# Measurement and Assessment of the Impact of Disinformation on State Cyber Sovereignty

Said Mammadov
Institute of Information Technology, Baku, Azerbaijan
said.mamadov@gmail.com

**Abstract— This article is a comprehensive overview of the strategic threats of the disinformation ecosystem to national cyber sovereignty and how such strategic threats can be analytically assessed. The specifics of various sovereignty models, including the US, the EU, Russia and China, are contrasted, and the capabilities of artificial intelligence and deepfake technologies to manipulate information are explored. The article outlines a three-tiered assessment system that consists of information flow indicators, outcomes based on societal impacts and sovereignty to quantify disinformation effects. Lastly, the need to implement a multi-criteria threat assessment model combining technical, institutional and social factors to manage strategic risks is highlighted.**

***Keywords— cyber sovereignty; disinformation ecosystem; AI, deepfakes; digital governance; information warfare.***

## I. INTRODUCTİON

The political communication in the digital era has radically transformed the power, authority, and sovereignty exercised in modern international politics. The information that was previously mediated by territorially based institutions in the form of state broadcasters, print media, and diplomatic channels now flows through a decentralised, transnational, and algorithmically controlled digital platform [1]. This shift has facilitated unparalleled connection and involvement, but it has also led to the exposure of states to new forms of political intrusion, the most notable one being disinformation.

Disinformation has become a major threat to state sovereignty as it attacks the cognitive and informational base on which political power is based. Sovereignty is maintained not only by the means of territorial control or coercive power but also by the means of legitimacy, trust, and the ability to create the meaning of collective importance [2]. These weaknesses are actively exploited through disinformation campaigns that distort the language of the people, reduce trust in institutions, and encourage the polarisation of society [3]. As opposed to classical propaganda, modern disinformation is enhanced by digital systems that offer speed to spread messages, micro-targeting, and plausible deniability.

The increased use of disinformation has made states think about their approach to cyberspace. At the beginning of the 21st century, the ideas of the internet were focused on openness, decentralisation, and international regulation.

Nevertheless, an increasing number of issues related to electoral manipulation, social unrest, and national safety have led to states claiming more command over the digital space.

This change has resulted in the emergence of the notion of cyber sovereignty, according to which cyberspace is an extension of national sovereignty that can be regulated and controlled by domestic authorities [4].

This literature review explores the dynamic nature of the disinformation and cyber sovereignty relationship as one of the reasons and rationalities behind the growing state control of cyberspace. It places cyber sovereignty in the context of more general discussions about international law, security studies, and political theory and evaluates critically what the sovereignty-based approach to digital governance would entail for democracy, human rights, and world politics.

## II. DİSİNFORMATİON ECOSYSTEM

Disinformation is false or misleading information created on purpose to achieve political or strategic goals. It differs from misinformation because it has harmful intent [5]. Misinformation is unintentionally relayed false information, driven by biases, low media literacy, and algorithmic amplification. Today's campaigns rarely rely on outright lies. Instead, they use partial truths, emotional appeals, and twisted contexts. This approach makes messages believable and easy to spread [6]. The goal is to undermine knowledge and trust by flooding society with conflicting stories [7].

Wardle and Derakhshan's [5] framework adds the concept of malinformation, which is true information used in harmful ways by framing it maliciously or sharing it out of context to cause damage to someone. Malinformation makes regulation harder because using real facts to threaten political stability forces states to balance control with protecting free speech [8]. Moreover, malinformation differs from disinformation and misinformation, as it relies on facts deliberately misused.

Modern disinformation depends on how internet platforms work. Recommended systems based on algorithms place more weight on engagement, emotional appeal, and novelty, unintentionally increasing misleading and polarising content [3]. Social networks enable fast, peer-to-peer sharing. As a result, private companies and non-state actors now control key digital infrastructure and information flows. This challenges the state's usual control over public discussion [9].

## III. CYBER SOVEREİGNTY AND İTS DİMENSİONS

Cyber sovereignty means a state applies its territorial authority to the digital world. States control infrastructure, data, laws, and online spaces [10, 11, 12]. As cyberspace becomes

DOI: 10.25045/cyber.gov.2026.21

key to governance, this broad concept helps protect national interests and data. It also helps stop foreign interference.

A key part of cyber sovereignty is having physical and legal control over data and digital systems. Data localisation, which is the practice of requiring data to be stored within a country's borders, and the development of domestic technology, help governments reduce reliance on foreign providers [13]. They also apply national laws to global online platforms. Cybersecurity laws and data protection laws, which aim to defend against digital threats and safeguard personal information, set limits and hold users accountable, even though cyberspace has no borders [12, 15].

Beyond infrastructure and laws, cyber sovereignty focuses on protecting information integrity and guarding against outside influence. States regulate online content and fight disinformation to maintain political legitimacy and a stable information environment. These actions often raise debates about free speech and civil rights [12, 14]. To become more independent, governments support domestic tech industries and set national security standards. This reduces the risks of spying, misinformation, and international conflict [13].

## IV. Models of State Cyber Sovereignty in the Digital Era

The pursuit of cyber sovereignty has led to various models that balance openness, control, and state power [16]. The EU uses regulations and rights-based laws, such as the GDPR and Digital Services Act, to protect user control, regulate platforms, and set global standards [16, 17, 18]. Critics argue the EU still relies on foreign technology for digital sovereignty [19]. The US prefers a market-driven, open model with little state control. It depends on big tech companies to represent digital sovereignty. However, this leaves gaps in the management of misinformation [16].

Russia and China focus on state-led control and strict digital borders. Russia’s 'sovereign internet' model centralises digital infrastructure, meaning essential online systems and services, under government control to maintain security and regime stability, using laws to build a national network [20]. China applies the strictest digital territorial control, combining state power with regulation of cyberspace. Its tools include the Great Firewall (a nationwide system that filters and blocks foreign websites), strict data rules, and tech oversight [21]. However, both face technical challenges as citizens often find ways to bypass censorship [20, 21].

Despite their differences, all four models face cross-border disinformation. National policies fragment the internet. Yet, harmful content still crosses borders. It exploits weak EU enforcement, US openness, and centralised controls in Russia and China [16, 21]. Experts say that strict territorial control online is limited. Effective digital threat responses demand cooperation and international governance.

## V. Undermining State Cyber Sovereignty: Emerging Threats

States’ ability to maintain cyber sovereignty is weakened by the borderless nature of modern digital technologies and global platforms [11]. They depend on foreign cloud systems, data centres, and algorithms controlled by multinational companies, which exposes them to outside control and surveillance [15]. This reliance grows as tech companies operate globally and often conflict with national laws, making enforcement difficult [22]. As a result, control over user data and content shifts from governments to private firms, greatly reducing state authority over online spaces [23].

State sovereignty and information integrity also face growing threats from AI-driven manipulation and cross-border disinformation campaigns. Deepfakes and automated propaganda let malicious actors create convincing content at scale. This weakens public trust in democratic institutions and poses serious legal and ethical challenges for regulators [24, 25]. Since digital platforms have global reach and low barriers to entry, they often avoid local laws. Using coordinated bots and influencers, these actors quickly spread propaganda across borders. Resulting in disrupted politics, damaged trust in institutions, and harmed governance and national innovation [26, 27].

## VI. Dangers Of Disinformation

Disinformation is a widespread structural threat in online communication systems. Its main danger is scale: automated accounts and engagement-focused algorithms spread false stories much faster than facts by appealing to emotions [28]. This automatic boost makes disinformation powerful enough to change community discussions and elections [26]. Platform algorithms are often opaque, hiding the origins of accounts and their spread patterns. This makes accountability difficult and allows hostile actors to manipulate information largely unnoticed.

This situation creates a significant imbalance in modern information warfare: attackers need few resources to spread false stories, while governments and platforms spend huge sums to monitor, detect, and regulate them [29]. Artificial intelligence makes this worse by enabling mass production of convincing deepfakes and synthetic media that current detection tools struggle to catch, blurring truth and lies [30,31]. These factors cause serious, lasting harm to knowledge security, create echo chambers, reduce public trust in institutions, and weaken state legitimacy. As a result, disinformation is no longer just a communication problem but a direct, systemic threat to democracy and cyber sovereignty.

## VII. How Disinformation Affects State Cyber Sovereignty

Disinformation directly threatens a state’s cyber sovereignty by attacking its cognitive foundations, institutional authority, and regulatory power, turning from a simple communication problem into a deep structural threat [32, 33]. At the cognitive level, targeted misinformation and algorithm-driven amplification use emotional stories to divide societies, disrupt informed democratic participation, and weaken knowledge security [34, 35]. This interference destabilises governance, especially during critical periods such as elections or pandemics, when coordinated campaigns spread confusion and erode trust in official messages [32]. Over time, exposure

to these narratives causes serious damage to institutions, reduces civic engagement, and undermines trust in key democratic institutions such as courts and elections [36].

Foreign actors also take advantage of these domestic weaknesses through "narrative hijacking," using organised networks of bots and influencers to reshape domestic policy debates and push foreign interests onto the national agenda [33]. This outside interference greatly reduces a state’s control over its information. Adding to the problem is regulatory incapacity; since digital platforms operate worldwide, disinformation campaigns easily exploit legal gaps to avoid local enforcement. The failure of national laws to effectively fight cross-border information operations ultimately prevents governments from punishing hostile actors and maintaining real sovereignty over their information spaces [37].

## VIII. Measuring Disinformation’s Impact on State Cyber Sovereignty

To evaluate the effects of the disinformation, it is necessary to go beyond counting the number of poor contents to multiplex measurement systems that relate information flows, social consequences, and sovereignty concerns. Modern studies highlight that disinformation is a structural issue that affects the opinion of people, the stability of the regimes, and the sovereignty of countries in the virtual world [33, 34]. Three-tier measurement architecture: information flow metrics, societal impact metrics and sovereignty impact metrics is a thorough framework that helps assess how disinformation is changing digital ecosystems and undermining cyber sovereignty [38].

### *A. Information Flow Metrics*

The first tier is concerned with the disinformation that is propagated using digital networks. To reveal the mechanisms of influence campaigns on the Internet, scholars emphasise that it is important to study the speed of dissemination, network centrality, or bot-induced amplification. Studies indicate that automated profiles and synchronised networks are able to spread deceptive content faster and have it reach large numbers within a short duration of time [39].

Network analysis tools may be used to define the influential nodes and organised actors, whereas natural language processing (NLP) tools will be used to categorise the narratives and recognise patterns of misinformation, as well as find semantic framing approaches [26]. The metrics can be used to map the structural features of disinformation campaigns and demonstrate how algorithmic ecosystems facilitate the speed with which manipulative stories are shared on the digital platform.

### *B. Societal Impact Metrics*

The second tier looks at the impact of disinformation on the attitudes and social behaviour of the people. Research indicates that misinformation exposure is associated with deteriorating institutional trust, rising political polarisation, and changes in the attitude of the population.

The indices of public trust, sentiment analysis and polarisation are indicators that are measurable, as regards how the society responded to disinformation campaigns. Behavioural indicators, including shifts in civic activity, voting, or online communication, provide additional data to support the transformation of manipulated information in the field of discourse [34]. Survey research and digital trace analysis are exposure mapping techniques that enable the researcher to determine the demographic groups most impacted so that policymakers can see the societal impacts in the bigger picture, rather than as a single post or a single story.

### *C. Sovereignty Impact Metrics*

The third tier assesses the impact of the disinformation on the sovereignty of cyberspace itself. Sovereignty metrics determine the degree to which external actors shape domestic discourse, misuse national communication lines, or curtail the freedom of policy. Studies of computational propaganda show that organised foreign influence campaigns can influence the national political agenda and undermine the government’s control of the information ecosystems [33].

The geopolitical implications of disinformation are indicated by external control over the dissemination of the narrative, the inconsistency between the official and the popular discourse, and the interference with the work of the national decision-making. The researchers believe that these metrics of sovereignty are necessary to explain how these operations of digital influence are converted into strategic vulnerabilities of states [37].

### *D. Connecting Content, Society, and Sovereignty*

These three levels should be considered in a holistic measurement structure to understand the extent to which disinformation is taking place. Instead of paying attention to the detection of fake posts only, effective analysis links the information dissemination pattern to the reaction of society and long-term governance consequences.

This comprehensive strategy will allow policymakers to determine causal relationships, such as the creation of content to the shift in public perception to the loss of sovereignty, which will give a more precise estimate of the digital threats and will serve as the foundation of the specific intervention.

## IX. State Sovereignty Risk Scoring

The increasing prominence of disinformation and cyber threats has prompted policymakers and academics to work on systematic systems of assessing risks to cyber sovereignty. The conventional cybersecurity measurements used to dedicate significant attention to technical weaknesses, but it is now essential to consider multi-criteria models that could include institutional capacity, geopolitical dynamics, and societal resilience [37, 40]. Sovereignty risk scoring system is a strategic way of evaluating information threats by incorporating a technical, institutional, geopolitical and societal aspect, whereas a Cyber Sovereignty Vulnerability Index (CSVI) is an instrument that is measurable to determine the actual effects of information operations on governance and national sovereignty.

### *A. Technical Dimension*

The technical aspect evaluates the strength of the digital infrastructure, the efficiency of detection tools, and the response mechanism speed towards cyber threats and disinformation campaigns. Research also points out that effective network security and sophisticated threat information systems will be needed to reduce coordinated information attacks and infrastructure vulnerabilities.

Moreover, AI-based anomaly detection and bot identification are automated detection technologies that are vital in the detection of malicious activities in online ecosystems [39]. The technical preparedness of a state thus has a direct effect on whether it can continue to keep its sovereignty on digital networks and the safety of its informational landscape against manipulation.

### *B. Institutional Dimension*

The other crucial parameter in scoring sovereignty risk is institutional strength. The success of a state in regulating digital platforms and dealing with cyber threats in their emergence depends on the effectiveness of the regulations, enforcement, and inter-agency cooperation. According to the research on internet governance, the issue of regulatory fragmentation and poor institutional frameworks damages national jurisdiction on cyberspace [37]. Resilience to disinformation is promoted with the help of strong institutional arrangements such as cybersecurity policies, data governance legislation, and coordinated crisis communication frameworks, as well as reinforcing the ability of states to impose digital norms [32].

### *C. Geopolitical Dimension*

The geopolitical aspect considers the vulnerability to aggressive powers, local war, and information warfare policies. Research on hybrid warfare suggests that cyber activities and information campaigns are frequently instruments of geopolitical rivalry, which aims at political stability and trust in the populace [35].

Those countries that are placed in conflicting geopolitical settings or are partners in disputes in the region are more susceptible to foreign information manipulation and organised cyberattacks. By evaluating the presence of geopolitical risk factors, the policy formulators can foresee possibilities of risks and formulate proactive measures of defence based on the national security priorities.

### *D. Societal Dimension*

Resilience in society is becoming one of the core determinants of cyber sovereignty. The disinformation campaigns affect national stability, which is affected by public media literacy, social cohesion, and susceptibility to manipulation. Studies indicated that digitally literate societies and those that have been better connected through trust networks are resistant to fake news and online propaganda. On the other hand, polarised societies and broken media landscapes provide an excellent environment to wage information warfare, increasing the threat of activities of foreign influence and undermining the capacity of the state to keep the public in coherent speech.

### *E. Cyber Sovereignty Vulnerability Index*

A Cyber Sovereignty Vulnerability Index (CSVI) incorporates data on all four dimensions to offer a quantitative assessment of the national risk exposure. The signs might be the size of the narrative, the external source of information, the change of trust among the citizens, and quantifiable effects on the decision-making process in the country.

The CSVI links the disinformation activity and the outcomes of governance by integrating technical indicators (network analysis), institutional performance metrics, geopolitical threat indicators, and sentiment analysis of society. Researchers believe that multi-layered indices allow policymakers to shift to proactive risk management and planning in digital governance, rather than just responding to these risks [33].

## CONCLUSİON

This literature review has explored the sophisticated association between disinformation, state sovereignty and digital governance. Disinformation is an inherent contradiction to fundamental conceptualisations of sovereignty, as it aims at the informational and cognitive basis of political power as opposed to the physical divide. States have responded by growingly interested in the concept of cyber sovereignty as a method of regaining control over the digital realm.

Although cyber sovereignty can be used to enhance the mitigation of disinformation and safeguard national security, it poses severe threats as well. Securitisation of information governance can facilitate censorship, surveillance and the reduction of democracy without restraint. The literature demonstrates that there is intense conflict between antagonistic modes of digital order that mirror more widespread ideological conflicts about the place of the state, the market, and individual rights.

Finally, the dilemma of modern governance is the necessity to maintain the legitimate requirement to counteract disinformation and the desire to maintain an open, pluralistic and rights-respecting information space. To solve this dilemma, it will be needed not only regulatory innovation but also theoretical clarity, cooperation among nations, and long-term normative discussion.

## ACKNOWLEDGEMENT

The author expresses his sincere gratitude to Academician Rasim Aliguliyev, Director General of the Institute of Information Technology, for his continuous support, valuable guidance, and strategic vision in advancing research in the field of cyber sovereignty and digital governance. His leadership, academic insight, and encouragement have significantly contributed to the development of this study and to the broader research environment in which it was conducted.

## REFERENCES


[1] M. Kelton, M. Sullivan, Z. Rogers, E. Bienvenue, and S. Troath, “Virtual sovereignty? Private internet capital, digital platforms and infrastructural power in the United States,” Int. Aff., vol. 98, no. 6, pp. 1977–1999, Nov. 2022, doi: 10.1093/ia/iiac226.

[2] J. Sarts, “Disinformation as a Threat to National Security,” in Disinformation and Fake News, S. Jayakumar, B. Ang, and N. D. Anwar, Eds., Singapore: Springer, 2021, pp. 23–33. doi: 10.1007/978-981-15-5876-4_2.

[3] Y. Benkler, R. Faris, and H. Roberts, Network propaganda: Manipulation, disinformation, and radicalization in American politics. Oxford University Press, 2018.

[4] M. L. Mueller, Networks and states: The global politics of Internet governance. MIT press, 2010. Accessed: Feb. 12, 2026. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=qH3TAvkAtsEC&oi=fnd&pg=PP1&dq=Networks+and+states:+The+global+politics+of+internet+governance&ots=3NmJAbpklY&sig=ghoyzZH_onr24uNugSVqR17Dc-I

[5] C. Wardle and H. Derakhshan, Information disorder: Toward an interdisciplinary framework for research and policymaking, vol. 27. Council of Europe Strasbourg, 2017. Accessed: Feb. 12, 2026. [Online]. Available: http://tverezo.info/wp-content/uploads/2017/11/PREMS-162317-GBR-2018-Report-desinformation-A4-BAT.pdf

[6] K. Starbird, “Disinformation’s spread: bots, trolls and all of us,” Nature, vol. 571, no. 7766, pp. 449–449, Jul. 2019, doi: 10.1038/d41586-019-02235-x.

[7] T. Rid, Active measures: The secret history of disinformation and political warfare. Profile Books, 2020.

[8] C. Tenove, “Protecting Democracy from Disinformation: Normative Threats and Policy Responses,” Int. J. Press., vol. 25, no. 3, pp. 517–537, Jul. 2020, doi: 10.1177/1940161220918740.

[9] L. DeNardis, The Internet in everything. Yale University Press, 2020. Accessed: Feb. 12, 2026. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=gy7EDwAAQBAJ&oi=fnd&pg=PP1&dq=DeNardis+The+internet+in+everything:+Freedom+&ots=H_48p_u_oT&sig=bASVC518eAz3CtBAxOfC7WzL66w

[10] S. Couture, S. Toupin, and A. Mayoral Baños, “Resisting and Claiming Digital Sovereignty: The Cases of Civil Society and Indigenous Groups,” Policy Internet, vol. 16, no. 4, pp. 739–749, Dec. 2024, doi: 10.1002/poi3.434.

[11] F. Pierucci, “Sovereignty in the Digital Era: Rethinking Territoriality and Governance in Cyberspace,” Digit. Soc., vol. 4, no. 1, p. 27, Apr. 2025, doi: 10.1007/s44206-025-00189-4.

[12] Q. Bu, “Data sovereignty in Africa: steering digital transformation between China and the West,” Int. Cybersecurity Law Rev., vol. 7, no. 1, pp. 131–145, Mar. 2026, doi: 10.1365/s43439-025-00165-1.

[13] Y. S. Yadav, “Agile Organizing of National Cyber Defence Capabilities: Decoupling External Dependencies and Catalysing Digital Sovereignty,” Procedia Comput. Sci., vol. 254, pp. 260–268, Jan. 2025, doi: 10.1016/j.procs.2025.02.085.

[14] H. T. Hung, “Exploring China’s cyber sovereignty concept and artificial intelligence governance model: a machine learning approach,” J. Comput. Soc. Sci., vol. 8, no. 1, p. 24, Feb. 2025, doi: 10.1007/s42001-024-00346-8.

[15] K. Balarabe, “Digital borders and beyond: Establishing normative grounds for cybersecurity and sovereignty in international law,” Comput. Law Secur. Rev., vol. 58, p. 106180, Sep. 2025, doi: 10.1016/j.clsr.2025.106180.

[16] O. Prokopyshyn and N. Trushkina, “THE GEOPOLITICS OF CYBERSECURITY: A COMPARATIVE ANALYSIS OF NATIONAL STRATEGIES FOR DIGITAL SOVEREIGNTY,” Polit. Secur., vol. 12, no. 2, pp. 59–71, Jul. 2025, doi: 10.54658/ps.28153324.2025.12.2.pp.59-71.

[17] S. Fratini, E. Hine, C. Novelli, H. Roberts, and L. Floridi, “Digital Sovereignty: A Descriptive Analysis and a Critical Evaluation of Existing Models,” Digit. Soc., vol. 3, no. 3, p. 59, Nov. 2024, doi: 10.1007/s44206-024-00146-7.

[18] G. Papyshev and K. J. D. Chan, “AI regulatory strategies for digital sovereignty: The role of geopolitics and technological disparities,” Electron. Mark., vol. 36, no. 1, p. 8, Jan. 2026, doi: 10.1007/s12525-025-00870-z.

[19] B. Hoeksema, “Digital Sovereignty, the Private Sector, and a Social Republican Alternative,” Digit. Soc., vol. 3, no. 3, p. 51, Oct. 2024, doi: 10.1007/s44206-024-00140-z.

[20] O. Bronnikova et al., “Circumventing the ‘Sovereignization’ of the Russian Internet: Toward an Infrastructure-Based Sociology of Digital Sovereignty and Its Resistances in Russia,” in Digital Sovereignty in the BRICS Countries: How the Global South and Emerging Power Alliances Are Reshaping Digital Governance, L. Belli and M. Jiang, Eds., in Communication, Society and Politics. , Cambridge: Cambridge University Press, 2025, pp. 167–189. doi: 10.1017/9781009531085.012.

[21] W. Ito, “The State-Oriented Model of Internet Regulation: The Case of China,” in Public and Private Governance of Cybersecurity: Challenges and Potential, T. Ishikawa and Y. Kryvoi, Eds., Cambridge: Cambridge University Press, 2023, pp. 40–68. doi: 10.1017/9781009374576.003.

[22] B. Cartwright, R. Frank, G. Weir, and K. Padda, “Detecting and responding to hostile disinformation activities on social media using machine learning and deep neural networks,” Neural Comput. Appl., vol. 34, no. 18, pp. 15141–15163, Sep. 2022, doi: 10.1007/s00521-022-07296-0.

[23] S. Alanazi, S. Asif, A. Caird-daley, and I. Moulitsas, “Unmasking deepfakes: a multidisciplinary examination of social impacts and regulatory responses,” Hum.-Intell. Syst. Integr., vol. 7, no. 1, pp. 131–153, Dec. 2025, doi: 10.1007/s42454-025-00060-4.

[24] S. Alanazi and S. Asif, “Exploring deepfake technology: creation, consequences and countermeasures,” Hum.-Intell. Syst. Integr., vol. 6, no. 1, pp. 49–60, Dec. 2024, doi: 10.1007/s42454-024-00054-8.

[25] F. Romero-Moreno, “Deepfake detection in generative AI: A legal framework proposal to protect human rights,” Comput. Law Secur. Rev., vol. 58, p. 106162, Sep. 2025, doi: 10.1016/j.clsr.2025.106162.

[26] K. Shu, “Combating disinformation on social media: A computational perspective,” BenchCouncil Trans. Benchmarks Stand. Eval., vol. 2, no. 1, p. 100035, Mar. 2022, doi: 10.1016/j.tbench.2022.100035.

[27] P. N. Vasist and S. Krishnan, “Powered by innovation, derailed by disinformation: A multi-country analysis of the influence of online political disinformation on nations’ innovation performance,” Technol. Forecast. Soc. Change, vol. 199, p. 123029, Feb. 2024, doi: 10.1016/j.techfore.2023.123029.

[28] D. Surjatmodjo, A. A. Unde, H. Cangara, and A. F. Sonni, “Information Pandemic: A Critical Review of Disinformation Spread on Social Media and Its Implications for State Resilience,” Soc. Sci., vol. 13, no. 8, p. 418, Aug. 2024, doi: 10.3390/socsci13080418.

[29] S. Khan and J. R. Wright, “Disinformation, Stochastic Harm, and Costly Effort: A Principal-Agent Analysis of Regulating Social Media Platforms,” Jun. 27, 2022, arXiv: arXiv:2106.09847. doi: 10.48550/arXiv.2106.09847.

[30] M. R. Shoaib, Z. Wang, M. T. Ahvanooey, and J. Zhao, “Deepfakes, Misinformation, and Disinformation in the Era of Frontier AI, Generative AI, and Large AI Models,” Nov. 29, 2023, arXiv: arXiv:2311.17394. doi: 10.48550/arXiv.2311.17394.

[31] M. Alkhatib, “A Multifaceted Deepfake Prevention Framework Integrating Blockchain, Post-Quantum Cryptography, Hybrid Watermarking, Human Oversight, and Policy Governance,” Computers, vol. 14, no. 11, p. 488, Nov. 2025, doi: 10.3390/computers14110488.

[32] W. L. Bennett and S. Livingston, “The disinformation order: Disruptive communication and the decline of democratic institutions,” Eur. J. Commun., vol. 33, no. 2, pp. 122–139, Apr. 2018, doi: 10.1177/0267323118760317.

[33] S. Bradshaw and P. N. Howard, “The global disinformation order: 2019 global inventory of organised social media manipulation,” 2019, Accessed: Feb. 12, 2026. [Online]. Available: https://digitalcommons.unl.edu/scholcom/207/

[34] J. A. Tucker et al., “Social media, political polarization, and political disinformation: A review of the scientific literature,” Polit. Polariz. Polit. Disinformation Rev. Sci. Lit. March 19 2018, 2018, Accessed: Feb. 12,

2026. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3144139

[35] C. Paul and M. Matthews, “The Russian ‘firehose of falsehood’ propaganda model,” Rand Corp., vol. 2, no. 7, pp. 1–10, 2016, Accessed: Feb. 12, 2026. [Online]. Available: https://www.jstor.org/stable/pdf/resrep02439.pdf

[36] B. Nyhan and J. Reifler, “Displacing misinformation about events: An experimental test of causal corrections,” J. Exp. Polit. Sci., vol. 2, no. 1, pp. 81–93, 2015, Accessed: Feb. 12, 2026. [Online]. Available: https://www.cambridge.org/core/journals/journal-of-experimental-political-science/article/displacing-misinformation-about-events-an-experimental-test-of-causal-corrections/69550AB61F4E3F7C2CD03532FC740D05

[37] L. DeNardis, The global war for internet governance. Yale University Press, 2014. Accessed: Feb. 12, 2026. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=jfxfAgAAQBAJ&oi=fnd&pg=PA1&dq=The+Global+War+for+Internet+Governance&ots=gDxyRhBmKY&sig=VX8QBQX3fvM-HT7dIBYTndzI-ik

[38] R. Recuero, “A systemic framework for disinformation on social media platforms,” Platf. Soc., vol. 2, p. 29768624251367199, Dec. 2025, doi: 10.1177/29768624251367199.

[39] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, “The rise of social bots,” Commun. ACM, vol. 59, no. 7, pp. 96–104, Jun. 2016, doi: 10.1145/2818717.

[40] J. S. Nye Jr, “Deterrence and dissuasion in cyberspace,” Int. Secur., vol. 41, no. 3, pp. 44–71, 2016, Accessed: Feb. 12, 2026. [Online]. Available: https://direct.mit.edu/isec/article-abstract/41/3/44/12147