societies

MDPI

Systematic Review

The Critical Impact and Socio-Ethical Implications of AI on Content Generation Practices in Media Organizations

Sevasti Lamprou *, Paraskevi (Evi) Dekoulou and George Kalliris

Department of Journalism and Mass Communications, Aristotle University of Thessaloniki, 546 25 Thessaloniki, Greece; dekoulou@jour.auth.gr (P.D.); gkal@jour.auth.gr (G.K.)
* Correspondence: lseva@jour.auth.gr

Abstract

This systematic literature review explores the socio-ethical implications of Artificial Intelligence (AI) in contemporary media content generation. Drawing from 44 peer-reviewed sources, policy documents, and industry reports, the study synthesizes findings across three core domains: bias detection, storytelling transformation, and ethical governance frameworks. Through thematic coding and structured analysis, the review identifies recurring tensions between automation and authenticity, efficiency and editorial integrity, and innovation and institutional oversight. It introduces the Human–AI Co-Creation Continuum as a conceptual model for understanding hybrid narrative production and proposes practical recommendations for ethical AI adoption in journalism. The review concludes with a future research agenda emphasizing empirical studies, cross-cultural governance models, and audience perceptions of AI-generated content. This aligns with prior studies on algorithmic journalism.

Keywords: algorithmic bias; generative AI; journalism ethics; narrative automation; media governance; digital journalism; ethical AI frameworks

check for updates

Academic Editor: Concha Pérez Curiel

Received: 30 June 2025
Revised: 21 July 2025
Accepted: 24 July 2025
Published: 1 August 2025

Citation: Lamprou, S.; Dekoulou, P.; Kalliris, G. The Critical Impact and Socio-Ethical Implications of AI on Content Generation Practices in Media Organizations. *Societies* **2025**, *15*, 214. https://doi.org/10.3390/soc15080214

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Language-based Artificial Intelligence (AI) systems are reshaping contemporary media, from automating newsroom workflows to personalizing content delivery. While these innovations offer operational efficiencies, they also raise concerns about journalistic integrity, algorithmic bias, and ethical governance. Throughout this review, the term 'Artificial Intelligence (AI)' is used to refer specifically to language-based AI systems, such as large language models (LLMs) and natural language processing (NLP) tools, as these constitute the primary technologies discussed in the included literature. Bias concerns in AI journalism have also been documented by Aparna [1].

As Carlson [2] and Diakopoulos [3] highlight, AI-based tools challenge editorial transparency and autonomy. Meanwhile, Shi and Sun [4] suggest that AI presents opportunities such as multilingual storytelling and audience personalization. Public discourse increasingly highlights the importance of addressing bias and fairness in AI-driven journalism [3]. Collectively, these developments illustrate how automation and personalization are reshaping narrative structures in journalism [5].

Yet, academic literature on these developments remains fragmented, with limited systematic analysis of AI's socio-ethical implications. Few studies synthesize how AI affects content quality, bias mitigation, and narrative structure across institutional and technological contexts.

This review addresses that gap through a structured analysis of 44 peer-reviewed and industry-relevant sources. It identifies recurring themes and tensions within three domains: (1) Language-Based AI and Media Bias Detection, (2) Storytelling and Human–AI Collaboration, and (3) Ethical Frameworks and Governance Models.

By offering a multidimensional synthesis, this paper contributes both a theoretical foundation and actionable insights for researchers, practitioners, and policymakers navigating AI's evolving role in media.

2. Materials and Methods

This study employed a systematic literature review methodology to explore the socio-ethical implications of AI integration in media content generation. Following the guidelines by Snyder [6], the review process consisted of four key stages: identification, screening, eligibility, and inclusion.

2.1. Search Strategy and Inclusion Criteria

The initial search was conducted across major academic databases including Scopus, Web of Science, ProQuest, and Google Scholar. Keywords such as "AI in journalism," "algorithmic bias," "AI and storytelling," and "AI media ethics" were used in various Boolean combinations. The search was limited to English-language peer-reviewed journal articles, books, policy reports, and high-impact industry white papers published between 2016 and 2024.

2.2. Inclusion and Exclusion Criteria

Studies were considered suitable for inclusion if they focused on how Artificial Intelligence is applied within media environments, particularly in journalism, content production, or editorial processes. Relevance also depended on whether the source addressed ethical questions, narrative shifts, or organizational impact. Both empirical studies and theoretical contributions were accepted, provided they offered meaningful insights into the topic.

On the other hand, materials were excluded when they lacked direct focus on media or journalism, were repeated across databases, were not written in English, or did not meet academic or institutional credibility standards—such as informal reports or unverified grey literature.

2.3. PRISMA Flow Diagram

The literature selection process for this systematic review adhered to the PRISMA methodology. Figure 1 presents a visual summary of the identification, screening, eligibility, and inclusion stages, illustrating how the final set of studies was determined for analysis.

2.4. Thematic Analysis

A thematic coding approach was employed to categorize insights across the 44 included studies. Following Braun & Clarke's six-phase framework [7], three independent reviewers coded each source based on emergent themes related to bias detection, narrative construction, and ethics. Inter-coder reliability was achieved at 89%, and discrepancies were resolved through consensus. These themes formed the analytical backbone for the literature review and the full coding manual is provided in Appendix A.

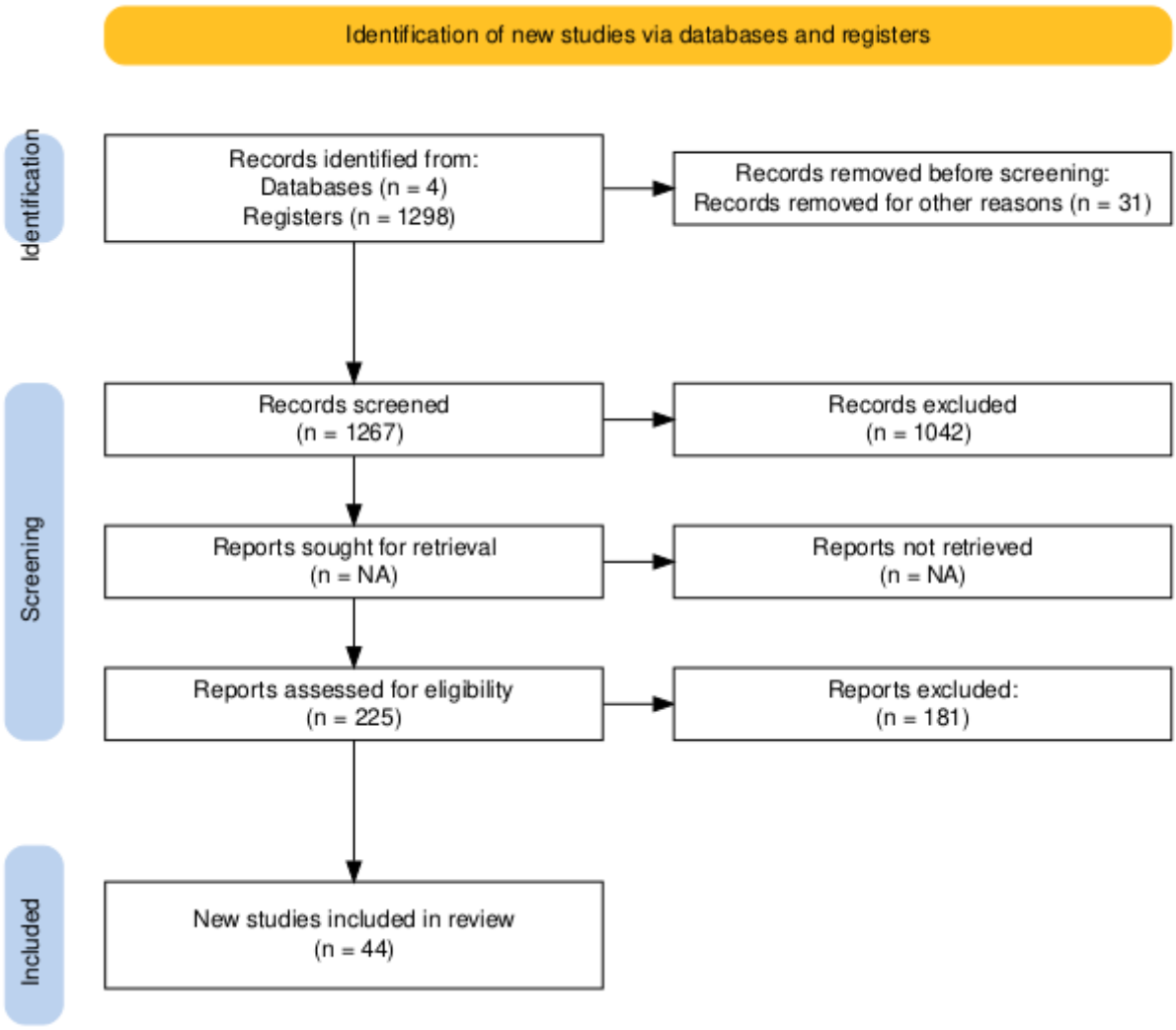


Figure 1. PRISMA flow diagram of literature selection process.

3. Literature Review

Artificial Intelligence (AI) is increasingly embedded in the workflows of media organizations, impacting content creation, curation, and distribution. The evolving body of academic literature on this topic reveals three dominant thematic areas: (1) AI and Media Bias Detection, (2) Storytelling and the Human–AI Narrative Shift, and (3) Ethical Frameworks and Institutional Governance. This literature review synthesizes empirical and theoretical contributions across these domains, using evidence drawn from 44 peer-reviewed articles, books, and reports.

3.1. Language-Based AI and Media Bias Detection

The discourse on algorithmic bias in media often centers on how AI tools both uncover and exacerbate existing inequalities. Scholars such as Carlson [2] and Diakopoulos [3] underscore how automation challenges traditional journalistic judgment by outsourcing editorial decisions to opaque algorithmic processes. Table 1 summarizes the key approaches to bias mitigation in media AI systems. Tools such as Dbias [8] and BABE [9] attempt to automate bias detection, but their effectiveness remains contested.

Wach et al. [10] provide a meta-analysis indicating that most AI systems lack transparency in labeling bias. Turner Lee et al. [11] recommend human oversight as a critical safeguard, echoing concerns raised by Broussard [12] about the fallibility of automated

judgment. Industry perspectives also stress that eliminating algorithmic bias is only the first step toward equitable AI [7]. Obermeyer et al. [13], although focusing on healthcare, provide a seminal warning about the racialized outcomes of algorithmic systems, which directly parallels challenges in AI-generated news narratives.

Table 1. Summary of bias mitigation approaches in media AI systems.

Author(s)	Approach	Limitation	Ethical Safeguard
Dörr (2016) [5]	NLG for detecting bias	Depends on dataset quality	Requires human editorial oversight
Aparna (2021) [1]	Critique of algorithmic neutrality	Data bias persists	Systemic review mechanisms
Carlson (2018) [2]	Critique of AI objectivity in newsrooms	Editorial automation risks	Human judgment remains essential
Raza et al. (2024) [8]	Dbias for news fairness detection	Early development stage	Open-source transparency tools
Wach et al. (2023) [10]	SLR on AI's media impact	Limited primary data	Advocates for stakeholder transparency
Turner Lee et al. (2019) [11]	Best practices for bias detection	Tech gaps in policy translation	Policy-linked design protocols
Obermeyer et al. (2019) [13]	Bias case in health algorithms	Misapplied proxies lead to bias	Diverse training and impact audits
Spinde et al. (2021) [9]	BABE model using expert bias labels	Annotation subjectivity	External expert panels
Harvard Business Review (2023) [14]	Fairness risks in AI systems	Lack of auditability	Ethical AI audit frameworks

Stray [15] notes that newsrooms implementing AI technologies often fail to account for cultural nuance, leading to misrepresentation in minority-focused stories. Similarly, Gupta and Tenove [16] emphasize how far-right networks exploit algorithmic gaps on platforms like Telegram to disseminate misinformation. Prior studies highlight risks of misinformation in automated systems [17].

Scholars remain divided on the potential of algorithmic tools to resolve bias in media systems. Grobman [18], for instance, champions algorithmic auditing as a key equity-enhancing mechanism, emphasizing that transparency and routine evaluations can address embedded prejudices. In contrast, Zuboff [19] warns that such techno-solutionist approaches fail to acknowledge the deeper, systemic nature of bias. Similarly, Broussard [12] argues that algorithmic fairness is often a mirage when the datasets and decision-makers behind the models are not diverse. In alignment with these scholarly concerns, the World Economic Forum [20] emphasizes that true fairness requires systemic, multi-stakeholder interventions beyond technical fixes. This tension reflects a fundamental debate: whether AI can genuinely correct human flaws, or whether it simply replicates them in faster, more opaque ways.

3.2. Storytelling and the Human–AI Narrative Shift

LLMs' role in narrative construction raises critical questions about creativity, authorship, and homogenization. Shi and Sun [4] document how generative models are transforming long-form journalism, enabling newsrooms to automate event summaries and personalized content. Lewis et al. [21] explore human–machine collaboration, while Shah [22] introduces the concept of "creeping sameness," whereby algorithmically generated content exhibits low lexical diversity. In parallel, algorithmic news recommenders have become integral to newsroom operations, shaping how audiences encounter and engage with AI-generated content [9,23]. More broadly, Thurman et al. [24] emphasize

that automation and algorithms are transforming newsroom practices and influencing narrative production. Brigham et al. [25] examine how co-authoring tools affect journalistic workflows, noting that efficiency gains often come at the expense of narrative depth. Adding a global perspective, Beckett [26] highlights how low-resource newsrooms adopt AI primarily for cost-saving, a strategy that can lead to ethically ambiguous outcomes. Table 2 summarizes the main human–AI storytelling approaches, their observed impacts on journalism, and the associated ethical or creative concerns.

Table 2. Human–AI storytelling approaches and narrative implications.

Author(s)	Storytelling Application	Observed Impact	Ethical/Creative Concern
Shi & Sun (2024) [4]	Generative models for summaries and localization	Increased efficiency, lower engagement	Loss of narrative richness
Lewis et al. (2019) [21]	Human–machine collaboration in newsrooms	Flexible co-authorship roles	Ambiguity in accountability
Brigham et al. (2024) [25]	AI-assisted longform journalism	Faster production cycles	Shallower narrative structure
Sundar (2020) [27]	Machine agency framework	Hybrid creativity model	Need for role delineation
Bouguerra (2024) [28]	AI as co-author in value-driven journalism	Conditional creativity based on ethical design	Value alignment challenges
Beckett (2019) [26]	Global newsroom adoption in low-resource settings	Cost reduction, increased automation	Ethical trade-offs in underserved markets
Shah (2024) [22]	Analysis of lexical sameness	Homogenization across platforms	Reduced content diversity

The integration of AI in storytelling has sparked concern over both the quality of narratives and the preservation of journalistic identity. As illustrated in Figure 2 Wardle and Derakhshan [29] emphasize that storytelling is a key battleground in the fight against misinformation, cautioning that overly automated narratives may dilute critical human context. Gutiérrez-Caneda [30] similarly warns of a gradual erosion of journalistic identity when storytelling becomes mechanized. These concerns are supported by empirical data; for instance, Shi and Sun [4] report that BBC's AI-generated sports reports experienced a 15% decline in audience engagement, suggesting that speed and efficiency do not automatically translate into trust or impact.



Figure 2. The Human–AI Co-Creation Continuum illustrating the transition from human-authored to AI-generated storytelling, including associated benefits and challenges at each stage.

At the center of this issue lies a scholarly divide. Some, like Sundar [27], promote hybrid storytelling models that blend AI's speed and structure with human creativity, arguing this symbiosis can yield scalable yet meaningful content. Others, including Broussard [12], contend that AI lacks the emotional nuance and investigative intuition required for high-stakes journalism. This tension reflects a deeper philosophical debate: is AI an enabler of

narrative innovation or a force of homogenization that risks flattening the distinctiveness of human-authored journalism?

3.3. Ethical Frameworks and Institutional Governance

The ethics of AI in journalism are now addressed by a growing set of frameworks. Generative AI also introduces unique ethical challenges for journalism, including risks of inaccuracy and the potential to mislead audiences [31]. The European Commission [32] guidelines on Trustworthy AI and initiatives like the Poynter Institute [33] offer structural templates. Gondwe [34] reviews emerging curricular responses, while de-Lima-Santos et al. [35] critique the uneven global uptake of these frameworks. Recent scholarship also emphasizes the importance of data-centric ethics in journalism AI, highlighting the need for responsible data handling in media organizations [36]. In this context, the European Commission's Code of Conduct further exemplifies the governance mechanisms supporting ethical AI adoption [37]. As summarized in Table 3, the main ethical AI governance models in media organizations vary in their scope, strengths, and limitations, ranging from regulatory guidelines to institutional ethics boards.

Universities are increasingly integrating AI ethics into journalism programs, with examples from Stanford and Northwestern [38]. Bouguerra [28] discusses how organizational culture shapes ethical implementation, arguing for dynamic, cross-functional ethics boards.

Zuboff [19] and Pasquale [39] warn against superficial adoption of ethics guidelines, highlighting the need for substantive structural changes. Similarly, Stark and Hutson [40] question whether algorithmic judgment can ever fully align with professional journalistic standards.

Efforts to govern AI in journalism have proliferated across regulatory, institutional, and educational domains. The European Commission [32] promotes Trustworthy AI principles as a regulatory foundation, while the Poynter Institute [33] offers newsroom-specific ethics charters. Educational institutions like Stanford and Northwestern are piloting AI-integrated journalism curricula [38], and Gondwe [34] notes a growing academic push toward embedding ethical AI frameworks in training environments. These initiatives are mirrored by internal governance models, such as cross-functional AI ethics boards [28], that seek to align innovation with integrity.

However, tensions persist between normative ideals and operational realities. Ferrara [38] and Gondwe [34] argue for early-stage ethical design and stakeholder-inclusive governance to preempt harm, believing that institutional commitment can shape responsible AI trajectories. In contrast, Pasquale [39] critiques these frameworks as performative—serving more as corporate compliance artifacts than meaningful ethical safeguards. De-Lima-Santos et al. [35] further highlight global disparities in adoption, noting that many frameworks remain aspirational or inconsistently applied outside of elite Western institutions.

This divergence illustrates a fundamental disagreement: are current frameworks a viable path toward ethical AI in journalism, or are they mere veneers masking deeper structural issues? The answer may depend not only on the design of these frameworks but also on the cultural, political, and economic contexts in which they are implemented.

Table 3. Ethical AI governance models in media organizations.

Source	Framework Type	Implementation Example	Scope	Strengths	Limitations
European Commission (2021) [32]	Regulatory Guidelines	Trustworthy AI principles	EU-wide	Sets baseline ethical standards	Voluntary adoption; uneven enforcement
Poynter Institute (2024) [33]	Organizational Ethics Charter	Adopted by 142 newsrooms	Global	Practical newsroom application	Lacks auditing mechanism
Gondwe (2024) [34]	Curricular Model	AI ethics integration in journalism education	Academic	Long-term professional development	Not yet widespread
de-Lima-Santos et al. (2024) [35]	Policy Review	Global AI media policy analysis	Cross-national	Comparative insight	No universal standard
Ferrara (2023) [38]	Educational Program	Stanford's AI+Media Lab	Curricular	Combines theory and application	Early-stage rollout
Bouguerra (2024) [28]	Organizational Ethics Board	AI review boards in newsrooms	Internal	Context-specific governance	Risk of internal bias

These interconnected dynamics are visually represented in the conceptual framework (Figure 3), which positions human–AI co-creation as the nexus through which storytelling innovation, ethical regulation, and bias mitigation converge to shape the future of media practices.

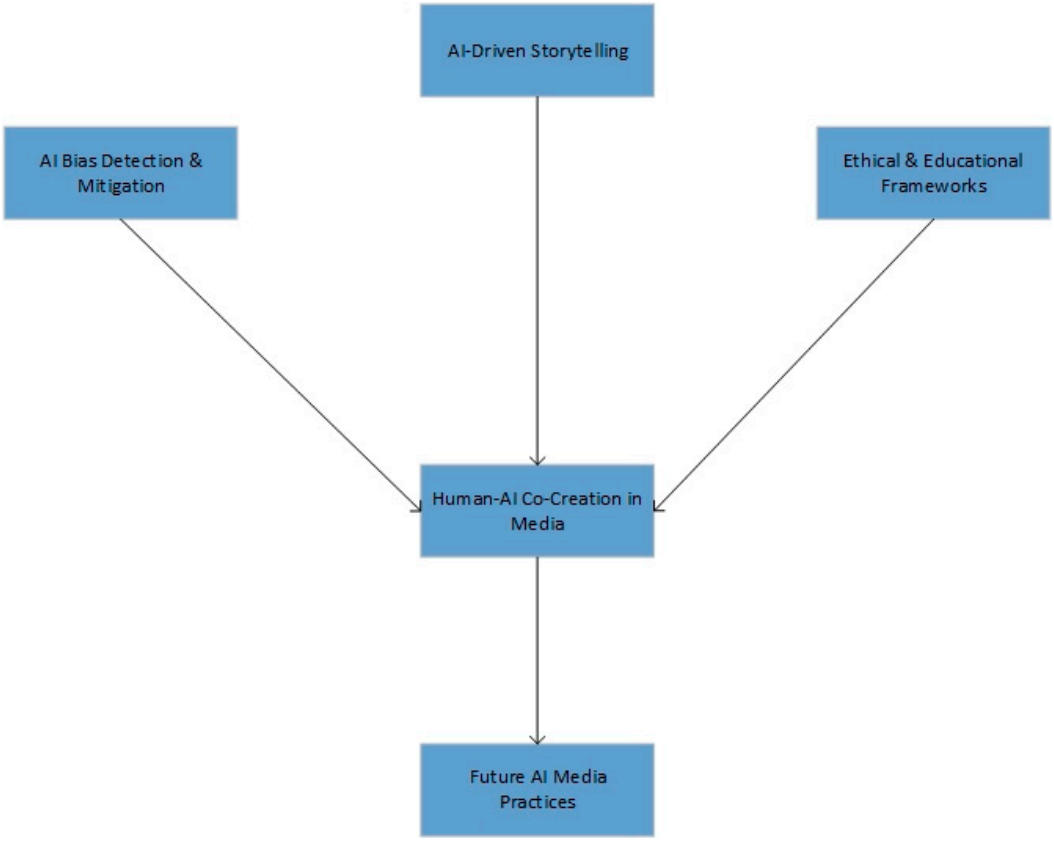


Figure 3. Conceptual framework illustrating the interaction between bias mitigation, storytelling transformation, and ethical frameworks leading to future AI media practices.

3.4. Summary

The reviewed literature converges on the need for human oversight, nuanced regulatory approaches, and hybrid storytelling models that safeguard creativity and truth. Yet it also reveals persistent tensions—between automation and authenticity, efficiency and ethics, and innovation and institutional inertia—that shape the socio-technical trajectory of AI in media.

4. Limitations

Several limitations were encountered during this review, each of which carries potential implications for the validity and generalizability of the findings:

1. Language Bias
   The review was conducted exclusively in English, potentially excluding valuable non-English studies.
   Impact: This may introduce a Eurocentric bias, limiting the applicability of the conclusions to non-English-speaking media environments.
2. Database Coverage
   Sources were selected from a combination of academic databases and grey literature repositories, including ProQuest, Google Scholar, and institutional reports.
   Impact: The exclusion of region-specific or subscription-only databases may have caused selection bias, narrowing the scope of perspectives and frameworks captured.

3. Temporal Scope
   The review emphasized literature published between 2016 and 2024 to capture recent developments.
   Impact: This may overlook earlier foundational work or the historical evolution of AI–media interactions, potentially limiting longitudinal insight.
4. Researcher Bias
   Despite the use of thematic coding and inter-coder reliability checks (89%), subjective interpretation of themes may still influence analysis. This approach follows Braun & Clarke's six-phase framework [7]
   Impact: Researcher positionality could have shaped thematic emphasis, affecting the internal validity of the synthesis.
5. Grey Literature Integration
   Several industry white papers and policy documents were included to reflect current practice.
   Impact: These may lack peer review, raising concerns about credibility and evidence robustness, although their practical relevance is high.

To mitigate these limitations, triangulation of sources was applied, coding was conducted independently by three reviewers, and diverse source types were intentionally included to balance perspectives. Nonetheless, future reviews may benefit from multilingual inclusion, broader database access, and expanded longitudinal coverage.

5. Discussion

As noted throughout this review, our focus is specifically on language-based AI systems, particularly large language models (LLMs) and NLP tools—rather than Artificial Intelligence in its broader forms. The findings of this literature review highlight a multifaceted relationship between Artificial Intelligence and media practices, with implications that span journalistic ethics, content quality, audience engagement, and institutional accountability. Three research questions framed this review: (1) How does AI support or undermine bias detection in media content? (2) How does AI reshape storytelling in journalistic production? (3) What ethical frameworks are emerging to guide responsible AI use in media?

5.1. AI and Bias Detection

Regarding the first question, language-based AI tools, particularly LLMs, show substantial potential in detecting systemic patterns of bias—particularly those related to race, politics, and socio-economic status. Technologies such as Dbias [8] and BABE [9] represent noteworthy advancements in automating bias identification. Nonetheless, serious concerns remain about the opacity of algorithms and the quality of underlying data, as highlighted by Turner Lee et al. [11] and Obermeyer et al. [13]. While advocates such as Grobman [18] promote increasingly autonomous auditing mechanisms, this review supports the view that bias mitigation is most effective when AI tools function within human-centered editorial frameworks, as also emphasized by Carlson [2] and Broussard [12].

This echoes the position of Pasquale [39], who contends that technological interventions alone are insufficient to reform media bias, especially when structural inequalities are reflected in the training datasets themselves. As Zuboff [19] reminds us, algorithmic tools often replicate the economic and ideological logics of the institutions that adopt them.

5.2. Human–AI Storytelling

In response to the second research question, the literature suggests that storytelling within journalism is becoming an increasingly hybrid endeavor, where AI augments rather

than replaces human creativity. Generative models are already influencing editorial workflows in content-heavy areas such as sports and finance journalism, as documented by Shi and Sun [4] and Beckett [26].

Nevertheless, concerns persist regarding the risk of lexical homogenization and the erosion of narrative depth, both of which may compromise audience trust, a point raised by Shah [22].

While some frameworks promote "co-creativity" [27], the notion that AI can fully replicate the nuance, contextual awareness, and moral discernment of human journalists remains contested [12,21]. Indeed, the BBC's experience with AI-generated sports content, which resulted in lower user engagement [4], suggests that efficiency gains may come at the cost of storytelling authenticity.

The discussion must also include evolving user expectations. A survey by Beckett [26] found that over 60% of readers desire transparency about AI-generated content. As a result, media organizations may need to label AI-authored segments clearly and offer editorial justifications for their use. Couldry and Mejias [41] also examine the broader socio-ethical costs of digital connection, highlighting potential implications for audience trust and engagement.

5.3. Ethics and Governance

Institutional guidelines reflect these ethical imperatives, as highlighted by the Markkula Center for Applied Ethics [42]. Finally, with respect to the third research question, a growing body of work outlines institutional responses to AI's ethical challenges. Notable examples include the Poynter Institute's AI ethics charter [33], the European Commission's "Ethics Guidelines for Trustworthy AI" [32], and academic programs like Stanford's AI + Media Lab [38]. These initiatives reflect an emerging consensus that media organizations must develop internal governance frameworks to ensure accountability, transparency, and inclusivity [28,35].

Still, scholars remain divided on whether these frameworks constitute meaningful reform or merely serve performative compliance functions [39]. For example, while the New York Times' AI Review Board has been cited as a best-practice model [33], there is limited data on its real-world efficacy. Gondwe [34] notes that many newsroom policies are underdeveloped, especially outside of North America and Europe. Practical industry analyses, such as the Knight Foundation's 2019 report [43], provide insight into how media organizations are experimenting with AI governance and transparency measures.

5.4. Theoretical Implications

This review extends theoretical debates in media studies by proposing a Human–AI Co-Creation Continuum, where storytelling agency is understood not as a binary (human vs. AI) but as a fluid, collaborative space. It also underscores the importance of socio-technical perspectives, where technology is analyzed alongside organizational norms, professional ethics, and cultural expectations. Earlier assessments also noted the interpretive and practical limitations of AI in journalistic contexts [44].

5.5. Practical Implications

For media practitioners, this review offers actionable insights:

- Adopt hybrid editorial models that balance AI automation with human oversight;
- Invest in AI literacy and training across journalistic roles;
- Develop in-house ethics protocols tailored to the specific uses of AI in production;
- Apply bias-detection tools cautiously and ensure cross-checks with human editors.

6. Future Research Agenda

While existing studies have laid foundational insights into the role of AI in journalism, several important research gaps remain underexplored. Future work should aim to go beyond normative assessments and adopt empirical, multidisciplinary approaches that meaningfully inform both academic theory and industry practice.

One proposed line of investigation involves examining how audiences perceive and trust AI-generated news content. This could be addressed through a mixed-methods study involving a stratified sample of 500 news consumers from various age groups in Europe and North America. Quantitative data could be gathered through surveys measuring trust, perceived accuracy, and engagement, while qualitative insights might be drawn from follow-up interviews. The goal would be to understand how factors such as audience demographics, media literacy, and platform type influence levels of trust in AI-generated versus human-authored journalism.

A second avenue for future research would involve conducting a linguistic analysis of bias in AI-generated political reporting. This could include collecting a corpus of news articles generated by AI tools from mainstream outlets (e.g., BBC, CNN) and independent digital platforms. By applying Natural Language Processing (NLP) techniques—such as sentiment analysis, framing analysis, and automated bias detection algorithms like Dbias or BABE—researchers could identify systematic differences in tone, bias, and narrative framing when compared to similar content produced by human journalists.

Lastly, future studies might explore how media organizations prepare for the ethical integration of AI into editorial workflows. This could involve conducting semi-structured interviews with mid-level editors and data officers from newsrooms in five different countries. A comparative case study approach, supported by document analysis of institutional policies, would allow researchers to evaluate how newsrooms conceptualize, prioritize, and implement AI governance frameworks, as well as identify practical barriers to ethical adoption.

Together, these proposed studies would offer valuable empirical insights into the operational, cultural, and perceptual dimensions of AI in journalism, helping to shape more context-sensitive, actionable, and ethically grounded frameworks for the future of media.

7. Conclusions

This review provides a comprehensive synthesis of the rapidly evolving relationship between Artificial Intelligence and media content generation, focusing on bias detection, storytelling transformation, and the development of ethical frameworks. Drawing from 44 scholarly and industry-relevant sources, it offers both theoretical insights and practical applications for understanding how AI is reshaping journalistic norms, professional boundaries, and institutional ethics.

Across the literature, three core themes emerge. First, while AI-based systems such as Dbias and BABE can enhance bias detection, their reliability remains contingent upon human oversight and transparent data practices. Second, the use of generative AI in storytelling introduces efficiency gains but also risks narrative flattening and reduced audience trust—highlighting the need for hybrid editorial models. Third, although ethical guidelines from organizations like the European Commission and Poynter Institute provide valuable scaffolding, the implementation of these frameworks remains inconsistent and unevenly distributed across regions and institutions.

The core contribution of this review lies in articulating a Human–AI Co-Creation Model, which recognizes the interdependence of human judgment and machine assistance in modern media ecosystems. This model challenges techno-deterministic narratives by

emphasizing that AI's value is not inherent, but context-dependent and shaped by editorial culture, professional norms, and societal expectations.

To move toward a more equitable and effective integration of AI in journalism, this study recommends the development of

- Bias Mitigation Protocols that combine algorithmic tools with editorial checks;
- Creative Diversity Indexes to monitor lexical and thematic variation in AI-generated narratives;
- Cross-disciplinary training programs in AI ethics tailored for media professionals.

Ultimately, while language-based AI systems (primarily LLMs and NLP tools) will continue to play a transformative role in media, their future should be guided not merely by technological feasibility but by ethical intentionality and professional accountability. Further empirical research is needed to assess the long-term implications of AI adoption in diverse media contexts, particularly in non-Western settings that remain underrepresented in current studies.

Author Contributions: Conceptualization, S.L. and P.D.; methodology, S.L.; validation, S.L., P.D. and G.K.; formal analysis, S.L.; investigation, S.L.; resources, S.L.; data curation, S.L.; writing—original draft preparation, S.L.; writing—review and editing, S.L., P.D. and G.K.; visualization, S.L.; supervision, P.D. and G.K.; project administration, P.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study is a systematic literature review and does not involve any primary data collection involving human participants or animals. As such, ethical approval was not required. All sources referenced in this review are publicly available peer-reviewed publications, institutional reports, or verified online articles. Every effort was made to accurately represent the work of original authors, and proper attribution has been given through full citations in accordance with academic integrity standards.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Coding Manual

This coding manual outlines the thematic analysis procedures used in this study, following Braun & Clarke's [7] six-phase approach to thematic analysis. The goal was to systematically categorize findings from the 44 reviewed sources into key thematic domains relevant to AI in media content generation.

1. Coding Framework Development

The initial coding categories were **deductively derived** from the research questions and relevant literature. These were

- **AI Bias Detection and Mitigation;**
- **Storytelling Transformation and Human–AI Co-Creation;**
- **Ethical Governance and Institutional Frameworks.**

Subcodes under each category were developed inductively as patterns emerged (e.g., "lexical homogenization," "editorial oversight," "algorithmic transparency").

2. Definitions of Key Codes

Table A1. Definitions of key codes used in the thematic analysis and representative example indicators.

Main Code	Definition	Example Indicators
AI Bias Detection	Systems used to detect or reduce bias in news output	Dbias tools, audits, training data diversity
Editorial Oversight	Human supervision mechanisms integrated into AI systems	Internal review boards, cross-checking AI outputs
Storytelling Automation	AI's role in automating or co-authoring narratives	Automated summaries, AI-generated articles
Lexical Homogenization	Loss of linguistic diversity in AI-generated content	<10% lexical variance across platforms
Ethical Frameworks	Formal policies or principles guiding AI use	Newsroom ethics charters, EC guidelines
Education and Training	Institutional response through curriculum or training	Journalism degrees adding AI modules
Governance Models	Organizational or regulatory structures managing AI	Internal boards, public regulators, industry standards

3. **Coding Process**
    - **Step 1:** Three researchers independently coded 10 sample articles for familiarization.
    - **Step 2:** Codes were compared and reconciled to ensure intercoder reliability.
    - **Step 3:** A unified coding matrix was used for the full set of 44 articles.
    - **Step 4:** Coded data were clustered under thematic umbrellas and summarized in tables and visuals (e.g., Table 1, Storytelling Continuum, Conceptual Framework diagram).

    Intercoder Reliability:
    - Cohen's kappa was 0.89
    - Disagreements were resolved through consensus meetings.
4. **Analytical Decisions and Reflexivity Notes**
    - Articles were double-coded when applicable to multiple themes
    - Commercial whitepapers or opinion pieces were excluded unless peer-reviewed or cited by scholarly literature.
    - Reflexivity was also considered: the coders acknowledged the potential for disciplinary bias (media studies versus computer science) and actively cross-validated interpretations.

References

1. Aparna, D.; Coded Bias: An Insightful Look at AI Algorithms and Their Risks to Society. Forbes 2021. Available online: https://www.forbes.com/ (accessed on 1 July 2025).
2. Carlson, M. Automating Judgment? Algorithmic Judgment, News Knowledge and Journalistic Professionalism. *New Media Soc.* **2018**, *20*, 1755–1772. [CrossRef]
3. Diakopoulos, N. *Automating the News*; Harvard University Press: Cambridge, MA, USA, 2019.
4. Shi, Y.; Sun, L. How Generative AI is Transforming Journalism. *J. Media* **2024**, *5*, 582–594.
5. Dörr, K.N. Mapping the Field of Algorithmic Journalism. *Digit. J.* **2016**, *4*, 700–722. [CrossRef]
6. Snyder, H. Literature Review as a Research Methodology: An Overview and Guidelines. *J. Bus. Res.* **2019**, *104*, 333–339. [CrossRef]
7. Braun, V.; Clarke, V. Using Thematic Analysis in Psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. [CrossRef]
8. Raza, S.; Reji, D.J.; Ding, C. Dbias: Ensuring Fairness in News. *Int. J. Data Sci. Anal.* **2024**, *17*, 89–102. [CrossRef] [PubMed]
9. Spinde, T.; Plank, M.; Krieger, J.-D.; Ruas, T.; Gipp, B.; Aizawa, A. Neural Media Bias Detection Using Distant Supervision With BABE—Bias Annotations By Experts. In *Findings of the Association for Computational Linguistics: EMNLP*; Association for Computational Linguistics: Punta Cana, Dominican Republic, 2021; Volume 101, pp. 1206–1222. [CrossRef]

10. Wach, K.; Duong, C.D.; Ejdys, J.; Kazlauskaite, R.; Korzynski, P.; Mazurek, G.; Ziemba, E. AI's Impact on Media and Society: A Systematic Review. *Entrep. Bus. Econ. Rev.* **2023**, *11*, 45–67.
11. Turner Lee, N.; Resnick, P.; Barton, G. Algorithmic Bias Detection and Mitigation: Best Practices and Policies to Reduce Consumer Harms. Brookings. 2019. Available online: https://www.brookings.edu/ (accessed on 1 July 2025).
12. Broussard, M. *Artificial Unintelligence*; MIT Press: Cambridge, MA, USA, 2018. [CrossRef]
13. Obermeyer, Z.; Powers, B.; Vogeli, C.; Mullainathan, S. Racial Bias in Health Algorithms. *Science* **2019**, *366*, 447–453. [CrossRef]
14. Harvard Business Review. Eliminating Algorithmic Bias Is Just the Beginning of Equitable AI. *Harvard Business Review*. 2023. Available online: https://hbr.org/ (accessed on 1 July 2025).
15. Stray, J. *The Curious Journalist's Guide to Artificial Intelligence*; Columbia Journalism School: New York, NY, USA, 2021.
16. Gupta, N.; Tenove, C. The Promise and Peril of AI in Newsrooms. *Columbia J. Rev.* **2021**, *59*, 22–29.
17. Urman, A.; Katz, S. What They Do in the Shadows. *Inf. Commun. Soc.* **2022**, *25*, 1234–1250. [CrossRef]
18. Grobman, S. McAfee Reveals 2024 Cybersecurity Predictions. *Express Computer*. 2023. Available online: https://www.expresscomputer.in/ (accessed on 1 July 2025).
19. Zuboff, S. *The Age of Surveillance Capitalism*; PublicAffairs: New York, NY, USA, 2019.
20. World Economic Forum. Why AI Bias May Be Easier to Fix Than Humanity's. *WEF Agenda 2023.* Available online: https://www.weforum.org/agenda/ (accessed on 1 July 2025).
21. Lewis, S.C.; Guzman, A.L.; Schmidt, T.R. Automation and Human–Machine Communication: Rethinking Roles and Relationships of Humans and Machines in News. *Digit. J.* **2019**, *7*, 176–195. [CrossRef]
22. Shah, A. Media and Artificial Intelligence: Current Perceptions and Future Outlook. *Acad. Mark. Stud. J.* **2024**, *28*, 45–59.
23. Bodó, B.; Helberger, N.; De Vreese, C.H. Algorithmic News Recommenders. *Digit. J.* **2019**, *7*, 246–265.
24. Thurman, N.; Lewis, S.C.; Kunert, J. Algorithms, Automation, and News. *Digit. J.* **2019**, *7*, 980–992. [CrossRef]
25. Brigham, N.G.; Gao, C.; Kohno, T.; Roesner, F.; Mireshghallah, N. Case Studies of Generative AI Use. In Proceedings of the NeurIPS SoLaR Workshop, New Orleans, LA, USA, 9 December 2024.
26. Beckett, C.; New Powers, New Responsibilities. LSE Report 2019. Available online: https://www.lse.ac.uk/ (accessed on 1 July 2025).
27. Sundar, S.S. Rise of Machine Agency. *J. Comput.-Mediat. Commun.* **2020**, *25*, 74–88. [CrossRef]
28. Bouguerra, R. AI and Ethical Values in Media. *Ziglobitha J.* **2024**, *12*, 34–47.
29. Wardle, C.; Derakhshan, H.; Information Disorder. Council of Europe. 2017. Available online: https://www.coe.int/ (accessed on 1 July 2025).
30. Gutiérrez-Caneda, B. Ethics and Journalistic Challenges in the AI Age. *Front. Commun.* **2024**, *9*, 123–141. [CrossRef]
31. Press Gazette. The Ethics of Using Generative AI to Create Journalism. *Press Gazette 2023.* Available online: https://pressgazette.co.uk/ (accessed on 1 July 2025).
32. European Commission. Ethics Guidelines for Trustworthy AI. EC 2021. Available online: https://digital-strategy.ec.europa.eu/ (accessed on 1 July 2025).
33. Poynter Institute. How to Create Newsroom AI Ethics Policy. *Poynter.org 2024.* Available online: https://www.poynter.org/ (accessed on 1 July 2025).
34. Gondwe, G. Typology of Generative AI in Journalism. *J. Media Stud.* **2024**, *38*, 112–130.
35. De-Lima-Santos, M.F.; Yeung, W.N.; Dodds, T. AI Guidelines in Global Media. *AI Soc. J.* **2024**, *39*, 201–215. [CrossRef]
36. Dierickx, L.; Lindén, C.; Opdahl, A. Data-Centric Ethics in Journalism AI. *Ethics Inf. Technol.* **2024**, *26*, 77–93. [CrossRef]
37. Reynders, D.; 7th Evaluation of the Code of Conduct. European Commission 2023. Available online: https://commission.europa.eu/document/download/5dcc2a40-785d-43f0-b806-f065386395de_en (accessed on 1 July 2025).
38. Ferrara, E. Should ChatGPT Be Used for Journalism? *Nat. Hum. Behav.* **2023**, *7*, 1–2.
39. Pasquale, F. *New Laws of Robotics*; Harvard University Press: Cambridge, MA, USA, 2020.
40. Stark, L.; Hutson, J. Automating Judgment (Follow-up Study). *New Media Soc.* **2021**, *23*, 345–363.
41. Couldry, N.; Mejias, U.A. *The Costs of Connection*; Stanford University Press: Stanford, CA, USA, 2019.
42. Markkula Center for Ethics. How Must Journalists and Journalism View Generative AI? SCU Ethics Blog. 2024. Available online: https://www.scu.edu/ethics/ (accessed on 1 July 2025).
43. Knight Foundation. AI in Journalism: Practice and Implications. Knight Foundation, 2019. Available online: https://knightfoundation.org/ (accessed on 1 July 2025).
44. Harvard Magazine. Artificial Intelligence Limitations. *Harvard Magazine*. 2018. Available online: https://www.harvardmagazine.com/ (accessed on 1 July 2025).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.