frontiers in Communication

TYPE Review
PUBLISHED 14 May 2026
DOI 10.3389/fcomm.2026.1823630



OPEN ACCESS

EDITED BY
Lambrini Papadopoulou,
National and Kapodistrian University of Athens, Greece

REVIEWED BY
Vian Bakir,
Bangor University, United Kingdom
Daniele Battista,
University of Salerno, Italy

*CORRESPONDENCE
Elsir Ali Saad Mohamed
drelsir.ali@uaqu.ac.ae

RECEIVED 06 March 2026
REVISED 02 April 2026
ACCEPTED 14 April 2026
PUBLISHED 14 May 2026

CITATION
Mohamed EAS, Mallek M, Alaawad HA, ELTinay IMO, Alrached SA and Alamin M (2026) Regulating artificial intelligence in digital media: governance, ethics, ownership, and democratic resilience (2023–2026 review).
Front. Commun. 11:1823630.
doi: 10.3389/fcomm.2026.1823630



# Regulating artificial intelligence in digital media: governance, ethics, ownership, and democratic resilience (2023–2026 review)

Elsir Ali Saad Mohamed[1]*, Mohamed Mallek[2], Haitham Abdulrahman Alaawad[1], Ikhlas Mustafa Omer ELTinay[3], Saleh Abied Alrached[1] and Mohanad Alamin[4]

[1]Umm Al Quwain University, Umm Al Quwain, United Arab Emirates, [2]University of Khorfakkan, Sharjah, United Arab Emirates, [3]University of Kalba, Kalba, United Arab Emirates, [4]City University Ajman, Ajman, United Arab Emirates

This review synthesizes scholarship published between 2023 and 2026 examining the intersection of artificial intelligence regulation and digital media. Drawing on systematic reviews, policy analyses, and empirical studies, the article maps the evolving regulatory landscape addressing AI-driven disinformation, examines ethical challenges posed by generative AI in communication, analyzes ownership patterns and power concentration among technology firms, and considers implications for media representation and democratic resilience. The review identifies competing priorities between regulatory interventions, ethical frameworks, and market dynamics, highlighting the need for integrated approaches that address both technical and sociopolitical dimensions of AI governance. Key findings reveal divergent regulatory philosophies across jurisdictions, persistent challenges in balancing innovation with protection of democratic values, and emerging multistakeholder frameworks that may offer pathways toward more resilient information ecosystems.

KEYWORDS
**artificial intelligence governance, democratic resilience, digital media regulation, disinformation, generative AI, platform power, political economy**

## 1 Introduction

The period between 2023 and 2026 has witnessed an unprecedented acceleration in the integration of artificial intelligence into digital media ecosystems. Generative AI technologies—particularly large language models and synthetic media tools—have fundamentally altered how information is created, disseminated, and consumed. This transformation has prompted urgent scholarly and policy attention to governance frameworks capable of addressing both the opportunities and risks associated with AI-mediated communication. Yet despite significant regulatory activity, including the implementation of the European Union's AI Act and Digital Services Act alongside proposed legislation in other jurisdictions, the governance landscape remains fragmented and contested.

This review addresses a critical gap in the literature by moving beyond descriptive accounts of regulatory developments to provide a critical synthesis of the structural

tensions that define the field. Rather than simply cataloging existing research, we ask: What are the unresolved debates and contradictions shaping AI governance in digital media? How do emerging patterns of infrastructural power challenge the assumptions underlying current regulatory frameworks? And what are the implications for democratic resilience, particularly for vulnerable communities?

Drawing on a systematic scoping review of scholarship published between 2023 and 2026, this article contributes to the literature in three ways. First, it critically evaluates the gap between regulatory philosophy and practical implementation, revealing how divergent approaches across jurisdictions create opportunities for regulatory arbitrage while failing to address the economic incentives driving platform behavior. Second, it synthesizes the emerging political economy literature on infrastructural power, demonstrating that control over computing infrastructure and proprietary datasets represents a qualitatively different form of influence than traditional content ownership. Third, it identifies a future research agenda focused on intersectional vulnerabilities and the "efficacy gap" in AI governance—the persistent disconnect between policy ambition and measurable impact on information ecosystems.

The review proceeds as follows. Section 2 outlines the scoping review methodology. Section 3 critically examines regulatory responses to AI-driven disinformation, highlighting contradictions between ambition and implementation. Section 4 analyzes the political economy of AI, focusing on structural asymmetries and infrastructural control. Section 5 explores ethical challenges and media representation, with particular attention to vulnerable groups. Section 6 examines the attention economy as a structural logic shaping AI governance outcomes. Section 7 identifies cross-cutting themes and proposes a future research agenda. Section 8 concludes with implications for democratic resilience.

TABLE 1 Comparative analysis of AI regulatory philosophies (2023–2026).

| Feature | European Union (EU) | United States (US) |
|---|---|---|
| Core philosophy | **Precautionary and Rights-Based:** Prioritizes fundamental rights and safety over market speed. | **Market-Oriented and Fragmented:** Prioritizes innovation, economic competition, and platform self-regulation. |
| Primary instruments | **EU AI Act** and **Digital Services Act (DSA)**. | Fragmented federal guidance and state-level initiatives. |
| Risk management | **Tiered Risk Classification:** Imposes strict requirements on "high-risk" AI systems. | **Self-Regulatory:** Relies on voluntary commitments and existing consumer protection laws. |
| Transparency mandates | Mandatory labeling for deepfakes and disclosure of AI-generated content. | Limited federal labeling mandates; largely dependent on platform-specific policies. |
| Compliance focus | **Ex-ante (Before Deployment):** Requires conformity assessments and data quality audits. | **Ex-post (After Harm):** Focuses on addressing harms after they occur through litigation or FTC oversight. |
| Market impact | Higher compliance costs; potential for slower innovation but higher public trust. | Lower barriers to entry; rapid innovation but higher risk of jurisdictional arbitrage. |
| Democratic aim | Safeguarding the "shared factual basis" and information integrity. | Protecting free speech while encouraging technological leadership. |

## 2 Methodology

This study employs a scoping review methodology following the framework of Arksey and O'Malley (2005) and the PRISMA-ScR guidelines (Tricco et al., 2018). Scoping review methodology is appropriate for this study given our aim to map the breadth of literature on AI governance in digital media, identify key concepts and gaps, and synthesize findings across diverse disciplinary perspectives including law, communication studies, political economy, and ethics.

### 2.1 Search strategy

Searches were conducted in three databases: Scopus, Web of Science, and Google Scholar. These databases were selected for their comprehensive coverage of peer-reviewed literature in social sciences, humanities, and policy studies. The search was limited to English-language publications between January 2023 and December 2026 to capture scholarship coinciding with the implementation of major regulatory frameworks, including the EU AI Act and Digital Services Act (DSA). The selection of January 2023 as the start date is methodologically significant as it captures the transition of these frameworks from legislative theory to practical implementation. Furthermore, this period encapsulates the scholarship responding to the rapid integration of generative AI into media ecosystems. While the DSA was proposed earlier, January 2023 marks a qualitative shift toward active enforcement and empirical scholarship focused on the real-world impact of these regulations. This specific temporal window is necessary to analyze the "efficacy gap"—the measurable disconnect between active policy ambitions and their actual influence on information ecosystems. Additionally, the 2023 start date captures the surge in scholarship following the widespread public adoption of advanced Large Language Models (LLMs), which fundamentally shifted the regulatory focus from general AI toward specific generative risks. A comparative overview of regulatory philosophies in the EU and US is presented in Table 1.

### 2.2 Inclusion and exclusion criteria

Publications were included if they met the following criteria:

(a) Addressed AI governance in relation to digital media, communication, or information ecosystems

(b) Presented original empirical research, systematic reviews, scoping reviews, or policy analyses
(c) Engaged explicitly with regulatory, ethical, ownership, or democratic implications
(d) Published in peer-reviewed journals, conference proceedings, or as authoritative policy reports
(e) Published between January 2023 and December 2026

Publications were excluded if they:

(a) Focused exclusively on technical AI development without social science or policy dimensions
(b) Were conference abstracts, opinion pieces, editorials, or non-peer-reviewed commentary without empirical grounding
(c) Were not available in English

## 2.3 Screening and selection

Initial searches produced 847 unique records after duplicate removal. Title and abstract screening against inclusion criteria reduced this to 124 full-text items for detailed assessment. Full-text assessment was conducted independently by two authors, with disagreements resolved through discussion. Following detailed assessment, the final corpus comprised 73 sources, including 12 systematic or scoping reviews, 48 empirical studies, and 13 policy documents.

## 2.4 Data extraction and synthesis

Data extraction focused on regulatory approaches, ethical issues, ownership dynamics, democratic impacts, and identified research gaps. Thematic synthesis was conducted iteratively, involving: (1) open coding to identify initial themes across sources; (2) organizing codes into descriptive themes; and (3) developing analytical themes that went beyond the content of individual studies to address the review's research questions (Thomas and Harden, 2008). This approach enabled us to identify patterns, contradictions, and gaps across the literature while maintaining fidelity to the source material.

## 2.5 Limitations

A recognized limitation of this scoping review is its reliance on English-language scholarship indexed within Scopus, Web of Science, and Google Scholar, a parameter that inherently emphasizes Global North perspectives. This methodological boundary serves as a deliberate and transparent rationale for the exclusion of a comprehensive analysis of the Chinese regulatory regime. To examine such a distinct governance model with sufficient academic rigor, a specialized linguistic and cultural corpus would be required, which falls outside the current study's scope. By maintaining this strategic focus on the EU-US axis, the review ensures a consistent and high-resolution analysis of the specific structural tensions and 'efficacy gaps' unique to democratic information ecosystems.

# 3 Regulatory responses to AI-driven disinformation: divergent philosophies and implementation aps

Scholarship on AI regulation in digital media reveals a complex landscape characterized not only by divergent philosophies but by a persistent gap between regulatory ambition and practical implementation. This section critically examines these tensions, moving beyond descriptive accounts to identify contradictions, debates, and unresolved questions in the literature.

## 3.1 Divergent regulatory philosophies: a contested terrain

The prioritization of the European Union and the United States as the focal points of this analysis is rooted in three strategic imperatives. First, these jurisdictions represent the two most influential, yet diametrically opposed, governance paradigms currently defining global digital media. While the EU's "Rights-Based" approach emphasizes fundamental safety, privacy, and the precautionary principle, the US "Market-Oriented" model consistently prioritizes rapid innovation and platform self-regulation.

Second, the concentration of transnational "infrastructural power" is uniquely localized within this geographic axis. Because the majority of "Big Tech" firms controlling the essential technical stack—ranging from cloud computing and foundation models to proprietary datasets—operate primarily under EU or US oversight, these regions serve as the most critical sites for examining how structural power is negotiated and contested.

Finally, the scope is intentionally restricted to these regions to analyze how divergent frameworks of liberal democracy attempt to safeguard "democratic resilience" and a "shared factual basis". Although state-led models, such as that of China, offer distinct regulatory alternatives, they function under fundamentally different sociopolitical assumptions regarding information integrity and public discourse that fall outside the democratic parameters of this review

## 3.2 The dual-use paradox: genuine dilemma or strategic narrative?

A significant debate in the literature concerns the characterization of generative AI as a "dual-use" technology—simultaneously a vehicle for harm and a tool for protection. Balabanova et al. (2025) employ game-theoretic analysis to demonstrate how platforms, regulators, and malicious actors engage in strategic interactions where detection tools prompt adaptive countermeasures. This perspective suggests that the dual-use nature of AI is an inherent technological property with real governance implications.

However, a critical reading of the literature reveals an alternative interpretation. Some scholars argue that the "dual-

use" framing is strategically deployed by platforms to justify self-regulation and resist binding obligations (Khanal et al., 2024). If AI is inherently both threat and solution, the argument goes, then regulators should defer to platform expertise in managing this balance. This debate remains unresolved but carries significant stakes: accepting the dual-use framing as purely technological forecloses questions about whether alternative governance arrangements—such as public infrastructure for detection tools—might shift the balance toward protection.

### 3.3 Implementation gaps and regulatory arbitrage

The literature consistently identifies jurisdictional variation as a structural weakness in current governance frameworks. Malicious actors exploit differences between jurisdictions, deploying AI-generated content that evades detection through subtle variations or hybrid human-AI workflows (Balabanova et al., 2025). More troubling, technology firms themselves engage in regulatory arbitrage, navigating between the EU's precautionary model and the US's market-oriented framework to dilute the impact of more rigid requirements (Chesterman, 2025).

This raises a critical question that the literature has only begun to address: can regulatory frameworks that operate at the nation-state level effectively govern technologies that are inherently transnational in their development, deployment, and impact? The Council of Europe's Framework Convention on Artificial Intelligence represents one attempt to establish binding international principles, yet its flexibility provisions may accommodate the very variation that enables arbitrage (Council of Europe, 2024). The literature suggests that without mechanisms for international coordination—and without addressing the underlying economic incentives driving platform behavior—even ambitious regulatory frameworks risk symbolic compliance without substantive impact.

## 4 Media oligarchs and the political economy of AI

While Section 3 examined formal regulatory frameworks, this section turns to a dimension that remains undertheorized in much governance scholarship: the political economy of AI infrastructure. The literature reviewed here reveals that power is increasingly concentrated not through content ownership but through control of the underlying technical stack—computing infrastructure, proprietary datasets, and distribution algorithms.

### 4.1 Structural asymmetry and infrastructural control

Rikap (2024) introduces the concept of "corporate governance beyond ownership" to describe how a handful of technology firms exert control through ownership of cloud computing infrastructure and exclusive access to massive, proprietary datasets. This represents a qualitative shift from earlier forms of media concentration. Where previous regulatory frameworks focused on content ownership limits to preserve media plurality, contemporary power operates at a deeper infrastructural level.

The implications for democratic resilience are profound and, according to the literature, underappreciated by current governance frameworks. Khanal et al. (2024) document how news organizations increasingly rely on platform-provided AI foundation models and cloud resources for both content production and distribution. This creates what we term infrastructural dependency: media organizations that once served as independent gatekeepers now compete for visibility within platform-controlled attention markets, forced to adopt AI tools that may prioritize virality over journalistic integrity.

### 4.2 Policy capture and the surge of Big Tech influence

The concentration of technical power translates directly into political influence, challenging the assumption that regulation can effectively constrain the firms it seeks to govern. Analysis of the policy process reveals that "Big Tech" power has surged dramatically during the review period; in 2025, corporations and trade groups accounted for over 80% of AI-related lobbying activity (Citizen.org, 2026).

This creates what the literature identifies as a feedback loop between technical concentration and political influence. Firms leverage their exclusive technical expertise and control over essential AI resources to frame regulatory debates, often steering interventions toward technical symptoms (e.g., content labeling requirements) rather than underlying economic incentives (e.g., engagement-based revenue models). The result, as Taeihagh (2025) notes, is regulatory capture by design: even well-intentioned frameworks may fail to address the structural conditions that generate harmful outcomes because those conditions remain outside the scope of regulatory consideration.

### 4.3 Implications for democratic resilience

The literature points to a sobering conclusion: current regulatory frameworks focusing on content moderation or labeling risk missing the deeper issue of private control over the infrastructure of public discourse. The EU AI Act, for all its ambition, operates within markets characterized by significant information asymmetries, where platforms possess the data regulators need to design effective interventions (EPRS, 2025). Without addressing these structural asymmetries, even the most progressive regulatory frameworks may produce at best symbolic compliance.

This analysis suggests a fundamental reorientation may be necessary. Rather than regulating AI applications as discrete products, future governance frameworks may need to address the infrastructural conditions—compute access, data ownership, algorithmic transparency—that shape the information environment. This remains an open question in the literature and a priority for future research.

# 5 Ethical challenges and media representation

While the political economy literature emphasizes structural power, a distinct but related body of scholarship examines how AI systems encode and amplify biases affecting marginalized communities. This section critically synthesizes this literature, moving beyond cataloging harms to examine underlying mechanisms and unresolved debates.

## 5.1 Algorithmic invisibility and the normative filter

A consistent finding across the literature is that AI systems do not merely reflect societal bias but structurally amplify it through automated distribution (Gutiérrez-Caneda et al., 2024; Peña-Alonso et al., 2025). Ricciardi Celsi and Zomaya (2025) identify bias as a persistent challenge rooted in training data, model architecture, and deployment contexts—a problem that cannot be resolved through technical fixes alone because it reflects deeper societal inequalities.

The concept of algorithmic invisibility captures how AI models trained on historical datasets often prioritize "normative" or Western-centric perspectives, leading to systematic marginalization of minority voices. Research indicates that generative AI tools frequently default to stereotypical portrayals of race, gender, and disability unless specifically prompted otherwise (López-Borrull and Lópezosa, 2025). This creates a "representation tax" where marginalized creators must work harder to achieve accurate self-representation—a dynamic that the literature suggests has significant implications for democratic inclusion.

## 5.2 Targeted harms and digital violence

A significant portion of scholarship focuses on specific, high-intensity risks for women and racialized communities that current regulatory frameworks struggle to mitigate. Park and Nan (2025) emphasize that the impact of generative AI on misinformation varies across communities, with marginalized groups potentially experiencing both disproportionate targeting and reduced access to verification resources.

The literature identifies non-consensual synthetic imagery (deepfakes) as a particularly concerning domain. These tools are disproportionately weaponized against women in the public eye to undermine professional credibility and psychological well-being (Singh, 2025). A critical question that emerges from the literature is whether current regulatory frameworks—with their emphasis on content labeling and removal—adequately address these forms of digital violence. The evidence suggests they do not, as such harms often fall below enforcement thresholds while causing significant individual and collective damage.

## 5.3 Intersectionality and differential impact

The concept of intersectionality—recognizing that individuals experience multiple, overlapping forms of marginalization—remains undertheorized in much of the AI governance literature. López-Borrull and Lópezosa (2025) note that deepfake technology raises particular concerns for vulnerable groups, yet systematic analysis of how these risks compound remains limited.

This represents a significant research gap. The literature suggests that communities with histories of misrepresentation by traditional media show higher levels of skepticism toward AI-generated content—a response that might be characterized as resilience rather than vulnerability (Hastuti, 2025). Yet policy discussions often overlook these differential experiences, treating audiences as undifferentiated. The need for inclusive governance is clear: as the literature argues, "Human-in-the-Loop" (HITL) requirements are insufficient if the "humans" in the loop do not reflect the diversity of the global population.

# 6 The attention economy as structural logic

The "attention economy"—the system in which human attention is treated as a scarce commodity to be captured and monetized—provides the underlying logic for much of contemporary digital media. This section examines how AI systems optimize for attention capture and how this optimization shapes information ecosystems and democratic discourse.

## 6.1 Engagement bias and algorithmic optimization

A core tension in digital media regulation lies in the conflict between public interest and platform profitability. AI-powered recommendation systems are mathematically optimized to maximize "time on site," with research indicating that emotionally resonant, polarizing, or sensationalist content consistently outperforms factual reporting in these automated auctions for attention (Balabanova et al., 2025; Vosoughi et al., 2018).

The literature identifies this as a structural contradiction: the business models that fund digital media infrastructure systematically incentivize content that may undermine democratic deliberation. As Park and Nan (2025) argue, generative AI does not create this dynamic but acts as a "force multiplier," enabling production of "infinite variations" of engaging content tailored to specific psychological profiles at near-zero marginal cost. The result is not simply more content but a qualitative shift in the scale and sophistication of engagement-optimized media.

## 6.2 Generative AI and the dilution of information quality

A significant development in 2024–2025 was the shift toward "on-demand" synthetic media, which allows for mass production of tailored content at scale. This has implications that the literature is only beginning to explore. Micro-targeted synthetic content enables platforms to create infinite variations of a single

narrative, each tailored to a specific demographic's psychological triggers (López-Borrull and Lópezosa, 2025). This "hyper-personalization" makes collective oversight and traditional fact-checking nearly impossible.

The literature identifies a dilution effect whereby the sheer volume of AI-generated content creates "noise" that overwhelms high-quality, human-led journalism. This "flooding" strategy is often deployed to drown out critical discourse or inconvenient truths, raising questions about the viability of traditional information quality interventions in environments characterized by synthetic content abundance.

### 6.3 Economic incentives vs. Regulatory mandates

Perhaps the most significant insight from the literature concerns the mismatch between regulatory focus and economic reality. Most current laws—including the EU AI Act—focus on content (e.g., labeling deepfakes) rather than the economic incentives that drive distribution. So long as platform revenue is tied to engagement metrics, AI will be tuned to amplify content that generates high arousal, whether or not that content is accurate or democratically beneficial.

Recent scholarship argues that for democratic resilience to be achieved, regulation must move toward structural decoupling—separating the ownership of AI infrastructure from control of recommendation algorithms (Taeihagh, 2025). This remains a provocative proposal rather than an established policy direction, but it points to a fundamental question: can the attention economy be reformed from within, or does democratic resilience require more fundamental restructuring of the economic relationships that shape digital media?

## 7 Cross-cutting themes and future research agenda

Synthesis of the scholarship reviewed above reveals three cross-cutting themes that define the current state of AI regulation in digital media. These themes highlight the tension between theoretical frameworks and practical realities, and they suggest priorities for future inquiry.

### 7.1 The efficacy gap: policy ambition vs. measurable impact

A recurring theme across the literature is the gap between regulatory ambition and measurable impact. Research indicates that while many platforms have adopted mandatory labeling for AI-generated content, these labels often fail to influence user trust or reduce the virality of disinformation (Park and Nan, 2025). Regulatory bodies often lack the technical infrastructure and "algorithmic auditing" capabilities required to verify platform compliance in real-time (EPRS, 2025).

This efficacy gap raises fundamental questions about the assumptions underlying current governance approaches. If transparency mandates do not produce transparency, and if labeling requirements do not reduce harm, what forms of intervention might be more effective? The literature has not yet provided definitive answers, but it points toward the need for independent, third-party algorithmic auditing and mechanisms for public accountability that go beyond platform self-reporting.

### 7.2 The shift from content to infrastructure

The literature reflects a significant evolution in analytical focus. Earlier scholarship (2023–2024) focused heavily on content moderation—deepfakes, hate speech, and disinformation as discrete problems. More recent work (2025–2026) emphasizes infrastructural power: who controls the "compute," the "foundation models," and the "distribution API" (Khanal et al., 2024; Rikap, 2024).

This shift has profound implications for governance. If democratic resilience depends less on individual pieces of content and more on who controls the underlying infrastructure, then regulatory frameworks focused on content will inevitably fall short. The emergence of "Sovereign AI" as a thematic response—the idea that nations must develop state-backed or public-interest AI infrastructure to remain independent of media oligarchs—reflects this recognition (Forum on Information and Democracy, 2025). Whether such sovereign infrastructure can be developed and maintained at sufficient scale remains an open question.

### 7.3 The fragility of shared reality

Perhaps the most urgent theme in the literature concerns the erosion of the "shared factual basis" necessary for democratic deliberation. The attention economy logic, combined with AI's ability to generate infinite content, creates a state where the public is not censored by the removal of information but by the overwhelming volume of conflicting synthetic narratives (López-Borrull and Lópezosa, 2025).

This leads to what scholars term epistemic fragmentation: different demographic groups exist in entirely different information realities, making consensus-building and policy debate nearly impossible (Chesterman, 2025). The implications for democratic resilience are severe. If citizens cannot agree on basic facts, the conditions for democratic accountability—including the ability to evaluate competing policy proposals and hold representatives accountable—may be fundamentally undermined.

### 7.4 Future research agenda

The literature reviewed here identifies several critical gaps that should guide future inquiry:

First, rigorous evaluation research is urgently needed. While the diversity of regulatory approaches has been documented, there is a scarcity of empirical evidence regarding which specific interventions successfully reduce disinformation or protect fundamental rights. Comparative effectiveness studies that track

outcomes across jurisdictions would move the field beyond theoretical design toward evidence-based policy.

Second, longitudinal studies on institutional trust are required. Rather than relying on snapshot analyses, future research should employ longitudinal designs to track changes in public discourse and trust alongside the phased deployment of AI. This is necessary to establish direct links between specific platform features or regulatory interventions and long-term measures of democratic health.

Third, qualitative attention to intersectionality is essential. While current literature identifies algorithmic bias as a technical risk, there is an urgent need for research that moves beyond quantitative measurements of bias to examine how marginalized communities—who experience overlapping forms of marginalization—personally perceive and navigate the risks of deepfakes and non-consensual synthetic media.

Fourth, continued scrutiny of the political economy is required to understand how ownership and control patterns respond to new regulatory interventions and whether these structures can be realigned to serve the public interest. The emergence of "Sovereign AI" models and public-interest infrastructure initiatives warrant particular attention.

## 8 Conclusion

The period 2023–2026 has been characterized by ambitious scope in both AI technology and its regulation. Yet this scoping review suggests that current governance models remain a work in progress, facing fundamental tensions that cannot be resolved through technical adjustments alone.

Several key findings emerge from the literature. First, regulatory frameworks are developing rapidly but remain fragmented. The contrast between the European Union's precautionary, rights-based approach and the United States' market-oriented framework creates significant challenges for global platforms and opportunities for regulatory arbitrage. This fragmentation suggests that international coordination—however difficult—may be necessary for effective governance.

Second, the political economy of AI reveals that power is increasingly concentrated through control over computing infrastructure and proprietary datasets. This structural asymmetry means that independent media plurality is threatened not just by algorithms but by dependency on leasing the very tools of production from Big Tech. Current regulatory frameworks focused on content ownership or transparency mandates do not adequately address this infrastructural power.

Third, the attention economy provides the underlying logic shaping AI deployment. Generative AI acts as a force multiplier for engagement-based business models, enabling production of "infinite variations" of sensational content at minimal cost. Regulatory frameworks that focus on content labeling without addressing the economic incentives driving distribution will likely produce at best symbolic compliance without substantive impact.

Fourth, the impact of AI-mediated media is not uniform. Deepfakes and synthetic media pose disproportionate risks to marginalized groups, yet regulatory frameworks often treat audiences as undifferentiated. The need for intersectional approaches that attend to differential impacts and incorporate diverse perspectives in governance processes is clear.

By bringing together literature on regulatory divergence, political economy, algorithmic bias, and the attention economy, this review reveals that the primary challenge for democratic resilience is not the regulation of individual AI applications but the governance of the infrastructure of public discourse. This requires moving beyond technical checklists to address the financial incentives that monetize automated engagement, transitioning from self-reported platform transparency to independent algorithmic auditing, and ensuring that democratic resilience includes the perspectives of those currently most vulnerable to algorithmic marginalization.

Ultimately, the regulation of AI in digital media is not a technical problem to be solved but a sociopolitical negotiation over the future of the public sphere. The period 2023–2026 represents a critical juncture in which foundational decisions about the structure of AI governance are being made. Only through sustained inquiry into regulatory efficacy, the economic incentives of the attention economy, and the lived experiences of marginalized communities can societies develop governance frameworks that ensure AI serves democratic values rather than undermining them.

## Author contributions

EM: Methodology, Formal analysis, Writing – review & editing, Writing – original draft. MM: Investigation, Methodology, Writing – review & editing. HA: Methodology, Writing – original draft, Data curation. IE: Writing – original draft, Resources, Data curation, Investigation. SA: Writing – original draft, Investigation, Methodology. MA: Resources, Investigation, Writing – review & editing, Methodology.

## Funding

The author(s) declared that financial support was not received for this work and/or its publication.

## Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. In a limited number of cases, generative AI was used to produce preliminary text fragments, which were then substantively rewritten, reorganized, and fact-checked by the authors to ensure originality, coherence, and compliance with journal policies.

## Publisher's note



## References

Arksey, H., and O'Malley, L. (2005). Scoping studies: towards a methodological framework. *Int. J. Soc. Res. Methodol.* 8 (1), 19–32. doi: 10.1080/1364557032000119616

Balabanova, N., Bashir, A., and Bova, P. (2025). Media and responsible AI governance: a game-theoretic and LLM analysis. *arXiv* [Preprint]. Available online at: https://arxiv.org/abs/2503.09858 (Accessed February, 2026).

Chesterman, S. (2025). Lawful but awful: evolving legislative responses to address online misinformation, disinformation, and mal-information in the age of generative AI. *arXiv* [Preprint]. Available online at: https://arxiv.org/abs/2505.15067

Citizen.org (2026). Generative influence: corporate lobbying and AI policy in 2025. Available online at: https://www.citizen.org/article/generative-influence/ (Accessed February, 2026).

Council of Europe (2024). Framework convention on artificial intelligence and human rights, democracy, and the rule of law. Council of Europe Treaty Series.

EPRS (2025). Information manipulation in the age of generative artificial intelligence. European Parliamentary Research Service. Available online at: https://www.europarl.europa.eu/RegData/etudes/BRIE/2025/779259/EPRS_BRI(2025)779259_EN.pdf (Accessed February, 2026).

Forum on Information and Democracy. (2025). Policy manual on digital platforms and information integrity. Available online at: https://informationdemocracy.org/wp-content/uploads/2025/10/PolicyManual_Digital_Final.pdf (Accessed February, 2026).

Gutiérrez-Caneda, B., Lindén, C.-G., and Vázquez-Herrero, J. (2024). Ethics and journalistic challenges in the age of AI: talking with professionals and experts. *Front. Commun.* 9, 1465178. doi: 10.3389/fcomm.2024.1465178

Hastuti, H. (2025). Algorithmic influence and media legitimacy: a systematic review. *Front. Commun.* 11. doi: 10.3389/fcomm.2025.1667471

Khanal, S., Zhang, H., and Taeihagh, A. (2024). Why and how is the power of Big Tech increasing in the policy process? The case of generative AI. *Policy Soc.* 44 (1), 52–69. doi: 10.1093/polsoc/puae012

López-Borrull, A., and Lópezosa, C. (2025). Mapping the impact of generative AI on disinformation: insights from a scoping review. *Publications* 13 (3), 33. doi: 10.3390/publications13030033

Park, S., and Nan, X. (2025). Generative AI and misinformation: a scoping review of the role of generative AI in the generation, detection, mitigation, and impact of misinformation. *AI. Soc.* 41, 1501–1515. doi: 10.1007/s00146-025-02620-3

Peña-Alonso, U., Peña-Fernández, S., and Meso-Ayerdi, K. (2025). Journalists' perceptions of artificial intelligence and disinformation risks. *arXiv* [Preprint]. Available online at: https://arxiv.org/abs/2509.01824 (Accessed December 25, 2025).

Ricciardi Celsi, L., and Zomaya, A. Y. (2025). Perspectives on managing AI ethics in the digital age. *Information* 16 (4), 318. doi: 10.3390/info16040318

Rikap, C. (2024). Dynamics of corporate governance beyond ownership in AI. Common wealth report. Available online at: https://www.commonwealth.org/publications/dynamics-of-corporate-governance-beyond-ownership-in-ai (Accessed December 25, 2025).

Singh, S. (2025). Unmasking digital deceptions: an integrative review of deepfakes. *MethodsX.*, doi: 10.1016/j.meex.2025.102476

Taeihagh, A. (2025). Governance of generative AI. *Policy Soc.* 44 (1), 1–22. doi: 10.1093/polsoc/puaf001

Thomas, J., and Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC. Med. Res. Methodol.* 8 (1), 45. doi: 10.1186/1471-2288-8-45

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., et al. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann. Intern. Med.* 169 (7), 467–473. doi: 10.7326/M18-0850

Vosoughi, S., Roy, D., and Aral, S. (2018). The spread of true and false news online. *Science* 359 (6380), 1146–1151. doi: 10.1126/science.aap9559