Artificial Intelligence in the Transformation of the Scientific Journal Publishing Business Model and Its Implications: Postprint

Authors: Wang Yahui, Wang Jing

Date: 2023-10-08T00:00:00+00:00

Abstract

【目的】To explore the potential impacts of artificial intelligence on the publishing ecosystem of scientific journals.【方法】Employing methods of trend analysis and case demonstration, this study re-examines and scrutinizes the content production, dissemination ecosystem, and value logic of traditional scientific journal publishing under the technological paradigm of artificial intelligence.【结果】It is argued that future publishing ecosystems will become more intelligent, efficient, customized, and scenario-based, specifically manifested in algorithmic topic planning, automated content creation, intelligent manuscript processing, personalized distribution and dissemination, scenario-based reading experiences, and community-oriented academic communication; the paper also introduces some practical bottlenecks and challenges in AI application, such as algorithmic discrimination, data insufficiency, and "information cocoons," and proposes adaptive governance approaches to address AI-related risks and challenges, such as balancing human-machine relationships and reshaping editorial values.【结论】The shaping and reconstruction of the scientific publishing ecosystem by artificial intelligence is full-chain and multi-perspective, requiring proactive embrace and active integration, while also necessitating prudent treatment of the risks and challenges associated with AI applications and strengthened regulatory governance.

Full Text

Preamble

Artificial Intelligence' s Transformation of and Implications for the Scientific Journal Publishing Ecosystem

Wang Yahui, Wang Jing

(Institute of Agricultural Information, Chinese Academy of Agricultural Sciences, Beijing 100081)

Abstract:
[**Objective**] This study explores the potential impacts of artificial intelligence on the scientific journal publishing ecosystem.
[**Methods**] Using trend analysis and case-based reasoning, we re-examine and scrutinize content production, dissemination ecology, and value logic in traditional scientific journal publishing through the lens of AI technology.
[**Results**] We argue that future publishing ecosystems will become more intelligent, efficient, customized, and contextualized, manifesting as algorithm-driven topic selection, automated content creation, intelligent manuscript processing, personalized distribution, contextualized reading experiences, and community-based academic exchange. We discuss practical bottlenecks and challenges in AI application, such as algorithmic bias, data insufficiency, and "information cocoons," and propose adaptive governance approaches that balance human-machine relationships and reshape editorial values.
[**Conclusion**] AI' s reshaping and reinvention of the scientific publishing ecosystem is comprehensive and multi-perspective, requiring proactive embrace and integration, yet also necessitating prudent treatment of risks and challenges through strengthened regulation and governance.

Keywords: scientific journals; artificial intelligence; intelligence; algorithms; publishing
CLC Number: G223
Document Code: A
Article ID: 1671-0134(2023)01-052-04
DOI: 10.19483/j.cnki.11-4653/n.2023.01.008
Citation Format: Wang Yahui, Wang Jing. Artificial Intelligence' s Transformation of and Implications for the Scientific Journal Publishing Ecosystem [J]. China Media Technology, 2023(01): 52-55.

Throughout publishing history, every major technological breakthrough—from "lead and fire" to "light and electricity" to "digital and network" —has brought tremendous changes to the publishing industry. The new wave of technological innovation represented by artificial intelligence will inevitably drive historic transformations in traditional scientific journal publishing and accelerate the development of intelligent publishing integration. Some scholars have conducted valuable research on the integration of AI and publishing. For instance, Wang Xiaoguang argues that AI, as a disruptive technology, has potential impacts on multiple publishing stages including content creation, editing, and distribution, though these impacts are difficult to predict [1]. Wu Feifei discusses the possibilities of AI integration in content production, book editing, marketing promotion, and user demand mining, suggesting that AI technology could be regarded as the ecological environment for future publishing [2]. However, most studies either address publishing in general terms with overly broad perspectives lacking specificity to particular media, or focus on news publishing, book

publishing, and professional education publishing with little attention to scientific journals. Even among the few studies on scientific journal publishing, most provide overview or speculative discussions about AI's potential impacts on publishing ecosystems or workflows, with insufficient introduction of specific cases and practical applications, or they revolve around a single process or stage such as topic selection, peer review, or content editing, resulting in somewhat narrow coverage. Overall, current research on the integration of AI and scientific journal publishing at the micro and concrete levels—particularly regarding full-process, full-chain fusion strategies, innovation pathways, and constraints—remains inadequate [3-4]. This study presents a panoramic and concrete depiction of AI's future publishing logic and its comprehensive, multi-link transformation and disruption of the scientific journal publishing ecosystem.

1. Transformation of the Scientific Journal Publishing Ecosystem

AI employs data mining, intelligent algorithms, machine learning, and other technical means for data collection, analysis, and processing. Through intelligent perception and intelligent services, it will substantially transform a series of business processes in scientific journals, including topic selection, content creation, manuscript processing, distribution, reading experience, and community building. This will shape a completely new ecosystem for content production, processing, distribution, service, and consumption in the publishing industry, achieving an intelligent, efficient, precise, customized, and contextualized publishing landscape.

1.1 Algorithm-Driven Topic Selection

With support from big data and cloud computing technologies, AI leverages data mining algorithms and deep learning to capture, mine, and analyze massive datasets. These may include research hotspots and core topics in target disciplines, as well as reader preferences and behavioral characteristics (comments, replies, downloads, shares, etc.), helping editors discover and filter more timely and forward-looking content, thereby enhancing the efficiency and precision of topic selection.

In the publishing domain, AI participation in topic selection has achieved some positive progress. Inkitt, an emerging German publisher that began as an e-book writing community platform, has accumulated over 200,000 e-books available for reading. It uses algorithmic models to analyze readers' online reading behaviors and interest preferences, predicting market potential for e-books, and then assigns e-book topics flagged by algorithms as bestsellers to traditional publishers for print publication [5]. Founder Electronics has launched a publishing big data solution that, after deduplication and denoising of data from e-commerce platforms, reading and comment platforms, and social networks, manages data availability evaluation, machine learning, sentiment analysis, and

user profiling to help editors identify research hotspots and optimize publishing topics. Springer Nature' s SciGraph linked open data platform integrates high-throughput, cross-domain content resources and knowledge data including journals, papers, projects, patents, topics, researchers, institutions, and usage data. Based on data fusion, knowledge discovery, and content computing, it constructs a large-scale knowledge graph enabling academic knowledge data integration and correlation. Through intelligent semantic search and thematic aggregation exploration based on the knowledge graph, it helps grasp disciplinary development trajectories and identify distinctive topics [6]. Clarivate Analytics' Web of Science platform employs the LDA (Latent Dirichlet Allocation) document topic generation model, an unsupervised machine learning technology that analyzes and processes disciplinary data in specific fields, filters key hotspots, extracts thematic vocabulary through algorithms, and automatically forms specific topics according to significance, providing convenience for journal topic selection, manuscript solicitation, and hotspot tracking [7].

1.2 Automated Content Creation

Automated content creation first emerged in news media, particularly achieving large-scale application in news reporting by major domestic and foreign media outlets, with typical examples including the Associated Press' s Wordsmith automated writing platform, Xinhua News Agency' s "Kuai Bi Xiao Xin," and Tencent' s Dream Writer. In academic and professional publishing, in April 2019, Springer Nature published the world' s first AI-authored academic book, *Lithium-Ion Batteries*, completed by an algorithm called "Beta Writer" developed in collaboration with the Applied Computational Linguistics Laboratory at Goethe University Frankfurt. Using a similarity-based clustering program, the algorithm sorted source documents into coherent chapters and generated concise paper abstracts, with citations provided as hyperlinks for reader convenience. By automatically summarizing over 53,000 academic documents in the lithium-ion battery field, it reviewed cutting-edge developments in the technology [8]. Scientists from Rensselaer Polytechnic Institute and other institutions developed an AI academic assistant called "PaperRobot" that, using natural language processing and natural language understanding technologies, deep-learned from large volumes of published papers in biomedicine to construct background knowledge graphs and assist in writing paper abstracts, key content, conclusions, and even further research suggestions on given topics. Turing test results on PaperRobot-generated papers showed that machine-created abstracts, conclusions, and research suggestions achieved higher acceptance rates than human-written ones [9].

It is worth noting that these AI experiments do not mean that machines will soon replace humans in completing logically clear, rigorously argued research papers with innovative discoveries. AI currently can only organize, link, and aggregate existing knowledge and concepts based on knowledge databases or background knowledge graphs. It does not and will not in the short term create

new knowledge or scientific discoveries, which are precisely the crystallization of wisdom from human continuous exploration, arduous 攻关, and collaborative innovation.

1.3 Intelligent Manuscript Processing

During the manuscript solicitation phase, AI can leverage knowledge graphs, association analysis, and text mining technologies to explore academic resources from major literature retrieval platforms and online databases, screen important experts and scholars in relevant research directions, track their research dynamics, evaluate their academic influence, and promptly send them invitations to contribute, thereby enhancing the relevance and matching degree of solicited manuscripts. AI-based academic search tools such as Semantic Scholar, Iris.ai, UNSILO, and Yewno employ machine learning and semantic analysis to help users quickly understand paper content, locate, classify, and filter valuable research, authors, and institutions, bringing great convenience to topic selection and manuscript solicitation [10-12]. For instance, UNSILO uses natural language processing and machine learning to analyze and extract paper content, identifying authors' main arguments and research findings, and can also capture full texts of millions of biomedical papers from the PubMed Central academic database, enabling editors to quickly gain a general understanding of research findings [11]. The research team from Tsinghua University's Department of Computer Science and Technology has built a big data analysis and service platform for scientific and technological resources called "AMiner," which encompasses over 230 million academic papers, patents, and 136 million scholars, integrating functions such as intelligent expert profile extraction, intelligent expert search, academic big data fusion, and academic evaluation, providing powerful search capabilities for scientific literature, experts, and academic activities.

During the peer review phase, addressing the limitations of existing plagiarism detection software that uses character-by-character matching and cannot recognize synonyms or similar sentences, some intelligent academic misconduct detection tools support recognition of entire sentences or partial paragraphs. Some anti-plagiarism tools have even developed chart detection capabilities to identify forged images, helping combat academic plagiarism more effectively. For example, Elsevier' s Evise automated editing system checks for plagiarism by retrieving and matching documents against the CrossCheck database. A research team from Syracuse University led by Dr. Daniel Acuna published an article on bioRxiv introducing a machine learning algorithm that used a keypoint-based detection method to examine over 2 million images from 760,000 open-access papers across 4,324 journals in the life sciences, finding that approximately 9% of images contained high degrees of duplication [13]. AI technology can also intelligently recommend suitable reviewers. For instance, Frontiers Publishers' AI Review Assistant (AIRA) combines internal custom algorithms and embeds functionalities from Google' s and CrossRef' s iThenticate (document originality checking tool) and Editage' s Ada (automated paper evaluation tool) to quickly

and accurately assess manuscript quality, match potential peer reviewers, and check for potential conflicts of interest among editors, reviewers, and authors.

During the editing and proofreading phase, large amounts of low-end, repetitive editorial work can be efficiently completed by AI-powered automatic inspection and error correction systems, including checking for errors in words, sentences, grammar, and rhetoric; verifying accuracy of terminology and measurement units; ensuring standardization of structure and formatting; and confirming appropriateness of statistical methods. For example, Elsevier' s Aries submission system employs the AI capabilities of StatReviewer software to verify the completeness and reliability of papers' experimental methods, statistical data, and research conclusions [11]. Founder Electronics' intelligent auxiliary proofreading system applies machine learning and deep learning technologies, utilizing methods such as word segmentation, entity recognition, syntactic analysis, and deep language models, and has initially developed 11 functions including checks for easily confused words, sensitive words, non-standard terminology, connectors, full-width/half-width characters, unit capitalization, and chart and formula numbering.

1.4 Personalized Distribution and Dissemination

Scientific journals have accumulated rich content resources including research outcomes and literature, as well as user information on authors, readers, experts, and editorial board members through long-term publishing practice, laying the data foundation for using intelligent algorithms to achieve precise content distribution. AI can track users' browsing, downloading, commenting, and sharing behaviors of academic resources across various data platforms and social media, depict precise user profiles, and complete adaptive and personalized content distribution based on users' reading needs and interest areas.

TrendMD, a Canadian publishing technology service company, provides personalized and precise recommendations for cross-platform related articles. By installing plugins on cooperating journal websites' backends, it indexes journals' historical metadata and uses collaborative filtering technology in its cross-platform content recommendation module to achieve precise recommendations for papers on cooperating journal platforms and third-party platforms based on readers' reading trends, delivering up to 800 million article recommendation links monthly to over 100 million readers. Chaoxing Group has launched the "Yuchuang" mobile publishing platform, providing intelligent learning resources for users. Its distinctive "fingerprint collection" function, combined with users' source data such as "favorites," "recently browsed," and "reading rankings" on the platform, can establish personal reading behaviors and curves and push customized content according to reading patterns.

1.5 Contextualized Reading Experience

Unlike traditional publishing' s static communication dominated by text and images, media communication in the contextual era pays greater attention to the immersive and visual experiences that content products create for users. Technologies such as VR (Virtual Reality), AR (Augmented Reality), and MR (Mixed Reality) enrich traditional publishing' s content presentation formats and user interaction methods, enabling more profound and intuitive expression of content and information, and delivering "on-site" experiences with deep immersion and active participation for users.

Introducing VR/AR technologies into scientific journal publishing can also enrich readers' reading formats and enhance their reading experience, particularly demonstrating significant application advantages in three-dimensional reproduction of instruments and equipment, experimental processes, and medical procedures in science, engineering, agriculture, and medical papers. *Journal of Shanghai University (Natural Science Edition)* exploratorily applied static and dynamic AR presentation to papers in its "Ternary Optical Computer" special issue. By downloading the paper-based AR app on intelligent terminals and scanning images in the text, readers can perfectly display ternary optical computer objects and experience videos on mobile devices [14].

1.6 Community-Based Academic Exchange

Leveraging big data technologies for classification, clustering, and association rule mining, and utilizing AI' s precise recommendation and effective matching, scientific journals can find industry peers with common academic interests to establish user communities, creating broader academic spaces and enhancing platform benefits and brand value through community dissemination. Within academic communities, users can share academic developments in real time, exchange latest research findings, and obtain personalized and precise academic resources on topics of interest.

Frontiers has consistently adhered to the concept and practice of community-driven publishing, developing the academic social platform Loop where scientists can publish research outcomes and academic activities on personal homepages, upload resources such as papers, videos, and news, follow each other, join academic groups, recommend valuable academic content to research peers, and even organize academic exchanges online. The Open Science Identity (OSID) platform initiated by the Key Laboratory of Publishing Integration Development (Wuhan) of the National Press and Publication Administration offers similar functions, supporting journals, editors, authors, and readers in building academic exchange communities for academic discussion, resource sharing, open exchange, and social interaction on journal papers and disciplinary hotspots within academic circles.

2. Bottlenecks and Challenges in Scientific Journal Publishing Applications

AI' s reshaping of traditional scientific journal publishing business processes and value innovation will spawn new ecosystems, models, and logic in publishing. However, fully applying AI in scientific journal publishing still faces practical difficulties and challenges, specifically manifested as serious homogenization in algorithm-driven topic selection and content creation, insufficient creativity, potential information narrowing from precise push notifications, and widespread difficulties in publishing data sharing and data barriers.

2.1 Homogenization in Algorithm-Driven Topic Selection and Content Creation

AI relies on machine intelligence to track and filter academic research hotspots and cutting-edge dynamics, pushing traditional publishing topic selection methods toward intelligent efficiency. However, algorithm-driven blind pursuit of hotspots and buzzwords may cause homogenization in topic direction and content across different journals in the same or related disciplines, lacking differentiation and innovation, and failing to demonstrate journal characteristics, which is detrimental to disciplinary innovation and the healthy development of the publishing ecosystem.

2.2 Insufficient Disciplinary Professionalism and Creativity in AI

Currently, AI remains in the stage of weak artificial intelligence and does not possess the reasoning and creative abilities unique to human consciousness, nor the capacity for analytical thinking and viewpoint generation. Automated creation is more suitable for templated and procedural content production. Due to the lack of professionalism and creativity, AI is not yet capable of composing scientific papers requiring specialized research, in-depth interpretation, and complex verification.

2.3 Potential Information Narrowing from Precise Push Notifications

Providing personalized recommendations based on user interests and preferences, while improving information distribution efficiency, can also create "information cocoons" [15], thereby narrowing users' information reception, limiting their perspectives, and solidifying their thinking. This is evidently detrimental to scientific journal audiences, primarily researchers, in broadening their information scope and knowledge base, particularly for interdisciplinary and cross-disciplinary learning and research.

2.4 Difficulties in Publishing Data Interaction and Sharing

Data is the "nourishment" for AI; without data support, AI application is impossible. However, in traditional publishing institutions, data value has not

received sufficient attention. Data accumulation, mining, and utilization regarding content production, content dissemination, publishing workflows, and user interaction are inadequate, and the potential value of this data has not been realized. Moreover, data across different publishing institutions and literature data platforms are fragmented, making open sharing difficult, with widespread data silos and barriers.

3. Response and Governance for AI Risks and Challenges

While AI brings a series of new opportunities for transforming the scientific journal publishing ecosystem, it inevitably accompanies a chain of new challenges. Addressing AI risks and challenges cannot focus solely on technical aspects; fundamentally, we must properly handle the relationship between AI and humans. Specifically regarding the regulation and governance of AI application risks and challenges in scientific journal publishing, the basic approach is: balance human-machine relationships, reshape editorial values, and enhance collaborative governance.

In the new era of AI, we need to deepen control over scientific journal publishing workflows, understanding of editorial activity patterns, and adherence to publishing ethics and cultural values. We must strengthen forward-looking prevention, weigh technological pros and cons, and reconcile value conflicts and ethical dilemmas. Algorithms should be made transparent, open, and understandable, enhancing algorithmic interpretability, comprehensibility, and predictability, while paying attention to and guarding against algorithmic bias, algorithmic power abuse, and "information cocoons" to solve problems such as utilitarian topic selection, templated content, and singular push notifications. We must strengthen editors' subjective consciousness, sense of responsibility, and gatekeeping awareness, allowing them to exert cultural leadership and subjective initiative in humanistic care and value judgment during topic selection and peer review, compensating for AI' s technical defects and value anomie, and building a publishing ecosystem where AI technology and editorial ethics coexist harmoniously. We must promote open sharing, standard unification, and business collaboration of industry publishing data. We need to intensify AI education and talent cultivation for the entire population, particularly improving editors' professional literacy in understanding and applying new AI technologies.

AI' s transformation and reinvention of the publishing ecosystem is comprehensive and multi-angle, representing further liberation of publishing productivity and deep-level transformation of production relations. We need to embrace AI with a developmental perspective and open mindset, actively integrating AI technology into the publishing industry' s development process. We must strengthen editors' leading position and scientific guidance responsibility, properly guiding and utilizing AI, while also carefully evaluating technical, legal, and ethical issues such as algorithmic bias, "information cocoons," data overreach, and infringement accountability.

References:

[1] Wang Xiaoguang. Artificial Intelligence and the Future of Publishing [J]. Technology and Publishing, 2017(11): 4-6.
[2] Wu Feifei. Integration and Application of Artificial Intelligence Technology in the Publishing Industry [J]. Publishing Wide Angle, 2018(1): 26-28.
[3] Dai Ni, Bu Zhaode. The Impact of Artificial Intelligence on Scientific Journal Publishing [J]. Publishing Wide Angle, 2021(11): 46-48, 51.
[4] Zhang Yifang, Zhang Ning. The Impact of Artificial Intelligence on the Editing and Publishing Industry of Medical Academic Journals in the All-Media Era and Coping Strategies [J]. China Media Technology, 2020(9).
[5] ALBAZAZ A. How Inkitt publishes your books: from preparation to promotion [EB/OL]. https://degfhkjgfjhf.inkitt.com/how-inkitt-publishes-your-books-from-preparation-to-promotion. 2017-05-27/2019-12-18.
[6] Song Ningyuan. Construction of Scientific Knowledge Graphs for Smart Data —Taking SciGraph as an Example [J]. Technology and Publishing, 2017(11): 18-19.
[7] An Qi. New Explorations in Academic Journal Development in the Big Data Era [J]. Chinese Editors, 2017(7): 58.
[8] Huanqiu.com. Publishing House Springer Nature Begins Using Machines to Write Books [EB/OL]. https://smart.huanqiu.com/article/9CaKrnKjvDR. 2019-04-03/2019-12-20.
[9] Wang Q, Huang L, Jiang Z, et al. PaperRobot: incremental draft generation of scientific ideas[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy: Association for Computational Linguistics, 2019: 1980-1991.
[10] Li Yuan. Digital Dissemination of Academic Journals in the Artificial Intelligence Era [J]. Chinese Journal of Scientific and Technical Periodicals, 2019(11): 1183-1190.
[11] HEAVEN D. AI peer reviewers unleashed to ease publishing grind [J]. Nature, 2018(7733): 609-610.
[12] Sun Henan. Integration Development of Artificial Intelligence Technology and Academic and Professional Publishing [J]. Publishing Wide Angle, 2018(7): 35-36.
[13] ACUNA D E, BROOKES P S, KORDING K P. Bioscience-scale automated detection of figure element reuse [J]. bioRxiv, 2018.
[14] Zhao Yu. AR Publishing Practice of *Journal of Shanghai University (Natural Science Edition)* [J]. Acta Editologica, 2019(1): 74.
[15] Chang Qin. Ethical Anomie Caused by Artificial Intelligence Technology in News Communication and Countermeasures [J]. China Media Technology, 2020(11): 28-30.

Author Biography: Wang Yahui (1986-), male, from Baofeng, Henan, Ph.D. candidate, Assistant Researcher, Director of the Editorial Department of Agricultural Communication and Journal Publishing, Institute of Agricultural

Information, Chinese Academy of Agricultural Sciences. Research direction: journal editing and publishing.
(Responsible Editor: Zhang Xiaojing)

Note: Figure translations are in progress. See original paper for figures.