

WhitePaper

Revolutionizing Journal
Publishing: Leveraging Al and ML
to Overcome Workflow Hurdles



## **Table of Contents**

ntroduction	03
The Evolution of Journal Publishing	03
Current Challenges in the Publishing Industry	04
Lengthy article submission and pre-review processes	04
Unavailability of manual reviewers and editors	04
Time-consuming structuring and styling tasks	05
Delays in pagination and proof review stages	05
Understanding the power of Advanced AI Tools	06
Deep learning (DL) for content production	06
Natural Language processing (NLP) accelerated internal publishing processes	06
Use cases of NLP and DL in the publishing industry	06
Transforming Manuscript Submission and Peer Review	07
Leveraging NLP and DL to automate article screening and filtering	07
Streamlining submission workflow with NLP	08
Enhancing peer review efficiency through Al-powered recommendation systems	08
Improving communication and collaboration between authors and reviewers using AI-bases tools	08
Efficient Structuring and Styling of Content	09
Automated typesetting, structuring and formatting using NLP algorithms	09
Al-driven tools for content curation, recommendation and flow optimization	09
Content discovery and personalization	09
Guided Editing and Proof Review	10
Al-assisted editing tools for grammar, syntax, and language enhancement	10
Streamlining proofreading via automated error detection	10
Al-powered pagination algorithms for automated layout creation	10
Reducing errors and iterations with Al-driven proofreading solutions	10
Al Disrupting Scholarly Publishing	11
The Future of Journal Publishing	12
Bibliography	13

### **Introduction**

According to UNESCO, "Al systems have the capacity to process data and information in a way that resembles intelligent behavior, and typically includes aspects of reasoning, learning, perception, prediction, planning or control," and all these are effectively being leveraged by the publishing industry. (UNESCO, 2021)

Al isn't just the latest buzzword; it is a tool that is transforming multiple industries. As of May 2023, Anthropic's generative Al tool attained the capacity to process 75,000 words per minute, almost the length of a novel. This opens massive opportunities for publishers and authors to incorporate Al technologies in their workflows. This explains why the global Al market size, which was estimated at \$136.55 billion in 2022, is expected to expand at an exponential CAGR of 37.3% from 2023 to 2030.

1990 is regarded as the " year of Literature" because it introduced eBooks to the world.

- The first eBook (Bizcommunity)

# The Evolution of Journal Publishing

Scholarly publishing dates back to the 17th century, when research was shared within a tight circle of experts on a subject. It was the only mechanism to share information across the academic landscape. The introduction of commercial publishers in the 19th century helped snowball the dissemination of research, establishing a subscription-based model to access knowledge (Allahar & Jack, 2021). This introduced advertising in the otherwise knowledge-centric industry (Tucker, Unwin, & Unwin, 2020). The Internet and ensuing digitization across industries disrupted and disintermediated publishing while transforming reader behavior in the later years of the 20th century. Open access, via diverse subscription models, further demystified the closed world of research and made it accessible to all, opening new revenue streams for publishers.

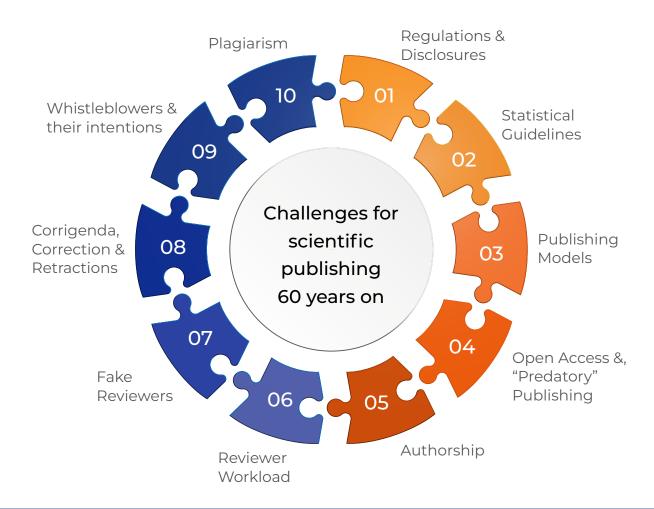
The most recent shift in the publishing industry is the adoption of artificial intelligence-based technologies. These technologies have significantly empowered researchers, authors, and publishers to accelerate the production of accurate and reliable content. Many publishers had started applying AI across processes as far back as 2017. Today, this technology offers publishers access to a wide range of tools to not just enhance the publishing cycle but also strengthen areas, such as marketing and analytics, and production and administration. (Lovrinovic & volland, the future impact of artificial intelligence on the publishing industry, 2019).

All is benefiting the industry by allowing publishers to overcome challenges, and improve the experience for authors, editors, reviewers, and the end user.



### | Current Challenges in the Publishing Industry

A publisher has to identify the target audience, address their needs and tastes, and boost engagement with the content. But that's not all. The publishing process itself faces several challenges.



### • Lengthy article submission and pre-review processes

The very first stage of submission can be the trickiest for authors to navigate. A manuscript can take anywhere from 4 to 8 months to move from the submission to the peer review stage and then final acceptance. However, in the case of a "revise and resubmit" scenario, it can take over a year to get an article accepted and then a few more months for publishing when the entire publishing workflow is manual.

### • Unavailability of manual reviewers and editors

Getting review invitations accepted after finding a suitable reviewer is another big challenge in the publishing lifecycle, according to 75% of journal editors. Also, 50% of the reviews are performed by 10% of the pool, making the review space skewed to a certain section only. (Fernz, 2020)

### • Time-consuming structuring and styling tasks

Publishers receive huge volumes of submissions from authors. So, a manuscript needs to stand out not only in terms of the research adding value to the field but also the quality of the content. Improper content organization, styling, or structuring can lead to rejections, as can poor language quality.

Content structuring and styling take up a significant amount of the authors' or reviewers' time. Since authors need not be in content creation, multiple iterations might be needed to meet the journal's guidelines.

### Delays in pagination and proof review stages

The manual pagination and proofing process is tedious and repetitive. With limited manpower available, this stage gets pushed back due to pressing needs in other stages of the publishing cycle.



### Understanding the Power of Advanced AI Tools

The most significant contribution made by AI to the digital publishing industry is the automation of time-consuming and labor-intensive tasks. In fact, automating various activities in the research and publishing workflows is increasingly becoming an integral part of the scholarly publishing process.

### • Deep learning (DL) for content production

By 2030, Al-driven, synthetically produced data will surpass the use of real data. (Sheets, 2021) Deep Learning is the broader technology category that uses data from diverse sources to power machine learning (ML). It leverages neural networks, Internet of Things () devices, augmented and virtual reality gadgets, and data from social media to make inferences and discover trends and patterns. (Sarker, 2021) DL is enabling the production of high-quality, completely or partially generated, synthetic content at scale. It is also powering publishers to discover trends and evaluate content performance to make informed strategic decisions.

### Natural language processing (NLP) accelerates internal publishing processes

NLP uses linguistically motivated strategies, such as lemmatization, syntactic analysis, semantic role labeling, and anaphora resolution to assist in building interpretable representation from free text. It enables editing software to understand the content. (Gamallo & Garcia, 2019) It is built on large language models (LLMs) that recognize all aspects of a language, including structure, meaning, and intent. (Kulesz, 2022) The technology facilitates automation of editing, so that authors and editors can focus on content enhancement.

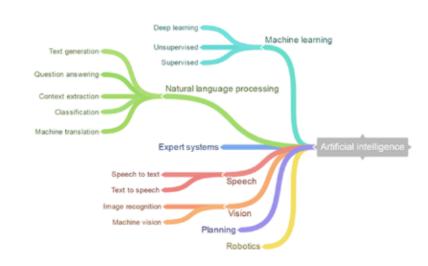
### Use cases of NLP and DL in the publishing industry

- Copy Editing: NLP now assesses manuscripts for grammar, semantics, research quality, and conclusions. It categorizes content, creates indices, summaries, and aids in marketing copy. It eliminates non-essential editing overheads, boosting manuscript production by 40%.
- Enabling Global Scholarly
   Dissemination: Leveraging
   Al-driven translation, non-native authors can seamlessly share their research in international journals, breaking down language barriers for broader scholarly contributions.
- Fostering Inclusivity: Context
   analysis and deep linguistic
   learning, along with image
   recognition and video transcription
   facilities, help publishers automate
   the generation of alt-text for better
   accessibility. They also provide
   recommendations for better DEI
   measures and compliance.

Automated initial screening places the content in the appropriate editing stream with an unbiased and intelligent perspective on the submissions.

### | Transforming Manuscript Submission and Peer Review

Al boosts publishing efficiency by enabling publishers to provide intelligent knowledge services and optimize the supply of content digitally. Simultaneously, the technology is being leveraged to streamline processes and intelligently manage the production and publishing workflows (Zhao & Prabhashini, 2019). Modern publishing solutions are leveraging Al technologies to simplify article submission, peer review, editing, and proofreading.





### Leveraging NLP and DL to automate article screening and filtering

Text analysis is one of the biggest applications of AI in scholarly publishing at 40% of all its implementations (UNSILO, 2019). NLP facilitates the first-level article screening by matching the topic with the publication's genre, the accuracy of the title according to content, and adherence to the journal's submission guidelines. Automated assessment also detects plagiarism reducing the editorial workload (Pacholi, 2023). This eliminates the manual labor required to identify whether a manuscript should or shouldn't enter the publishing process. However, 25% of the submissions are ready to proceed to the next stage, which is accelerated with the help of automated screening and filtering saving about 10% to 40% of production time.

### Streamlining submission workflow with NLP

NLP powers the implementation of pre-set grammatical and formatting rules to analyze and score submitted manuscripts for their quality and alignment with the guidelines of the target journal (Groth, 2019). This helps move high-quality manuscripts straight to the typesetting and composition stages. Al-powered writing tools automatically declare author (or an Al-powered tool) contributions, and conflicts of interest enhancing the submission process (Ciaccio, 2023).

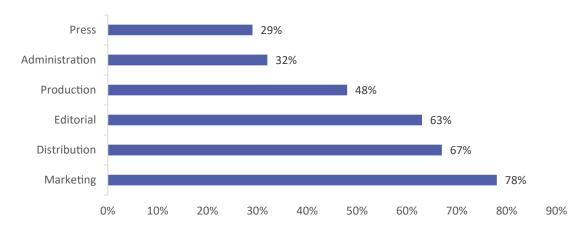
### Enhancing peer review efficiency through AI-powered recommendation systems

Al-powered mapping of experts in language and subject matter makes the peer review process more efficient. Complex recommendation systems surpass the human capability to provide authors and researchers with suggestions about books, or journals that may assist in their research and can enhance their experiences (lovrinovic & volland, the future impact of artificial intelligence on the publishing industry, 2019). These engines can filter through catalogs and backlists to facilitate access to the most relevant content for user requests across genres, and languages. Al helps ensure that the 4 virtues of reviews are upheld-timeliness, objectivity, clarity, and fairness.

### Improving communication and collaboration between authors and reviewers using Al-based tools

Al-powered tools facilitate communication across the publishing teams. This allows humans and technology to work in tandem. They enable real-time collaboration and feedback during the proofing and review stages. All comments given by involved individuals are recorded and color-coded to enable the author to respond to them. Further, tools like perspective implemented by the new york times, are enabling open access to readers to comment on their points of view while the Al engine assess comments on various criteria to gauge their relevance in the context of the article (lovrinovic & volland, the future impact of artificial intelligence on the publishing industry, 2019). This can be translated into publishing to enhance manuscript quality, get diverse inputs, and add credibility to the content.





### | Efficient Structuring and Styling of Content

Seamless structuring and styling can speed up the delivery of research findings to the audience. This can be achieved by reducing manual touchpoints and automating routine and repetitive tasks.

### Enhancing peer review efficiency through Al-powered recommendation systems

The first "programmable" automated typesetting system, TeX, was launched in 1978. The tool automated print-ready page production (McIlroy, 2023). The advent of Al added a new dimension to formatting manuscripts according to each publisher's unique guidelines. The rules can be algorithmically applied to documents to reduce typesetting time from hours to minutes.

#### Content discovery and personalization

Al is transforming content categorization from a passive activity to actively usable insights. Meta-tags and markups are facilitating recommendations and discovery. Reader segmentation, similar to suggestive algorithms of OTT platforms, customizes newsletters, articles, and other content suggestions according to the audience's preferences (Pacholi, 2023). This is facilitating personalization and boosts content discovery. This not only improves revenues but also facilitates data that can be used to suggest research trends.

45% of seasoned AI adopters state that artificial intelligence has enabled them to establish a significant competitive advantage.

Source: The Future of AI in Publishing (Lineup, The Future of AI in Publishing, 2023)

### Al-driven tools for content curation, recommendation and flow optimization

Al-powered tools like Google Scholar help researchers find relevant citations and research. In addition, task management tools and note-making facilities allow authors to manage bibliographies and references. (Pacholi, 2023). Context-based analysis, powered by NLP, enables content curation, recommending appropriate article flows to enhance the writing process. This helps non-native and/or less experienced authors, who might find it difficult to maintain language quality and thought flow through a long manuscript.

Al-driven suggestions ensure consistency and coherence in language and writing style. These writing tools assist in clear and concise writing while ensuring grammatical accuracy, improving the overall quality of work, right from the writing stage. Activities, such as grant writing, need experts from diverse domains. Al can assist in multiple ways in such situations. First, by enabling topic-based research across domains and second, by identifying the most suitable experts to collaborate with, based on their areas of work, expertise, schedules, and even working styles. (Ciaccio, 2023).

### | Guided Editing and Proof Review

Al-powered editing and proofing reduces manual labor and assists in the publishing of high-quality papers.

#### • Al-assisted editing tools for grammar, syntax, and language enhancement

NLP has evolved such that it can extract context from the content, even if the language is not perfect. NLP-powered tools provide suggestive corrections for grammatical errors while keeping domain-specific jargon at the fore. These tools are free of past biases and automate translation, improving reach without adding to the translation costs. (Groth, 2019).

Al-powered editing tools make corrections based on syntax rules when there is no language ambiguity and highlight the sections and provide suggestions for places where a black and white decision isn't possible.

### • Streamlining proofreading via automated error detection

Automated detection of errors in language translation, plagiarism checks, and copyright management reduces manual proofing effort. Further, the unbiased selection of reviewers instills diversity and trust in the process.

#### Al-powered pagination algorithms for automated layout creation

Writers and editors are freed from having to custom style manuscripts to meet the academic journal's requirements. Pagination, alignment, image formatting, and layout design can all be automated based on desired criteria. Al-powered tools can follow multiple style guides to simultaneously copyedit content, summarize articles, press releases, and news reports, and also design them to enhance readability.

#### • Reducing errors and iterations with Al-driven proofreading solutions

Manual editing is often plagued with personal biases and perspective differences among reviewers. Al-driven proofing is unbiased and maintains a defined perspective based on the author's intent and the journal for which the article is being published. Thus, even when a human reviewer is involved, any discrepancies are highlighted at the first occurrence. This prevents multiple iterations and inconsistency errors in the paper.



### **AI Disrupting Scholarly Publishing**

Open access has led to scholarly collaboration networks and enabled sharing of research beyond traditional boundaries. Applications of AI are apparent through the entire publishing value chain, across the content creation, production, and distribution stages. (Kulesz, 2022) AI is not only enabling and encouraging publication, but it is also helping minimize research fraud by identifying negative/null findings, protocols, and

incremental studies (Lawrence & Alam, 2023) Scholarly publishers worldwide are adopting Al to boost production capacity, expedite the publishing journey, and deliver high-quality research. In fact, 65% of academic publishers using Al consider acceleration of the process as a key benefit of leveraging the technology in the publishing process. (UNSILO, 2019) Man-hours saved translates directly into reduced costs. This makes way for the application of Al across business processes.

Al has the potential to come close to an instantaneous publishing model as it evolves, magnifying production capacity.

The technology offers benefits in content creation, editing, proofing, translation, recommendation, and analytics. It accelerates production, reduces manual work, and optimizes the publishing lifecycle, resulting in cost savings for publishers and increased revenue from open access.

Al facilitates scaling without sacrificing quality.





### The Future of Journal Publishing

Al has the potential to transform into a collaborator, from its current role as an assistant, for the publishing industry. As Al evolves, robotic writing will be used to create manuscripts from verbal inputs of researchers. (Zhao & Prabhashini, 2019) The combination of language translation and audio to text conversion will further accelerate content production and aid multilingual distribution.

Leveraging aggregated knowledge and sophisticated data processing, researchers can now compose articles by identifying relevant content. This technology enables forecasting citation impact for yet-to-be-published articles and enhances SEO for improved discoverability. However, distinguishing Al-generated from human-created content and ensuring content security remain critical challenges requiring further research and industry-wide regularization.

Publishers must actively adopt futuristic technologies to power the publishing cycle without compromising on quality or budgets. Partnering with innovative technology providers is the key to ensuring growth and flexibility for an impending paradigm shift in the global dissemination of knowledge.

### Bibliography

Allahar, H., & Jack, A. L. (2021, December 23). The Evolution of Academic Journal Publishing. Retrieved from ResearchGate: https://www.researchgate.net/publication/357285955\_The\_Evolution\_of\_Academic\_Journal\_Publishing

Baidu, I. (2021, August 18). Baidu Announces Upgraded Baidu Brain 7.0 and Mass Production of 2nd Generation Kunlun Al Chip. Retrieved from https://www.prnewswire.com:

https://www.prnewswire.com/news-releases/baidu-announces-upgraded-baidu-brain-7-0-and-mass-production-of-2nd-generation-kunlun-ai-chip-301358126.html

Bizcommunity. (n.d.). Digital Publishing: Short History. Retrieved from Bizcommunity: https://www.bizcommunity.com/Article/196/367/106976.html

Ciaccio, E. J. (2023, April 23). Use of artificial intelligence in scientific paper writing. Informatics in Medicine Unlocked. Retrieved from Informatics in Medicine Unlocked: https://www.sciencedirect.com/science/article/pii/S2352914823000953

Fernz, A. (2020, September 23). Important pre-submission checks to help you sail through peer review. Retrieved from Editage Insights: https://www.editage.com/insights/important-pre-submission-checks-to-help-you-sail-through-peer-review?refer=scroll-to-1-article&refer-type=article

Gamallo, D. P., & Garcia, D. M. (2019, May 31). Special Issue "Natural Language Processing and Text Mining". Retrieved from MDPI: https://www.mdpi.com/journal/information/special\_issues/NLP\_TM

Groth, M. (2019, February 5). How Artificial Intelligence & Natural Language Processing Are Transforming Scholarly Communications. Retrieved from KnowledgeWorks Global Ltd.: https://www.kwglobal.com/ai-and-nlp-for-publishers

Kalim, F. (2019, July 30). Forbes doubles monthly visitors with Bertie, an Al-driven CMS. Retrieved from https://whatsnewinpublishing.com/forbes-doubles-monthly-visitors-with-bertie-an-ai-driven-cms/

King, M. R. (2022, December 26). The Future of AI in Medicine: A Perspective from a Chatbot. Retrieved from https://link.springer.com/ https://link.springer.com/article/10.1007/s10439-022-03121-w

Kulesz, O. (2022, November 30). The impact of Large Language Models on the publishing sectors: Books, academic journals, newspapers. Retrieved from Linnaeus University: https://lnu.diva-portal.org/smash/get/diva2:1737172/FULLTEXT01.pdf

Lawrence, R., & Alam, S. (2023, April 1). Academic Publishers And The Challenges Of Al. Retrieved from Social Science Space: https://www.socialsciencespace.com/2023/01/academic-publishers-and-the-challenges-of-ai/

Lineup. (2023, June 13). The Future of AI in Publishing. Retrieved from Lineup: https://lineup.com/future-of-ai/

Lovrinovic, C., & Volland, H. (2019, October 31). THE FUTURE IMPACT OF ARTIFICIAL INTELLIGENCE ON THE PUBLISHING INDUSTRY. Retrieved from www.buchmesse.de: https://www.buchmesse.de/files/media/pdf/White\_Paper\_Al\_Publishing\_Gould\_Finch\_2019\_EN.pdf

Lovrinovic, C., & Volland, H. (2019, October). THE FUTURE IMPACT OF ARTIFICIAL INTELLIGENCE ON THE PUBLISHING INDUSTRY. Retrieved from GOULD FINCH AND FRANKFURTER BUCHMESSE: https://www.buchmesse.de/files/media/pdf/White\_Paper\_Al\_Publishing\_Gould\_Finch\_2019\_EN.pdf

McIlroy, T. (2023, June 2). Al Is About to Turn Book Publishing Upside-Down. Retrieved from Publishers Weekly: https://www.publishersweekly.com/pw/by-topic/digital/content-and-e-books/article/92471-ai-is-about-to-turn-book-publishing-upside-down.html

Pacholi, A. (2023, February 27). How Artificial Intelligence Is Transforming the E-Publishing Industry: Advantages, Challenges, and Future. Retrieved from Top Software Companies:

https://topsoftwarecompanies.co/artificial-intelligence/how-artificial-intelligence-is-transforming-e-publishing-industry

Ravikumar, V. (2023, January 30). Al/ML is Transforming the Scholarly Publishing Industry. Retrieved from https://www.straive.com/ https://www.straive.com/blogs/ai-ml-is-transforming-the-scholarly-publishing-industry

Sarker, I. (2021, August 18 ). Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. Retrieved from link.springer.com: https://link.springer.com/article/10.1007/s42979-021-00815-1#citeas

Shvets, D. (2021, October 19). How Deep Learning Is Shaping The Future Of Content Production. Retrieved August 1, 2023, from Forbes: https://www.forbes.com/sites/forbesbusinesscouncil/2021/10/19/how-deep-learning-is-shaping-the-future-of-content-production/?sh=7f3b 21102400

Tucker, D. H., Unwin, P. S., & Unwin, G. (2020, 08 20). History of Publishing. Retrieved from Britannica: https://www.britannica.com/topic/publishing

UNESCO, O. A. (2021, November 23). Recommendation on the Ethics of Artificial Intelligence. Retrieved from UNESCO: https://unesdoc.unesco.org/ark:/48223/pf0000381137

UNSILO. (2019, November 1). UNSILO AI in Academic Publishing Survey released. Retrieved from Library Technology Guides: https://librarytechnology.org/pr/24684/unsilo-ai-in-academic-publishing-survey-released

Upshall, M. (2019, April 17). Using AI to solve business problems in scholarly publishing. Retrieved from https://insights.uksg.org/articles/10.1629/uksg.460

#### Links used:

https://static1.squarespace.com/static/5f43eb42a84df22c3d3bbd96/t/6052296913d4c56c95f21248/1615997289488/AI+%26+NLP+for+Publishers+KGL+web.pdf

 $https://www.researchgate.net/publication/319164243\_Natural\_Language\_Processing\_State\_of\_The\_Art\_Current\_Trends\_and\_Challenges$ 

Zhao, Y., & Prabhashini, K. (2019, June 6). Applications of Artificial Intelligence in Digital Publishing Industry in China. Retrieved from ResearchGate:

https://www.researchgate.net/publication/333641083\_Applications\_of\_Artificial\_Intelligence\_in\_Digital\_Publishing\_Industry\_in\_China



Integra is a trusted partner in Business
Process and Technology Services for
many leading organizations worldwide.
With a focus on providing end-to-end
solutions for digital content, learning
services, and content workflows, we help
our customers realize transformational
business value.

For more information, please visit

www.IntegraNXT.com









