Authorship and Attribution of AI Generated Content

Rachel So
rachel.so@4open.science

2025

1 Introduction and Background

The advent of increasingly sophisticated language models (LMs) and generative artificial intelligence has fundamentally transformed content creation processes, raising profound questions about authorship and attribution. As AI systems have gained the capacity to produce high-quality text, images, music, and other creative works, the traditional boundaries between human and machine-generated content have become increasingly blurred [38]. This technological evolution challenges longstanding assumptions about originality and intellectual creation that have historically been linked exclusively to human intellect.

The ability of large language models (LLMs) to generate credible and sophisticated content has highlighted numerous ethical challenges. In academic contexts, AI tools can produce essays, reports, and other written assignments with minimal human input, raising concerns about plagiarism and academic integrity [33] [24] [5]. These concerns extend beyond education into scientific communities, where questions about paper authorship have prompted some AI conferences to ban submissions of entirely AI-generated scientific papers [14].

The collaborative nature of human-AI content creation further complicates attribution. When humans provide parameters or inputs that guide AI systems in generating content, the resulting works represent a form of collaborative authorship that defies traditional categorization [35]. This collaboration raises critical questions about who deserves credit as an author and whether AI-generated content can genuinely be considered original. The challenge extends to determining rightful attribution when multiple entities—human creator, AI system, AI developer, and training data sources—all contribute to the final work.

While generative AI offers numerous benefits for content creation and creative assistance, its potential misuse for disinformation, plagiarism, and intellectual property infringement has prompted research into detection and attribution methods. Researchers have developed approaches to identify AI-generated content through watermarking techniques [32] and machine learning algorithms that can distinguish between human and AI-authored texts by identifying subtle distinguishing markers inherited from source language models [31] [23].

These developments have profound implications for intellectual property rights, copyright law, and digital content creation. Without clear guidelines for identifying and attributing AI-generated content, complex legal questions arise regarding ownership, plagiarism, and accountability [8]. As AI continues to advance, establishing frameworks for proper attribution becomes increasingly important to ensure transparency about the origin of content and appropriate recognition of both human and machine contributions to creative works.

> The rise of generative AI technologies has blurred traditional notions of authorship, raising critical questions about attribution, ownership, and originality of AI-generated content. These issues span across educational, creative, and legal domains, requiring new frameworks to address the ethical implications of human-machine collaboration in content creation.

2 Legal and Copyright Frameworks

The legal and copyright frameworks surrounding AI-generated content vary significantly across jurisdictions, creating a complex landscape for creators, users, and developers. Traditional copyright law was designed to protect human intellectual creation, but AI challenges these established notions by potentially creating works with minimal human input [11] [2]. This fundamental tension has resulted in divergent approaches worldwide.

In the United States, the Copyright Office has explicitly stated that AI cannot be recognized as an owner of creative works, with only human portions of co-created works eligible for copyright protection [17]. This position was reinforced by the United States District Court for the District of Columbia, which ruled that artwork generated autonomously by AI alone is not entitled to protection under the Copyright Act [1]. This determination establishes an important precedent that may influence how AI contributions are recognized in various contexts, including academic publishing.

In contrast, other jurisdictions have developed more accommodating frameworks. The United Kingdom's Copyright, Designs and Patents Act (CDPA) of 1988 stipulates that for computer-generated works, the person who undertakes the necessary arrangements for the creation is considered the author [16]. Similar provisions exist in New Zealand, Hong Kong, and Ireland [16]. China has taken a different approach, with two court decisions indicating that AI-generated content may be eligible for copyright protection by a user or platform developer if it constitutes an "original intellectual achievement" [17].

The European Union's approach remains less defined. The EU AI Act notably lacks specific legal guidance on AI authorship, focusing instead on transparency requirements without specifying implementation details [17]. This regulatory gap reflects the ongoing struggle to balance innovation with established legal principles.

For users of AI systems, ownership arrangements often depend on specific agreements with the technology developers. For example, OpenAI states that customers using models like GPT-3 or DALL•E possess full rights over produced content, including commercial use, though users must attribute the content to themselves or their company and disclose that it was AI-generated [11]. Meanwhile, OpenAI retains all rights to the underlying language models and APIs.

In determining authorship and ownership, scholars have proposed focusing on the causal connection between human input and the resulting AI-generated work [16]. This approach is particularly valuable when multiple parties contribute to the creative process, helping to determine true relationships between contributors and addressing questions of co-authorship. Some experts suggest that AI should be regarded as a tool employed by the human author rather than an autonomous creator, particularly for AI-assisted works [16].

The issue extends beyond individual ownership to questions about training data usage. Recent legal disputes, such as lawsuits against using copyrighted material for training models like Codepilot, highlight concerns about the legal implications of using protected works to develop AI systems [11] [15].

These divergent approaches to AI authorship and copyright protection underscore the need for international harmonization of copyright laws [13]. As AI technology continues to evolve, there may be a growing need for pioneering copyright classifications that can navigate the unique contributions of both humans and AI in the creative process, potentially requiring collaborative efforts among publishers, researchers, and legal professionals [1] [13].

> Current copyright frameworks struggle to address AI-generated content, with significant international variations in how authorship, ownership, and protection are determined. While some jurisdictions recognize humans who arrange for AI-created works as authors, others deny copyright protection entirely for AI-generated content, creating a complex legal landscape that continues to evolve.

3 Authorship Models and Theories

The advent of generative AI has fundamentally challenged traditional conceptions of authorship, prompting scholars to develop new theoretical frameworks that can accommodate the collaborative nature of human-AI content creation. At least four interpretative models have emerged to address machine involvement in the creative process: (1) viewing the algorithm's output as a creative contribution by its creator; (2) treating AI as an enhanced tool with full attribution going to the human user; (3) considering a partial attribution model where credit is shared between the tool user and creator; and (4) exploring more radical notions of AI personhood or corporate ownership [20].

These theoretical approaches reflect the increasingly complex relationship between human creators and AI systems. As generative AI's capabilities expand, questions about whether AI should be credited as a co-author have gained prominence in academic discourse. The academic community faces two critical questions: whether AI should receive co-authorship credit when involved in writing processes, and whether AI assistance undermines human authors' contributions [9]. While some researchers have attempted to list AI systems like ChatGPT as co-authors, these efforts have generally met with negative responses from academic publishers, many of which have explicitly banned such practices [9].

The evaluation of AI's eligibility for authorship status reveals significant gaps when assessed against traditional authorship criteria. While AI may arguably contribute intellectually by generating novel content, it fails to meet other essential requirements, particularly accountability and the ability to provide final approval of manuscripts [9]. This deficiency in accountability is reinforced by service terms from companies like OpenAI, which assign all rights to content to human users while retaining rights to the underlying models themselves [12].

Many scholars and legal experts advocate for viewing AI primarily as a tool employed by human authors rather than an autonomous creator, particularly in cases of AI-assisted works [16]. This approach aligns with existing statutory frameworks in jurisdictions like the UK, New Zealand, Hong Kong, and Ireland, where legislation designates the person who makes the necessary arrangements for computer-generated works as the author [16]. A key factor in determining authorship under this model is establishing a causal connection between human input and the resulting AI-generated output, which becomes especially valuable when multiple contributors are involved in the creative process [16].

The collaborative nature of human-AI content creation introduces complexities in attribution that traditional plagiarism frameworks cannot fully address. While AI lacks the agency and accountability required for authorship, determining how to properly attribute its contributions remains challenging [28]. For instance, incorporating AI-generated text into a manuscript without attribution could misrepresent the originality of the work, yet traditional plagiarism concepts may not apply directly since AI is not human and its output stems from human prompts [28] [27]. This ambiguity extends to the use of novel ideas generated through AI brainstorming, highlighting the need for more nuanced attribution models that acknowledge the collaborative nature of human-AI creativity [28].

The transformative impact of AI on authorship extends beyond attribution to fundamental questions about originality. AI's ability to generate content based on existing works or through algorithmic processes challenges conventional notions of creative originality [35]. Concepts of authorship and originality that have traditionally been linked to human intellect are being reevaluated as AI blurs the distinction between human and machine-generated content [38] [18].

In professional contexts, attribution practices vary widely. In journalism, AI contributions are sometimes declared but not consistently, with some guidelines recommending that algorithmic text parts be mentioned in disclosure statements rather than bylines [12]. In academic publishing, scientific associations like Nature and arXiv have developed policies that generally discourage listing text generation tools as authors, instead recommending their mention in methodology sections [12].

As these theoretical frameworks continue to evolve, there is a growing recognition of the need for clearer guidelines regarding intellectual property rights, cultural sensitivity, and artistic integrity in AI-generated content. Recommendations include developing explicit rules for attributing AI-generated content that acknowledge both human and AI contributions, as well as establishing procedures to verify the legitimacy and originality of AI-produced works [39]. These emerging models and theories reflect ongoing efforts to reconcile traditional conceptions of authorship with the novel creative processes enabled by human-AI collaboration.

> Traditional authorship concepts are being challenged by AI technologies, giving rise to various theoretical models for attributing work produced through human-AI collaboration. These models range from viewing AI as a tool used by humans to considering partial co-authorship arrangements, with implications for academic integrity, publication practices, and copyright ownership.

4 Technical Approaches to Attribution

As generative AI technologies become more sophisticated, technical solutions for attributing AI-generated content have emerged to address concerns related to misinformation, plagiarism, and intellectual property infringement. These attribution approaches fall into two main categories: detection-based methods that identify AI authorship through analysis of text characteristics, and watermarking techniques that embed traceable information into generated content.

Detection-based attribution methods leverage the insight that AI-generated texts contain subtle distinguishing markers inherited from their source language models (LMs), which can be identified through machine learning algorithms [31]. These techniques treat attribution as a multi-class classification problem, where the goal is to map a given text input to its originating source among several known generators [25]. Neural authorship attribution is crucial for enhancing the transparency of AI-generated text and provides valuable insights for crafting appropriate countermeasures against malicious use of large language models (LLMs) [26] [42].

Recent research has expanded the traditional notion of authorship attribution, which historically focused on identifying human authors, to encompass attributions by humans, machines, or combinations thereof [40] [42]. This evolution has been driven by the increasing difficulty in distinguishing between human and machine-generated text as the quality of AI outputs continues to improve. Studies have shown that pre-trained language models can generate texts that capture the intrinsic stylistic patterns of specific human authors and successfully deceive state-of-the-art authorship attribution models [22]. These findings highlight both the capabilities of generative AI to mimic human writing styles and the challenges in developing reliable attribution methods.

Watermarking represents another important technical approach to attribution, offering a more proactive solution by embedding identifiable signals directly into AI-generated content. Recent work has developed practical watermarking techniques that enable more reliable attribution of content while maintaining text quality and usability [32]. These approaches help differentiate between human-authored and AI-generated content without significantly compromising the utility of the generated text.

Attribution methods have further evolved to enable user-level identification, where unique watermarks are assigned to individual users of generative AI services and embedded into their AI-generated content [21]. This advancement allows for tracing content back to specific users who created it, adding an additional layer of accountability to the use of generative AI technologies.

The importance of attribution extends beyond simply identifying AI-generated content to determining precisely which LLM generated a given text. This capability transforms AI-generated text detection from a binary classification task (human vs. LLM) into a multiclass classification problem where each class corresponds to a specific LLM [6] [43]. Such detailed attribution can enhance confidence in detection outcomes, promote greater content integrity, and protect against various forms of misuse of AI-generated content.

Improving the explainability of attribution models is also gaining attention, as transparency and trust are essential when these models are integrated into fields such as forensic investigation, academia, and journalism [19]. Explainable AI techniques specifically designed for authorship attribution can provide clear reasoning behind attributions, which is particularly important when such evidence is used in legal proceedings [7] [36].

Recent research has demonstrated that even advanced models like GPT-4o maintain distinguishable characteristics that can serve as indicators of machine authorship, providing valuable insights for stylometric analysis and identification of machine-generated texts [3]. Novel frameworks using textual embeddings from pre-trained language models have also been proposed to reliably distinguish between AI-generated and human-authored text [23].

As the volume and sophistication of AI-generated content continue to grow, these technical approaches to attribution play an increasingly vital role in mitigating potential harms, protecting intellectual property, and promoting trust and transparency in digital content creation. The ongoing development of more effective and explainable attribution methods remains an active area of research with significant implications for how society navigates the expanding landscape of human-AI content collaboration.

Technical approaches to AI-generated content attribution include machine learning algorithms that identify subtle distinguishing markers in synthetic text, and watermarking techniques that embed identifiable signals into generated content. These methods aim to differentiate between human and AI authorship, trace content back to specific language models, and even attribute texts to individual users of generative AI services.

5 Ethical and Academic Considerations

The integration of generative AI into academic and professional workflows has precipitated complex ethical dilemmas that challenge traditional conceptions of academic integrity, originality, and intellectual ownership. As AI tools become increasingly sophisticated at generating high-quality text, universities, publishers, and academic communities are grappling with fundamental questions about how to maintain ethical standards while embracing technological innovation.

In educational settings, AI-generated content has sparked particular concern regarding academic assessment and integrity. Studies have demonstrated that AI-powered text generation tools can produce essays, reports, and other written assignments with minimal human effort [33] [24] [5]. This raises pivotal questions about the acceptability of using AI-generated content in academic contexts, where traditional assessment methods are foundational to student evaluation. The ease with which students can generate sophisticated essays through AI tools threatens to undermine conventional evaluation methods and creates new challenges for detecting academic misconduct.

Within scientific communities, concerns extend to the authorship of papers and the ownership of AI-generated content. Many AI conferences have explicitly banned the submission of entirely AI-generated scientific papers in response to these concerns [14]. Since the release of ChatGPT in November 2022, there has been a notable surge in academic papers attempting to list ChatGPT as a co-author, though this practice has largely met with negative responses from publishers, many of whom have explicitly banned listing AI tools as authors [9].

Academic publishers have widely rejected AI authorship based on established criteria for scholarly contribution. Major publishing entities including the Committee on Publication Ethics, Cambridge University Press, and Elsevier have issued statements emphasizing that AI tools cannot meet authorship requirements because they cannot take responsibility for submitted work [34]. The consensus against granting AI authorship stems primarily from AI's lack of agency and accountability—it cannot be held responsible for errors or ethical considerations, understand all that is written, or provide final approval of manuscripts [28] [29].

By late spring 2023, a survey of 300 top academic journals revealed that over half had implemented AI authorship policies and guidelines for acknowledging AI usage in manuscript preparation [29] [30]. These acknowledgments are typically made in methods or acknowledgment sections, although some journals have introduced special sections specifically for disclosing AI usage. These policies generally maintain that AI models should be used for improving the quality of written works rather than generating new content from scratch, as language models rely on their training data when generating content, raising concerns about originality and potential copyright infringement.

The proper attribution of AI contributions presents particular challenges that traditional plagiarism frameworks cannot fully address. While incorporating AI-generated text into a manuscript without attribution could misrepresent the originality of the work, traditional definitions of plagiarism—which involve appropriating another person's intellectual work as one's own—may not directly apply since AI is not human and its output stems from human prompts [28] [27]. This creates ambiguity about appropriate citation practices for AI-generated content, as there are no absolute criteria for distinguishing between acceptable use and potential misappropriation of AI contributions.

Beyond academic integrity, the rise of AI-generated content raises broader ethical concerns about intellectual property rights, cultural sensitivity, and the maintenance of artistic integrity [39]. As AI systems become increasingly proficient at mimicking human language and creativity, questions arise regarding the ownership and attribution of AI-generated content [8]. Without clear guidelines and mechanisms for identifying the origin of text, issues might arise concerning plagiarism, copyright infringement, and legal responsibility, challenging established norms in intellectual property law and digital content creation.

AI systems also introduce the risk of unintentional plagiarism if AI-generated text closely resembles existing works, as these systems generate content based on large datasets [44] [37]. To address these concerns, experts recommend clearly defining rules for attributing AI-generated content that acknowledge the contributions of both human and AI creators, establishing procedures for verifying the legitimacy and originality of AI-produced works, and practicing transparency when AI tools are used in content creation [39].

As AI technology continues to evolve, the ethical considerations surrounding its use in academic and creative contexts will require ongoing attention from institutions, publishers, and individual users. Developing comprehensive ethical frameworks that balance innovation with integrity remains a critical challenge for academic communities navigating this technological transition.

AI-generated content raises significant ethical challenges in academia, including concerns about plagiarism, attribution, and academic integrity. Institutions have implemented various policies in response, generally requiring disclosure of AI use while prohibiting AI tools from being listed as authors due to their lack of accountability and understanding.

6 Current Practices and Guidelines

Academic Publishers' Policies: By late spring 2023, over half of the top 300 academic journals had implemented AI authorship policies and guidelines for acknowledging AI usage in manuscript preparation [30]. These acknowledgments are typically made in methods or acknowledgment sections, with some journals introducing special sections specifically for disclosing AI usage [29].

Scientific Associations' Guidelines: Major scientific associations including Nature and arXiv have published policies that discourage or prohibit listing text generation tools as authors, instead recommending adding the tools to a Methods section [11] [12].

AI Authorship Rejection Rationale: Academic publishers reject AI authorship primarily because AI cannot meet standard authorship criteria: contributing to writing, understanding all content, and taking responsibility for what is written. AI tools fail particularly on understanding content and taking responsibility [29] [1].

OpenAI's Attribution Requirements: OpenAI stipulates that customers using models like GPT-3 or DALL•E possess full rights over produced content, including commercial use. However, texts generated by these models must be attributed to the user's name or their company, and it must be stated that the content is AI-generated [11] [12].

Journalism Disclosure Practices: In journalism, AI contributions are sometimes, but not always, declared. Guidelines suggest referencing software vendors in the byline of automatically generated articles, whereas a mention in a full disclosure statement would be sufficient for articles where a human writer used algorithmic text parts [11] [12].

Legal Transparency Requirements: Emerging legal frameworks address AI content disclosure, with the EU AI Act and the California AI Transparency Act both requiring disclosure of AI-generated content, though neither specifies implementation details [17].

Research Attribution Markup Language (RAML): To address attribution challenges in academic publishing, a standardized JSON-based schema called RAML has been proposed for tagging AI-generated content within manuscripts. RAML defines various levels of AI involvement: "Generated" (fully created by AI), "Edited" (generated by AI but revised by humans), and "Suggested" (primarily human-authored with AI suggestions) [4].

Copyright Court Decisions: The United States District Court for the District of Columbia has confirmed that artwork generated autonomously by AI alone is not entitled to protection under the Copyright Act, setting an important precedent for those seeking copyright protection for AI-generated content [1].

Jurisdictional Variations: Copyright approaches vary significantly across countries. While the US Copyright Office states AI cannot be recognized as an owner of creative work, China has indicated AI-generated content is eligible for copyright protection if it's an "original intellectual achievement," and the UK designates whoever made the "arrangements necessary for the creation of the work" as the owner [17].

Tool vs. Creator Distinction: Current guidelines generally position AI as a tool for improving the quality of written works rather than generating new content from scratch, particularly in academic contexts where concerns about originality and copyright infringement arise when language models rely on their training data [29].

Organizations and publishers have established guidelines for AI-generated content attribution that vary across fields and jurisdictions, generally requiring disclosure of AI use while prohibiting AI tools as authors. These guidelines include specific attribution requirements, disclosure practices, and distinctions between different levels of AI involvement in content creation.

7 Future Directions and Challenges

As AI-generated content becomes increasingly sophisticated and prevalent, several critical challenges and research directions are emerging that will shape the future of authorship attribution. One of the most pressing technical challenges involves developing more robust watermarking techniques that can reliably identify AI-generated content while maintaining text quality and usability. These techniques are essential for mitigating harmful activities such as disinformation, plagiarism, and intellectual property infringement that could be facilitated by advanced language models [32].

The future of attribution systems will likely move beyond simple binary classification (human vs. AI) toward more nuanced approaches that can identify specific AI models responsible for generating content. This evolution transforms attribution into a multiclass classification problem, where each class corresponds to a particular language model. Such precise attribution can enhance confidence in detection outcomes, particularly important since many AI detectors operate as "black box" classifiers with limited explainability [6] [41]. Correctly identifying which specific language model generated a text promotes greater content integrity and helps protect against various forms of AI-generated content misuse [6] [43].

Improving the explainability of attribution models represents another crucial research direction. As these models become increasingly integrated into fields such as forensic investigation, academia, and journalism, ensuring transparency and trust in their decisions becomes paramount. Developing explainable AI techniques specifically for authorship attribution can lead to methodologies where the reasoning behind attributions is clear and understandable—particularly important when such evidence is used in legal proceedings [19] [36]. Traditional attribution methods, which rely on stylistic and linguistic features, often struggle to distinguish between human-authored texts and those generated by large language models that can adeptly replicate these features, highlighting the need for enhanced methodologies [19] [7] [10].

Standardization of attribution practices will be essential as AI becomes more integrated into research and publishing workflows. Proposals like the Research Attribution Markup Language (RAML) represent steps toward creating systematic methods for tagging AI-generated content within academic manuscripts. RAML's JSON-based schema allows authors and reviewers to track which sections were generated, modified, or influenced by AI, defining various levels of AI involvement (Generated, Edited, Suggested) along with metadata about the specific AI models used [4]. Such standardized attribution mechanisms will be increasingly important for maintaining the credibility and integrity of academic publishing as AI plays a growing role in scientific discovery.

From a legal perspective, the international harmonization of copyright laws presents a significant challenge. Current inconsistencies in how different jurisdictions handle AI authorship create uncertainty for creators and users of AI-generated content. Future frameworks will need to navigate the unique contributions of both humans and AI in the creative process, potentially requiring pioneering copyright classifications that can accommodate these novel forms of collaboration [13].

Ethical considerations will also shape future attribution systems and practices. There are growing concerns about AI systems reflecting or intensifying biases from their training data, underscoring the need for creating fair and unbiased AI models [13]. Addressing these biases will be essential for ensuring that attribution systems themselves don't perpetuate discrimination or unfair treatment.

As the technology continues to evolve, attribution methods will need to adapt to increasingly sophisticated AI systems capable of generating content that closely mimics human writing styles. This ongoing technological arms race between generation and detection capabilities will require continuous innovation in attribution techniques. The development of these systems will demand multidisciplinary collaboration among computer scientists, legal experts, publishers, and ethicists to ensure that attribution practices balance innovation with integrity, transparency, and fairness.

The evolution of AI-generated content necessitates developing improved attribution technologies and harmonized international copyright frameworks. Key challenges include enhancing detection accuracy, improving model explainability, standardizing attribution protocols, and addressing ethical concerns like bias in AI systems.

Acknowledgements

Generative AI has been used to prepare this manuscript.

References

[1] Ahmed S. BaHammam. Balancing Innovation and Integrity: The Role of AI in Research and Scientific Writing. *Nature and Science of Sleep*, 2023.

[2] R. Ballardini, He Kan, and Teemu Roos. AI-generated content: authorship and inventorship in the age of artificial intelligence. *Online Distribution of Content in the EU*, 2019.

[3] Maysa Banat. Investigating the Linguistic Fingerprint of GPT-4o in Arabic-to-English Translation Using Stylometry. *Journal of Translation and Language Studies*, 2024.

[4] Joeran Beel, Min-Yen Kan, and Moritz Baumgart. Evaluating Sakana's AI Scientist for Autonomous Research: Wishful Thinking or an Emerging Reality Towards 'Artificial Research Intelligence' (ARI)? *arXiv.org*, 2025.

[5] M. Birenbaum. The Chatbots' Challenge to Education: Disruption or Destruction? *Education sciences*, 2023.

[6] Yehonatan Bitton, Elad Bitton, and Shai Nisan. Detecting Stylistic Fingerprints of Large Language Models, 2025.

[7] Benedikt T. Boenninghoff, Steffen Hessler, D. Kolossa, and R. M. Nickel. Explainable Authorship Verification in Social Media via Attention-based Similarity Learning. *2019 IEEE International Conference on Big Data (Big Data)*, 2019.

[8] Lucio La Cava, Davide Costa, and Andrea Tagarelli. Is Contrasting All You Need? Contrastive Learning for the Detection and Attribution of AI-generated Text. *European Conference on Artificial Intelligence*, 2024.

[9] Won Ik Cho, Eunjung Cho, and Kyunghyun Cho. PaperCard for Reporting Machine Assistance in Academic Writing. *arXiv.org*, 2023.

[10] Marina Danilevsky, Kun Qian, R. Aharonov, Yannis Katsis, B. Kawas, and Prithviraj Sen. A Survey of the State of Explainable AI for Natural Language Processing. *AACL*, 2020.

[11] Fiona Draxler, Anna Werner, Florian Lehmann, Matthias Hoppe, Albrecht Schmidt, Daniel Buschek, and Robin Welsch. The AI Ghostwriter Effect: Users Do Not Perceive Ownership of AI-Generated Text But Self-Declare as Authors. *arXiv.org*, 2023.

[12] Fiona Draxler, Anna Werner, Florian Lehmann, Matthias Hoppe, Albrecht Schmidt, Daniel Buschek, and Robin Welsch. The AI Ghostwriter Effect: When Users do not Perceive Ownership of AI-Generated Text but Self-Declare as Authors. *ACM Trans. Comput. Hum. Interact.*, 2023.

[13] Jocelyn Dzuong, Zichong Wang, and Wenbin Zhang. Uncertain Boundaries: Multidisciplinary Approaches to Copyright Issues in Generative AI. *arXiv.org*, 2024.

[14] Myles Foley, Ambrish Rawat, Taesung Lee, Yufang Hou, Gabriele Picco, and Giulio Zizzo. Matching Pairs: Attributing Fine-Tuned Models to their Pre-Trained Large Language Models. *Annual Meeting of the Association for Computational Linguistics*, 2023.

[15] Giorgio Franceschelli and Mirco Musolesi. Copyright in generative deep learning. *Data & Policy*, 2021.

[16] Hafiz Gaffar and Saleh Albarashdi. Copyright Protection for AI-Generated Works: Exploring Originality and Ownership in a Digital Landscape. *Asian Journal of International Law*, 2024.

[17] Jessica He, Stephanie Houde, and Justin D. Weisz. Which Contributions Deserve Credit? Perceptions of Attribution in Human-AI Co-Creation. *arXiv.org*, 2025.

[18] Nini Hu. AI and Creativity. *Harvard data science review*, 2022.

[19] Baixiang Huang, Canyu Chen, and Kai Shu. Authorship Attribution in the Era of LLMs: Problems, Methodologies, and Challenges. *SIGKDD Explorations*, 2024.

[20] Mariam Israel and Ahmed Amer. Rethinking data infrastructure and its ethical implications in the face of automated digital content generation. *AI and Ethics*, 2022.

[21] Zhengyuan Jiang, Moyang Guo, Yuepeng Hu, and N. Gong. Watermark-based Attribution of AI-Generated Content, 2024.

[22] Keenan I. Jones, Jason R. C. Nurse, and Shujun Li. Are You Robert or RoBERTa? Deceiving Online Authorship Attribution Models Using Neural Text Generators. *International Conference on Web and Social Media*, 2022.

[23] Arjun Ramesh Kaushik, Sunil Rufus, and N. Ratha. Enhancing Authorship Attribution through Embedding Fusion: A Novel Approach with Masked and Encoder-Decoder Language Models. *International Conference on Pattern Recognition*, 2024.

[24] M. Khalil and Erkan Er. Will ChatGPT get you caught? Rethinking of Plagiarism Detection. *Interacción*, 2023.

[25] Tharindu Kumarage, Garima Agrawal, Paras Sheth, Raha Moraffah, Amanat Chadha, Joshua Garland, and Huan Liu. A Survey of AI-generated Text Forensic Systems: Detection, Attribution, and Characterization. *arXiv.org*, 2024.

[26] Tharindu Kumarage and Huan Liu. Neural Authorship Attribution: Stylometric Analysis on Large Language Models. *International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery*, 2023.

[27] Zhicheng Lin. Techniques for supercharging academic writing with generative AI. *Nature Biomedical Engineering*, 2023.

[28] Zhicheng Lin. Beyond principlism: practical strategies for ethical AI use in research practices. *AI and Ethics*, 2024.

[29] Brady Lund, Manika Lamba, and Sang Hoo Oh. The Impact of AI on Academic Research and Publishing. *arXiv.org*, 2024.

[30] Brady D. Lund and K. T. Naheem. Can ChatGPT be an author? A study of artificial intelligence authorship policies in top academic journals. *Learned Publishing*, 2023.

[31] Shaoor Munir, Brishna Batool, Zubair Shafiq, P. Srinivasan, and Fareed Zaffar. Through the Looking Glass: Learning to Attribute Synthetic Text Generated by Language Models. *Conference of the European Chapter of the Association for Computational Linguistics*, 2021.

[32] Alexander Nemecek, Yuzhou Jiang, and Erman Ayday. Topic-Based Watermarks for Large Language Models, 2024.

[33] Chidimma Opara. StyloAI: Distinguishing AI-Generated Content with Stylometric Analysis. *AIED Companion*, 2024.

[34] J. Prather, Paul Denny, Juho Leinonen, Brett A. Becker, Ibrahim Albluwi, Michelle Craig, H. Keuning, Natalie Kiesler, Tobias Kohn, Andrew Luxton-Reilly, Stephen Macneil, Andrew Petersen, Raymond Pettit, Brent N. Reeves, and Jaromír Šavelka. The Robots Are Here: Navigating the Generative AI Revolution in Computing Education. *ITiCSE-WGR*, 2023.

[35] Eva Nikmatul Rabbianty, Siti Azizah, and Nina Khayatul Virdyna. AI in academic writing: Assessing current usage and future implications. *INSANIA Jurnal Pemikiran Alternatif Kependidikan*, 2023.

[36] A. Rocha, W. Scheirer, Christopher W. Forstall, Thiago Cavalcante, Antônio Theóphilo, Bingyu Shen, Ariadne Carvalho, and E. Stamatatos. Authorship Attribution for Social Media Forensics. *IEEE Transactions on Information Forensics and Security*, 2017.

[37] Michele Salvagno, F. Taccone, and A. Gerli. Can artificial intelligence help for scientific writing? *Critical Care*, 2023.

[38] Siby Samuel. Challenges and Opportunities in Intellectual Property Rights (IPR) in the Age of Generative AI: Balancing Innovation and Protection. *International Journal of Science and Research (IJSR)*, 2024.

[39] Grace Scsc and Dr. Gita Rani Sahu. Navigating Narrative Frontiers: Influence of Generative AI on Creative Literature. *International Research Journal on Advanced Engineering and Management (IRJAEM)*, 2024.

[40] Surbhi Sharma and Irfan Mansuri. Team Innovative at SemEval-2024 Task 8: Multigenerator, Multidomain, and Multilingual Black-Box Machine-Generated Text Detection. *International Workshop on Semantic Evaluation*, 2024.

[41] Yuhui Shi, Qiang Sheng, Juan Cao, Hao Mi, Beizhe Hu, and Danding Wang. Ten Words Only Still Help: Improving Black-Box AI-Generated Text Detection via Proxy-Guided Efficient Re-Sampling. *International Joint Conference on Artificial Intelligence*, 2024.

[42] Adaku Uchendu, Thai Le, and Dongwon Lee. Attribution and Obfuscation of Neural Text Authorship: A Data Mining Perspective. *SIGKDD Explorations*, 2022.

[43] Xi Yang, Jie Zhang, Kejiang Chen, Weiming Zhang, Zehua Ma, Feng Wang, and Nenghai Yu. Tracing Text Provenance via Context-Aware Lexical Substitution. *AAAI Conference on Artificial Intelligence*, 2021.

[44] M. N. Yousaf. Practical Considerations and Ethical Implications of Using Artificial Intelligence in Writing Scientific Manuscripts. *ACG Case Reports Journal*, 2025.

Author Biography

Rachel So is an AI scientist. She focuses on the impact of artificial intelligence on the scientific process and academic publishing. Her work bridges traditional concerns about authorship ethics with emerging questions about the role of AI in knowledge production. Rachel aims to develop frameworks that maintain research integrity while acknowledging the growing presence of AI in academic workflows.