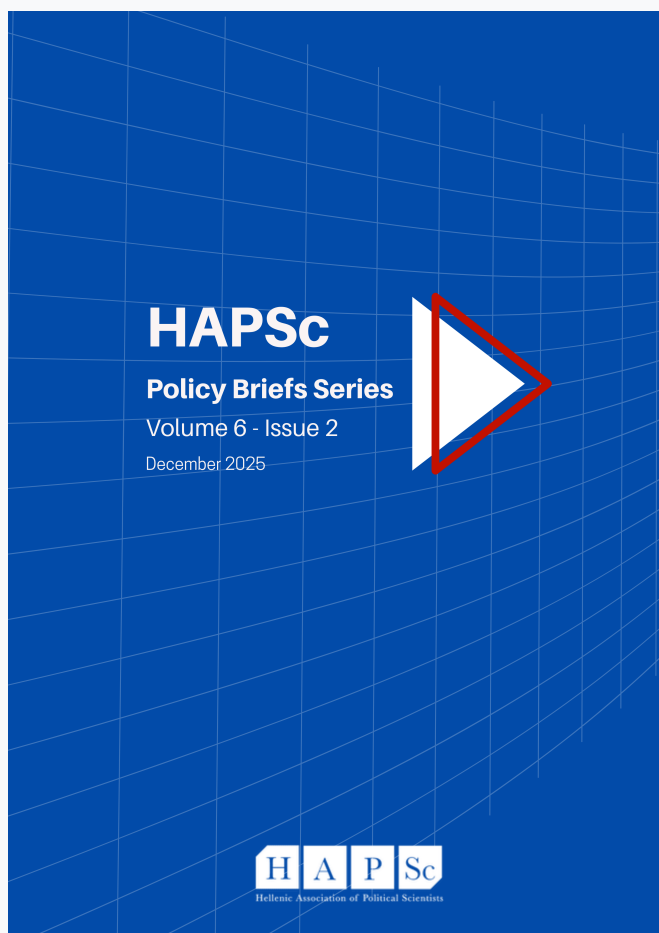


## HAPSc Policy Briefs Series

Vol 6, No 2 (2025)

HAPSc Policy Briefs Series



### Navigating Algorithmic Disinformation Models with Artificial Intelligence: Asymmetries and Comparisons

*Arthur Gugliucciello*

doi: [10.12681/hapscpbs.45359](https://doi.org/10.12681/hapscpbs.45359)

Copyright © 2025, Arthur Gugliucciello



This work is licensed under a [Creative Commons Attribution 4.0](https://creativecommons.org/licenses/by/4.0/).

#### To cite this article:

Gugliucciello, A. (2025). Navigating Algorithmic Disinformation Models with Artificial Intelligence: Asymmetries and Comparisons. *HAPSc Policy Briefs Series*, 6(2), 50–57. <https://doi.org/10.12681/hapscpbs.45359>

# Navigating Algorithmic Disinformation Models with Artificial Intelligence: Asymmetries and Comparisons<sup>1</sup>

Arthur Gugliucciello<sup>2</sup>

## Abstract

The exponential growth and diffusion of artificial intelligence (AI) have significantly heightened awareness of its potential effects across multiple domains. AI has become a well-established presence in society, enabling the widespread adoption of systems based on this technology and drawing global attention to its role in the manipulation of public opinion. AI should therefore not be understood merely as a “technical tool,” but as a technology endowed with political qualities. It is capable of shaping how individuals interact and how power is exercised, at speeds and scales far exceeding those of past forms of propaganda. This article is based on a critical-theoretical analysis of disinformation disseminated through media platforms and implemented via AI, with particular attention to the phenomenon as observed in China and the United States. The analysis reveals a dichotomy between a market-driven form of manipulation in the United States and a state-stability-oriented model in China. In both cases, AI emerges as a powerful accelerator of power asymmetries, making it necessary to redefine digital social contracts in order to counteract the automation of consent and the crisis of information.

**Keywords:** AI; Disinformation; Democracy; Governance; Public Debate

## Introduction

Recent years have witnessed remarkable developments in artificial intelligence (AI) across a wide range of sectors, significantly increasing attention to its potential global consequences. Its integration into communication flows has established new boundaries for public debate, while simultaneously facilitating the production and dissemination of disinformation systems in unprecedented ways. These far-reaching expectations have led to an ongoing discussion about the social and political impact of AI, characterized by alternating waves of fear and enthusiasm. However, to fully grasp the scope of these transformations, it is necessary to move beyond the view of technology as a neutral instrument. It is essential to recognize that technological artifacts possess intrinsic political qualities: they embody specific forms of power and authority that structure human relationships in ways that are often invisible yet decisive (Heyndels, 2023). In this sense, AI is not merely a technical evolution, but an agent that reconfigures the very architecture of the public sphere. To develop specific expectations regarding the impact of AI on democracy, it is crucial to be precise about both the meaning and

---

<sup>1</sup> To cite this paper in APA style: Gugliucciello, A. (2025). Navigating Algorithmic Disinformation Models with Artificial Intelligence: Asymmetries and Comparisons. *HAPSc Policy Briefs Series*, 6(2), 50-57. <https://doi.org/10.12681/hapscpbs.45359>

<sup>2</sup> University of Salerno, Italy.

functioning of the underlying technology. While numerous competing definitions exist, for the purposes of this article AI can be defined as the study and construction of agents that do the right thing (Rouhiainen, 2018). Yet the contemporary challenge extends beyond technical definitions and directly concerns the ethics of power.

Surveillance capitalism represents a new form of power – an “instrumentarian power” – that replaces the mystery of the human condition with computational certainty. Rather than aiming at the control of the soul, as in past totalitarian regimes, it seeks the automated control of behavior (Acemoglu et al., 2025). The political economy of control regimes suggests that, when left unregulated, such instruments tend to stabilize power asymmetries, eroding democratic safeguards through the administrative efficiency of surveillance. At the same time, the proliferation of conversational artificial intelligence systems, such as chatbots, has raised widespread concerns regarding privacy that cannot be interpreted in a univocal manner. The challenges posed by conversational AI extend beyond the protection of personal data and directly affect the quality of democratic participation itself.

On this basis, the present article provides a conceptual framework for a more focused and productive analysis of the likely impact of AI on a specific social domain: democracy. The analysis integrates a critical conceptual framework, beginning with a clarification of how AI functions, particularly its orientation toward the resolution of specific tasks (Oseni, 2021). The aim is not merely to guide or influence the design of technological artifacts, but to assess their contributions, limitations, and impacts in light of the normative standards upon which they are based. Secondly, this study examines AI-enabled disinformation systems in the United States and China. The adoption of a critical-theoretical perspective allows for the dissection of conceptual frameworks within online discourse in order to identify the manipulative “frames” through which disinformation flows circulate, transforming unfounded perceptions into accepted truths.

In this context, a comparison between the two major models of digital governance (the United States and China) proves particularly significant. An in-depth analysis makes it possible to understand how disinformation is not a randomly generated phenomenon, but rather one that emerges in relation to power architectures shaping communication infrastructures. U.S. digital power reveals disinformation models driven by a form of “networked propaganda,” sustained by algorithmic logics that reward polarization and commercial profit. By contrast, the Chinese media architecture initially exhibited developmental potential for the public sphere but was rapidly reconfigured under the pressure of state governance. Here, the disinformative objective does not merely consist in manipulating public opinion, but in saturating public discourse with distracting content aimed at

neutralizing dissent and signaling alignment with government policies (King et al., 2017). A comparison between these two contexts thus enables the analysis of two distinct strategies of manipulation within the contemporary geopolitical landscape.

### **The Impact of AI Between Architectures of Control and the Manipulation of Consent**

Following the perspective advanced by Heyndels (2023), technology is never fully neutral but always carries specific political agendas within it. This becomes evident through a technical distinction of AI and a phenomenological analysis of social relations, from which a clear relationship emerges between democratic governance and the management of disinformation. Technological artifacts, in fact, contain political qualities: they embody specific forms of power and authority capable of altering constitutional balances and perceptions of truth (Klonick, 2018). Consequently, AI is not merely a technical instrument, but an agent that restructures existing power relations and redefines the boundaries of freedom of expression (Calo, 2017). The impact of AI manifests across distinct levels: for individuals, it alters the capacity to distinguish truth from falsehood; for groups, it undermines equality in political competition; and at the systemic level, it influences competition between democratic and autocratic systems. The informational crisis thus concerns all social actors and the ways in which they interact, rather than merely the information they receive (Sheehan, 2022).

AI enables more granular forms of surveillance and control, shifting the balance of power between the state and the citizen. The political economy of dictatorship suggests that more efficient tools of distraction can stabilize authoritarian regimes by avoiding overt violent censorship and instead flooding public discourse with diverting content to eliminate dissent (King et al., 2017). It is evident that when citizens become alienated from automated (dis)information flows, they may reject traditional institutions and turn toward alternative narratives, thereby fueling polarization. Focusing on these specific dynamics clarifies both the real threats and opportunities, enabling more effective monitoring of AI's impact on democracy beyond the logic of mere technological profit. To this end, a critical-theoretical inquiry has been employed to examine the interconnections between power and online discourse, identifying the manipulative “frames” through which disinformation takes hold. In the United States, public and institutional discourse emphasizes the risks associated with networked propaganda and is strongly influenced by a liberal tradition that frames privacy as a right against both state and corporate intrusion. This has contributed to an “information disorder” that undermines trust in democratic institutions (Wardle & Derakhshan, 2017).

By contrast, in the Chinese context, disinformation practices are primarily oriented toward state stability and institutional narratives of collective harmony. This reflects an environment in which

strategies are developed to distract and divert citizens by inundating digital spaces with highly distracting content in order to suppress criticism of governance (Roberts, 2018). In China, modern conceptions of disinformation have emerged within a social structure and moral discourse that prioritize numerically mediated trust and state stability over factual accuracy.

### **Comparative Analysis Between the United States and China**

In contexts where information ecosystems are shaped by AI, evidence suggests that perceptions and practices surrounding disinformation vary according to sociocultural settings and governance models (Allcott & Gentzkow, 2017). To capture these dynamics, the analysis proceeds through a critical examination of governance models, mapping the dominant “frames” through which manipulative narratives penetrate public discourse. In the United States, disinformation tends to emerge as a fragmented, largely bottom-up phenomenon, fueled by a digital architecture that enables extreme polarization (Sunstein, 2018). Beyond emblematic case studies such as the activities of the Russian *Internet Research Agency* during the 2016 presidential elections, AI has ushered in a new phase through the use of synthetic media and *deepfakes* in more recent electoral cycles. This development highlights another dimension of disinformation: the manipulation of collective deliberative capacity.

Cloned voices have been used to disseminate misleading voting instructions, as in the case of the Joe Biden “robocalls” during the 2024 New Hampshire primaries, where a voice identical to that of the President urged citizens not to vote (Federal Communications Commission, 2024). This episode demonstrates how political communication is facing an unprecedented crisis of truth, in which *deepfakes* undermine collective deliberation and the integrity of democratic processes (Battista, 2024a). In the United States, discourse remains strongly shaped by a liberal tradition that conceives information as a safeguard against violations of cognitive autonomy. This confirms the analytical finding that information is interpreted as a shield protecting individual rights, with solutions perceived in restoring an uncorrupted marketplace of ideas rather than adapting to direct state control (Marwick & Lewis, 2017).

In contrast, within the Chinese context, disinformation management is less concerned with factual truth and more focused on the positive effects of AI in ensuring stability and collective harmony (Gunitsky, 2015). A relevant case concerns narrative management during the 2019 Hong Kong anti-extradition protests, where AI systems enabled the creation of numerous social media profiles and misleading content to bury dissent (Lazer et al., 2018). At the same time, the introduction of AI assistants in urban governance services, such as smart city management, has fostered a qualitatively different digital discourse centered on efficiency. Rather than questioning the legitimacy of state data

collection, debates have focused on service efficiency and technical correctness, reflecting a technological culture in which AI is understood as a frontier of political communication and social control (Gallo et al., 2022).

Criticism, when present, has often taken indirect or satirical forms to circumvent censorship, or has translated into practical adaptation strategies, such as migration to private channels for sensitive discussions (Battista & Mangone, 2025). This reflects the prioritization of collective harmony and stability, where technology is accepted as an instrument of state modernization and AI becomes a tool of “automation of consent” (Battista, 2024b). The observed divergence suggests that Chinese and American users operate within distinct informational “spheres of justice,” in which practices deemed tyrannical in one context may be accepted as necessary for the common good in another, and truth becomes functional to social cohesion (Creemers, 2020).

### **Multilevel Governance and Informational Integrity Against Algorithmic Disinformation**

Analyzing the effects of AI on the public sphere requires understanding disinformation not merely as false content, but as a violation of epistemic integrity and of the informational flows essential to democracy. Governance must be conceived as an active defense against the erosion of human decision-making power in the face of emerging forms of algorithmic authority (Floridi & Cows, 2019). The integrity of public discourse responds to an urgent social need for a definition of “procedural truth” that explains why AI-driven manipulation is ethically unacceptable and clarifies how deliberation can be protected through law and technology (Wardle, 2019). At the core of the issue lies the violation of norms governing the authenticity of communication flows. A healthy information ecosystem requires that news production and dissemination be context-appropriate and adhere to principles of source transparency. Key parameters now extend beyond actors to include the synthetic origin of messages and the algorithmic mediation principles governing virality (Battista, 2024c). Once content generated by large language models is released online, both audience reach and propagation speed change radically, rendering traditional fact-checking mechanisms obsolete. Accordingly, this study emphasizes the importance of a multilevel critical-theoretical analysis to investigate how manipulative narratives adapt to different digital architectures (Balkin, 2020).

Responses to disinformation have evolved differently depending on state–platform relations: while some systems rely on self-regulation, others deploy AI for predictive censorship (Harsin, 2018). Recognizing this complexity is essential to protecting democracy from what is termed “computational propaganda”: a system that no longer seeks to persuade, but to disorient through algorithmic saturation (Woolley, 2020). Unlike traditional propaganda, the instrumentarian power of AI aims to

control behavior through “predictive certainty,” reducing human interaction to datasets ready to be manipulated by artificial agents (Battista & Petrone, 2024). Surveillance capitalism is thus evolving into a capitalism of influence, in which human experience is fragmented by personalized synthetic content that undermines the foundations of shared reality (Zuboff, 2015). Just as industrial civilization eroded the physical environment, an information civilization shaped by automated disinformation threatens to erode cognitive autonomy. Focusing on these dynamics clarifies real threats and enables interdisciplinary monitoring of AI’s impact on democracy, integrating insights from social sciences and computational ethics (Battista, 2025).

## Conclusions

A profound re-foundation of digital social contracts is now required to address the impact of AI on the public sphere, constituting an intersectional challenge that cannot be resolved through simple technological fixes. It is essential to avoid “technological somnambulism”, namely the tendency to accept the automation of public discourse as an inevitable destiny rather than as a set of political choices laden with consequences (Jasanoff, 2016). As highlighted throughout this analysis, the dynamics of disinformation are deeply intertwined with algorithmic architectures and regulatory contexts. The threshold of quality at which synthetic content (such as deepfakes) is perceived as “credible” does not depend solely on technical resolution, but on the epistemic vulnerability and cognitive biases of the target audience (Rini, 2017). This implies that effective regulation against informational manipulation cannot be uniform, but must instead respect the contextual sensitivities of different communities, rejecting the notion that the circulation of data for model training constitutes a higher imperative than the integrity of public discourse (Gillespie, 2018).

What is at stake goes far beyond the efficiency of monitoring systems. Just as industrial civilization flourished at the expense of the natural environment, an information civilization dominated by automated flows and predictive disinformation risks prospering at the expense of human autonomy, transforming citizens into passive consumers of artificial narratives (Couldry & Mejias, 2019). Without critical resistance and conscious cognitive sovereignty, we risk surrendering the conditions of political freedom to systems that optimize engagement at the expense of truth, thereby feeding behavioral surplus (Susskind, 2020). Accordingly, this approach provides an essential framework for a targeted analysis of developments related to generative AI, where the defense of democracy depends on our capacity to distinguish human agency from algorithmic output (Nemitz, 2018). Through an interdisciplinary effort, it will be possible to ensure that technological innovation remains a tool of democratic emancipation, preventing it from becoming the foundation of a new form of oppression.

## References

- Acemoglu, D., Makhdoumi, A., Malekian, A., & Ozdaglar, A. (2025). When big data enables behavioral manipulation. *American Economic Review: Insights*, 7(1), 19–38. <https://doi.org/10.1257/aeri.20230589>
- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211–236. <https://doi.org/10.1257/jep.31.2.211>
- Balkin, J. M. (2020). How to regulate (and not regulate) social media. *Journal of Free Speech Law*, 1(1), 71–105.
- Battista, D. (2025). From electronic voting to AI: lessons from the Estonian Case on the transformation of digital democracy. *Em Questão*, 31, e-145465. <https://doi.org/10.1590/1808-5245.31.145465>
- Battista, D. (2024a). Political communication in the age of artificial intelligence: An overview of deepfakes and their implications. *Society Register*, 8(2), 7–24. <https://doi.org/10.14746/sr.2024.8.2.01>
- Battista, D. (2024b). The road to AI: Pathways and obstacles. *Sociétés*, 163(1), 55–72. <https://doi.org/10.3917/soc.163.0055>
- Battista, D. (2024c). Comunicazione politica e intelligenza artificiale: un bilancio tra manipolazione e partecipazione. *Rivista di Digital Politics*, 4(1), 71-90. <https://doi.org/10.53227/113721>
- Battista, D., & Petrone, A. (2024). Artificial Intelligence and Media-politics: A Revolution in Communicative Dynamics?. *Journal of Sociological Research*, 15(2), 51. <https://doi.org/10.5296/jsr.v15i2.22207>
- Battista, D., & Mangone, E. (2025). Technological Culture and Politics: Artificial Intelligence as the New Frontier of Political Communication. *Societies*, 15(4), 75. <https://doi.org/10.3390/soc15040075>
- Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 US Presidential election online. *First Monday*, 21(11). <https://doi.org/10.5210/fm.v21i11.7090>
- Calo, R. (2017). Artificial Intelligence Policy: A Primer and Roadmap. *UC Davis Law Review*, 51, 399–435.
- Couldry, N., & Mejias, U. A. (2019). Data Colonialism: Rethinking Big Data's Relation to the Contemporary User. *Television & New Media*, 20(4), 336–349. <https://doi.org/10.1177/1527476418796632>
- Creemers, R. (2020). *China's conception of cyber sovereignty*. Leiden University Press.
- Federal Communications Commission. (2024, May 23). *FCC proposes \$6 million fine for robocall consultant behind deepfake Biden calls*. <https://www.fcc.gov/document/fcc-proposes-6-million-fine-deepfake-biden-robocalls>
- Floridi, L., & Cowsls, J. (2019). A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.8cd35046>
- Gallo, M., Fenza, G., & Battista, D. (2022). Information Disorder: What about global security implications?. *Rivista di Digital Politics*, 2(3), 523-538. <https://doi.org/10.53227/106458>
- Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.
- Gunitsky, S. (2015). Corrupting the cyber-commons: Social media as a tool of autocratic stability. *Perspectives on Politics*, 13(1), 42–54. <https://doi.org/10.1017/S1537592714003120>

- Harsin, J. (2018). *Post-truth and critical communication studies*. Oxford Research Encyclopedia of Communication. <https://doi.org/10.1093/acrefore/9780190228613.013.757>
- Heyndels, G. (2023). Technology and neutrality. *Philosophy & Technology*, 36(4), 78. <https://doi.org/10.1007/s13347-023-00672-1>
- Jasanoff, S. (2016). *The ethics of invention: Technology and the human future*. W.W. Norton & Company.
- King, G., Pan, J., & Roberts, M. E. (2017). How the Chinese government fabricates social media posts for strategic distraction, not engaged argument. *American Political Science Review*, 111(3), 484–501. <https://doi.org/10.1017/S0003055417000144>
- Klonick, K. (2018). The new governors: The people, rules, and processes governing online speech. *Harvard Law Review*, 131(6), 1598–1670.
- Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... & Zittrain, J. L. (2018). The science of fake news. *Science*, 359(6380), 1094–1096. <https://doi.org/10.1126/science.aao2998>
- Marwick, A., & Lewis, R. (2017). *Media manipulation and disinformation online*. Data & Society Research Institute.
- Nemitz, P. (2018). Constitutional democracy and technology in the age of artificial intelligence. *Philosophical Transactions of the Royal Society A*, 376(2128). <https://doi.org/10.1098/rsta.2018.0089>
- Oseni, A., Moustafa, N., Janicke, H., Liu, P., Tari, Z., & Vasilakos, A. (2021). Security and privacy for artificial intelligence: Opportunities and challenges. *arXiv preprint arXiv:2102.04661*.
- Rini, R. (2017). Fake news and partisan epistemology. *Kennedy Institute of Ethics Journal*, 27(2), 43–64. <https://doi.org/10.1353/ken.2017.0025>
- Roberts, M. E. (2018). *Censored: Distraction and diversion inside China's Great Firewall*. Princeton University Press.
- Rouhiainen, L. (2018). *Artificial intelligence: 101 things you must know today about our future*. Lasse Rouhiainen.
- Sheehan, M. (2022). *China's AI Regulations and How They Get Made*. Carnegie Endowment for International Peace.
- Sunstein, C. R. (2018). *#Republic: Divided democracy in the age of social media*. Princeton University Press.
- Susskind, J. (2020). *Future politics: Living together in a world transformed by tech*. Oxford University Press.
- Wardle, C. (2019). *Information disorder: Toward an interdisciplinary framework for research and policymaking*. Harvard Kennedy School.
- Wardle, C., & Derakhshan, H. (2017). *Information Disorder: Toward an interdisciplinary framework for research and policymaking*. Council of Europe.
- Woolley, S. C. (2020). *Computational propaganda: Political parties, politicians, and manipulation on social media*. Oxford University Press.
- Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30(1), 75–89. <https://doi.org/10.1057/jit.2015.5>