AI-DRIVEN DISINFORMATION AS A GLOBAL CYBERSECURITY THREAT TO DEMOCRATIC SYSTEMS

Murat EMEÇ[*]
ORCID ID: 0000-0002-9407-1728

Declaration[*]

Abstract

The proliferation of generative artificial intelligence has fundamentally transformed our current information landscape, facilitating the widespread production and consumption of photorealistic yet fictitious media. So, misinformation is no longer just a communication issue; it is an absolute cybersecurity threat, especially for democracies that depend on trust, transparency, and an informed public. This paper views AI-fueled disinformation as a cognitive cyber-attack aimed at modifying perceptions, shaping belief formation, and undermining the legitimacy of institutions, rather than solely targeting technical systems. Based on literature in cybersecurity, political communication, and AI governance, the research investigates how generative AI augments disinformation, which systemic failures in democracies it exploits, and why current interventions fall short. The results, he says, are that AI-supported disinformation damages public trust, increases social and political fragmentation, and distorts electoral and governmental processes in ways that are often hard to detect and even harder to put right. The research suggests expanding existing cybersecurity strategies to protect democracies in the age of generative AI by considering not only information integrity and cognitive security but also societal resilience to disinformation and propaganda, as well as technical protections.

Keywords: AI-driven disinformation, cybersecurity, democratic systems, cognitive security, generative artificial intelligence

[*] Computer Science Application and Research Centre, Istanbul University, murat.emec@istanbul.edu.tr
[*] AI tools were utilized to enhance linguistic clarity, refine structure, and support the overall organization of the manuscript. All conceptual insights, analytical interpretations, and final conclusions were independently developed and verified by the author. All analysis and conclusions are solely the author's responsibility.

Introduction

The accelerating advance of artificial intelligence has decidedly altered the world's information landscape, redefining how it is produced, distributed, and understood. Whilst AI-based technology offers great promise in automation, communication, and decision-making, it also introduces new and pervasive risks to the integrity of the informational fabric of digital spaces. Chief among these risks is AI-fueled disinformation, one of the most disruptive and potentially damaging threats, especially to democratic systems that require informed citizens, transparent institutions, and trust-based governance.

The difference between AI-fueled disinformation and previous varieties of misinformation is the former's scale and its ability to adapt in real time — like Russia's so-called firehose of falsehood and, eventually, its personalised pitch. It is not just you or me, however: Big language models are now able to spew forth coherent, contextually relevant narratives in the blink of an eye. Meanwhile, deepfake technology means even experts can fall for unnatural-looking, sounding clips. These technologies, when used alongside automated bot networks and algorithmic pushes on social media, enable propagandists to influence operations orders of magnitude faster, more effectively, and more widely than in the past. As a result, disinformation has evolved from an opportunistic informational weapon to a tool in an organised, technology-enabled attack strategy. From a cybersecurity perspective, AI-generated disinformation represents a new level of attack that democratic societies must oppose. Conventional cybersecurity principles focus on protecting networks, including data and critical infrastructure. By contrast, AI disinformation attacks these layers at the cognitive level: it hones perceptions, alters belief systems, and plays with emotions. Such campaigns can undermine trust, skew public discourse, and sway electoral results without ever hacking any technical system. This change requires a new definition of cybersecurity that includes cognitive security alongside traditional technical defences. The cognitive threats are specifically pernicious to democratic systems, given their open and pluralistic nature – largely dependent on free information flow. Elections, media ecosystems and public deliberation are at risk when disinformation can organise across social divisions, identity-based tensions and political polarisation. These threats are further exacerbated by AI, which can facilitate customised messaging to specific groups based on their views, demographic profiles, or psychological triggers. It is therefore not just a matter of pockets of untruths, but of the erosion over time of democratic legitimacy and trust.

The world-spanning, border-blurring design of digital platforms only exacerbates the geopolitical dimensions of AI-generated disinformation. Influence operations can occur beyond national borders, making the problem of attribution and responsibility even more complicated. AI can be exploited by state and non-state actors to meddle in other nations' politics, engage in information warfare, or cause societies to unravel without ever pushing a key on their keyboards through traditional cyberattacks. This intersection of cybersecurity, information warfare, and hybrid threats has significant implications for traditional governance and defence arrangements. While policymakers and academics are increasingly assuming AI-based disinformation as a serious social threat, they also find that this phenomenon is often mishandled in cybersecurity discussions. There is a place for technical tools — or detection algorithms and content moderation systems, but they cannot solve the problem on their own. Effective responses require a combination of technological protections, regulatory efforts, institutional coordination, and societal resilience. Considering AI-based disinformation a cybersecurity problem provides a consistent framework for compiling these components. Against this background, the current work identifies AI-generated disinformation as a key security threat to democracy. It synthesises the best available knowledge, explains how AI-enabled disinformation attacks on democracy work, and identifies traditional cybersecurity strategies. By focusing on cognitive security and democratic resilience, the study seeks to contribute to the ongoing debate about how democracies can defend themselves in an age of generative AI.

Literature Review

With the advent of AI and content-generation tools, there is a new form of disinformation. Previous studies on digital media and democratic life highlighted that algorithmic systems curate what people see online, shape their political behaviour at scale, and, in turn, influence citizen participation in democratic processes (Lorenz-Spreen, 2022; Vaccari, 2020). AI-powered disinformation is distinctive because, unlike traditional misinformation, this technology is characterised by automation that seamlessly adapts to craft personalised messages that can reach millions. Recent research has characterised AI-facilitated disinformation as a hybrid socio-technical phenomenon that emerges from the interplay between algorithmic behaviours, platform infrastructure and human biases. Computational studies confirm that generative models have the potential to generate continuous narratives that are coherent over time and aligned with a specific ideology, and to adapt to an audience's attention rather than relying on occasional falsehoods (Romănishyn, 2025; Saeidnia, 2025).

This understanding has led scholars to regard disinformation not as an isolated anomaly but as a structural, ongoing risk. Generative AI techniques like large language models and synthetic media systems have made it much less resource-intensive to create realistic-looking fake content. A study shows that AI-generated text can closely replicate human writing style, rhetorical devices, and emotional tone (Drolsbach, 2025; Olanipekun, 2025). This has enabled bad actors to astroturf convincing but false narratives on a scale never before seen on digital platforms. The emergence of deepfake audio and video technology heightens the risks. Research on deepfakes has also shown that these recordings can even provoke disbelief among citizens who regard themselves as digitally literate, particularly in imitations of political figures, journalists, or public institutions (Ratnawita, 2015; Kaczmarek, 2015). As a result, the distinction between real and fake facts becomes fuzzier, eroding the credibility of visual and audio evidence in democratic discourse. An increasing number of authors now cast AI-powered disinformation as a cybersecurity problem, even though it doesn't break into technical systems. Instead of relying on hardware or software weaknesses, AI-powered disinformation attacks the human/cognitive layer: by influencing perceptions, trust, and decision-making (Mazurczyk, 2023; Mirzoyan, 2023). This reconceptualisation puts disinformation in line with the broader cyber threat model, which acknowledges that psychological and social manipulation is part and parcel of these types of threats. According to these scholars, such cognitive attacks may have strategic effects similar to those of conventional cyber operations, undermining political institutions or destroying social unity (Kreps, 2023; Singh, 2025). It is hard to attribute their mysterious origins, reflecting legal and jurisdictional grey areas in which cyber-extortion artists can operate. This imbalance, therefore, contributes to the appeal of AI-generated disinformation in hybrid and information warfare campaigns. Numerous empirical studies have demonstrated the negative impacts of AI-powered disinformation on democratic processes. Elections are especially vulnerable: there is evidence that AI narratives can alter voters' perceptions, suppress turnout, and undermine electoral legitimacy (Bennett, 2025; Yilmaz, 2024). More generally, continued exposure to disinformation increases polarisation and undermines the coherence of the public sphere. At the institutional level, AI-generated misinformation erodes confidence in journalism, scientific authority, and democratic institutions. When individuals can no longer trust the veracity of information, public deliberation diminishes, and trust in public institutions wanes (Marsden, 2021; Pawelec, 2022). Longitudinal studies indicate that such loss of trust is cumulative, eroding democratic resilience (Schipper, 2025; Wong, 2025). Increased attention to regulatory and governance responses has been driven by growing

awareness of these risks. Comparative studies reveal diverse efforts – from platform-led campaigns to state-driven approaches to cybersecurity strategies that incorporate information integrity into the national security agenda (Shoaib, 2023; Zhou, 2023). However, academics say such interventions also have to strike a careful balance between protecting democracy and fundamental rights, including the freedom of expression. Policy analysis is also illuminating the limitations of purely technical remedies. Content discovery tools may lag what the most sophisticated generative AI technologies can do (MacDonald, 2025; Schroeder, 2025). Therefore, scholars are increasingly promoting a multi-level counter-strategy that introduces technical measures combined with institutional collaboration and social resilience. Education, digital skills and cross-border cooperation are often mentioned as key elements of long-term defence mechanisms (Bontridder, 2021; Corsi, 2024; Imam, 2025).

Combined, this body of work illustrates that disinformation powered by AI poses a serious challenge to democratic societies, encompassing not only the provision of misinformation but also new cognitive and cybersecurity threats. Despite important strides in charting the contours of its mechanisms and effects, significant shortfalls persist — most notably in consolidating insights from the cybersecurity literature, governance science, and democratic theory into a comprehensive framework (Aditya, 2025; Sophia, 2025). These gaps need to be addressed by integrating interdisciplinary efforts to adapt to an evolving, increasingly intricate threat (Sambur, 2025; Wahab, 2025).

Methodology

This research is based on a qualitative, conceptual, analytical, and empirical design to investigate AI-generated disinformation as a global cyber threat to democracies. Given that no single empirical dataset captures all the dimensions of extent, pace, and complexity of AI-facilitated disinformation, a conceptual approach is warranted and legitimate. It allows insights from cybersecurity, political science, media studies, and AI governance to be systematically incorporated.

The methodological approach is based on the premise that AI-facilitated disinformation is hybrid, spanning technological, cognitive, and institutional dimensions. Rather than testing a particular hypothesis, the study's objective is to identify standard mechanisms, structural weaknesses, and cross-contextual structures that recur across different democratic contexts (Lorenz-Spreen & Endres, 2022; Mazurczyk et al., 2023). This design facilitates theory-building and policy-relevant analysis, which are both central to the study's objectives.

The insights are based solely on peer-reviewed academic and policy papers and interdisciplinary studies, referenced in the endnotes. The study uses a controlled corpus to maintain concept consistency and reduce the risk of arbitrary selection of data sources.

The literature conveys various disciplinary viewpoints, including:

- Artificial intelligence and generative models
- Cybersecurity and information warfare
- Democratic procedures and political communication
- Regulation of platforms and digital policy

It specifically prioritised references to AI-generated content, automation, synthetic media, and algorithmic amplification in democratic environments (Vaccari, 2020; Shoaib, 2023; Bennett et al., 2025).

The research uses thematic content analysis and cybersecurity-focused threat-mapping. Each publication was read and coded using predetermined analytical categories, developed in the literature and adjusted throughout the coding process.

The analysis was conducted in 3 stages:

1. First Coding: Repeated themes on AI capabilities, disinformation methods and democratic effects.
2. Thematic Clustering: We grouped codes that seemed similar to one another into overarching thematic categories resembling cybersecurity threat models.
3. Integrated Interpretation: Combining themes to analyse how AI-based disinformation operates as a systemic cybersecurity risk.

This allows for an overview of studies in a systematic comparison and of context-specific variation (Mirzoyan, 2023; Singh, 2025). To bridge disinformation research with cybersecurity theory, the work maps classical threat-modelling constructs, attack vectors, targets, vulnerabilities and impacts to AI-powered disinformation. Democracies are seen as complex systems in which human cognition is a key security layer.

Table 1. Cybersecurity-Oriented Threat Mapping Framework

Threat Component	Description	Application to AI-Driven Disinformation
Threat Actors	State and non-state entities	Political groups, foreign actors, coordinated networks
Attack Vectors	Means of attack	AI-generated text, deepfake audio/video, automated bots
Targets	Assets under threat	Voters, public trust, and democratic institutions
Vulnerabilities	System weaknesses	Cognitive biases, polarisation, platform algorithms
Impacts	Consequences	Electoral interference, trust erosion, and democratic instability

This mapping illustrates the similarities between AI-enabled disinformation and established cyber threats, strengthening the claim that information integrity must be recognised as a core cybersecurity asset (Kreps, 2023; Zhou, 2023).

Based on the thematic analysis, three analytical dimensions were consistently applied across all sources.

Table 2. Core Analytical Dimensions Used in the Study

Dimension	Focus	Analytical Purpose
Technological	Generative AI, automation, synthetic media	Identify enabling capabilities
Cognitive–Cyber	Perception, trust, and belief manipulation	Assess cognitive attack mechanisms
Democratic–Institutional	Elections, governance, public discourse	Evaluate systemic impacts on democratic systems

These dimensions facilitated a systematic review of each publication, which, in turn, enabled the examination to capture micro-level levers and consider how they contributed to broader systemic impact.

To increase analytical validity, this study adopts a triangulation approach across various academic disciplines and publication types. A result is only highlighted when it is consistently observed across studies. Explicit coding criteria and analytical dimensions provide reliability through clear definitions, allowing for the replication of this study in future research.

The qualitative nature of the study means that transferability, but not statistical generalisation, is a feature, yet conceptually straightforward and explanatorily deep analyses are necessary to engage with fast-moving threats (Schipper, 2025; Wong, 2025).

The study acknowledges several limitations. It is based on secondary work and does not include experimental or large-scale empirical evidence. Moreover, secondly, progress in generative AI technologies often moves so rapidly that some of the particular tools mentioned may themselves be outdated. However, by examining underlying mechanisms and structural patterns, the results remain relevant in the long term (MacDonald, 2025; Schroeder, 2025).

Combining cybersecurity threat modelling with democratic theory and AI governance research, this methodological approach provides a new lens for studying AI-mediated disinformation. It broadens a traditional cybersecurity architecture to include the realms of cognition and institution, supporting both further empirical work and future policy development and analysis.

Discussion

This article contributes to and extends this work by highlighting that AI-enabled disinformation should not be conceptualised merely as a communication challenge, but rather as a core cybersecurity threat to democratic societies. Results postulate that generative AI-powered disinformation has become a scalable, adaptable, and asymmetric attack surface at the cognitive layer of socio-technical systems, rather than at their technical infrastructure (Mazurczyk, 2023; Mirzoyan, 2023).

This reframing has profound implications for cybersecurity theory. The traditional cybersecurity model is focused on protecting the confidentiality, integrity, and availability of digital assets, but has proven insufficient at preventing threats that undermine cognitive trust or information integrity. AI-generated disinformation plays on just these kinds of gaps. The debate thus complements recent calls elsewhere to broaden cybersecurity regimes to encompass cognitive and information security, especially in democratic settings that require public trust (Kreps, 2023; Singh, 2025).

One of the most important observations from the analysis is a structural asymmetry between attackers and defenders. Generative AI dramatically lowers the threshold for influence operations by enabling even small groups and non-state actors to wield political influence on a level hitherto reachable only by states (Drolsbach, 2025; Olanipekun, 2025). Democratic institutions, on the other hand, are bound by law and ethics to react at the same tempo and with the same immediacy.

This imbalance is similar to that in other cybersecurity areas, but is amplified by the difficulty of attribution and jurisdictional complications. “The whole idea of disinformation is it is like a virus,” he said. These perceptions alone, in the absence of any technical violation, can undermine democratic legitimacy (Bennett, 2025; Yilmaz, 2024).

The conversation also underscores that, though machine-generated disinformation accumulates significant resistance over the long term, it undermines democratic resilience. Perpetual exposure to synthetic content not only produces misinformed individuals but also undermines trust in all sources of information, in general, real and authoritative (Vaccari, 2020; Marsden, 2021). That slow erosion of epistemic trust is especially hard to recover once it is established.

Table 3. Key Discussion Themes and Cybersecurity Implications

Theme	Key Insight	Cybersecurity Implication
Cognitive attacks	Disinformation targets perception and trust	Extend threat models beyond technical infrastructure
Asymmetry	AI enables low-cost, high-impact operations	Attackers gain a strategic advantage
Trust erosion	Long-term democratic harm	Trust becomes a central security asset
Platform amplification	Algorithms intensify the spread	Requires shared governance responsibility

From an ecosystem perspective, trust is a fundamental system precondition for democratic governance. Further, it undermines accountability and democracy, reduces citizen participation, and deepens political polarisation. The results are consistent with the literature, which argues that disinformation is an emerging cause of democratic fatigue (Schipper, 2025) and reduced political participation (Wong, 2025).

Another layer of the debate has revolved around the relationship between AI-generated disinformation and platforms' algorithms. Engagement-driven recommendation systems tend to spread emotionally heightened or divisive synthetic narratives (Lorenz-Spreen, 2022; Jaidka, 2024). Integrated into generative AI, these mechanisms generate self-reinforcing feedback loops that amplify the visibility and spread of harmful content. These issues bring to the fore acute questions of platform accountability and government regulation. While platforms have the technical power to reduce amplification, commercial motivations and the international scope of their business make enforcement challenging. The conversation supports a collaborative (platforms-regulators-civil society) shared-risk approach to information integrity (Zhou, 2023; MacDonald, 2025).

It also highlights the folly of relying solely on technical rectification. Detection mechanisms, watermarking systems, and automated moderation are necessary but can be easily bypassed and do not address the underlying social structures that sustain disinformation (Kaczmarek, 2025; Schroeder, 2025). Heavy reliance on automated moderation can also result in (over)blocking incidents and a risk of suppressing freedom of expression (Bontridder, 2021; Marsden, 2021). Given these constraints, the multi-layered nature of defence strategies that fuse technological detractors with institutional coordination and social policies is a necessity. Education, public enlightenment and media literacy stand out as crucial long-term factors contributing to the resilience of democracy (Corsi, 2024; Aditya, 2025).

The exchange also highlights important implications for cybersecurity policy and governance. Integrating AI-fueled disinformation into national cybersecurity strategy can improve readiness to address hybrid threats and foster coherence among government institutions (Shoaib, 2023; Imam, 2025). That kind of integration would better link election security, platform governance and information integrity efforts. Crucially, the report also argues that democratic resilience should be seen as a security objective in its own terms. This entails a shift from reactive tactics to proactive initiatives, such as capacity-building, international cooperation, and norm-building, aimed at maintaining trust and transparency in digital public spheres.

Table 4. Comparative Assessment of Countermeasures

Countermeasure Type	Strengths	Limitations
Technical detection	Scalable and automated	Easily bypassed; risk of errors
Platform moderation	Allows rapid intervention	Incentive misalignment; limited transparency
Regulation	Provides accountability and deterrence	Jurisdictional constraints
Societal resilience	Builds long-term capacity	Slow to develop; resource-intensive

The overall debate highlights that AI-enabled disinformation is a multidimensional cybersecurity threat, with effects that extend well beyond proximate political events and contribute to the erosion of the structural foundations of liberal democracy. Dealing with this threat will require rethinking cybersecurity in ways that include cognitive and institutional dimensions, informed by coherent policy action supported by technological innovation and societal-wide engagement.

Conclusion

AI-fueled disinformation has become a worldwide cybersecurity threat that strikes at the heart of democratic systems: public trust and decision-making. In contrast to traditional cyber-attacks, which focus on digital infrastructure, AI-driven influence operations target the cognitive and social foundations of democracy by amplifying persuasive falsehoods, generating counterfeit credibility, and hastening the proliferation of manipulative narratives across the net. This pivot widens the democratic attack surface, from networks and systems to perceptions, legitimacy and institutional credibility.

The argument developed here is that the strategic threat posed by AI-generated disinformation is much deeper than merely convincing citizens of particular falsehoods. Its damage is greater in eroding the foundations of democratic governance itself — common understandings, accountability, and informed participation. This legitimacy crisis — which does not disappear when false frames are eventually debunked — trains us to internalise artificial and conflicting information, normalise uncertainty, increase division amongst us, and wane confidence in the

media, elections, and public institutions. In that sense, the greatest danger isn't any one fake news story but the slow disintegration of epistemic stability itself.

The dynamics of these realities have a significant bearing on their cybersecurity strategy. Ensuring democratic integrity, in sum, means going beyond the narrow life-and-death technicality of security toward a broader approach that also encompasses information integrity, institutional coordination, and social robustness. This means that mitigation needs to work across several layers: improving transparency and accountability by online platforms, enhancing the rapid-response and strategic communication capabilities of public institutions, and investing in sustainable civic resilience through education and media literacy. These measures should, of course, be conceived as safeguards for democracy, not as undermining factors that control too much, take decisions behind closed doors, or impose unjustified limitations on freedom of expression.

In sum, the challenge to democracy posed by generative AI is not just a matter of protecting technical infrastructure but of defending democratic reality. Policymakers, cybersecurity experts, digital platforms, and civil society need to consider AI-powered disinformation as a significant security threat that requires coordinated, flexible, rights-respecting responses that adapt to the changing threat environment.

References

Aditya, S. (2025). The misinformation epidemic: combating AI-generated fake content and deepfakes. World Journal of Advanced Research and Reviews. https://doi.org/10.30574/wjarr.2025.26.2.1752

Bennett, W. L. & Livingston, S. (2025). Platforms, Politics, and the Crisis of Democracy: Connective Action and the Rise of Illiberalism. Perspectives on Politics. https://doi.org/10.1017/s1537592724002123

Bienvenue, E. (2020). Computational propaganda: political parties, politicians, and political manipulation on social media. International Affairs. https://doi.org/10.1093/ia/iiaa018

Bontridder, N. & Poullet, Y. (2021). The role of artificial intelligence in disinformation. Data & Policy. https://doi.org/10.1017/dap.2021.20

Corsi, G., Marino, B., & Wong, W. (2024). The spread of synthetic media on X. Harvard Kennedy School Misinformation Review. https://doi.org/10.37016/mr-2020-140

Demartini, G., Mizzaro, S., & Spina, D. (2020). Human-in-the-loop Artificial Intelligence for Fighting Online Misinformation: Challenges and Opportunities. IEEE Data Eng. https://consensus.app/papers/humanintheloop-artificial-intelligence-for-fighting-demartini-mizzaro/6fe75dfbff

Drolsbach, C. & Pröllochs, N. (2025). Characterising AI-Generated Misinformation on Social Media. ArXiv. https://doi.org/10.48550/arxiv.2505.10266

Imam, M. (2025). The Threats of AI and Disinformation in Times of Global Crises. Bulletin of Islamic Research. https://doi.org/10.69526/bir.v3i4.394

Jaidka, K., Chen, T., Chesterman, S., Hsu, W., Kan, M., Kankanhalli, M., Lee, M., Seres, G., Sim, T., Taeihagh, A., Tung, A., Xiao, X., & Yue, A. (2024). Misinformation, Disinformation, and Generative AI: Implications for Perception and Policy. Digital Government: Research and Practice.

Kaczmarek, K., Karpiuk, M., & Melchior, C. (2025). Disinformation as a Threat to State Security. Przegląd Nauk o Obronności. https://doi.org/10.37055/pno/205780

Kreps, S. E. & Kriner, D. (2023). How AI Threatens Democracy. Journal of Democracy. https://doi.org/10.1353/jod.2023.a907693

Lorenz-Spreen, P., Oswald, L., Lewandowsky, S., & Hertwig, R. (2022). A systematic review of worldwide causal and correlational evidence on digital media and democracy. Nature Human Behaviour. https://doi.org/10.1038/s41562-022-01460-1

MacDonald, E. (2025). Digital Authoritarianism and the Erosion of Democratic Norms: A Comparative Study of State Surveillance in Hybrid Regimes. OTS Canadian Journal. https://doi.org/10.58840/2an15325

Marsden, C., Brown, I., & Veale, M. (2021). Responding to Disinformation. Regulating Big Tech. https://doi.org/10.1093/oso/9780197616093.003.0012

Mazurczyk, W., Lee, D., & Vlachos, A. (2023). Disinformation 2.0 in the Age of AI: A Cybersecurity Perspective. Communications of the ACM. https://doi.org/10.1145/3624721

Miller, M. L. & Vaccari, C. (2020). Digital Threats to Democracy: Comparative Lessons and Possible Remedies. The International Journal of Press/Politics. https://doi.org/10.1177/1940161220922323

Mirzoyan, A. (2023). Digital Authoritarianism as a Modern Threat to Democratic Stability: Restriction of Freedom or Network Politicisation?. Journal of Political Science: Bulletin of Yerevan University. https://doi.org/10.46991/jops/2023.2.6.062

Olanipekun, S. O. (2025). Computational propaganda and misinformation: AI technologies as tools of media manipulation. World Journal of Advanced Research and Reviews. https://doi.org/10.30574/wjarr.2025.25.1.0131

Ratnawita, R. (2025). Cybersecurity in the AI Era Measures Deepfake Threats and Artificial Intelligence-Based Attacks. Journal of the American Institute. https://doi.org/10.71364/s3emxx77

Românishyn, A., Malytska, O., & Goncharuk, V. (2025). AI-driven disinformation: policy recommendations for democratic resilience. Frontiers in Artificial Intelligence. https://doi.org/10.3389/frai.2025.1569115

Saeidnia, H. R., Hosseini, E., Lund, B., Tehrani, M. A., Zaker, S., & Molaei, S. (2025). Artificial intelligence in the battle against disinformation and misinformation: a systematic review of challenges and approaches. https://doi.org/10.1007/s10115-024-02337-7

Sambur, B. (2025). Artificial intelligence (AI) and institutional religion. Cyberpolitik Journal, 10(19), 94–97.

Schipper, T. (2025). Disinformation by design: leveraging solutions to combat misinformation in the Philippines' 2025 election. Data & Policy. https://doi.org/10.1017/dap.2025.18

Schroeder, D. T., Cha, M., Baronchelli, A., Bostrom, N., Christakis, N., Garcia, D., Goldenberg, A., Kyrychenko, Y., Leyton-Brown, K., Lutz, N., Marcus, G., Menczer, F., Pennycook, G., Rand, D. G., Schweitzer, F., Summerfield, C., Tang, A., Bavel, J. V., Linden, S. V. D., Song, D., & Kunst

Shoaib, M. R., Wang, Z., Ahvanooey, M. T., & Zhao, J. (2023). Deepfakes, Misinformation, and Disinformation in the Era of Frontier AI, Generative AI, and Large AI Models. 2023 International Conference on Computer and Applications (ICCA). https://doi.org/10.1109/icca59364.2023.10401723

Singh, K. & Kumar, H. (2025). Futuristic Media Information Literacy to Counter AI Generative Deepfake Media Content and Its Implications. Jharkhand Journal of Development and Management Studies. https://doi.org/10.70994/jjdms.10689.10703

Sophia, L. (2025). The Social Harms of AI-Generated Fake News: Addressing Deepfake and AI Political Manipulation. Digital Society &amp";" Virtual Governance. https://doi.org/10.6914/dsvg.010105

Vaccari, C. & Chadwick, A. (2020). Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. Social media + Society. https://doi.org/10.1177/2056305120903408

Wong, W. (2025). Trends in Political Science Research: Artificial Intelligence and Voter Disinformation. International Political Science Abstracts. https://doi.org/10.1177/00208345251339324

Yılmaz, I., Akbarzadeh, S., Abbasov, N., & Bashirov, G. (2024). The Double-Edged Sword: Political Engagement on Social Media and Its Impact on Democracy Support in Authoritarian Regimes. Political Research Quarterly. https://doi.org/10.1177/10659129241305035

Zhou, J., Zhang, Y., Luo, Q., Parker, A. G., & Choudhury, M. D. (2023). Synthetic Lies: Understanding AI-Generated Misinformation and Evaluating Algorithmic and Human Solutions. Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. https://doi.org/10.1145/3544548.

Wahab, A. A. (2025). Artificial intelligence (AI) and cybersecurity. Cyberpolitik journal, 10(19), 85–93.