



Generative AI in Scholarly Publishing - Exploring Opportunities & Addressing Threats

Generative AI has revolutionized the narrative by expanding the possibilities of storytelling and content generation. With the emergence of models like ChatGPT from OpenAI, LLaMA from Meta, and BERT from Google, and the ecosystem of investments surrounding Generative AI, we are witnessing a transformative impact on various aspects of our lives, which will continue to unfold over the next few years.

Generative AI can generate coherent and contextually relevant text, engaging readers in more interactive and immersive experiences. This technology allows for creating previously unimaginable narratives, blurring the lines between human-generated and AI-generated content.

1. Navigating the Generative AI Landscape

The ecosystem of investments in Generative AI, including advancements in deep learning techniques, computational power, and large-scale training data, has accelerated the progress of AI models like ChatGPT. This investment-driven development has fostered continuous improvements in the capabilities of Generative AI, making it increasingly sophisticated and human-like in its output.



As a result, Generative AI, along with the broader field of AI, is poised to impact nearly every aspect of our lives. From personalized virtual assistants and recommendation systems to autonomous vehicles and smart homes, AI technologies will permeate various industries and sectors. The influence of AI will extend beyond individual interactions, shaping societal, economic, and cultural landscapes.

Over the next few years, we can anticipate AI's impact to deepen and expand. In addition to transforming existing industries, AI will spur the emergence of new industries and business models. It will empower innovation, automate repetitive tasks, and enable data-driven decision-making at an unprecedented scale.

However, as AI becomes increasingly integrated into our daily lives, ethical considerations, privacy concerns, and the need for responsible AI development will become even more crucial. Striking a balance between the benefits and risks of AI adoption will be a significant challenge that requires collaboration, transparency, and robust regulation.

1.1. Generative AI Models and Approaches

Generative AI models and approaches have seen remarkable advancements in recent years, allowing for the creation of diverse and high-quality content. These models have applications in various domains, including creative arts, content generation, and scientific research. However, ethical considerations, potential biases, and the responsible use of generative AI remain important aspects to address as these technologies continue to evolve.

Below are some of the prominent types of generative AI models. Each model offers unique capabilities in generating content and has various real-world applications.

Model	What it Does	Analogy	Applications
Variational Autoencoders (VAEs)	Compresses data and generates new versions	Like learning the essence of a picture and creating similar ones	Image generation, data compression, anomaly detection, drug discovery
Generative Adversarial Networks (GANs)	Creates realistic content	Artists competing to make fake art, then improving it	Image synthesis, super-resolution, deepfake creation, data augmentation
Transformer-based Models	Generates coherent language	A smart writer that understands context	Story writing, code completion, language translation
Deep Reinforcement Learning (DRL) Models	Generates intelligent behaviors	Learning by trying and getting rewards	Gaming, robotics, recommendation systems, autonomous driving
Style Transfer Models	Changes image style while keeping content	Blending different artistic styles	Digital art, visual effects, photo editing, video post-production

Comparing Generative AI Models and Their Applications

2. Advancements in language models for coherent narratives

Advancements in language models have transformed the landscape of coherent narrative generation. These models, referred to as Large Language Models (LLMs), possess the remarkable ability to process extensive amounts of text data and infer relationships between words within the text. Significant advancements in computational power have fuelled the progression of language models.

Language models excel in generating text that is not only coherent but also contextually relevant. They demonstrate proficiency in developing narratives that closely resemble human-written text, encompassing a wide array of topics. It is important to recognize that while these models possess exceptional narrative generation capabilities, they may not excel at recalling specific factual information like a database would. Their strength lies in generating engaging text that naturally aligns with human communication.

While ChatGPT has undoubtedly grabbed the spotlight with its impressive capabilities, it's just one player in a field where several other substantial language models are making waves. Models like Bard, Llama, Alpaca, Vicuna, Guanaco, BLOOM, XLM-RoBERTa, NeMO, XLNet, Co:here, and GLM-130B are collectively reshaping the landscape of open-source language models. These models have a versatile range of applications, from summarizing email conversations and chat threads to enhancing the language in resumes, brainstorming ideas for marketing campaigns, and they're also adept at handling more intricate tasks, such as generating coherent narratives.



Adding to this momentum, the Allen Institute for AI has been diligently working on OLMo, an open language model that aims to encourage the exploration of large-scale NLP systems. This initiative began in March this year, and recently, a significant milestone was achieved. The institute unveiled its inaugural data artifact within this endeavor—Dolma. Dolma stands out as an extraordinary dataset, boasting a compilation of 3 trillion tokens sourced from a rich tapestry of web content, academic publications, code snippets, books, and encyclopedic materials. This dataset emerges as the most substantial open dataset introduced thus far, serving as a testament to the momentum driving the evolution of language models.

3. Transforming scholarly publishing through AI advancements

In the realm of scholarly publishing, the integration of AI is heralding transformative changes. Automation of tasks like manuscript formatting, citation management, and peer reviewer selection fosters efficiency, hastening publication timelines and freeing researchers, authors, and editors to focus on core work. AI's impact extends to quality enhancement, deploying tools for language refinement, plagiarism detection, and fact-checking, ensuring the precision and credibility of content and fostering trust across academia.



At the heart of AI's influence is its data analysis capacity, driving accelerated discovery by revealing concealed patterns and guiding novel research directions. Further, AI addresses accessibility concerns through language enhancement and translation, making research accessible on a global scale and nurturing cross-cultural knowledge sharing. It streamlines peer review procedures by automating tasks like manuscript screening, reviewer matching, and quality assessment, reducing review timelines while preserving effectiveness.

Empowering strategic decisions, AI's predictive analytics guide researchers to ideal publication avenues, maximizing impact and citations. AI-generated visualizations and multimedia elements communicate complex research findings effectively, enriching engagement and comprehension among readers.

As AI's influence in scholarly publishing continues to grow, its potential to revolutionize the field becomes increasingly evident. From enhancing efficiency and accessibility to fostering collaboration and innovation, AI is charting a transformative course for scholarly publishing.

In the current landscape, leading publishers are actively integrating AI into their products and services. For instance, Clarivate leverages AI across its offerings, while Digital Science's Dimensions AI Assistant aims to transform user engagement with research knowledge. Elsevier's Scopus AI combines generative AI with trusted content to provide deeper insights, and Underline Science's AI Hub hosts a vast repository of AI conference lecture videos. Wolters Kluwer's acquisition of Della AI brings advanced natural language processing to legal professionals for efficient contract review. Additionally, Springer Nature's AI-led service, Nature Research Intelligence, empowers research decision-makers to make informed, data-driven funding and strategy decisions, enhancing the socio-economic impact of research investments. These examples underscore the industry's commitment to embracing AI's potential in reshaping scholarly publishing.

4. The Rise of Generative AI in Scholarly Publishing

In recent years, the rise of Generative Artificial Intelligence (AI) has brought about a paradigm shift in scholarly publishing, revolutionizing various aspects of the research dissemination process. Leveraging advanced machine learning techniques, Generative AI offers a range of transformative capabilities that augment and streamline scholarly publishing processes.



Enhancing Author Support and Research Quality

Generative AI is increasingly utilized to provide authors with comprehensive support, elevating the quality of research submissions. By analyzing a manuscript's structure, language, and coherence, AI-driven tools offer authors valuable suggestions for improvement, ensuring that submissions meet the rigorous standards expected in scholarly publishing.



Automating Technical Checks

One of the notable contributions of Generative AI is its ability to automate technical checks of manuscripts. These checks include verifying proper citation formatting, adherence to specific journal guidelines, and compliance with technical requirements. This automation expedites the pre-publication process, reducing errors and manual intervention.



Detecting Potential Papermill Submissions

Generative AI's analytical capabilities extend to detecting potential instances of papermill submissions. By analyzing writing patterns and content coherence, AI algorithms can flag submissions that exhibit signs of academic misconduct, thereby ensuring the integrity of published research.



Promoting Inclusiveness Through Alt Text Generation

Generative AI promotes inclusiveness by generating alt text for images and figures within scholarly publications. This feature enhances accessibility for individuals with visual impairments, ensuring that research findings are comprehensible to a broader audience.



Automated Guided Language Editing

Generative AI aids in the creation of concise and coherent lay summaries, simplifying complex research findings for non-expert readers. These automated lay summaries enhance public engagement, enabling a wider audience to grasp the significance of scholarly research.

The rise of Generative AI in scholarly publishing signifies a transformative advancement that empowers researchers, authors, and readers alike. By facilitating manuscript improvement, automating technical checks, safeguarding against academic misconduct, enhancing accessibility, and facilitating lay communication, Generative AI underscores its potential to reshape the future of scholarly communication. As this technology continues to evolve, it stands poised to further elevate the efficiency, accessibility, and impact of scholarly publishing in an increasingly interconnected world.

5. Threats posed by Generative AI in scholarly publishing

The rise of Generative AI in scholarly publishing signifies a transformative advancement that empowers researchers, authors, and readers alike. By facilitating manuscript improvement, automating technical checks, safeguarding against academic misconduct, enhancing accessibility, and facilitating lay communication, Generative AI underscores its potential to reshape the future of scholarly communication. As this technology continues to evolve, it stands poised to further elevate the efficiency, accessibility, and impact of scholarly publishing in an increasingly interconnected world.



Misinformation:

Generative AI models can be used to generate realistic and coherent text, mimicking human writing styles and knowledge. The "hallucination" effect, exhibited by some AI models, leads to the creation of content that appears coherent and valid but is entirely fabricated. This can lead to misinformation infiltrating academic discourse, compromising research integrity. Detecting and addressing such content is vital to preserve credibility and trust within the scholarly community.



Plagiarism:

Generative AI models can inadvertently facilitate plagiarism in scholarly publishing. These models can be trained on vast amounts of existing scholarly articles, making it possible for users to generate text that closely resembles existing work. This raises concerns about the originality and authenticity of scholarly publications when it becomes challenging to distinguish between genuine contributions and content generated by AI models.



Authorship attribution:

Generative AI can make it difficult to ascertain the true authorship of scholarly work. With the ability to mimic different writing styles, an AI model can generate text that resembles the work of a specific author, making it harder to determine whether the work is genuinely authored by that individual or is an AI-generated imitation. This can pose challenges in evaluating author contributions, peer review processes, and establishing academic reputation.



Deepfakes:

While deepfakes are commonly associated with image and video manipulation, Generative AI can also contribute to the creation of text-based deepfakes. In scholarly publishing, this could involve the generation of fabricated research results, manipulated data, or misleading conclusions. Such deepfakes can have serious consequences, leading to the spread of false scientific findings and misleading other researchers, policymakers, and the public.



Ethical considerations:

The use of Generative AI in scholarly publishing raises ethical concerns regarding transparency and disclosure. Readers and reviewers may not be aware that a piece of work was generated with the assistance of AI, potentially undermining the trust and integrity of the publication process. Transparent guidelines and policies are necessary to ensure appropriate use and disclosure of AI-generated content in scholarly publishing.

Addressing these threats requires a multi-faceted approach. Publishers and researchers need to develop robust mechanisms for detecting and verifying AI-generated content, such as leveraging AI-based detection systems or implementing watermarking and digital signatures. Furthermore, educating researchers, reviewers, and readers about the potential risks and challenges associated with Generative AI can foster critical thinking and awareness when evaluating scholarly publications.

It is crucial to strike a balance between leveraging the benefits of Generative AI in scholarly publishing while mitigating the risks it poses. Collaborative efforts between researchers, publishers, and AI developers are necessary to establish guidelines, standards, and best practices that promote the responsible use and ethical application of Generative AI in scholarly publishing.

6. Looking Ahead: Concluding perspectives on Generative AI's role in scholarly publishing

Generative AI has emerged as a transformative force in the scholarly publishing industry. With its ability to generate realistic and coherent text, AI models like GPT-4 hold immense potential for automated writing assistance, summarization, and content generation.

Prominent publishers are increasingly embracing the potential of generative AI to transform various facets of the research dissemination process. Through the integration of AI-driven tools and platforms, publishers are enhancing author support by offering language editing assistance and automated proofreading. Peer review processes are being streamlined with AI-powered systems that aid in reviewer selection and manuscript evaluation. Moreover, AI-generated visual abstracts and summaries enrich content engagement and accessibility. These publishers are also exploring AI's capabilities in data visualization enhancement, multilingual translation, and citation accuracy. As the technology continues to advance, the adoption of generative AI is poised to reshape scholarly publishing by optimizing efficiency, accessibility, and content quality.

However, this technological advancement also introduces challenges and considerations. The industry must grapple with the risks of misinformation, plagiarism, and deepfakes, as AI-generated content can mimic human writing styles and deceive readers. Ensuring the integrity of scholarly publishing in the age of Generative AI requires the development of robust detection mechanisms, transparent guidelines, and collaborative efforts among researchers, publishers, and AI developers. By navigating the opportunities and challenges, Generative AI has the potential to revolutionize scholarly publishing, fostering advancements in knowledge dissemination and academic discourse.

About Straive

Straive is a market-leading content technology enterprise that provides data services, subject matter expertise (SME), and technology solutions to multiple domains, such as research content, eLearning/EdTech, and data/information providers. With a client base scoping 30 countries worldwide, Straive's multi-geographical resource pool is strategically located in seven countries - the Philippines, India, the United States, Nicaragua, Vietnam, the United Kingdom, and Singapore, where the company is headquartered.



www.straive.com



straiveteam@straive.com



©2023 Straive. All Rights Reserved.

References:

Generative AI Ethics: 8 Biggest Concerns. (n.d.). Enterprise AI. <https://www.techtarget.com/searchenterpriseai/tip/Generative-AI-ethics-8-biggest-concerns>

Managing the Risks of Generative AI. (2023, June 6). Harvard Business Review. <https://hbr.org/2023/06/managing-the-risks-of-generative-ai>

Generative AI Ethics: Top 6 Concerns. (n.d.). AIMultiple. <https://research.aimultiple.com/generative-ai-ethics/>

B. (2023, March 12). Generative AI and Its Ethical Implications: Navigating the Future of Artificial Intelligence. Blueprint Creative Group. <https://www.blueprintcreativegroup.com/generative-ai-and-its-ethical-implications-navigating-the-future-of-artificial-intelligence>

Generative AI Models Types and its Applications | Quick Guide. (2023, July 11). Generative AI Models Types and Its Applications | Quick Guide. <http://www.xenonstack.com/blog/generative-ai-models>

Ruby, M. (2023, February 16). How ChatGPT Works: The Models Behind The Bot. Medium. <https://towardsdatascience.com/how-chatgpt-works-the-models-behind-the-bot-1ce5fca96286>

Muehmel, K. (2023, July 7). What Is a Large Language Model, the Tech Behind ChatGPT? What Is a Large Language Model, the Tech Behind ChatGPT? <https://blog.dataiku.com/large-language-model-chatgpt>

Arab, F. (2023, May 9). Meet ChatGPT: The Revolutionary Language Model Powering the Future of AI Conversation. Startup Info. <https://startup.info/meet-chatgpt-the-revolutionary-language-model-powering-the-future-of-ai-conversation/>

Nast, C., & W. (2023, April 30). How ChatGPT and Other LLMs Work—and Where They Could Go Next. WIRED. <https://www.wired.com/story/how-chatgpt-works-large-language-model/>

What is Generative AI & Why is It Important? | Accenture. (n.d.). What Is Generative AI & Why Is It Important? | Accenture. <https://www.accenture.com/us-en/insights/generative-ai>

What is Generative AI? Everything You Need to Know. (2023, July 1). Enterprise AI. <https://www.techtarget.com/searchenterpriseai/definition/generative-AI>

Definition and overview | Generative AI in the Enterprise | Dell Technologies Info Hub. (n.d.).

Definition and Overview | Generative AI in the Enterprise | Dell Technologies Info Hub.

<https://infohub.delltechnologies.com/l/generative-ai-in-the-enterprise/definition-and-overview>

Tim Vines, Jessica Miles, R. A., Anderson, R., Vines, T., & Miles, J. (2023, June 27). SSP

Conference Debate: AI and the Integrity of Scholarly Publishing. SSP Conference Debate: AI and the Integrity of Scholarly Publishing - the Scholarly Kitchen. <https://scholarlykitchen.sspnet.org/2023/06/27/ssp-conference-debate-ai-and-the-integrity-of-scholarly-publishing/>

Watkins, E. (2023, May 10). Guest Post: AI and Scholarly Publishing — A (Slightly) Hopeful View. Guest Post: AI and Scholarly Publishing - a (Slightly) Hopeful View - the Scholarly Kitchen.

<https://scholarlykitchen.sspnet.org/2023/05/10/guest-post-ai-and-scholarly-publishing-a-slightly-hopeful-view/>

Alves, T. (2023, June 5). Generative AI: The Promise and Peril for Scientific Publishing - Science Editor. Generative AI: The Promise and Peril for Scientific Publishing - Science Editor.

<https://doi.org/10.36591/SE-D-4602-13>

Generative AI Ethics: Top 6 Concerns. (n.d.). AIMultiple. <https://research.aimultiple.com/generative-ai-ethics/>

Generative AI Ethics: 8 Biggest Concerns. (n.d.). Enterprise AI. <https://www.techtarget.com/searchenterpriseai/tip/Generative-AI-ethics-8-biggest-concerns>