciences*, *13*(8), 418. https://doi.org/10.3390/socsci13080418

6. Azzimonti, M., & Fernandes, M. (2023). Social media networks, fake news, and polarization. *European journal of political economy*, *76*, 102256.

7. Brezatis, I. (2023). *Weaponization of social media in conflict-The high level of coordination and integration between those engaged in physical battlefield warfare and those engaged in social media operations* (Master's thesis, Πανεπιστήμιο Πειραιώς). http://dx.doi.org/10.26267/unipi_dione/3515

8. Rodič, B. (2025). Social Media Bot Detection Research: Review of Literature. *arXiv preprint arXiv:2503.22838*. https://doi.org/10.48550/arXiv.2503.22838

9. Marigliano, R., Ng, L. H. X., & Carley, K. M. (2024). Analyzing digital propaganda and conflict rhetoric: a study on Russia's bot-driven campaigns and counter-narratives during the Ukraine crisis. *Social Network Analysis and Mining*, *14*(1), 170.

10. Rodilosso, E. Filter Bubbles and the Unfeeling: How AI for Social Media Can Foster Extremism and Polarization. *Philos. Technol.* 37, 71 (2024). https://doi.org/10.1007/s13347-024-00758-4

11. Fahad, A., & Mustafa, S. E. (2025). Locked in echoes: unveiling the dynamics of social media echo chambers and Hindu radicalization targeting Muslim youth in Delhi. *Humanities and Social Sciences Communications*, *12*(1).https://doi.org/10.1057/s41599-025-04638-w

12. Xu, G., Qian, M., & Meng, L. (2025). Misinformation dissemination on social media: key research themes and evolutionary paths between 2013 and 2023. *Humanities and Social Sciences Communications*, *12*(1), 1775. https://doi.org/10.1057/s41599-025-06067-1

13. Soroush Vosoughi *(2018),* The spread of true and false news online.*Science***359**,1146-1151(2018).DOI:10.1126/science.aap9559

14. Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, *19*(1), 22-36. https://doi.org/10.1145/3137597.313760

15. Gillespie, T. (2014). The relevance of algorithms. *Media technologies: Essays on communication, materiality, and society*, *167*(2014), 167.

16. Shao, C., Ciampaglia, G.L., Varol, O. *et al.* The spread of low-credibility content by social bots. *Nat Commun* **9**, 4787 (2018). https://doi.org/10.1038/s41467-018-06930-7

17. Frame, A., & Brachotte, G. (2018, June). Engineering victory and defeat: the role of social bots on Twitter during the French PresidentialElections. In *Comparing two outsiders' 2016-17 wins: Trump & Macron's campaigns*.

18. Bradshaw, S., & Howard, P. N. (2019). The global disinformation order: 2019 global inventory of organised social media manipulation. Oxford Internet Institute, University of Oxford, https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/09/CyberTroop-Report19.pdf

19. Quattrociocchi, W., Scala, A., & Sunstein, C. R. (2016). Echo chambers on Facebook. *Available at SSRN 2795110.*

20. Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the national academy of sciences*, *118*(9), e2023301118. https://doi.org/10.1073/pnas.2023301118

21.   Aulisio, G. J. (2018). # republic: Divided Democracy in the Age of Social Media. https://doi.org/10.1080/10848770.2018.1449394

22.   Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of economic perspectives*, *31*(2), 211-236.DOI: 10.1257/jep.31.2.211

23.   Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of economic perspectives*, *31*(2), 211-236.DOI: 10.1257/jep.31.2.211

24.   Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., ... & Nyhan, B. (2018). Social media, political polarization, and political disinformation: A review of the scientific literature. *Political polarization, and political disinformation: a review of the scientific literature (March 19, 2018)*. http://dx.doi.org/10.2139/ssrn.3144139

25.   Lewandowsky, S., Ecker, U. K., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the "post-truth" era. *Journal of applied research in memory and cognition*, *6*(4), 353-369.https://doi.org/10.1016/j.jarmac.2017.07.008

26.   Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. F., ... & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, *115*(37), 9216-9221. https://doi.org/10.1073/pnas.1804840115