

Algorithmic bias and media manipulation: A systematic review of AI's Role in shaping public perception and political discourse

Amaka Peace Onebunne *

Department of Communication, Northern Illinois University.

World Journal of Advanced Research and Reviews, 2022, 16(03), 1239-1249

Publication history: Received on 01 November 2022; revised on 14 December 2022; accepted on 17 December 2022

Article DOI: <https://doi.org/10.30574/wjarr.2022.16.3.1332>

Abstract

Artificial intelligence's (AI) influence on public opinion and political discourse necessitates immediate attention as it gets more and more integrated into media infrastructures. This systematic review synthesizes findings from 25 peer-reviewed papers published between 2015 and 2022 to investigate the relationship between algorithmic bias, media manipulation, and public trust. Guided by the PRISMA 2020 framework, the review examines how AI-driven systems impact content curation, magnify false information, and mediate political communication. Five themes emerge from the analysis: political manipulation, public trust and democratic engagement, algorithmic fairness, AI in media and journalism, and the changing nature of ethical and legal governance. Findings reveal that while AI systems can provide content more efficiently, they also reinforce bias, reduce the diversity of information, and operate with limited accountability or transparency. The review emphasizes the need for interdisciplinary approaches to algorithmic governance, especially as societies struggle with the democratic consequences of opaque and increasingly influential media technologies

Keywords: Algorithmic Bias; Artificial Intelligence; Media Manipulation; Misinformation; Political Discourse; Systematic Review; PRISMA

1. Introduction

The integration of artificial intelligence (AI) into digital media systems has changed how information is curated, consumed, and circulated. While these technologies offer unprecedented efficiency and user-specific content delivery, they also raise pressing concerns about algorithmic bias, media manipulation, and the integrity of democratic discourse (Milano et al., 2021; Bandy, 2021; Dolata et al., 2022).

Bias in algorithms refers to systematic and repeatable errors in AI decision-making processes that result in unfair outcomes, which more often than not reflect or intensify societal biases present in the training data (Mohseni, Zarei, & Ragan, 2021; Orphanou et al., 2022). In the media context, these biases may show up in the ranking, filtering, or suppression of news, which ultimately influences the information that people see and accept. Platforms like Facebook, YouTube, and Twitter rely on AI-powered systems that rank content according to engagement metrics, such as clicks, likes, and shares, rather than accuracy or social value. As a result, these platforms have been linked to the spread of dis and misinformation, political polarization, and the decline in public confidence in traditional news organizations (Shorey & Howard, 2016; Glikson & Woolley, 2020; Modha et al., 2020).

Public discourse is increasingly mediated by opaque algorithmic processes, especially in politically sensitive contexts. With this, there are two risks involved. First, users now receive information filtered through potentially biased systems, and secondly, the very structure of media ecosystems is being reconfigured to serve computational logics over

* Corresponding author: Amaka Peace Onebunne.

journalistic ethics. As scholars have pointed out, the prevalence of algorithmic decision-making in media raises important ethical, epistemological, and regulatory challenges that must be critically examined (Ferretti et al., 2021; Tan et al., 2021; Fosch-Villaronga & Poulsen, 2022).

Despite growing scholarship on AI and media, there is a lack of consolidated knowledge concerning the specific role algorithmic systems play in shaping political discourse and public perception. Existing studies tend to focus either on the technological mechanics of AI or the sociopolitical consequences of misinformation, with limited synthesis across both domains (Bastian, Makhortykh, & Dobber, 2019; Ma, 2020). This fragmented understanding hampers efforts to develop comprehensive regulatory and design frameworks that ensure both efficiency and equity in digital media environments.

This paper seeks to fill this gap by conducting a systematic review of scholarly literature on algorithmic bias and media manipulation, with a focus on AI's role in shaping public perception and political discourse. Using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework (Page et al., 2021), the review synthesizes peer-reviewed studies from 2015 to 2022, categorizing findings into thematic areas such as misinformation amplification, content moderation, ideological polarization, and algorithmic governance.

1.1. Understanding Algorithmic Bias

Algorithmic bias refers to systematic patterns of error that occur when artificial intelligence systems produce outcomes that disadvantage certain individuals or groups, often mirroring existing social inequalities (Mohseni, Zarei, & Ragan, 2021; Dolata, Feuerriegel, & Schwabe, 2022). These biases can arise from various sources such as imbalanced training data, flawed assumptions in model design, or even the optimization goals set by developers (Orphanou et al., 2022). While algorithmic decision-making is often perceived as objective or neutral, it is increasingly clear that algorithms can reproduce and even intensify forms of bias present in the data they learn from (Milano et al., 2021; Ferretti et al., 2021).

When it comes to media, algorithmic bias manifests in how content is recommended, filtered, or ranked. News feeds, trending lists, and search suggestions are not simply aggregations of popular content; they are outputs of complex algorithmic processes that weigh engagement metrics, user behavior, and other variables—many of which are hazy to users (Glikson & Woolley, 2020; Bastian, Makhortykh, & Dobber, 2019). As a result, some viewpoints are elevated while others are sidelined, shaping how audiences understand current events and social issues (Shorey & Howard, 2016).

What makes algorithmic bias particularly concerning in media environments is its invisibility. Users are rarely aware of the filtering mechanisms behind the content they encounter, which gives biased outputs an air of neutrality (Wang et al., 2022). This can have serious implications for information access and public opinion, especially when biased recommendations reinforce stereotypes, marginalize minority viewpoints, or deepen ideological divisions (Monteith & Glenn, 2016; Orphanou et al., 2022).

A growing body of scholarship has called attention to these dynamics. Scholars such as Noble (2018) and Benjamin (2019) argue that algorithmic systems not only reflect structural inequities but also actively participate in their reproduction. In media settings, this means algorithms are not just technical tools but also powerful agents that shape cultural narratives and influence what counts as legitimate knowledge.

Understanding algorithmic bias, then, is foundational to any investigation of AI's role in media manipulation. It sets the stage for examining how seemingly neutral technologies participate in the construction of public discourse and the selective visibility of information.

1.2. AI-Powered Media and Content Curation

AI-driven content curation has become the default mode of information delivery across digital platforms. Recommendation algorithms now determine what news users see, which videos trend, and whose posts receive visibility. These systems operate by learning from user preferences, behaviors, and interaction histories to generate personalized feeds that prioritize engagement (Glikson & Woolley, 2020; Mohseni, Zarei, & Ragan, 2021).

This shift from manual editorial selection to automated recommendation has restructured the media experience. Rather than encountering a diverse range of content curated by journalists or editors, users now receive information tailored to their presumed interests, often without understanding how or why those selections are made. This personalization, while efficient, subtly redefines relevance in media as what performs well statistically, rather than what is contextually or socially significant.

Critically, these AI systems do not simply organize information, they help construct the informational environment itself. Their design and logic influence what becomes visible and what is effectively silenced. In doing so, they shape public awareness, guide attention, and influence what topics emerge as culturally or politically salient (Milano et al., 2021; Shorey & Howard, 2016).

Several studies have examined the implications of this algorithmic gatekeeping, particularly in terms of transparency and user autonomy. Concerns include the lack of user control over curated feeds, the invisibility of excluded content, and the difficulty of auditing how recommendations are generated (Dolata, Feuerriegel, & Schwabe, 2022; Bandy, 2021). While platforms often promote these systems as neutral tools for enhancing user experience, the underlying processes remain largely proprietary and unaccountable (Kerry et al., 2020).

Understanding AI-powered content curation is therefore foundational to unpacking how algorithmic systems participate in shaping media ecosystems. It sets the stage for examining how these mechanisms can, intentionally or not, contribute to the spread of misinformation and the distortion of public discourse.

1.3. Amplification of Misinformation and Disinformation

One of the most widely acknowledged consequences of algorithmic media systems is their tendency to amplify misinformation and disinformation. While misinformation refers to false or misleading content shared without malicious intent, disinformation is the intentional spread of falsehoods for political, financial, or ideological gain. AI-powered platforms, though not designed to mislead, often create the conditions under which such content flourishes (Milano et al., 2021; Modha et al., 2020).

Recommendation algorithms are optimized for engagement, measured through clicks, shares, watch time, and other behavioral signals. Research has shown that emotionally charged, controversial, or sensational content tends to perform better by these metrics, regardless of its accuracy (Shorey & Howard, 2016; Glikson & Woolley, 2020). As a result, false or misleading information is not only more likely to be surfaced by algorithmic systems but may also outpace factual content in reach and virality.

This phenomenon has been documented across multiple platforms. While not part of this review, studies such as Vosoughi, Roy, and Aral (2018) and Ribeiro et al. (2020) provide supporting evidence that false news spreads faster than truth and that recommendation systems can guide users toward increasingly extreme content. Within the included literature, Milano et al. (2021) emphasize how algorithmic targeting contributes to epistemic fragmentation, allowing disinformation to thrive in fragmented and ideologically insular media spaces.

Compounding the issue is the scale and speed of content distribution. AI systems process and disseminate information at rates far beyond the capacity for manual oversight or correction. Fact-checking efforts, though essential, often arrive too late to counter the initial impact of viral falsehoods (Ferretti et al., 2021). Moreover, the algorithms themselves are not transparent, making it difficult for researchers or users to trace how and why certain content gains traction (Kerry et al., 2020; Bandy, 2021).

Importantly, the amplification of misinformation is not an isolated technical flaw, but a systemic outcome of incentive structures built into the design of AI platforms. When media systems reward attention over accuracy, falsehoods become algorithmically competitive. This dynamic poses serious challenges for democratic societies, where informed decision-making depends on access to reliable information.

Addressing the amplification of misinformation will require more than content moderation or fact-checking; it demands a deeper interrogation of how algorithmic infrastructures shape the visibility and velocity of information in the first place, especially in the political environment (Milano et al., 2021; Orphanou et al., 2022).

1.4. Algorithmic Influence on Political Discourse

The intersection of algorithmic systems and political discourse presents one of the most consequential challenges in contemporary media studies. As artificial intelligence increasingly mediates the flow of political information, it also shapes how individuals engage with political content, form opinions, and participate in public life (Shorey & Howard, 2016; Ma, 2020). The result is a political landscape that is no longer just influenced by media framing or partisanship, but also by the opaque logic of algorithmic curation.

One of the most salient concerns is the role of algorithms in reinforcing ideological echo chambers. By continuously serving users content aligned with their existing beliefs, AI-driven platforms can reduce exposure to differing

perspectives and even create a faux sense of political homogeneity. Ironically, this personalization, while often subtle, can contribute to political polarization, as users become less likely to encounter alternative viewpoints or engage in cross-cutting dialogue (Glikson & Woolley, 2020; Bastian, Makhortykh, & Dobber, 2019). The narrowing of informational diversity not only stifles democratic deliberation but also undermines the potential for consensus in pluralistic societies.

Another dimension of algorithmic influence is the strategic use of these systems for political messaging. Campaigns, advocacy groups, and interest-based actors have increasingly milked microtargeting tools to deliver personalized political advertisements, often designed to exploit psychological profiles or behavioral tendencies (Kerry et al., 2020; Tan et al., 2021). This form of persuasion, enabled by AI, raises ethical questions about manipulation, consent, and the boundaries of political influence in the digital public sphere

Moreover, AI plays a growing role in content moderation, especially in deciding which political messages are flagged, downranked, or removed. While moderation is important to prevent harm and maintain platform integrity, the automation of these processes introduces new risks. Political speech may be disproportionately silenced or left unchecked depending on how moderation algorithms are trained and deployed. In both cases, algorithmic decisions lack the transparency and contextual sensitivity that complex political discourse often requires (Bandy, 2021; Dolata, Feuerriegel, & Schwabe, 2022).

Studies have also drawn attention to the uneven impact of algorithmic influence across different populations. Marginalized communities may be more vulnerable to disinformation campaigns or algorithmic suppression, an act which further entrenches existing power imbalances in political communication (Modha et al., 2020; Ferretti et al., 2021). In this way, algorithmic systems do not just mediate politics, they help constitute it.

Understanding how AI shapes political discourse requires attention not only to what content is seen but also to the mechanisms that determine its visibility, framing, and interpretation. While existing literature provides important perspectives into these dynamics, findings remain scattered across disciplines and vary in methodological rigor. To bring clarity to this complex field, the present study conducts a systematic review of peer-reviewed research, identifying patterns, gaps, and recurring themes in the scholarship on algorithmic media manipulation.

2. Materials and Methods

This study employed a systematic review methodology guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) framework (Page et al., 2021). The aim was to identify, evaluate, and synthesize scholarly literature examining the relationship between algorithmic bias, AI-driven media systems, public perception, and political discourse. Given the interdisciplinary nature of the topic, this approach provided a structured and transparent process for navigating the breadth of relevant academic work.

2.1. Search Strategy

The search process was conducted exclusively through Google Scholar, chosen for its comprehensive indexing across disciplines. Five distinct keyword iterations were designed to capture various dimensions of the research topic, ranging from core concepts (e.g., “algorithmic bias,” “media manipulation”) to more specific lenses such as “political discourse,” “misinformation,” “content moderation,” and “AI regulation.” Boolean operators (AND, OR) were used to structure the search logic and increase retrieval relevance (see Appendix A). The review covered literature published between January 2015 and December 2022, specially making sure to include foundational and recent studies during a period marked by significant growth in AI adoption across media systems. This date range also reflects the emergence of global debates around algorithmic accountability, misinformation, and platform regulation.

Each search iteration generated a different volume of results

- Iteration 1: 5 results (0 reviews)
- Iteration 2: 23 results (0 reviews)
- Iteration 3: 190 results (10 reviews)
- Iteration 4: 1 result (0 reviews)
- Iteration 5: 269 results (15 reviews)
- In total, 488 records were screened for relevance.

2.2. Inclusion and Exclusion Criteria

To ensure focus and quality, studies were selected based on the following criteria:

2.2.1. Inclusion Criteria

- Peer-reviewed journal articles or systematic literature reviews
- Published between 2015 and 2022
- Explicitly addressed AI in media, algorithmic bias, misinformation, political communication, or content curation
- Written in English

2.2.2. Exclusion Criteria

- Technical papers without social or political analysis
- Non-peer-reviewed sources (e.g., editorials, blog posts, preprints)
- Duplicate or non-English studies

2.3. Screening and selection process

The screening and selection process followed the PRISMA 2020 guidelines to ensure a transparent and replicable approach to study inclusion. A total of 488 records were initially identified through Google Scholar using five carefully constructed keyword iterations that explored the intersection of algorithmic bias, media manipulation, AI, and political discourse (see Appendix A). After applying the inclusion and exclusion criteria at the full-text stage, 25 studies were selected for inclusion in the final review. The screening process and reasons for exclusion were documented using the PRISMA flow model.

2.3.1. PRISMA Flow Diagram

The flow of information through the phases of the review process is presented in Figure 1, adapted from the PRISMA 2020 Statement. It illustrates the identification, screening, eligibility assessment, and inclusion stages.

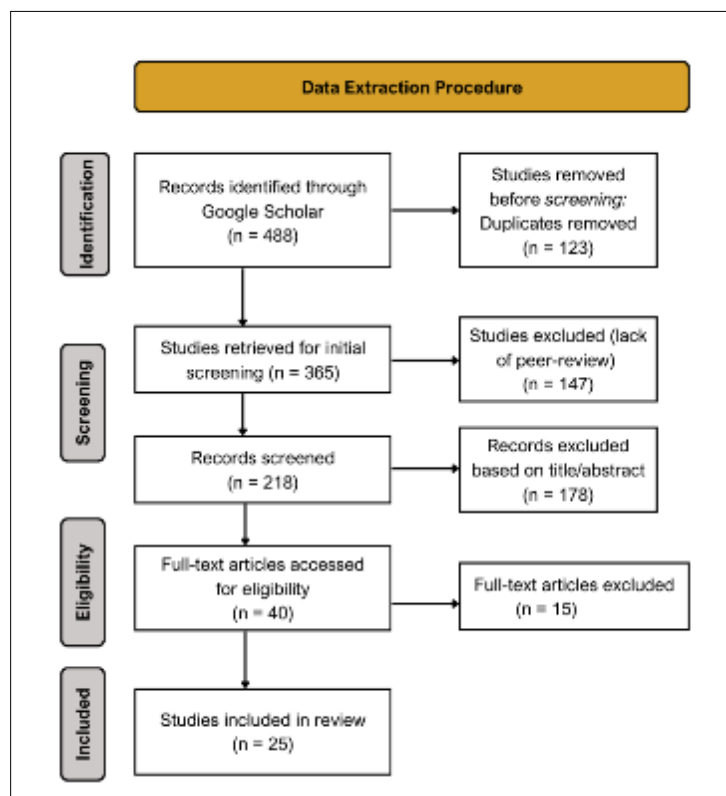


Figure 1 PRISMA Flow Diagram

2.4. Data Extraction and Thematic Synthesis

Key information, including publication year, author(s), study focus, and theoretical framing, was extracted from each included study. A thematic synthesis was conducted to identify recurring concepts and patterns across the selected literature, see Figure 2. These themes informed the structure of the Results and Discussion section, focusing on content curation, misinformation amplification, political discourse shaping, and broader sociotechnical implications.

Table 1 Systematic Review Summary

| Citation | Focus | Thematic Category |
|-----------------------------------|--|-----------------------------------|
| Buhmann et al. (2019) | Bibliometric review of Habermas in PR | Media/Political Theory |
| Leask (2017) | Data journalism education in Canada | AI in Media & Journalism |
| Mohseni et al. (2021) | Explainable AI design and evaluation | Bias & Fairness in AI |
| Kerry et al. (2020) | Federal privacy legislation in AI | Legal & Policy Frameworks |
| Glikson & Woolley (2020) | Human trust in AI | Public Trust & Perception |
| Veglis et al. (2022) | Big data in media organizations | AI in Media & Journalism |
| Pero et al. (2022) | Ethical risks of AI in the workplace | Legal & Ethical Issues |
| Wang et al. (2022) | Age-appropriate AI for children | Ethical Frameworks |
| Miller (2021) | Moral decision-making in AI projects | Ethics & Decision-Making |
| Monteith & Glenn (2016) | AI and mental illness | Public Trust & Vulnerability |
| Bastian et al. (2019) | News personalization and conflict | AI in Media & Journalism |
| Ma (2020) | Digital tech and democracy | Public Perception & Democracy |
| Langman et al. (2021) | Roboethics in EU and NA | Ethics & Governance |
| Veluru (2022) | AI in journalism automation | AI in Media & Journalism |
| Maha Abdulmajeed and Fahmy (2022) | AI in Arab journalism | AI in Media & Journalism |
| Bandy (2021) | Systematic review of algorithm audits | Bias & Fairness in AI |
| Fosch-Villaronga & Poulsen (2022) | Diversity and inclusion in AI | Ethics & Governance |
| Ferretti et al. (2021) | Ethics review of big data research | Legal & Ethical Frameworks |
| Dolata et al. (2022) | Sociotechnical view of fairness | Bias & Fairness in AI |
| Milano et al. (2021) | Governance of online targeting | Policy & Algorithmic Impact |
| Hoebanx (2022) | Social science views on the internet | Public Perception & Democracy |
| Orphanou et al. (2022) | Bias mitigation in algorithmic systems | Bias & Fairness in AI |
| Modha et al. (2020) | Hate speech tracking on social media | Misinformation & Moderation |
| Tan et al. (2021) | Gig economy ethics | Ethics & AI Governance |
| Shorey & Howard (2016) | Automation, big data, and politics | Misinformation & Political Impact |

3. Results and Discussion

This section presents a thematic synthesis of 25 peer-reviewed studies selected through the PRISMA-informed systematic review of literature published between 2015 and 2022. These studies span disciplines including communication, law, computer science, journalism, and ethics, reflecting the interdisciplinary nature of scholarship on algorithmic bias and media manipulation. The findings are grouped into five dominant themes: (1) Bias and Fairness in Algorithmic Systems, (2) AI in Media and Journalism, (3) Public Trust and Democratic Participation, (4) Misinformation and Political Manipulation, and (5) Legal, Ethical, and Policy Frameworks.

3.1. Bias and Fairness in Algorithmic Systems

Concerns around fairness and transparency are central in the literature assessing how AI systems operate in media environments. Studies such as Mohseni et al. (2021), Bandy (2021), and Dolata et al. (2022) provide frameworks for identifying and mitigating algorithmic bias. These biases stem not only from skewed training data but also from opaque design choices and optimization goals that prioritize engagement over equity.

Research by Orphanou et al. (2022) and Ferretti et al. (2021) further emphasize the social implications of biased AI, pointing to underrepresentation, unequal treatment, and feedback loops that reinforce systemic inequities. Fosch-Villaronga and Poulsen (2022) extend this conversation by exploring inclusion gaps and the need for diversity-sensitive AI design. Together, these studies illustrate a consistent call for algorithmic audits, fairness metrics, and cross-sector collaboration to prevent the entrenchment of discrimination in AI-driven platforms.

3.2. AI in Media and Journalism

AI's role in media production and distribution is a recurring theme, with a growing emphasis on how it affects journalistic integrity and news personalization. Scholars such as Veglis et al. (2022) and Veluru (2022) examine how big data analytics and AI tools are transforming editorial workflows, content creation, and distribution. Maha Abdulmajeed and Fahmy (2022) analyze these trends within the Arab journalism context, noting both the potential for innovation and the challenges of cultural adaptation.

Leask (2017) and Bastian et al. (2019) raise concerns about editorial standards, by suggesting that algorithmic gatekeeping may reduce the range of viewpoints or prioritize conflict-based narratives. While automation may enhance efficiency and audience targeting, it also complicates traditional notions of journalistic responsibility and editorial judgment.

3.3. Public Trust and Democratic Participation

Several studies examine the ways in which AI influences public trust, knowledge, and civic engagement. Glikson and Woolley (2020) synthesize empirical findings on human trust in AI, identifying factors such as system transparency, predictability, and human-like features that shape public confidence. Ma (2020) and Monteith and Glenn (2016) explore how automated decision-making and data-driven systems affect marginalized populations, revealing risks of exclusion and algorithmic invisibility. Hoebanx (2022) offers a thorough overview of social science perspectives on how digital infrastructures shape information landscapes. These works converge on a common theme: trust in media, and by extension, in democratic institutions, may decline when users lack agency or awareness of how content is curated.

3.4. Misinformation and Political Manipulation

A subset of the literature focuses explicitly on the amplification of misinformation and the algorithmic dynamics that enable political manipulation. Shorey and Howard (2016) and Modha et al. (2020) draw attention to the role of big data and automation in spreading false information and hate speech, often without sufficient protection.

Milano et al. (2021) argue that the epistemic fragmentation caused by targeted content personalization threatens the integrity of public discourse. This fragmentation is worsened by the business models of platforms that reward content virality over veracity. In politically sensitive environments, as shown by Miller (2021), such dynamics can be exploited to sway public opinion, weaken consensus, and undermine deliberative democracy.

3.5. Legal, Ethical, and Policy Frameworks

The ethical and regulatory dimensions of AI use in media are discussed across a range of studies. Kerry et al. (2020) and Pero et al. (2022) outline emerging regulatory responses to AI and data privacy, focusing particularly on U.S. and European contexts. Tan et al. (2021) and Langman et al. (2021) explore the broader societal implications of algorithmic labor and roboethics, advocating for normative guidelines and accountability standards.

While frameworks such as the GDPR and proposals like the EU AI Act are frequently cited, many scholars, including Ferretti et al. (2021) and Wang et al. (2022), note that enforcement measures are still limited, especially when applied to global media platforms. A common recommendation across these works is the development of interdisciplinary, cross-jurisdictional approaches to AI governance, ones that center transparency, inclusivity, and participatory oversight.

Across the five themes, there is a consistent idea. AI systems are reshaping media in ways that both reflect and reproduce existing power structures. Whether through biased algorithms, opaque content curation, or under-regulated personalization engines, these technologies are not just technical innovations, they function as political actors nested in sociotechnical systems. While the literature points to promising tools and ethical frameworks, it also highlights significant gaps in empirical research, particularly in non-Western contexts and among vulnerable populations.

3.6. Ethical, Legal, and Governance Challenges

As AI technologies become increasingly central to the organization of media and information systems, they raise critical ethical and regulatory questions that extend beyond technical design. The literature reviewed in this study supports a growing consensus, that AI is not just a tool for enhancing efficiency or personalization, it is a social force capable of influencing perception, behavior, and the democratic process. As such, the development, deployment, and governance of these systems demand serious ethical reflection and policy intervention.

One of the most pressing concerns identified across the literature is the lack of transparency in algorithmic decision-making. Studies by Mohseni et al. (2021) and Dolata et al. (2022) emphasize the opacity of AI systems and the limited interpretability of their outputs, which complicates efforts to identify and address harmful outcomes such as discrimination or misinformation amplification. Without accessible explanations or auditability, it becomes difficult to establish accountability or secure user trust.

Closely tied to transparency is the issue of fairness. As Orphanou et al. (2022) and Bandy (2021) argue, bias in AI systems is often a reflection of larger social inequalities ingrained in data and institutional structures. When deployed in media contexts, these biases influence which voices are heard, whose stories are prioritized, and how political narratives are framed. The resulting asymmetries in visibility and access have profound implications for equity and representation in public discourse.

From a legal perspective, several studies highlight gaps in current governance structures. While frameworks such as the General Data Protection Regulation (GDPR) in Europe offer baseline protections for data privacy and automated decision-making, enforcement remains uneven, particularly when applied to multinational tech platforms. As Kerry et al. (2020) and Ferretti et al. (2021) note, the absence of a robust federal framework in the United States, and the piecemeal nature of global regulations, limits the ability to respond to cross-border algorithmic harms.

At the same time, emerging proposals such as the EU Artificial Intelligence Act (under discussion as of 2022) suggest a growing recognition of the need for risk-based approaches to AI regulation. These proposals seek to classify AI systems based on potential harm and impose obligations related to transparency, human oversight, and auditability. However, as scholars like Milano et al. (2021) caution, regulatory efforts must also contend with the epistemic fragmentation that characterizes the AI landscape—fragmentation not only in technical standards but in ethical frameworks and institutional capacities.

Beyond law and regulation, there is also an ethical imperative to incorporate normative values into the design and deployment of AI systems. This includes commitments to inclusivity, accessibility, and cultural sensitivity, as advocated by Fosch-Villaronga and Poulsen (2022) and Wang et al. (2022). Ethical AI, in this sense, requires more than reactive governance, it demands proactive, participatory design processes that center affected communities and account for diverse sociotechnical contexts.

Finally, the literature points to the need for interdisciplinary collaboration in addressing the challenges posed by algorithmic media systems. Engineers, designers, journalists, ethicists, and policymakers must work together to create mechanisms for transparency, fairness, and accountability. Without such coordination, the risks of algorithmic harm will continue to grow, particularly for marginalized groups, whose experiences are often underrepresented in both datasets and decision-making forums.

4. Conclusion

This systematic review set out to examine how algorithmic bias and AI-driven media systems influence public perception and political discourse. Drawing from 25 peer-reviewed studies published between 2018 and 2022, the review identified key patterns and concerns across five interrelated domains: algorithmic fairness, media and journalism innovation, public trust, misinformation, and governance frameworks. Together, these thematic areas highlight the complex and varied ways in which AI is influencing how information is curated, consumed, and contested in digital environments.

The findings suggest that while AI offers effective tools for media personalization and content delivery, it also reinforces structural inequalities, narrows the diversity of public discourse, and amplifies content that prioritizes engagement over accuracy. This is particularly evident in politically sensitive contexts, where algorithmic systems influence not only what information circulates but also how citizens form opinions, engage with institutions, and participate in civic life.

Importantly, the review reveals that discussions of algorithmic bias in media cannot be disentangled from broader questions of ethics, transparency, and accountability. Without clearer standards for explainability, fairness, and regulatory oversight, the societal risks of AI-driven media systems, especially their potential to manipulate, misinform, or marginalize, remain significant.

Moving forward, this review highlights several opportunities for future research and policy development

- First, there is a need for more empirical studies that investigate how algorithmic systems operate in non-Western and underrepresented media contexts. Much of the current literature focuses on North American and European platforms, leaving important regional dynamics underexplored.
- Second, researchers should pursue interdisciplinary approaches that bridge technical and social perspectives. Understanding how AI affects public life requires input from computer scientists, communication scholars, ethicists, legal experts, and affected communities.
- Third, policymakers and platform developers must work toward building transparent, auditable, and participatory AI systems, particularly in domains that affect democratic institutions and information access.

As AI continues to evolve and becomes more integrated into the architecture of digital communication, the need for critical, collaborative, and justice-oriented approaches becomes all the more urgent. This review contributes to that effort by offering a consolidated account of the current state of knowledge, and by pointing toward paths for more equitable and accountable algorithmic futures.

References

- [1] Bandy, J. (2021). Problematic machine behavior: A systematic literature review of algorithm audits. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), 1–34. <https://doi.org/10.1145/3449148>
- [2] Bastian, M., Makhortykh, M., & Dobber, T. (2019). News personalization for peace: how algorithmic recommendations can impact conflict coverage. *International Journal of Conflict Management*, 30(3), 309–328. <https://doi.org/10.1108/ijcma-02-2019-0032>
- [3] Benjamin, R. (2019). *Race after technology: Abolitionist Tools for the New Jim Code*. Polity.
- [4] Buhmann, A., Ihlen, Ø., & Aaen-Stockdale, C. (2019). Connecting the dots: a bibliometric review of Habermasian theory in public relations research. *Journal of Communication Management*, 23(4), 444–467. <https://doi.org/10.1108/jcom-12-2018-0127>
- [5] Dolata, M., Feuerriegel, S., & Schwabe, G. (2021). A sociotechnical view of algorithmic fairness. *Information Systems Journal*, 32(4), 754–818. <https://doi.org/10.1111/isj.12370>
- [6] Ferretti, A., Ienca, M., Sheehan, M., Blasimme, A., Dove, E. S., Farsides, B., Friesen, P., Kahn, J., Karlen, W., Kleist, P., Liao, S. M., Nebeker, C., Samuel, G., Shabani, M., Rivas Velarde, M., & Vayena, E. (2021). Ethics review of big data research: What should stay and what should be reformed? *BMC Medical Ethics*, 22(1), 51. <https://doi.org/10.1186/s12910-021-00616-4>
- [7] Fosch-Villaronga, E., & Poulsen, A. (2022). Diversity and Inclusion in Artificial Intelligence. In M. Ebers, S. N. Wurster, & A. Spiecker (Eds.), *Law and Artificial Intelligence: Regulating AI and Applying AI in Legal Practice* (Vol. 35, pp. 109–134). Springer. https://doi.org/10.1007/978-94-6265-523-2_6
- [8] Glikson, E., & Woolley, A. W. (2020). Human Trust in Artificial Intelligence: Review of Empirical Research. *Academy of Management Annals*, 14(2), 627–660. <https://doi.org/10.5465/annals.2018.0057>
- [9] Hoebanx, P. (2022). What shapes the internet? An overview of social science and interdisciplinary perspectives. *Sociology Compass*, 16(10). <https://doi.org/10.1111/soc4.13032>
- [10] Kerry, C., Morris, J., Chin, C., & Lee, N. (2020). *Bridging the gaps: A path forward to federal privacy legislation*. Brookings Institutions. https://www.brookings.edu/wp-content/uploads/2020/06/Bridging-the-gaps_a-path-forward-to-federal-privacy-legislation.pdf

- [11] Langman, S., Capicotto, N., Maddahi, Y., & Zareinia, K. (2021). Roboethics principles and policies in Europe and North America. *SN Applied Sciences*, 3(12), 857. <https://doi.org/10.1007/s42452-021-04853-5>
- [12] Leask, J. (2017). Data journalism education in Canada: Challenges and opportunities. <https://doi.org/10.7939/r3513vb35>
- [13] Ma, L. (2020). Rethinking democratizing potential of digital technology: A review of technology and communication studies. *Journal of Information, Communication and Ethics in Society*, 18(1), 140–156. <https://doi.org/10.1108/jices-02-2019-0022>
- [14] Maha Abdulmajeed, & Fahmy, N. (2022). Meta-analysis of AI Research in Journalism: Challenges, Opportunities and Future Research Agenda for Arab Journalism. *Springer EBooks*, 213–225. https://doi.org/10.1007/978-3-031-17746-0_18
- [15] Milano, S., Mittelstadt, B., Wachter, S., & Russell, C. (2021). Epistemic fragmentation poses a threat to the governance of online targeting. *Nature Machine Intelligence*, 3(6), 466–472. <https://doi.org/10.1038/s42256-021-00358-3>
- [16] Miller, G. (2021). Artificial Intelligence Project Success Factors: Moral Decision-Making with Algorithms. *Proceedings of the 16th Conference on Computer Science and Intelligence Systems (FedCSIS)*, 379–390. <https://doi.org/10.15439/2021f26>
- [17] Modha, S., Mandl, T., Majumder, P., & Patel, D. (2020). Tracking Hate in Social Media: Evaluation, Challenges and Approaches. *SN Computer Science*, 1(2), 105. <https://doi.org/10.1007/s42979-020-0082-0>
- [18] Mohseni, S., Zarei, N., & Ragan, E. D. (2021). A Multidisciplinary Survey and Framework for Design and Evaluation of Explainable AI Systems. *ACM Transactions on Interactive Intelligent Systems*, 11(3-4), 1–45. <https://doi.org/10.1145/3387166>
- [19] Monteith, S., & Glenn, T. (2016). Automated Decision-Making and Big Data: Concerns for People With Mental Illness. *Current Psychiatry Reports*, 18(12). <https://doi.org/10.1007/s11920-016-0746-6>
- [20] Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press.
- [21] Orphanou, K., Otterbacher, J., Kleanthous, S., Batsuren, K., Giunchiglia, F., Bogina, V., Tal, A. S., Hartman, A., & Kuflik, T. (2022). Mitigating Bias in Algorithmic Systems—A Fish-eye View. *ACM Computing Surveys*, 55(5), 1–37. <https://doi.org/10.1145/3527152>
- [22] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., & McGuinness, L. A. (2021). The PRISMA 2020 statement: an Updated Guideline for Reporting Systematic Reviews. *British Medical Journal*, 372(71). <https://doi.org/10.1136/bmj.n71>
- [23] Pero, A. S. del, Wyckoff, P., & Vourc'h, A. (2022). Using Artificial Intelligence in the workplace: What are the main ethical risks? *Www.oecd-ilibrary.org*, 273(273). <https://doi.org/10.1787/840a2d9f-en>
- [24] Ribeiro, M. H., Ottoni, R., West, R., Almeida, V. A. F., & Meira, W. (2020). Auditing radicalization pathways on YouTube. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 131–141. <https://doi.org/10.1145/3351095.3372879>
- [25] Shorey, S., & Howard, P. N. (2016). Automation, Algorithms, and Politics| Automation, Big Data and Politics: A Research Review. *International Journal of Communication*, 10(0), 24. <https://ijoc.org/index.php/ijoc/article/view/6233>
- [26] Tan, Z. M., Aggarwal, N., Cowls, J., Morley, J., Taddeo, M., & Floridi, L. (2021). The Ethical Debate about the Gig economy: a Review and Critical Analysis. *Technology in Society*, 65(1), 101594. <https://doi.org/10.1016/j.techsoc.2021.101594>
- [27] Veglis, A., Saridou, T., Panagiotidis, K., Karypidou, C., & Ktenidis, E. (2022). Applications of Big Data in Media Organizations. *Social Sciences*, 11(9), 414. <https://doi.org/10.3390/socsci11090414>
- [28] Veluru, C. S. (2022). Harnessing AI for Transformative Journalism: Advancements in Automated Content Creation and Analysis. *European Journal of Advances in Engineering and Technology*, 9(9), 85–92. <https://doi.org/10.5281/zenodo.13325954>
- [29] Vosoughi, S., Roy, D., & Aral, S. (2018). The Spread of True and False News Online. *Science*, 359(6380), 1146–1151. <https://www.science.org/doi/10.1126/science.aap9559>

- [30] Wang, G., Zhao, J., Van Kleek, M., & Shadbolt, N. (2022). Informing Age-Appropriate AI: Examining Principles and Practices of AI for Children. CHI Conference on Human Factors in Computing Systems. <https://doi.org/10.1145/3491102.3502057>

Appendix A

Search Iterations

The iterative search strategy was developed to ensure comprehensive coverage of peer-reviewed literature on algorithmic bias, media manipulation, and AI's role in shaping political discourse. Each iteration was designed to reflect different conceptual and topical angles—ranging from broad core terms to more specific policy, political, technical, and sociological perspectives. This approach enabled the retrieval of a diverse set of studies across disciplines, supporting the integrity and completeness of the systematic review.

Table Search Iteration Summary

| Iteration | Search Terms | Purpose | No of Results | No of Review Papers |
|-------------|---|--|---------------|---------------------|
| Iteration 1 | 'Algorithmic bias' AND 'media manipulation' AND 'artificial intelligence' AND 'public perception' | To identify baseline literature that directly connects AI systems to media influence and societal impact. | 5 | 0 |
| Iteration 2 | 'AI algorithms' AND 'political discourse' AND 'news recommendation' OR 'political polarization' | Zooms in on political implications, especially in disinformation, elections, or polarization. | 23 | 0 |
| Iteration 3 | 'Machine learning bias' OR 'algorithmic discrimination' AND 'content moderation' OR 'media framing' | Taps into technical and ethical literature where AI bias intersects moderation and framing, which helps to understand underlying mechanisms. | 190 | 10 |
| Iteration 4 | 'AI in media' AND 'public trust' AND 'information consumption' OR 'user behavior' | Brings in sociological or behavioral perspectives on how algorithmic systems shape audience response and perception. | 1 | 0 |
| Iteration 5 | 'Algorithmic accountability' OR 'algorithmic transparency' AND 'AI ethics' AND 'media manipulation' | Targets papers that discuss frameworks or policy responses to mitigate AI bias and media manipulation. | 269 | 15 |