AGORA International Journal of Juridical Sciences, http://univagora.ro/jour/index.php/aijjs
ISSN 1843-570X, E-ISSN 2067-7677
Vol. 20, No. 1 (2026), pp. 146-154

LEGAL REGULATION OF COUNTERING DISINFORMATION IN THE AGE OF AI

T. KORTUKOVA, A. MAZARAKI

Tamara Kortukova[1], Anatolii Mazaraki[2]
[1] [2] State University of Trade and Economics, Kyiv (Ukraine)
[1] https://orcid.org/0000-0002-3376-0159, E-mail: t.kortukova@knute.edu.ua
[2] https://orcid.org/0000-0003-1817-0510, E-mail: rector@knute.edu.ua

Abstract: *The rapid proliferation of generative Artificial Intelligence (hereinafter - AI) has revolutionized the creation of disinformation, rendering traditional, static legal frameworks insufficient. This study examines the multifaceted legal challenges posed by AI-driven threats, specifically focusing on deepfakes and automated manipulative content that jeopardize democratic stability. Methodology. The research utilizes a comparative and doctrinal analysis of the European Union's regulatory landscape. Specifically, it evaluates the efficacy of the EU AI Act and the Digital Services Act (hereinafter – DSA) concerning transparency obligations, technical watermarking, and algorithmic accountability. The analysis reveals that while current mechanisms provide a foundation for oversight, significant legal gaps remain in addressing the automated nature of synthetic media and the jurisdictional challenges of anonymous creators. Findings indicate that rigid mandates often struggle to keep pace with evolving technological capabilities. The study concludes that a resilient, adaptive regulatory framework is necessary to preserve national security and democratic integrity. It argues for a living legal system that balances technical enforcement with the protection of fundamental digital rights, ensuring that efforts to counter disinformation do not result in over-censorship or the infringement of freedom of expression.*

Keywords: *Artificial Intelligence, disinformation, deepfakes, EU AI Act, information security, digital rights.*

1. Introduction

In democratic societies, concern about the consequences of our growing reliance upon AI is rising (Eileen & Megan, 2019). The advent of AI era has fundamentally reshaped the global information landscape (Kortukova et al., 2025; Ağtaş, 2025; Abbes, 2025). Traditionally, the creation of "fake news" required significant time and resources; however, modern generative models now enable the production of highly persuasive disinformation, ranging from manipulative text to hyper-realistic deepfakes, in a matter of seconds.

Statistics indicates that 83% of people globally believe disinformation represents a direct threat to the stability of democracy. This concern is grounded in high levels of exposure: 51% of Europeans report having been exposed to disinformation online, while 63% of younger Europeans – the demographic most integrated into digital spaces – encounter "fake news" more than once a week (European Commission, 2022).

Domestic and foreign actors increasingly employ mass disinformation to trigger societal tension and undermine security. These campaigns, often orchestrated by third countries, function as hybrid threats, frequently paired with cyberattacks, to destabilize internal security and electoral integrity. A prominent example is russia's military doctrine, which formally classifies information warfare as a core operational domain (European Commission, 2018).

The Brexit and Cambridge Analytica scandals previously illustrated how the illicit collection and analysis of Big Data allows for the creation of sophisticated "psychological

profiles". (Oleart & García, 2025) AI now automates this process, enabling the mass production of hyper-personalized disinformation designed to exploit the specific emotional and cognitive vulnerabilities of individual users.

AI facilitates the illusion of broad public consensus or total social chaos, directly eroding trust in electoral institutions. A prime example is the 2024 "Biden robocall" case, where AI-generated voices were used to misinform voters at a critical moment before the polls (Enyong, 2025). Such incidents demonstrate how synthetic technology can obstruct the fundamental right of citizens to receive accurate information and exercise free will in the voting booth.

For policymakers, these events revealed a critical vulnerability: the ease with which democratic processes could be destabilized through the coordinated manipulation of digital platforms. Consequently, in democratic societies, concern about the consequences of our growing reliance upon AI and digital curation is rising.

This technological leap presents an unprecedented challenge for global legal systems: finding a way to safeguard the truth without infringing upon the fundamental rights. AI-based disinformation has created a complex legal vacuum. While accessible AI tools make it simple to produce and broadcast false narratives, they make it nearly impossible for authorities to monitor or regulate the source. Pinning legal responsibility on anonymous creators, who often operate in different countries, is a major logistical challenge. This globalized, automated nature of deception makes traditional national laws increasingly obsolete in providing real accountability.

The objective of this article is to conduct a comprehensive analysis of the legal challenges posed by AI-driven disinformation, specifically focusing on how generative technologies like deepfakes necessitate a shift in traditional regulatory frameworks. By evaluating mechanisms such as the EU AI Act and the DSA, the study aims to propose a resilient, adaptive legal system that protects democratic integrity and national security without infringing upon the fundamental right to freedom of expression.

2. Methods

The methodology employed in this article is dictated by the specific objectives of the study, aimed at understanding the objective reality of legally regulating AI-driven disinformation within the digital frameworks of the EU.

The research was conducted on the basis of a dialectical materialist methodology that reflects the intrinsic relationship between theory and practice. The methodological basis of the study was formed by a dialectical approach to understanding the evolving nature of generative AI and its impact on democratic processes. The scientific tools used were grounded in the principles of objectivity and pluralism, allowing for a comprehensive knowledge of the intersection between technical AI capabilities and legal constraints.

The author relied on modern methods of cognition, defined by contemporary legal science and tested through the observation of digital enforcement. In particular, both general and special scientific methods were utilized.

Methods of analysis and synthesis were used to break down the complex components of the "Regulatory Triangle" and subsequently reassemble them to identify gaps in the interplay between the EU AI Act and DSA.

Methods such as abstraction and generalization were instrumental in deriving conclusions regarding the enforcement paradox of Article 50.

Furthermore, the research involved the method of normative modeling and legal forecasting to anticipate how the "satire loophole" might be exploited in future disinformation campaigns.

Formal-legal method assisted in defining the essence of limited risk systems and transparency obligations under the EU AI Act.

The use of these methods facilitated a deep exploration of the subject, revealing that the regulation of AI-driven disinformation is not merely a theoretical concern but one of immense practical importance for the stability of democratic institutions.

The normative basis of the work includes the EU AI Act, the DSA, the European Media Freedom Act (hereinafter – EMFA), and the relevant EU acquis concerning digital services and human rights.

3. Results

3. 1. Definition of Disinformation

The formal codification of a legal definition of disinformation is a critical prerequisite for the rule of law in the information age. However, there has been surprisingly little attention paid in legal scholarship to the specific legal definitions of disinformation, and how these definitions relate to current legal frameworks (Fathaigh et al., 2021).

While the terms "misinformation" and "disinformation" are often used interchangeably in casual discourse, with some even using disinformation as an umbrella term, international agencies such as the United Nations Educational, Scientific and Cultural Organization (hereinafter – UNESCO) and the Office of the United Nations High Commissioner for Human Rights (hereinafter – OHCHR) maintain a strict distinction based on intent (UNESCO, 2025).

To provide clarity, scholars and international bodies generally categorize these phenomena into the types: 1) misinformation - the inadvertent sharing of false or inaccurate information (such as rumors, pranks, or outdated news) without the intention to cause harm; 2) disinformation - the strategic and deliberate dissemination of false content (such as hoaxes, propaganda, or spear-phishing) created with the specific malice to deceive the public, spread fear, or undermine social trust (UNHCR, 2022); 3) malinformation - the use of factual, true information that is shared out of context or with the explicit intent to cause harm to a person, organization, or country (such as the leaking of private information for political sabotage) (Maret, 2025).

Historically, the European Union’s approach to disinformation has a clear distinction between illegal content (such as hate speech or defamation) and harmful but legal content (such as disinformation) (Fathaigh et al., 2021).

According to the European Commission’s High Level Expert Group, the legal threshold for disinformation is met when false or deceptive content is strategically crafted and amplified with the explicit objective of generating profit or inflicting harm upon the public. This definition emphasizes that the "information disorder" is defined as much by its malicious intent as by its factual inaccuracy (Final report, 2018).

As established by the European Commission, disinformation is fundamentally understood as verifiably false or misleading information that is created, presented and disseminated for economic gain or to intentionally deceive the public, and may cause public harm (European Commission, 2018).

While the EU has primarily treated disinformation as a non-binding guideline (soft law), individual Member States are beginning to codify it into formal legislation. A primary example is the Portuguese Charter of Human Rights in the Digital Era, which adopted the European Commission’s 2018 definition almost verbatim in its Article 6.

However, this transition from guideline to law sparked intense debate, centered on two main issues: critics argued that terms like "verifiably false or misleading narrative" are too ambiguous for legal enforcement; concerns arose over a provision allowing the State to empower specific entities to label the quality of media outlets, creating a risk of government-sanctioned censorship.

Consequently, the Portuguese President requested a constitutional review of these norms. This pending case before the Constitutional Court represents a landmark test: can the EU's definition of disinformation survive a rigorous constitutional challenge regarding fundamental rights?

To preempt a strike-down by the Constitutional Court, the Portuguese Parliament passed Law 15/2022, which repealed the most controversial segments of Article 6, effectively stripping the "right" of its enforcement mechanisms and shifting the focus toward broader EU alignment and digital literacy.

In conclusion, the formal codification of a legal definition of disinformation is a vital but complex prerequisite for the rule of law. While international frameworks like those of the European Commission, UNESCO, and OHCHR provide essential distinctions based on malicious intent and public harm, the case of the Portuguese Charter demonstrates the fragility of these concepts when shifted from soft law to enforceable legislation. The eventual repeal of Article 6's most controversial provisions highlight a fundamental tension: without extreme linguistic precision, any legal attempt to regulate "verifiably false narratives" risks being struck down as an unconstitutional expansion of state power. Ultimately, the quest for a legal standard proves that protecting the information ecosystem requires a delicate balance between mitigating coordinated deception and safeguarding the fundamental right to free expression.

3.2. The Role of Generative AI in the Technological Acceleration of Disinformation

The emergence of generative AI, specifically Large Language Models (hereinafter – LLMs) and Generative Adversarial Networks (hereinafter – GANs), has fundamentally altered the mechanics of creating and distributing disinformation (López-Borrull & Lopezosa, 2025). While traditional influence operations once required significant human capital, such as "troll farms," contemporary processes are increasingly automated, enabling the production of deceptive content at an industrial scale.

In its March 2025 report on foreign information manipulation and interference, the European External Action Service concluded that generative AI has significantly lowered the financial and technical barriers for hostile actors. By streamlining and automating content generation, these technologies allow threat actors to execute large-scale influence operations with greater efficiency and reduced costs (European External Action Service, 2025).

Generative AI accelerates the degradation of the information ecosystem through three primary vectors: scalability and content automation large language models (hereinafter – LLMs), hyper-realism and synthetic media generative adversarial network (hereinafter – GANs), psychological micro-targeting.

LLM allow for the instantaneous generation of thousands of unique variations of a single narrative. This enables a strategy known as "flooding the zone", where a high volume of synthetic content overwhelms verified facts, making it difficult for the public to discern truth (Sallami et al., 2024). Furthermore, AI eliminates linguistic barriers; bad actors can now produce grammatically flawless content in dozens of languages, removing a primary obstacle to cross-border influence operations.

GAN technology facilitates the creation of deepfakes, synthetic media that is nearly indistinguishable from authentic recordings. This technology does more than just fabricate statements from public figures; it erodes the overall evidentiary value of real media. This phenomenon, often called the "liar's dividend", occurs when individuals can plausibly dismiss genuine events or recordings as AI-generated products.

AI algorithms can analyze the digital footprints of users to craft personalized disinformation. Models can automatically adapt the tone, vocabulary, and emotional framing of a message to exploit the specific vulnerabilities and cognitive biases of different social groups. This transforms disinformation from a "broadcast" tool into a precision-guided

instrument of manipulation. Consequently, generative AI does not merely create fake news; it constructs a self-sustaining ecosystem of synthetic reality (Mazaraki & Kortukova, 2026). In a legal context, this makes a precise definition of disinformation even more critical. The velocity of algorithmic dissemination far outpaces traditional regulatory responses and manual fact-checking, necessitating a framework that can address the intent and the systemic nature of these AI-driven campaigns.

3.3. The EU Regulatory Triangle, as a Shield Against Disinformation

In the face of hybrid threats, the European Union has moved from a voluntary code of conduct approach to a robust active defense strategy. This is manifested in a Regulatory Triangle, which include the EU AI Act, the DSA and the European Media Freedom Act, where three landmark legislative acts create a synergy designed to disrupt the lifecycle of disinformation: creation, distribution, and environmental resilience.

The EU AI Act, which entered into force on August 1, 2024, establishes a risk-based hierarchy to regulate synthetic content and safeguard democratic integrity. Under this framework, generative tools like chatbots and deepfakes are generally classified as limited risk, requiring mandatory transparency such as watermarking and user disclosure (Article 50). However, systems designed to manipulate human behavior or influence elections are elevated to "high-risk" or prohibited entirely under Article 5 if they employ deceptive techniques that impair informed decision-making.

Complementing these rules, the European AI Office oversees general-purpose AI (hereinafter – GPAI) models, which Recital 110 identifies as potential sources of systemic risk, including the dissemination of disinformation and threats to democratic processes. To bridge the gap until full implementation, the Commission introduced the GPAI Code of Practice (2025) and the AI Pact, providing developers with ethical guidelines for safety, intellectual property, and the exclusion of pirated training data.

In the context of European digital sovereignty, the DSA represents a paradigm shift from a "laissez-faire" approach to a regime of strict algorithmic accountability. While the AI Act focuses on the technical creation of content, the DSA targets the amplification and distribution mechanisms used by Very Large Online Platforms (hereinafter – VLOPs) and Very Large Online Search Engines (hereinafter – VLOSEs). The main tasks of the DSA are to counteract the spread of illegal content on the Internet and reduce societal risks arising from disinformation and other forms of harmful content (Prymak & Kortukova, 2025). It operates on the principle that what is illegal offline is illegal online, providing the European Commission with unprecedented oversight into how digital gatekeepers manage the flow of information during sensitive periods, such as elections or public health crises (Frosio & Geiger, 2024).

A cornerstone of the DSA's fight against disinformation is the mandatory systemic risk assessment. Under Articles 34 and 35, platforms with more than 45 million monthly active users in the EU must proactively identify and mitigate risks related to the dissemination of illegal content and "negative effects on democratic processes". This forces tech giants to move beyond simple content moderation and instead audit their own recommendation algorithms to ensure they do not prioritize sensationalism or "engagement at any cost" over factual accuracy. If a platform's architecture is found to be facilitating coordinated inauthentic behavior or bot-driven manipulation, the Commission can impose fines of up to 6% of annual global turnover (Husovec, 2024).

Furthermore, the DSA institutionalizes the Code of Practice on Disinformation, transforming it from a voluntary set of guidelines into a co-regulatory framework (Husovec, 2025). Platforms are now required to provide researchers with access to their internal data to monitor the effectiveness of their anti-disinformation measures. This transparency creates a "public interest audit" where civil society and academia can hold platforms accountable for

how their tools are exploited by foreign actors for foreign information manipulation and interference. By mandating rapid response mechanisms and clear labeling of political advertising, the DSA ensures that the digital public square remains a space for informed debate rather than a laboratory for psychological manipulation (Caruso, 2024).

While the DSA focuses on the "pipes" through which information flows, the European Media Freedom Act (hereinafter – EMFA), which became fully applicable in August 2025, addresses the "source" (European Commission, 2024). The EMFA recognizes that a pluralistic, independent media ecosystem is the most effective structural antidote to disinformation. By protecting journalists from state interference and ensuring the transparency of media ownership, the Act seeks to build a resilient information environment where factual reporting can outcompete manipulated narratives.

A key pillar of the EMFA in the fight against Foreign Information Manipulation and Interference is the coordination of national regulators through the newly established European Board for Media Services (Article 12). Under the Act, if a media service provider from outside the EU, often state-controlled, systematically engages in disinformation that threatens public security, the Board can coordinate collective restrictive measures across all Member States (Husovec, 2025). This prevents "forum shopping," where a hostile actor might exploit a weaker regulatory regime in one country to broadcast propaganda across the entire Union.

One of the most debated aspects of the EMFA is Article 17, which creates a specific procedural shield for recognized media outlets on VLOPs. To prevent tech giants from inadvertently silencing legitimate journalism in their quest to purge disinformation, Article 17 requires platforms to: allow media outlets to declare their independent status based on transparent editorial standards; provide a 24-hour warning before removing or suspending content from these outlets, unless the content poses a systemic risk as defined by the DSA; establish a "fast-track" mechanism to resolve disputes, ensuring that timely news, crucial during elections or crises, is not buried by slow moderation processes (EPD, 2024).

The EMFA further mandates that all news and current affairs media must make their ownership structures public (Article 6). This requirement is designed to expose hidden conflicts of interest or "capture" by political or foreign entities who may be using a local outlet as a proxy for disinformation campaigns.

To operationalize this triangle, the EU has moved toward a centralized intelligence model. The European Board for Media Services works in tandem with the European AI Office and the European Centre for Algorithmic Transparency to provide a holistic view of emerging threats. This is particularly critical in addressing foreign information manipulation and interference, where malicious actors often bypass traditional moderation by using "hybrid" tactics, combining AI-generated personas with coordinated bot networks and hijacked local media proxies (EEAS, 2026). Under this unified framework, a systemic risk identified by a media regulator can trigger an immediate audit of a platform’s algorithm under the DSA, ensuring that the response is as agile as the threat itself.

This regulatory architecture represents a significant evolution of the Brussels effect. By moving beyond the binary choice of "censorship vs. free-for-all", the EU is pioneering a third way: algorithmic and institutional transparency. The shift from voluntary codes to legally binding obligations means that digital sovereignty is no longer a theoretical concept but a functional reality. As other jurisdictions look to the EU’s model, the Regulatory Triangle serves as a blueprint for how democratic societies can protect their cognitive sovereignty without sacrificing the fundamental right to freedom of expression (Frosio & Geiger, 2024). This structural resilience ensures that even as disinformation tactics evolve, from simple "fake news" to sophisticated AI-driven psychological operations, the European digital space remains anchored in verifiable facts and pluralistic debate.

4. Discussion

While the Regulatory Triangle provides a sophisticated defense, the legal regulation of AI-driven disinformation faces unique, evolving hurdles. These "regulatory friction points" stem from the technical nature of generative AI and the legal complexities of enforcing transparency in a decentralized digital environment. A primary challenge is the temporal gap between the evolution of AI capabilities and the enactment of legal standards. As documented in early 2026, while the AI Act's transparency obligations (Article 50) set high standards for watermarking and metadata, malicious actors have quickly pivoted to "low-fidelity" but high-impact AI models that bypass these built-in safeguards. This creates a "policy response lag" where regulations effectively address "yesterday's threats" while attackers leverage open-source models that lack mandatory watermarking protocols. Under the AI Act, systems creating deepfakes or synthetic text for public information are classified as limited risk, requiring disclosure rather than prohibition. However, enforcing this at scale presents a significant hurdle (Christakis, 2026). Article 50 (4) provides exceptions for content that is "evidently artistic, creative, satirical, or fictional" (Raman et al., 2025). In practice, disinformation campaigns often cloak themselves as "satire" or "parody" to evade mandatory labeling, placing an immense interpretive burden on platform moderators and courts.

Current watermarking techniques are not yet interoperable or tamper-proof. Research indicates that simple post-processing, such as re-encoding a video or taking a screenshot of AI-generated text, can strip away the machine-readable markers required by law, rendering the legal obligation technically unenforceable in many real-world scenarios.

A critical legal gray zone exists in the attribution of liability between the provider of an AI model and the deployer who uses it to spread disinformation. While the AI Act focuses on the upstream developers (e.g., ensuring a model is safe), the DSA focuses on the downstream distribution (e.g., the platform where it goes viral).

GPAI models with systemic risk are subject to stricter oversight, yet determining when a model's propensity to "hallucinate" or generate biased content crosses the threshold into a "systemic threat to democratic processes" remains legally ambiguous (European Parliament, 2025). Furthermore, the rise of "agentic AI", autonomous bots capable of managing social media accounts and engaging in real-time debates, tests the limits of traditional agency law. If an AI "agent" autonomously spreads disinformation without direct human instruction, the current legal framework struggles to assign culpability. Finally, there is the persistent risk of regulatory overreach. By mandating that platforms mitigate "negative effects on democratic processes" (DSA Art. 34), the EU effectively asks private tech giants to act as arbiters of political truth. Critics argue that this creates a chilling effect, where platforms may over-censor legitimate but controversial political speech to avoid the risk of massive fines, potentially undermining the very media pluralism the EMFA seeks to protect.

5. Conclusion

The legal regulation of AI-driven disinformation represents a defining challenge of the digital age, requiring a precise calibration between technological innovation, legal certainty, and the preservation of fundamental rights. Based on the analysis of the current regulatory landscape, several critical findings emerge.

First, a structural "policy response lag" persists. Despite the high standards set by the EU AI Act, malicious actors increasingly bypass legislative safeguards by utilizing "low-fidelity" or open-source models that circumvent mandatory watermarking protocols. This technical evasion renders static legal standards vulnerable to the rapid evolution of decentralized AI development.

Second, the enforcement paradox of Article 50 highlights a disconnect between legal intent and technical reality. The classification of synthetic content as limited risk is

fundamentally undermined by the "Satire Loophole" and the absence of tamper-proof watermarking. These factors allow disinformation to be strategically re-categorized as creative expression or "scrubbed" of metadata, complicating the task of judicial and platform-level attribution. Third, the threat of regulatory overreach poses a risk to democratic discourse. To mitigate the severe financial penalties prescribed by the Digital Services Act, platforms may lean toward aggressive over-censorship. Such a "chilling effect" threatens the media pluralism and freedom of expression that these regulations are ostensibly designed to protect.

In summary, while the current legal framework provides a vital foundation, it must be treated as a living regulatory organism. Effective governance requires a transition from static mandates to interoperable, resilient technical standards and a more agile, risk-based approach to liability. Only through such adaptation can the law remain a robust shield against the sophisticated threats of the synthetic information era.

REFERENCES

1. Eileen, D. & Megan, M. (2019). Artificial Intelligence and Human Rights. *Journal of Democracy, 30 (2),* 115-126.
2. Kortukova, T., Kudin, V., Dei, M., Onyshchenko, A., Kravchuk, P. (2025) Legal Challenges of Artificial Intelligence in the European Union's Digital Economy. *International Journal of Informatics and Communication Technology, 4,* 960-971. 10.11591/ijict.v14i3.pp960-971.
3. Ağtaş, E. (2025). The double-edged impact of Artificial Intelligence on corporate reputation: Opportunities and threats. *AGORA International Journal of Economical Sciences, 19*(2), 16–32. http://univagora.ro/jour/index.php/aijes
4. Abbes, N. (2025). The impact of innovation on economic growth in Algeria: An econometric analysis using the ARDL model (1996–2021). *AGORA International Journal of Economical Sciences, 19*(2), 1–15. http://univagora.ro/jour/index.php/aijes
5. European Commission. (2022). *A strengthened EU Code of Practice on Disinformation*. https://commission.europa.eu/topics/countering-information-manipulation/strengthened-eu-code-practice-disinformation_en
6. European Commission. (2018). *Tackling online disinformation: A European approach* (COM/2018/236 final). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52018DC0236
7. Oleart, A., & García, L. B. (2025). From self to co-regulation in the EU's approach to disinformation: The framing power of Big Tech business lobbies in the lead to the Digital Services Act. *International Review of Public Policy, 7*(2). https://doi.org/10.4000/14rse
8. Enyong, A. B. (2025). *A report on the role Cambridge Analytica played in the latest American presidential election and the Brexit referendum in the UK*. University of East London. https://doi.org/10.13140/RG.2.2.34497.62560
9. Ó Fathaigh, R., Helberger, N., & Appelman, N. (2021). The perils of legally defining disinformation. *Internet Policy Review, 10(4).* https://doi.org/10.14763/2021.4.1584
10. United Nations Educational, Scientific and Cultural Organization. UNESCO (2025). *Misinformation*. https://www.unesco.org/en/query-list/m/misinformation
11. United Nations High Commissioner for Refugees (UNHCR). (2022). *Factsheet 4: Types of misinformation and disinformation*. https://www.unhcr.org/innovation/wp-content/uploads/2022/02/Factsheet-4.pdf
12. Susan, M. (2025). Mal-Information, the Anatomy of an "Information Disorder". *Secrecy and Society 3(2)*. https://doi.org/10.55917/2377-6188.1090
13. Final report of the High Level Expert Group on Fake News and Online Disinformation (2018) https://digital-strategy.ec.europa.eu/en/library/final-report-high-level-expert-group-fake-news-and-online-disinformation

14. European Commission. (2018). *Tackling online disinformation: A European approach* (COM/2018/236 final). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52018DC0236
15. López-Borrull, A., & Lopezosa, C. (2025). Mapping the Impact of Generative AI on Disinformation: Insights from a Scoping *Review. Publications, 13(3),* 33. https://doi.org/10.3390/publications13030033
16. European External Action Service. (2025). 3rd EEAS report on foreign information manipulation and interference threats: Exposing the architecture of FIMI operations. https://www.eeas.europa.eu/sites/default/files/documents/2025/EEAS-3nd-ThreatReport-March-2025-05-Digital-HD.pdf
17. Sallami, D., Chang, Y.-C., & Aïmeur, E. (2024). From Deception to Detection: The Dual Roles of Large Language Models in Fake News, *Arxiv*, https://doi.org/10.48550/arXiv.2409.17416
18. Mazaraki N. & Kortukova T. (2026) Towards Responsible AI in Legal Services: the Case for Regulating Chatbots. *Bulletin of the Taras Shevchenko National University of Kyiv. Legal Sciences, 130(2)*. https://doi.org/10.17721/1728-2195/2025/2.130-9
19. Prymak, B. & Kortukova, T. (2025). Ensuring the Rights of Online Platforms in the Mechanism of the Digital Services Act. *Problems of Modern Transformations. Series: Law, Public Management and Administration, (18).* https://doi.org/10.54929/2786-5746-2025-18-01-25
20. Frosio G., Geiger C. (2024) Towards a Digital Constitution How the Digital Services Act Shapes the Future of Online Governance. *Verfassungsblog*. https://verfassungsblog.de/towards-a-digital-constitution/
21. Husovec, M. (2024). The Digital Services Act's red line: what the Commission can and cannot do about disinformation. *Journal of Media Law*, *16*(2), 270-346.
22. Husovec, M. (2025). Framing the role of experts in platform governance: Negotiating the code of practice on disinformation as a case study. *Internet Policy Review*, *14*(1). https://doi.org/10.14763/2025-1-1823
23. Caruso, C. (2024). Towards the Institutions of Freedom: The European Public Discourse in the Digital Era. *German Law Journal*, 1-24. https://doi.org/10.1017/glj.2024.68
24. European Commission. (2024). The European Media Freedom Act (EMFA): Questions and Answers. European Union. https://ec.europa.eu/commission/presscorner/detail/en/qanda_22_5505
25. European Partnership for Democracy (EPD). (2024). Joint Statement on Article 17 of the European Media Freedom Act. https://epd.eu/news-publications/joint-statement-on-article-17-of-the-european-media-freedom-act/
26. European External Action Service (EEAS). (2026). 3rd Annual Report on Foreign Information Manipulation and Interference Threats. https://www.eeas.europa.eu/eeas/3rd-eeas-report-foreign-information-manipulation-and-interference-threats-0_en
27. Frosio, G., & Geiger, C. (2024). The Digital Services Act and the Brussels Effect on Platform Content Moderation. *Chicago Journal of International Law, 24(1).*
28. Christakis, T. (2026). You trust your chatbot with everything. Should you? Part 1: How the controller uses your chat data. *AI Regulation Papers, 26(3).*
29. Raman, D., Madkour, N., Murphy, E. R., Jackson, K., & Newman, J. (2025). Intolerable risk threshold recommendations for Artificial Intelligence. *Arxiv*. https://doi.org/10.48550/arxiv.2503.05812
30. European Parliament. (2025). The European Union's AI code of practice (EPRS_ATA(2025)775890). European Parliamentary Research Service. https://www.europarl.europa.eu/RegData/etudes/ATAG/2025/775890/EPRS_ATA(2025)775890_EN.pdf