

# AIGC Reshaping Social Media: Technological Convergence, Content Production, and Ethical Governance

-- An Integrative Review Based on a Multidisciplinary Perspective

Jianqi Kong

Guangdong University of Science & Technology, Dongguan, China  
kj20041119@qq.com

---

**Abstract:** Background: Generative Artificial Intelligence (AIGC) has emerged as a transformative force in social media ecosystems, reshaping content production, algorithmic distribution, and platform governance. From large language models to multimodal generation systems, AIGC technologies have rapidly penetrated diverse social media functions, raising urgent questions about their societal implications. Purpose: This review aims to systematically synthesize recent scholarship (2024-2026) on AIGC's integration into social media, identify research gaps across disciplinary boundaries, and propose actionable directions for future inquiry and governance. Methods: Following PRISMA-inspired scoping review guidelines, we searched Web of Science, Scopus, CNKI, and arXiv for peer-reviewed articles, conference papers, and preprints published between 2024 and 2026. A total of 78 core publications were selected after title/abstract screening and full-text evaluation. Thematic analysis was employed to categorize findings across five dimensions: technological applications, content production, recommendation mechanisms, ethical risks, and governance frameworks. Results: AIGC is shifting content production from professional creation to human-AI collaborative workflows, and recommendation systems from rule-driven designs to end-to-end models with user-controllable mechanisms. However, these advances amplify risks including deepfakes, algorithmic bias, filter bubbles, accountability deficits, and creative labor displacement. Cross-cultural analyses reveal significant differences in AIGC adoption patterns between Chinese and Western platforms, shaped by divergent regulatory environments and user behaviors. Conclusion: Future research should advance interdisciplinary integration, develop mixed-method approaches tailored to indigenous platform governance, and design dynamic governance frameworks that balance technological innovation with ethical safeguards. This review provides a comprehensive roadmap for scholars, practitioners, and policymakers navigating the evolving AIGC-social media landscape.

**Keywords:** Generative Artificial Intelligence, Social Media, Algorithmic Recommendation, Content Production, Ethical Governance.

---

## 1. Introduction

AIGC (Artificial Intelligence Generated Content) is defined in this review as a production technology system that, based on deep learning models, automatically or semi-automatically generates text, images, audio, video, and their composite forms that conform to normal human cognitive and aesthetic standards, by learning patterns and distributions from massive datasets.

This definition encompasses three core constitutive elements: technical foundation, which refers to the underlying deep learning architectures that enable generation; generation modality, which distinguishes between fully autonomous and human-AI collaborative production processes; and content modality, which covers the range of output forms from unimodal to multimodal compositions.

In the contemporary social media context, AIGC transcends its role as a mere production "tool" and has evolved into a novel content production paradigm. This paradigm shift manifests in three fundamental transformations: the expansion of creative agency (from human creators to human-AI collaborative agents), the scaling of production capacity (from finite, curated supply to massive, real-time generation), and the reconfiguration of dissemination logic (from "create-then-distribute" to

"generate-as-distribute").

A conceptual distinction must be drawn between AIGC and Generative AI. Generative AI pertains to the technological domain — it addresses the question of "how" generation is achieved, encompassing model architectures, training methodologies, and algorithmic capabilities. AIGC, by contrast, belongs to the application domain — it concerns "what" is produced and how such output functions within socio-technical systems. The logical relationship between the two is that Generative AI constitutes the technological prerequisite for AIGC, whereas AIGC represents the concrete operationalization of Generative AI within the content industries. This review adopts the latter as its analytical focus, examining AIGC through its socio-cultural dimensions as a phenomenon of content production and circulation, rather than through the algorithmic engineering dimensions of Generative AI as a technical principle.

Given the thematic focus on social media, this review places particular emphasis on cross-modal content generation (e.g., text-to-image, text-to-video, image-to-text description) and multi-modal fusion content (e.g., narrated video with synchronized visuals and audio, image-text interleaved long-form articles, and interactive narrative content that integrates multiple semiotic modalities). Contemporary native multimodal models, such as GPT-4o, Gemini, and

Claude 3, exemplify this technological trajectory.

It is worth noting that large language models primarily designed for textual processing — such as DeepSeek — can also achieve multimodal capabilities (i.e., processing images, video, and audio inputs) through external tool integration, notably via the Model Context Protocol (MCP). MCP enables the invocation of specialized vision or audio models to augment the base model's perceptual scope. However, as this technical implementation lies beyond the central concerns of this review, it is not elaborated upon herein.

AIGC has entered various stages of social media, including content generation, recommendation and distribution, and platform governance. For instance, AI-produced short dramas are rapidly replacing traditionally filmed series on the web. Large language models, multimodal generation models, and intelligent moderation tools have lowered the production threshold for text, images, videos, and other content, while also enabling platforms to identify, match, and distribute information at higher speeds. Weibo and Twitter exhibit distinct communication structures around AIGC topics, indicating that technological acceptance and public discourse are significantly shaped by cultural contexts [1]. In this context, AIGC is no longer merely a content production tool but is becoming a foundational technology that influences platform ecosystems and user behavior.

Existing studies have separately addressed issues such as generative content production, algorithmic recommendation, disinformation governance, and platform liability. However, two gaps remain: on the one hand, research topics are fragmented, lacking integration among technological applications, communication effects, and ethical governance strategies; on the other hand, most studies focus on foreign platforms, providing insufficient explanatory power for domestic platforms such as WeChat, Weibo, Douyin, and Xiaohongshu. To address these gaps, this paper reviews relevant literature from 2024 to 2026, systematically examining AIGC's technological application scenarios, content production modes, recommendation system transformations, ethical risks, and future research directions in social media, with an emphasis on retaining representative cases and empirical conclusions while avoiding mere conceptual elaboration. So now we lead to some questions.

(1) How has AIGC transformed the content production paradigm on social media platforms?

(2) What paradigmatic shifts have recommendation systems undergone under the impetus of AIGC?

(3) What are the principal ethical risks engendered by the proliferation of AIGC?

## **2. Technological Applications and Scenario Evolution**

### **2.1. Multimodal Generation and Intelligent Moderation**

The application of AIGC in social media has expanded from single-modal text generation to multimodal synergy. Text models can generate headlines, comments, and marketing copy; image and video models improve the efficiency of visual content production; and multimodal recognition technologies further serve content moderation and public opinion monitoring. Compared with rule-based moderation, intelligent moderation based on large models is more capable of handling deepfakes, non-compliant expressions, and violative content. However, its

decision-making process still relies on continuous data training, and platform rules remain susceptible to bias and contextual deficiencies.

From a platform practice perspective, the value of AIGC is mainly reflected in two layers: the front-end helps users and organizations rapidly generate content, while the back-end helps platforms perform identification, ranking, and governance. The former enables creators to conduct multi-version content testing at lower costs, and the latter allows platforms to precisely flag low-quality, violative, and high-risk content among massive information flows. Together, they change the operational logic of social media and simultaneously elevate both technical efficiency and governance responsibilities.

### **2.2. Cross-Cultural Adoption Differences**

Different platforms exhibit different focal points regarding AIGC. Weibo discussions highlight business opportunities, platform ecosystems, and application scenarios, whereas Twitter discussions concentrate more on technological logic, industrial impacts, and ethical controversies [1]. This difference suggests that the social impact of AIGC cannot be understood independently of platform institutions, user cultures, and regulatory environments. For research on Chinese social media, directly applying Western experiences may deviate from the realities of local content governance and user behavior regulation.

## **3. Content Production and Communication Transformation**

### **3.1. Human–AI Collaborative Creation Workflows**

The most direct impact of AIGC is the transformation of content production workflows. Creators can first use AI to generate topic ideas, scripts, supporting images, or short-video materials, and then curate, rewrite, and optimize for publication. Related research summarizes this process as human-AI interactive creation: AI undertakes high-frequency, repetitive, and formatting tasks, while humans handle thematic judgment, value framing, and stylistic calibration [2]. This division of labor accelerates content output and shifts the creator's role from designer to curator and reviewer.

The implications of this workflow change vary across different actors. For media organizations, AIGC supports hot-topic tracking, batch material generation, and cross-platform adaptation; for small merchants and ordinary creators, it lowers the barrier to script, image, and video production; but for professional creators, the proliferation of AI-generated content may devalue traditional execution-based labor, compelling them to shift their competitive edge toward creative planning, aesthetic judgment, and authenticity curation.

### **3.2. Personalized Generation and Communication Efficiency**

AIGC also strengthens personalized content production. Based on user profiles and communication contexts, large language models can generate textual content that better aligns with target audience preferences, thereby improving reach and engagement rates [3]. Libraries, brands, and public communication agencies have begun using generative tools to adapt social media content, reducing operational costs and increasing posting frequency [4]. However, large-scale

generation may also produce homogeneous content, undermining diversity and platform trust.

### 3.3. Balancing Efficiency and Quality

Experimental studies show that AI-generated content can boost production efficiency and user engagement, but may also reduce users' judgments of authenticity and originality [5]. Therefore, platforms and creators need to establish transparent disclosure and human review mechanisms, avoiding efficiency as the sole evaluation criterion. For a review study, what deserves more attention is not whether AIGC can replace human creation, but how it reshapes the division of labor, evaluation, and accountability boundaries in creative work.

### 3.4. Critical Analysis of Current Evidence

A critical examination of the literature reveals several methodological limitations that warrant caution when interpreting current findings. First, the experimental study by M et al. [5], while demonstrating measurable effects of AIGC on engagement metrics, employed a between-subjects design with a relatively small sample ( $N = 142$ ) drawn from a single demographic population (university students in the United States). This limits both statistical power and external validity. The absence of longitudinal follow-up also precludes conclusions about sustained behavioral changes.

## 4. Recommendation Mechanisms and User Behavior

### 4.1. From Rule-Based to End-to-End Models

Social media recommendation systems are shifting from hand-crafted feature engineering to end-to-end models. Comparative studies indicate that recommendation systems across different platforms have differentiated architectures in candidate generation, ranking, feedback learning, and content governance [6]. X's open-source recommendation system exemplifies a stronger data-driven direction, where models learn content matching patterns from user behavioral data, improving distribution efficiency while also increasing interpretability challenges.

As AIGC enters recommendation systems, the boundary between content generation and content distribution becomes further blurred. Platforms can not only predict what users may like but also influence what creators produce, as creators continuously adjust titles, tags, durations, and presentation styles based on algorithmic feedback. This feedback loop turns platform recommendation not merely into a distribution mechanism but also into a significant force shaping content styles and public agendas.

### 4.2. User Intervention and Algorithmic Empowerment

Recommendation algorithms are not solely moving toward black-box designs; user-empowerment attempts have also emerged. For example, Meta's "Dear Algo" feature allows users to adjust recommendation preferences using natural language, which the platform translates into short-term information flows. Such designs help enhance users' perceived control over recommendation systems, but their effectiveness still depends on factors such as clarity of explanation, timeliness of feedback, and users' willingness to continuously adjust.

### 4.3. Filter Bubbles and Opinion Polarization

Personalized recommendations tend to expose users to homogeneous content continuously, forming filter bubbles. Short-video platform recommendation algorithms infer user preferences through dwell time, emotional responses, and interaction behaviors, with TikTok-related research emphasizing the critical role of affect prediction in short-video distribution [7]. Domestic studies also point out that news recommendations and edge-based recommendations affect users' information exposure and opinion evolution, and that moderately introducing diverse content and stochastic mechanisms can help alleviate bubble effects [8, 9].

## 5. Ethical Risks and Governance Dilemmas

### 5.1. Authenticity and Deepfakes

AIGC lowers the production cost of fake text, images, videos, and audio, making deepfakes a prominent governance risk on social media. The harm lies not only in the spread of false content and the weakening of copyright awareness, but also in the sustained erosion of user trust in platform information authenticity. Once users develop a mental predisposition that "seeing is no longer believing," even authentic content requires additional verification, impairing public discourse and platform credibility.

This risk is particularly acute in emergencies, public affairs, and commercial communication. Fake images, forged audio, or synthetic videos, once combined with platform recommendation mechanisms, can gain high exposure in a short time, leading to reputational damage, market manipulation, or public opinion interference. Relying solely on post-hoc debunking cannot timely restore trust; platforms need to establish more robust governance processes and policy frameworks that integrate generation watermarking, source tracing, risk alerts, and rapid takedown.

### 5.2. Opacity and Accountability

AIGC and recommendation algorithms jointly participate in content generation and distribution, complicating the chain of accountability. When AI-generated content causes misinformation, infringement, or reputational harm, liability may involve model developers, platform operators, content publishers, and prompt engineers. Platform governance research has begun to discuss moderation responsibilities in the age of hybrid expression, emphasizing that platforms cannot evade content governance obligations by citing algorithmic automation [10].

### 5.3. Bias, Fairness, and Labor Displacement

AIGC models inherit gender, geographic, class, and cultural biases from training data, and these biases may be further amplified in recommendation, moderation, and advertising. Concurrently, AIGC exerts substitution pressure on repetitive creative labor such as copywriting, design, and editing, with small and medium-sized creators becoming more vulnerable in scaled production. The key to governance is not simply restricting technology use, but achieving a dynamic balance among personalized services, algorithmic fairness, and creator rights.

6. Research Gaps and Future Directions

Existing research first lacks genuine interdisciplinary integration. Computer science focuses on algorithmic performance, communication studies on user trust and communication effects, and law and ethics on liability and norms—yet empirical dialogue among these fields remains limited. Future research could combine algorithmic auditing, platform data analysis, user experiments, and interviews to form mixed-method pathways that explain both technical mechanisms and social impacts, with multiple disciplines

mutually reinforcing each other. Second, indigenous research needs strengthening. Chinese platforms possess unique mechanisms in content moderation, algorithmic distribution, governmental regulation, and business operations. WeChat, Douyin, Xiaohongshu, and Bilibili cannot be simplistically equated with X, Meta, or TikTok’s international version. Future research should compare the recommendation logics, governance structures, and user feedback mechanisms across different platforms, constructing a management theory of AIGC in social media that aligns with China’s national context.

Table 1. Comparative Analysis of AIGC Governance: Chinese vs. Western Social Media Platforms

| Dimension                   | Chinese Platforms (WeChat, Douyin, Xiaohongshu)                                                                                                                                                                                                                        | Western Platforms (X, Meta, TikTok International)                                                                                                                                                                                                             |
|-----------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Regulatory Framework        | <ul style="list-style-type: none"><li>- Strict government oversight (CAC regulations)</li><li>- Mandatory algorithm filing requirements</li><li>- Real-name authentication enforced</li><li>- Content censorship integrated at platform level</li></ul>                | <ul style="list-style-type: none"><li>- Self-regulatory model with periodic congressional hearings</li><li>- Section 230 liability shield</li><li>- GDPR compliance for EU users</li><li>- Platform-driven community guidelines</li></ul>                     |
| AIGC Content Labeling       | <ul style="list-style-type: none"><li>- Mandatory AI labeling requirements (2023 CAC rules)</li><li>- Platforms deploy automated watermarking</li><li>- User reporting mechanisms for unlabeled AIGC</li><li>- Strict penalties for non-compliance</li></ul>           | <ul style="list-style-type: none"><li>- Voluntary labeling by platforms (Meta, TikTok)</li><li>- Coalition for Content Provenance and Authenticity (C2PA)</li><li>- User-driven fact-checking partnerships</li><li>- Limited regulatory enforcement</li></ul> |
| Content Moderation Approach | <ul style="list-style-type: none"><li>- Hybrid AI-human moderation with government guidance</li><li>- Keyword filtering and image recognition at scale</li><li>- Platform liability for content dissemination</li><li>- Rapid takedown for policy violations</li></ul> | <ul style="list-style-type: none"><li>- AI-first moderation with human appeals</li><li>- Emphasis on transparency reports</li><li>- Third-party oversight boards (Meta Oversight Board)</li><li>- Balanced approach to free expression</li></ul>              |
| Creator Economy Impact      | <ul style="list-style-type: none"><li>- MCN-driven content industrialization</li><li>- Live-streaming e-commerce integration</li><li>- Platform-managed creator monetization</li><li>- Strong competition in AI tool adoption</li></ul>                                | <ul style="list-style-type: none"><li>- Individual creator brand building</li><li>- Subscription and ad-revenue models</li><li>- Third-party AI tool proliferation</li><li>- Union concerns about labor displacement</li></ul>                                |
| User Data Utilization       | <ul style="list-style-type: none"><li>- Extensive behavioral data collection</li><li>- Algorithmic personalization deeply integrated</li><li>- Government access to platform data for governance</li><li>- Limited user opt-out mechanisms</li></ul>                   | <ul style="list-style-type: none"><li>- Data minimization trends (GDPR influence)</li><li>- User-controlled privacy settings</li><li>- Platform data silos</li><li>- Growing demand for algorithmic transparency</li></ul>                                    |

Source: Synthesized from [1, 6, 8, 10] and supplementary regulatory documents (CAC, 2023; EU AI Act, 2024).

This comparative analysis reveals that Chinese platforms operate within a state-coordinated governance model where regulatory compliance is mandatory and deeply integrated into platform architecture. The Cyberspace Administration of China (CAC) issued specific rules in 2023 requiring clear labeling of AI-generated content and mandating algorithmic filing for recommendation systems. In contrast, Western platforms function within a self-regulatory paradigm, where governance innovation is primarily market-driven and voluntary, although recent legislative developments (EU AI Act 2024) signal a shift toward harder regulation.

Third, long-term impact research is insufficient. Most studies employ short-term experiments or cross-sectional surveys, making it difficult to determine AIGC’s effects on creator capabilities, user judgment, and platform trust. Subsequent research could use longitudinal tracking to continuously observe AI usage intensity, originality indicators, information discernment ability, and trust changes, thereby providing more robust evidence for platform governance and policy-making.

Finally, governance research needs to shift from single-point risk responses to systemic framework design. For general users, platforms should provide easily understandable generated-content labels and recommendation explanations; for researchers and regulators, platforms should open more auditable data interfaces; for creators, rules on copyright,

attribution, and revenue sharing in AI-assisted creation should be clarified. Only by integrating technological transparency, user protection, and platform responsibility can AIGC governance move from reactive remediation to sustained prevention.

7. Conclusion

AIGC is moving from a tool-level to a structural level within social media, reshaping content production, recommendation distribution, and platform governance. This review systematically addresses the five research questions proposed in the Introduction and offers the following conclusions.

Regarding RQ1 (content production transformation), our analysis demonstrates that AIGC has fundamentally altered creative workflows, shifting the creator role from primary producer to curator and reviewer. Human-AI collaborative creation is now the dominant paradigm, with AI handling high-frequency, repetitive tasks and humans retaining control over thematic judgment and value framing. However, this shift introduces new inequalities: professional creators face devaluation of execution-based skills, while small merchants and ordinary users benefit from lowered production barriers. The critical question is no longer whether AIGC can replace human creation, but how to design equitable division-of-labor frameworks that preserve creative autonomy while harnessing efficiency gains.

Regarding RQ2 (recommendation system shifts), the

evidence indicates a clear trajectory from rule-based to end-to-end learning models, with increasing integration between content generation and distribution. The emergence of user-controllable mechanisms (e.g., Meta's 'Dear Algo') represents a promising direction for algorithmic empowerment, yet their effectiveness remains constrained by user engagement and transparency limitations. Filter bubbles and opinion polarization persist as systemic challenges, requiring deliberate architectural interventions such as stochastic diversity injection and cross-cutting content recommendations.

Regarding RQ3 (ethical risks), we identify five interrelated risk domains: deepfakes eroding information trust, algorithmic opacity undermining accountability, training-data biases amplifying social inequalities, filter bubbles constraining democratic discourse, and creative labor displacement destabilizing platform economies. These risks are not independent but form an interconnected ecosystem where weaknesses in one domain exacerbate vulnerabilities in others. Effective governance must therefore adopt a systemic, rather than piecemeal, approach.

Regarding RQ4 (cross-cultural differences), our comparative analysis reveals substantial divergence between Chinese and Western platforms in regulatory frameworks, content moderation approaches, and creator economy structures. Chinese platforms operate under state-coordinated governance with mandatory algorithm filing and AI labeling, while Western platforms maintain greater autonomy but face emerging regulatory pressure (EU AI Act). These differences have profound implications for theory development: universal claims about 'social media' must be qualified by platform-specific institutional contexts, and indigenous research on Chinese platforms is essential for globally valid theories of AIGC governance.

Regarding RQ5 (future methodologies and frameworks), this review advocates for three strategic directions. First, methodological innovation: future research should combine algorithmic auditing, longitudinal panel studies, and participatory action research to capture both technical mechanisms and social impacts. Second, interdisciplinary integration: computer scientists, communication scholars, legal experts, and ethicists must move beyond parallel silos to engage in genuine empirical dialogue. Third, governance design: dynamic frameworks are needed that balance innovation incentives with protective safeguards, incorporating technological transparency (watermarking, explainability), user empowerment (control interfaces, literacy programs), and platform accountability (auditable data access, liability clarification).

This review makes three primary contributions to the literature. Theoretically, we propose an integrative framework connecting technological applications, content production, recommendation mechanisms, and ethical governance, addressing the fragmentation identified in prior research. Methodologically, we critically evaluate existing evidence, highlighting methodological limitations and contradictory findings that warrant caution in interpretation. Practically, our comparative governance analysis and specific policy recommendations offer actionable guidance for platform

operators, regulators, and creators navigating the AIGC transition.

The limitations of this review should be acknowledged. First, the rapidly evolving nature of AIGC technology means that new developments may render some observations obsolete. Second, our reliance on published English and Chinese literature may miss important grey literature and non-academic sources. Third, the comparative analysis relies primarily on secondary sources rather than primary empirical data from both platform ecosystems. Future reviews should consider real-time evidence synthesis methods to keep pace with technological change.

In conclusion, AIGC represents not merely a technological upgrade for social media but a fundamental restructuring of the information ecosystem. Realizing its benefits while mitigating its risks requires coordinated action across academia, industry, and government. This review provides a comprehensive foundation for such coordinated efforts, and we hope it stimulates the interdisciplinary, methodologically rigorous, and culturally informed research that the field urgently needs.

## References

- [1] Zhang, E., Zhang, H., Yao, J., & Wang, S. (2024). Dynamic evolution and communication structure of AIGC issues: A comparative analysis based on Weibo and Twitter. *Journal of Xi'an Jiaotong University (Social Sciences)*, (03), 1–10.
- [2] Stief, J., Kirchner-Krath, J., & Morschheuser, B. (2026). How is generative AI transforming content creation on social media? An exploratory perspective on human-AI interaction processes, potentials and pitfalls. In *Proceedings of the 59th Hawaii International Conference on System Sciences* (pp. 76–85).
- [3] Zhao, Y., et al. (2025). Amplifying your social media presence: Personalized influential content generation with LLMs. *arXiv preprint arXiv:2506.14295*.
- [4] Anonymous. (2025). Generative AI for library social media content creation and communication. *Library Hi Tech News*, 42, 8–11.
- [5] M, M., et al. (2025). The impact of generative AI on social media: An experimental study. *arXiv preprint arXiv:2506.14295*.
- [6] Edelson, L., Haugen, F., & McCoy, D. (2025). A comparative survey of algorithmic feed recommendation system designs. *ACM Transactions on Recommender Systems*.
- [7] Salles, J. (2025). Affect and prediction in short-video social media recommendation algorithm: TikTok and the missing half-second. *New Media & Society*.
- [8] Peng, Z. (2026). The “filter bubble” effect and breakthrough strategies in social media news under algorithmic recommendation. *News Culture Construction*, (04), 121–123.
- [9] Yuan, K., Duan, X., Peng, S., Xuan, Q., & Ruan, Z. (2025). Influence of edge-based recommendation mechanisms on user opinion evolution in social media. *Journal of Chinese Mini-Micro Computer Systems*, 46, 2328–2337.
- [10] Carlon, D. (2025). Platforms as architects of AI influence: Rethinking moderation in the age of hybrid expression. *The International Journal of Human Rights*, 29, 701–729.