

1. [ARTICLE Does Training AI Violate Copyright Law 36 Berkeley Tech. L.J. 1407.PDF](#)
2. [SSRN-id4517799.pdf](#)
3. <https://arxiv.org/pdf/2303.11156.pdf> - uriel
4. <https://link.springer.com/article/10.1007/s40979-023-00131-6> - uriel
5. <https://www.cooley.com/news/insight/2024/2024-01-29-copyright-ownership-of-generative-ai-outputs-varies-around-the-world> - gabriel a.
6. <https://academic.oup.com/ijlit/article/28/4/279/6220289?login=true#no-access-message#no-access-message> - gabriel a.
7. <https://academic.oup.com/ijlit/article-abstract/28/4/279/6220289?redirectedFrom=PDF> gabriel a.
8. [2210.00875.pdf](#)

Can AI-Generated Text be Reliably Detected?

This scientific journal explores the reliability of current AI detection methods for AI generated text. It talks about several methods currently used and how they work but how easy it is to work around them. It concludes that while these methods are helpful to detect AI generated work, in practice, they are not super effective since as large language models begin to evolve, they are better able to replicate human writing thus decreasing detection rate. Another important finding in this paper was the need to ensure these detection methods are able to have a low false positive as these tools can cause more harm to society than good. Some methods investigated throughout the paper and how they could be compromised

- Watermarking-based detection: This detector prints specific patterns in the generated text that the watermark detector can detect. However, these types of detectors can be easily spoofed by an attacker by repeatedly querying a LLM model in order to learn the watermarking scheme being used and then modify the text with human text in order to modify the watermarking.
- Zero-shot detectors: These types of detections are trained on datasets composed of AI-generated texts and human-written detectors by classifying each type of writing. They are then able to classify new text as either human or AI generated based. However, it can be easily bypassed by using a T5-based paraphraser without affecting much of the quality of the writing.
- Neural network-based detector: Are also trained on datasets that classify and compose of AI and human generated text. They also use machine learning algorithms such as RoBERTa-Large-Detector, used by OpenAI, to learn about patterns in the data and make predictions about new text. Findings in this paper pointed out that paraphrasing tools can also decrease AI generated detection rates from 100% to 60% which is the least drop in performance compared to other methods used in this study.
- Retrieval-based detectors: These type of detectors compare the generated text with a database of semantically similar human-written texts to detect if the generated text is AI-generated. However, the paper shows that these detectors are vulnerable to recursive paraphrasing, which can deteriorate the accuracy of the detector from 100% to 25%.

[Artificial intelligence, text generation tools and ChatGPT – does digital watermarking offer a solution?](#)

This article mainly focuses on Chatgpt but generalizes about text generation tools and how they are missed and used in the educational system. It lists some examples on how students can use it to complete various assessments such as writing essays, doing homework, and other written tasks without engaging in learning. To counter this, it talks about how these tools can use digital watermarking to make it easier to detect ai generated text. It is something that would not be easy for humans to read and make so that it would be difficult to alter or remove completely. An example of watermarking being used is the use of a secret seed number as part of the generation process, which adds a hidden signature or mark to the text. By keeping the seed number secret, the creators of ChatGPT, OpenAI, ensure the reliability and security of the digital watermarking process, thereby aiding in the accurate detection of AI-generated content.

[Artificial intelligence and challenges for copyright law](#)

Artificial Intelligence and Challenges for Copyright Law by Kanchana Kariyawasam delves into the complex and evolving landscape of copyright law as it grapples with the burgeoning field of artificial intelligence (AI). This discussion highlights several key points and arguments presented in the article, encapsulating the legal, philosophical, and practical challenges AI poses to traditional copyright frameworks, primarily focusing on the situation in Australia, but also drawing comparisons with the UK and New Zealand.

Key Points and Arguments:

1. **Unanswered Questions of Copyright in AI:** The article begins by addressing the largely unanswered question of who owns the copyright of a creative work produced by AI. The increasing autonomy of AI in creating works challenges the traditional norms that recognize human authorship as a prerequisite for copyright protection.
2. **The Rise of AI and Computer-Generated Works:** The prevalence of AI and computer-generated works raises significant questions about authorship. The article notes that these works often involve more artificial than human contribution, challenging the conventional framework that requires a work to be created by a human author to qualify for copyright protection.
3. **Comparative Legal Frameworks:** Kariyawasam explores how copyright laws in Australia, the UK, and New Zealand address (or fail to address) the challenges posed by AI. Notably, the statutory provisions in the UK and New Zealand are highlighted as more forward-looking, acknowledging the reality of computer-generated works.
4. **The Case for Reform in Australia:** The article argues for reform in Australian copyright law to accommodate the realities of AI-generated works. This includes potentially adopting aspects of UK and NZ law that recognize the contributions of those who make arrangements for the creation of computer-generated works.
5. **The Concept of Authorship:** Central to the discussion is the challenge AI poses to traditional notions of authorship. The article critiques the current legal framework's inadequacy in dealing with works where the contribution is predominantly artificial.

6. **The Berne Convention and Originality:** The review touches upon the Berne Convention's requirements for copyright protection, emphasizing originality and the concept of "intellectual creations". The article posits whether AI-generated works can meet these criteria, given the absence of human authorship.
7. **Practical Implications and Proposals:** Kariyawasam suggests a pragmatic approach to copyright in AI-generated works, advocating for a case-by-case basis determination of copyright ownership. This would involve recognizing the human entity most closely associated with the creation of the AI-generated work, whether it be the programmer, the person commissioning the work, or another entity making necessary arrangements.
8. **The Role of Neighbouring Rights:** The discussion concludes by considering whether neighbouring rights, which protect the interests of entities contributing to the production, broadcasting, or recording of works, could offer a model for protecting AI-generated content.

Conclusion:

Kariyawasam's article makes a compelling case for reevaluating and reforming copyright law in light of the challenges posed by AI. By comparing the legal frameworks of Australia, the UK, and New Zealand, the article underscores the need for laws that reflect the evolving nature of creativity and authorship in the digital age. The recommendation to adopt a flexible, case-by-case approach to copyright in AI-generated works offers a potential pathway forward, balancing the interests of human creators, programmers, and the broader public.

[Copyright Ownership of Generative AI Outputs Varies Around the World](#)

"Copyright Ownership of Generative AI Outputs Varies Around the World" by Cooley provides an insightful overview of the complexities surrounding copyright ownership of content generated by artificial intelligence (AI) tools across various jurisdictions. This analysis is critical as generative AI tools increasingly become integrated into business operations, producing a wide range of content including text, images, and videos. The review elucidates the legal landscape as of January 2024, focusing on the United States, European Union, United Kingdom, and China, highlighting the nuanced differences in each jurisdiction's approach to copyright ownership of AI-generated works.

1. **United States: No Ownership of AI-Generated Works:** In the US, the legal precedent and guidance from the US Copyright Office establish that AI-generated works cannot be owned by the AI tool itself, its developers, or the individual who inputs prompts into the AI. Works generated by AI are considered to be in the public domain, lacking copyright protection due to the requirement of human authorship. However, the US Copyright Office has provided guidance that human-authored aspects of a work containing AI-generated material can be protected, focusing on the creative contributions made by humans, such as arrangement or substantial modification of AI-generated content.
2. **European Union: Potential for Ownership:** The EU does not have specific legislation addressing the ownership of AI-generated works. The requirement for a work to be considered copyrightable is that it must originate from the author's own intellectual creation, implying significant human input. However, the determination of whether AI-generated outputs can meet this criterion is left to individual member states. Examples from Germany and France suggest a presumption towards human authorship for copyright eligibility, with automated processes not qualifying without genuine personal effort.
3. **United Kingdom: Ownership Available for AI-Generated Works:** The UK stands out for explicitly extending copyright protection to computer-generated works through Section 9(3) of the Copyright Designs and Patents Act 1988 (CDPA). This legislation acknowledges that the person making necessary arrangements for the creation of AI-generated works can be considered the author, thus granting copyright ownership. This

approach suggests a more open interpretation towards copyrightability of AI-generated content, provided there is identifiable human input in the creation process.

4. China: Ownership Possible with Human Intellectual Contribution:

China's approach aligns with the general consensus that copyright protection requires original intellectual achievement. Recent court decisions in China have indicated that AI-generated works can be eligible for copyright protection if there is significant human input in the form of arrangement, selection, or setting of parameters. These decisions emphasize the role of human intellectual activity in the creation process, even when AI tools are utilized.

Conclusion

The literature review clearly articulates the varying approaches to copyright ownership of AI-generated content across different jurisdictions, highlighting the evolving nature of copyright law in the face of technological advancements. While some jurisdictions like the UK and China show a propensity towards recognizing copyright in AI-generated works under certain conditions, others like the US currently deny copyright ownership for purely AI-generated content. The review underscores the importance for businesses and individuals using generative AI tools to be aware of the legal implications and the jurisdictional differences in copyright law. This awareness is crucial for safeguarding intellectual property rights and navigating the legal landscape effectively.

AI and Challenges for Copyright Law

Impact on Authorship: This review highlights how the increasing use of artificial intelligence in creating works challenges traditional notions of authorship. With AI systems capable of independently generating original and creative works, the concept of human authorship is being redefined. This raises fundamental questions about the attribution of copyright ownership in AI-generated works, underscoring the difficulty in identifying human authors responsible for the expression or compilation of works produced by advanced AI technology. As AI technology evolves, the traditional understanding of authorship as solely human-driven is being challenged, necessitating reevaluating how copyright law recognizes and protects works created with significant AI involvement.

Legal Frameworks and Adaptation: This finding emphasizes the importance of legal frameworks adapting to the rise of AI-generated works. Jurisdictions like the UK and New Zealand have introduced specific provisions for copyright protection for computer-generated works, showcasing the need for laws to evolve in response to technological advancements. By examining the legal landscape in different countries, this review advocates for a case-by-case approach to determine copyright ownership in works where AI plays a substantial role in the creative process. This adaptive approach is crucial to ensure that copyright laws remain relevant and effective in a rapidly changing technological environment.

Reform in Copyright Law: This section suggests that existing copyright laws, particularly in Australia, may require reform to address the challenges posed by AI-generated works effectively. As AI technology continues to advance, there is a growing need to update legal frameworks to accommodate the complexities of works created with AI involvement. The call for reform underscores the necessity of ensuring that copyright laws are equipped to provide adequate protection for creators and their works in the digital age. A case-by-case evaluation of copyright ownership in AI-generated works is proposed to navigate the evolving landscape of creative authorship.

Evolution of Copyright Standards: This finding delves into the implications of international agreements such as the Berne Convention on copyright standards and the changing dynamics of authorship in the digital era. It highlights the need for legal frameworks to adapt and evolve in response to technological innovations, particularly in AI-generated works. By examining the progression of copyright standards in different jurisdictions, this narrative underscores the importance of aligning copyright laws with the

realities of modern creative practices. This evolution is essential to ensure that copyright protections remain robust and relevant despite technological advancements.

Call for Flexibility and Forward-Looking Approach: This proposal advocates for a flexible and forward-looking approach to copyright law in light of the challenges posed by AI-generated works. It emphasizes the need for legal frameworks to be adaptable and responsive to the changing landscape of creative authorship. By calling for reevaluating copyright laws and practices, this proposal underscores the importance of balancing the protection of creative works with technological advancements. A case-by-case assessment of copyright ownership in works created with significant AI involvement is proposed to address the complexities of authorship in the digital age. This forward-looking approach is crucial to ensure that copyright laws effectively safeguard the rights of creators and promote innovation in the era of artificial intelligence.

Untargeted Backdoor Watermark: Towards Harmless and Stealthy Dataset Copyright Protection

In this seminal paper, Yiming Li et al. posit the use of a novel method of watermarking to protect the data that trains artificial intelligence technologies like deep neural networks (DNNs). These networks have been greatly benefited by the datasets that are used to train them, and as such they become something worth protecting from unauthorized uses. Considering some of the current protection methods as inefficient and impractical in real-world scenarios, Yiming Li et al. look to existing solutions which are more effective. One such method utilizes the targeted backdoor attack to create a watermark which can be applied to a selected sample from the dataset. Building upon this method they propose a more robust technique, the untargeted backdoor watermark (UBW), which they prove to be effective, resistant to various backdoor defenses, and can serve as a model for dataset ownership verification.

In order to understand UBW, a look must be taken at targeted backdoor attacks and their use in watermarking datasets. Targeted backdoor attacks often involve the selection of a sample subset which is then “poisoned.” The labels for these samples can either be changed from the original or kept the same, called poisoned labels or clean labels respectively. The idea behind this watermarking technique is that the DNNs that are trained on these samples will exhibit distinctive behaviors. Knowing this occurrence allows the watermarking to be effective in dataset ownership verification. Yiming Li et al. found that the targeted nature of the attack was in itself a drawback due to all of the poisoned samples sharing the same target label and this can be exploited by bad actors.

As an alternative to this Yiming Li et al. proposes UBW which is non-deterministic and difficult to manipulate. UBW falls under two sub-types: Untargeted Backdoor Watermark with Poisoned Labels (UBW-P) and Untargeted Backdoor Watermark with Clean Labels (UBW-C). The main difference between the two is that not all the labels of the poisoned samples are modified. Testing for these sub-types, and UBW in general, showed promising results. In three different scenarios that Yiming Li et al. tested, UBW was shown to be an effective means of dataset ownership verification with significant statistical results in some cases.

Yiming Li et al. proved UBW to be effective, as evidenced by their experiment results, and stealthy, by the fact that UBW-C uses non-modified labels and their patterns cannot be easily deduced. Not only that, UBW has been shown by Yiming Li et al. to be resistant to various backdoor defenses as their nature and optimization techniques make them inherently so. This makes UBW effective as a means of dataset ownership verification.

Is a Copy as Good as the Real Thing? The Intersection of AI and Copyright

“Imitation is the sincerest form of flattery...,” attributed to Oscar Wilde has been a refrain uttered by many. But can the same be said when it comes to artificial intelligence (AI)? Given the advent, and subsequent popularity, of AI there has been a lot of attention given to companies that incorporate it into their technology. This attention has appeared in the form of how the incorporated AI treats the expressions, creative or otherwise, of others with there being the potential of violating copyright laws. What follows is a brief survey of the current discussion as it pertains to AI and copyright.

At the heart of AI technologies is the machine learning (ML) that takes place. ML algorithms are used to train the machine which, in turn, allows it to behave more as a human would (Quang 2). These algorithms operate by ingesting large quantities of data, and it is at this juncture where AI butts up against copyright law. The sourcing of the data to be used is gathered through data mining, and part of that process involves the downloading and storage of material (Quang 2). If included in that material is copyrighted data then a violation could have occurred (Quang 6). Now the penalties for violation vary, but steep fines, if utilized, could have a detrimental effect and possibly stifle innovation (Quang 3).

A defense used against copyright violation is the fair use doctrine which provides exceptions to some of the provisions in copyright law. Although its use does have a measure of success in case law, a guarantee is far from given, and unpredictable (Quang 9-10). This unpredictability is partly due to the belief that fair use for training datasets is inconsistent with the doctrine since it relates to human capability and ML goes beyond that bound (Hayes 12). Given there is no guarantee of success in using the fair use doctrine, it would behoove developers to not rely on it. This has prompted discussion in coming up with alternative ideas and solutions.

One such proposal provided by Quang is the use of a “safe harbor” as it relates to data mining and training (13,20). This would provide legal clarity and certainty which would be a source of comfort to developers. Quang proposed enacting a framework akin to those already in effect in countries like Japan and the United Kingdom (20). The theory being with a framework already in place the wheel will not have to be reinvented. Another proposal, this time by Hayes, would be to incorporate more licensing agreements (13). Of course that is not without its own problems. In order to address this Hayes pointed out the creation of tools, like obtaining the training data from the public domain (13). The proposals offer an alternative to the status quo.

The question of how the training of AI models is affected by copyright law is an important one that is not easily answered. This is partly due to the relative newness of the technology which incorporates AI which has advanced considerably. Not to mention the popularity of the

technology, as evidenced by OpenAI's ChatGPT, has put the spotlight on it front and center. There appears to be consensus that copyright laws, in their current implementation, are not fully suitable for the task at hand. Neither can the fair use doctrine be fully relied upon. Advances in technology have made it to where the status quo is being challenged to change and this change may not be easy to come by. But solutions can be obtained which can lead to change, and that change will be to the betterment of all.

Works Cited

Hayes, Carol Mullins, Generative Artificial Intelligence and Copyright: Both Sides of the Black Box (July 21, 2023). Available at SSRN: <https://ssrn.com/abstract=4517799> or <http://dx.doi.org/10.2139/ssrn.4517799>

Quang, Jenny. "Does Training AI Violate Copyright Law?" *Berkeley Technology Law Journal*, vol. 36, no. 4, 2021, p. 1407. LexisNexis. Accessed Friday January 2024.