

An Analysis of the Publishing Industry's "Labyrinth of Responsibility" and Governance Paths in the Era of Generative Artificial Intelligence: From the Perspective of Actor-Network Theory

Yuwen Cao^{1*}, Jiahua Chen²

¹*School of Journalism and Communication, Beijing Normal University, Beijing, China*

²*School of Philosophy, Beijing Normal University, Beijing, China*

**Corresponding Author. Email: cyw990121@163.com*

Abstract. The rapid advancement of generative AI is reshaping the publishing industry, enhancing efficiency while triggering a severe responsibility and ethics crisis. It disrupts the traditional linear responsibility chain among authors, editors, and publishers, creating a "responsibility vacuum" and "responsibility gap". This study employs actor-network theory and distributed responsibility theory to analyze publishing as a heterogeneous network of human and non-human actors. It reveals how responsibility is dynamically "translated" and "distributed" within this network, identifying three mechanisms behind the "responsibility maze": asymmetric human-AI dependence, responsibility shirking among actors, and AI's unexpected autonomy. A multi-layered governance path is proposed, centered on forward-looking, proportional, and traceable responsibility, advocating for standards across micro, meso, and macro levels to foster a trustworthy human-machine collaborative publishing ecosystem.

Keywords: Generative artificial intelligence, responsibility, publishing ethics, actor network theory, distributed responsibility

1. Introduction

Generative artificial intelligence, particularly large language models, is fundamentally reshaping the global publishing ecosystem. While AI enhances efficiency across content creation, editing, and distribution [1], it simultaneously introduces profound ethical and governance challenges. The traditional publishing model, built upon a linear chain of responsibility with clearly defined roles for authors, editors, and publishers as "gatekeepers" [2], is being disrupted. The "black box" nature and autonomous capabilities of generative AI create a "responsibility vacuum" where accountability becomes ambiguous, and a "responsibility gap" where the link between human intent and AI output is obscured [3].

This erosion of clear responsibility attribution threatens the foundational trust between readers, authors, and publishers. To address this "responsibility maze," this study employs actor-network

theory and distributed responsibility ethics as its core analytical framework. It conceptualizes publishing as a heterogeneous network of human and non-human actors (e.g., authors, algorithms, data, developers) to investigate three research questions: how responsibility relationships are configured within this network (RQ1); how the "responsibility vacuum" and "gap" emerge and evolve (RQ2); and how a multi-layered, operational governance framework can be constructed for effective responsibility allocation and positive AI development in publishing (RQ3).

2. Literature review

This section reviews traditional publishing ethics and its challenges under AI, examines key debates on AI responsibility, and introduces an integrated theoretical framework of actor-network theory and distributed responsibility.

2.1. Challenges to traditional publishing ethics from AI

Traditional publishing ethics operates on a clear chain of responsibility, where authors ensure content originality and accuracy, while editors and publishers act as "gatekeepers" for review and quality control [4]. This model remained stable in the print era. However, generative AI represents a "paradigm shift" [5], blurring the lines between human and machine creation. The traditional binary author-gatekeeper structure becomes inadequate as AI participates throughout content production. The responsibility entity expands from media institutions and practitioners to include intelligent systems themselves, along with developers, operators, and users [6]. Core ethical concerns—plagiarism, inaccuracy, bias, and defamation—now originate from opaque algorithmic systems, creating a crisis for existing ethical frameworks in addressing AI-generated misinformation [7]. Issues like lack of transparency, unpredictability, and ambiguous accountability highlight the urgent need for new ethical constructions in the AI era.

2.2. The responsibility dilemma in AI ethics

The publishing industry's responsibility challenge reflects the broader AI "responsibility gap" [3]. This gap arises when autonomous systems cause harm that cannot be attributed to any specific human agent due to the unpredictability of AI behavior. Wang Tianen notes this as a phenomenon of unresolved responsibility, urging higher-level solutions [8]. Platforms often evade accountability by citing "technological neutrality," creating an imbalance between social risk and economic benefit.

Scholars propose solutions primarily focused on human actors. One approach insists on "human-centered" responsibility, tracing accountability to ultimate human decision-makers [9]. The "Meaningful Human Control" (MHC) framework ensures human responsibility through technological design [10], applied in contexts like AI social experiments [11]. Another approach advocates forward-looking responsibility [13], emphasizing anticipation of long-term technological consequences, such as systemic risks like cultural homogenization.

While valuable, these perspectives often focus narrowly on identifying ultimate human bearers of responsibility. They are less suited to analyzing the complex, mesh-like networks of human and non-human actors in modern publishing, necessitating a more descriptive theoretical tool.

2.3. Core theoretical framework: integrating actor-network theory and distributed responsibility

This study integrates Actor-Network Theory (ANT) and distributed responsibility to construct its analytical framework.

2.3.1. Actor-Network Theory (ANT)

ANT views the social world as a dynamic network of human and non-human actors linked through associations [14]. Its principle of "generalized symmetry" places humans (authors, editors) and non-humans (AI models, algorithms, data) on equal analytical footing. The key concept of "translation" describes how actors continually define each other's roles and interests [15]. In publishing, responsibility is not static but constantly translated. For example, delegating fact-checking to AI translates the editor's responsibility, while the AI's built-in filters represent another translation. The "responsibility vacuum" often emerges precisely where such translations fail or go unacknowledged.

2.3.2. Distributed responsibility theory

Luciano Floridi's theory of distributed morality [16] argues that in an information society, moral actions and consequences are distributed across multiple intelligent agents. This aligns with practical efforts like the EU's "Responsible Innovation" (RRI) and frameworks for "Responsible AI" design [17]. Chen Xiaoping's three-tiered AI ethics framework also provides a systematic reference for responsibility construction [18].

Combining ANT with distributed responsibility yields this study's core thesis: Responsibility in AI-era publishing has not vanished but exists in a distributed form across the entire actor-network. Developers bear responsibility via data and algorithm design; publishers through review guidelines; editors and authors through prompt engineering and final gatekeeping. The "responsibility gap" and "vacuum" arise not from an absence of responsibility, but from the opacity and inequity of its distribution and a lack of effective traceability and accountability mechanisms.

3. Method

This study primarily employs a qualitative research methodology that combines theoretical construction with case studies. By creatively integrating actor networks and distributed responsibility theory, viewing publishing activities as a dynamic network composed of both human and non-human actors. This provides an innovative theoretical perspective for analyzing the responsibility dilemmas brought about by generative AI. Based on this framework, specific publishing application scenarios reveal the process of the "translation" and "distribution" of responsibility among various actors within the actor network, constructing a systematic governance path encompassing three core principles and a multi-layered framework.

4. Research findings

This chapter applies the previously constructed theoretical framework to analyze the use of AI-assisted tools in publishing, tracing the trajectory of how responsibility is "translated" and "distributed" within the actor network. It systematically reveals the specific generative mechanisms behind issues such as the "responsibility vacuum" and "responsibility gap," and based on this, proposes a multi-level governance pathway.

4.1. Deconstructing the publishing industry's "responsibility maze" from an actor-network perspective

Using the example of a historian employing an AI writing assistant to produce a popular science book, a heterogeneous actor-network can be constructed, comprising human actors (author, editor, publisher) and non-human actors (large language model, training dataset, user interface, content filtering rules) [14]. Within this network, responsibility is not statically attached to a single node but is continuously defined, allocated, and transformed during human-computer interaction—a process of "translation." First, when an editor delegates fact-checking tasks to an AI tool, their responsibility undergoes its initial "translation": from "I must know all the answers myself" to "I need to be skilled at using the tool and judging its output." However, this delegation also entails a "shift" or "exemption" of responsibility, risking the dilution of the editor's core gatekeeping role by technology, as Pasquale warned in *The Black Box Society*, where automated systems can lead to "passing the buck" [19]. Secondly, AI models are not value-neutral tools. When designing and training models, developers pre-embed their value judgments and ethical considerations into algorithms and content filtering rules [20]. This means developers "translate" their social responsibility into the proxy actions of a non-human actor (the model), potentially creating a form of implicit censorship or bias, thereby pre-distributing some content responsibility. Finally, the "black box" nature of generative AI—the high opacity of its internal decision-making logic—is a key link in the "translation" of responsibility among various stakeholders. When AI generates content that seems plausible but is actually fabricated, authors, editors, and even developers can reasonably claim an inability to fully predict or explain the model's specific output, causing the chain of responsibility to break and become blurred, creating a "vacuum zone" for attribution.

Based on a micro-analysis of the responsibility "translation" process, three core generative mechanisms constituting the "responsibility maze" can be summarized. First, asymmetric dependence and accountability. Human actors (publishers, editors) increasingly rely on AI-assisted tools for efficiency. However, when ethical issues arise, existing laws and social norms cannot directly hold AI accountable, creating a contradiction of being "reliable but not accountable," which is the most direct source of the "responsibility vacuum." Second, the "ball game" of responsibility within the network. Because responsibility is distributed across the entire actor network, when problems occur, different human actors instinctively shift blame, creating a situation of "buck-passing," ultimately resulting in no one being held accountable for the overall consequences. Third, the "unintended autonomy" of non-human actors. The emergent capabilities of generative AI often exceed the expectations of its creators and users [5]. It generates innovative but factually flawed viewpoints based on hidden correlations in training data, or its outputs cannot be directly and linearly attributed to the original intent of any single human actor. This "unintended autonomy" swallows clear chains of causality, rendering traditional fault-based liability principles ineffective.

4.2. Distributed responsibility governance pathways for the publishing industry under generative AI

The preceding analysis indicates that the responsibility dilemma triggered by generative AI in the publishing industry stems from the dispersion, translation, and ambiguity of responsibility within complex actor networks. The core of governance is not to attempt to eliminate the distributed nature of responsibility, but to resolve the "responsibility maze" by proactively, visibly, and traceably managing and governing the distributed responsibility inherent in generative AI.

The governance pathway is constructed upon three core principles. First, proactive responsibility, necessitating a shift from passive "firefighting" to active "fire prevention." Philosopher Hans Jonas emphasized that a forward-looking model of responsibility must be adopted for technologies with far-reaching impact [13]. Stakeholders in the publishing industry must anticipate the potential risks of AI and take preventive measures. Second, proportional responsibility. Since responsibility is distributed, its allocation should be relatively fair. The responsibility of each (human) actor should be commensurate with their influence, control, and level of expertise within that network. Third, traceability. In the field of AI-assisted academic publishing, traceability is far more than a technical requirement; it is the ethical cornerstone for maintaining academic integrity and ensuring the verifiability of knowledge. A traceability mechanism that runs through the entire process of AI-assisted research, writing, and publishing must be established, recording key decision points and traces of AI involvement, thereby laying a technical and ethical foundation for clarifying distributed responsibility and rebuilding academic trust.

Based on the above principles, this study constructs a multi-level governance framework ranging from micro-level practices to macro-level regulation. At the micro-level, the focus is on specific practices within publishing workflows, establishing two solid lines of defense: firstly, implementing mandatory transparent disclosure, giving AI-generated content an "identity card" to protect the right to know of readers and the academic community; secondly, clearly defining "human responsibility" in AIGC creation, ensuring that "human-in-the-loop" retains the core responsibility for the final review, verification, and approval of AI output, transforming editors and authors into genuine "critical collaborative creators" and "gatekeepers." At the meso-level, relying on industry ethical norms and laws and regulations, the distributed responsibility of the publishing industry is standardized. National publishing associations should take the lead in formulating industry-wide ethical guidelines for AI application and establishing AI ethics review committees. Simultaneously, in collaborations between authors, publishers, and technology providers, the principle of proportional responsibility should be applied to clearly define the rights, obligations, and liabilities for breach of contract of all parties through contracts. Professional insurance products for AI-related liabilities should be actively explored to match distributed responsibility with distributed risk-bearing capacity. At the macro-level, government and industry organizations need to collaborate to establish clear rules and long-term governance mechanisms. Firstly, a "graded management" system should be implemented for different types of publishing content, imposing strict supervision on serious content while adopting relatively flexible management strategies for creative content. Secondly, ensuring that ethical rules keep pace with technological development, dynamically updating industry guidelines, piloting "AI Innovation Zones," and introducing a third-party ethics audit system to conduct regular "check-ups" on the fairness and transparency of AI systems. In summary, the goal of macro-management is to draw clear "tracks" and "safety red lines" for the application of AI in the publishing industry. Through layered management and dynamically updated rules, risks can be effectively controlled, safeguarding the lifelines of truth in journalism and academic integrity, while also encouraging bold innovation so that technology truly serves the prosperity of publishing culture.

5. Discussion and conclusions

This study identifies how generative AI leads the publishing industry into a "responsibility labyrinth" characterized by ambiguous accountability chains and dispersed moral actors. By applying actor-network theory (ANT) and distributed responsibility ethics as our core analytical framework, we have deconstructed the mechanisms behind this labyrinth and proposed a multi-

layered governance path for proactively managing ethical responsibility distribution in AI-driven publishing.

Our analysis demonstrates that generative AI's intervention disrupts the traditional linear "gatekeeper" model, creating both a "responsibility vacuum" where no single actor claims accountability and a "responsibility gap" where the connection between human intent and AI output becomes obscured [3]. By modeling publishing as a heterogeneous network of human and non-human actors, we reveal that responsibility doesn't disappear but undergoes continuous "translation" and "distribution" throughout the network. The ambiguity stems from three key failures in these micro-level translations: the misalignment in responsibility delegation, the pre-programmed values embedded by developers [20], and the broken decision chains caused by algorithmic black boxes. These failures manifest as three distinct mechanisms: asymmetric human-AI dependence and accountability, responsibility shirking among network actors, and the unexpected autonomy of non-human actors [5].

Rather than seeking a single responsible party or avoiding responsibility's inherent ambiguity, we argue the solution lies in embracing and managing responsibility's dynamic, distributed nature. Our proposed governance framework, built on principles of forward-looking responsibility [13], proportional accountability, and traceability, operates across micro, meso, and macro levels to construct a trustworthy human-machine collaborative publishing ecosystem.

The theoretical contributions of this study are twofold. First, we systematically integrate ANT and distributed responsibility ethics into the interdisciplinary domain of publishing studies and AI ethics, providing an analytical perspective that transcends anthropocentrism and enables more precise examination of responsibility flows in technologically mediated societies. Second, we advance understanding of digital-age publishing ethics by demonstrating how its focus must expand from content authenticity verification to long-term governance of entire content production and dissemination chains, opening new directions for future research.

Practically, this study offers clear guidelines for publishing stakeholders, while the multi-level governance framework provides theoretical foundation and practical reference for developing industry standards, ethical guidelines, and adaptive regulations.

This study has several limitations that indicate fruitful directions for future research. As a conceptual and theoretical inquiry, our framework requires empirical validation and refinement through qualitative methods such as in-depth interviews with publishing professionals navigating responsibility translation challenges. Furthermore, while we focus on text-generating AI, future research should address the more complex responsibility issues emerging from multimodal AIGC applications. Finally, the cultural specificity of our governance framework necessitates comparative studies exploring the adaptability and localization of distributed responsibility models across global publishing markets.

For the publishing industry, generative AI represents neither pure blessing nor absolute nightmare, but rather a prism revealing the insufficient institutional and ethical resilience of traditional industries facing technological disruption. We conclude that while we cannot prevent responsibility from distributing across human-machine networks, we can—through rational design and collaborative effort—establish effective governance systems to create a smarter, more diverse, and truly responsible future for publishing in the AI age.

References

- [1] Qin Yanhua, Li Yifan, Yan Lingling, Fu Jianing, Hou Yuli. A Glimpse into the Application of Artificial Intelligence Large Language Model in the Publishing Industry [J]. China Media Technology, 2024, 31(1): 34-40.

- [2] SHOEMAKER P J, VOS T P. Gatekeeping theory [M]. New York: Routledge, 2009.
- [3] MATTHIAS A. The responsibility gap: ascribing responsibility for the actions of learning automata [J]. *Ethics and Information Technology*, 2004, 6(3): 175-183.
- [4] WECKERT J, ADENEY D. Computer and information ethics [M]. Santa Barbara: Greenwood Press, 2013.
- [5] BOZKURT A. Generative artificial intelligence (AI) powered conversational educational agents: the inevitable paradigm shift [J]. *Asian Journal of Distance Education*, 2023, 18(1): 1-13.
- [6] Geng Xiaomeng, Yu Guoming. The Basis and Action Roadmap for the Construction of Ethics in Intelligent Media: Technological Reality, Ethical Framework and Value Adjustment [J]. *Modern Communication (Journal of Communication University of China)*, 2020, 42(01): 12-16.
- [7] JI Z, LEE N, FRIESKE R, et al. Survey of hallucination in natural language generation [J]. *ACM Computing Surveys*, 2023, 55(12): 1-38.
- [8] Wang Tianen. The ethical bridging of the "responsibility gap" in artificial intelligence applications: a case study of autonomous vehicles [J]. *Philosophical Analysis*, 2022, 13(01): 15-30+196.
- [9] TADDEO M, FLORIDI L. How AI can be a force for good [J]. *Science*, 2018, 361(6404): 751-752.
- [10] SANTONI DE SIO F, MECACCI G. Four Responsibility Gaps with Artificial Intelligence: Why they Matter and How to Address them [J/OL]. *Philosophy & Technology*, 2021, 34(4): 1057-1084.
- [11] Yu Ding, Li Zhengfeng. An experiment in intelligent society: the responsibility gap and governance of scenario innovation [J/OL]. *Studies in Science of Science*, 2022: 1-15.
- [12] Xia Yonghong, Li Jianhui. Research on the ethical responsibility of artificial intelligence medical diagnostic systems—focusing on the issue of responsibility voids [J]. *Chinese Journal of Medical Ethics*, 2020, 33(7): 796-802.
- [13] JONAS H. The imperative of responsibility: in search of an ethics for the technological age [M]. Chicago: University of Chicago Press, 1984.
- [14] LATOUR B. Reassembling the social: an introduction to actor-network-theory [M]. Oxford: Oxford University Press, 2005.
- [15] CALLON M. Some elements of a sociology of translation: domestication of the scallops and the fishermen of St Brieuc Bay [J]. *The Sociological Review*, 1984, 32(1_suppl): 196-223.
- [16] FLORIDI L. Distributed morality in an information society [J]. *Science and Engineering Ethics*, 2013, 19(3): 727-743.
- [17] VIRGINIA D. Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way [M]. Switzerland: Springer Nature Switzerland AG, 2019.
- [18] Chen Xiaoping. The ethical system of artificial intelligence: basic architecture and key issues [J]. *Journal of Intelligent Systems*, 2019, 14(4): 605-610.
- [19] PASQUALE F. The black box society: the secret algorithms that control money and information [M]. Cambridge: Harvard University Press, 2015.
- [20] SEAWARD L. The values encoded in the code: a guide to understanding AI ethics [J/OL]. *MIT Science Policy Review*, 2023 [2024-06-01]. <https://doi.org/10.38105/spr.3607402>.