Special Conference on Artificial Intelligence (SpCf. on AI)

Copyright in the era of AI

Research Report

Leiden Model United Nations 2024

"Out of this world"

Forum:	Special Conference on Artificial Intelligence (SpCf. on AI)
Issue:	Copyright in the era of AI
Position:	President

Introduction

The development of digital content has been completely transformed by artificial intelligence (AI), especially generative models like GPT-4, DALL-E, and Stable Diffusion, which allow machines to produce text, images, music, and other media with little assistance from humans. Despite being revolutionary, this quick technology advancement has created difficult legal, moral, and economic challenges that are difficult for current copyright frameworks to handle. In the age of artificial intelligence, traditional copyright law—which was created to safeguard human innovation and originality—faces previously unheard-of challenges: Who is the owner of stuff produced by AI? Is an AI product just a reflection of the data it was trained on, or can it ever be genuinely unique? Importantly, if original creators' copyrighted works are utilised to train these models without their permission, how can their rights be upheld?

Text and data mining (TDM), a method essential to AI model training but rife with legal ambiguity, is at the centre of this copyright debate. To learn and produce new content, AI systems rely on enormous datasets, which frequently contain copyrighted information. A global patchwork of laws that makes international compliance more difficult has resulted from the use of protected works without authorisation, which has led to different approaches to copyright restrictions in places like the US and the EU. Concerns over the possible economic effects of AI on creative industries are also voiced by writers, singers, artists, and other creators, who worry that machine-generated content could undermine human-made works and cause livelihood disruptions.

Definition of Key Terms

Generative AI

Artificial intelligence models that create new material using patterns discovered in massive datasets are referred to as generative AI. Through "learning" from enormous volumes of pre-existing data, including protected content, these models can produce text, photos, music, and other media types. Examples include DALL-E for photos, ChatGPT for text, and other AI programs that can produce original works on their own.

Text and Data Mining (TDM)

TDM is the process by which artificial intelligence (AI) systems gather, examine, and identify patterns in vast amounts of data, usually from the internet or other sources. Because it enables models to process and learn from large datasets, TDM is essential for AI training. However, when the data contains copyrighted content that might be utilised without consent, thereby violating producers' rights, TDM presents copyright issues. varied jurisdictions have varied regulations for TDM; the US and the EU have noteworthy policies.

Copyrighted Content

Any original work of authorship, including text, pictures, music, and videos, that is legally protected to grant the creator the sole right to use, distribute, and reproduce their creation is considered copyrighted content. Unauthorised use of copyrighted content for training by AI models may violate these rights, sparking discussions over whether AI developers ought to get licences or pay for the usage of such content.

"Opt-Out" Model

The opt-out model is a regulatory approach where rights holders must take active steps to prevent their copyrighted works from being used by AI systems, often by making them machine-readable or by other specific means. This model, used in the EU under the DSM Directive, places the burden on creators to control their content's use in TDM, allowing AI developers to use copyrighted content for training unless creators explicitly opt-out.

Intellectual Property (IP)

Trade secrets, patents, trademarks, copyrights, and other legal protections for works of art are all included in intellectual property. IP rules guarantee that firms, inventors, and creators can maintain control over and profit from their works. Since autonomous systems like AI might create new content based on IP-protected data without the creator's authorisation, using copyrighted IP in AI training creates substantial issues because current rules were not created for these systems.

Sui Generis Rights

Specialised, distinct types of intellectual property protection that are not covered by conventional copyright laws are known as sui generis rights. Sui generis rights are suggested by some academics and decision-makers as a potential solution for AI-generated works that require a substantial amount of human input. This strategy would balance public access to completely autonomous AI outputs by putting them in the public domain while offering AI-generated content modest protection.

General Overview

The use of copyrighted content in AI training

At the heart of the copyright debate in the era of AI lies the question of how AI models are trained. Large volumes of data, frequently including copyrighted content from the internet, are needed to train an AI. Large-scale data collection is necessary to create generative AI, which produces models that can learn from copyrighted texts, photos, videos, music, and other media without necessarily obtaining consent from the original authors or copyright holders. Although it is crucial for AI development, text and data mining (TDM), the process by which AI systems collect and examine web data to identify trends and produce new content, poses serious copyright issues.

In some regions, AI developers are restricted in how they can use copyrighted content for training. In the European Union (EU), the 2019 Copyright in the Digital Single Market (DSM) Directive includes two TDM exceptions that govern AI training: Article 3 allows non-commercial TDM for scientific research, but Article 4 extends TDM exemptions to commercial entities if the content is legally accessible, and rights holders have not opted out by using a machine-readable restriction. By imposing the responsibility of protection on rights holders, who must actively forbid AI from exploiting their content, the EU's "opt-out" model permits authors to maintain control over their creations.

On the other hand, copyright law in the US is based on the "fair use" theory, which permits restricted, unlicensed use of copyrighted content for things like research and commentary without requiring consent from the owners of the rights. This approach permits certain forms of TDM without explicit permission, providing a more lenient environment for AI developers to gather data. The US fair use doctrine does not directly address AI but has traditionally favoured access to data for innovation. This disparity between the EU and the US means that AI developers face different legal standards, with EU-based developers potentially facing stricter limitations, as they must comply with copyright restrictions that do not apply under US fair use.

The lack of consistency in TDM policies between the EU, US, and other jurisdictions creates challenges for AI developers who operate internationally. For AI companies using copyrighted content for training, the risk of legal repercussions is higher in regions with stricter TDM requirements, which can hinder innovation. The EU and other jurisdictions have voiced concerns about ensuring a fair competitive landscape, as more restrictive copyright policies in one region could place its developers at a disadvantage compared to those in more permissive regions.

Ownership and authorship of content generated by AI

Since traditional copyright systems were created with human authors in mind, the issue of who owns AI-generated content is at the heart of copyright discussions. Authors of original works are usually granted exclusive rights under copyright laws, which give them authority over the works' public performance, distribution, and reproduction. These laws are predicated on the idea that human authorship is the source of originality and creativity. However, by producing content on its own, generative AI calls into question this idea and raises concerns regarding the ownership and protectability of AI-generated content.

In the United States, the Copyright Office and courts have upheld the requirement for human authorship, ruling that AI-generated works do not qualify for copyright protection. A prominent 2023 case ruled that a visual artwork created solely by an AI system could not be copyrighted, as "human authorship is a bedrock requirement for copyright." Similarly, EU law requires that copyrighted works reflect the creator's personal, intellectual input, meaning that only human-generated works meet this standard. These jurisdictions maintain that AI lacks the capacity for creative intent or personal expression and therefore cannot hold copyright.

Different techniques to AI authorship are being investigated in other regions. The Beijing Internet Court in China decided that, in certain situations, AI-generated content should be granted copyright. In one instance, Tencent's Dreamwriter AI, which produced news items on its own, was found to be protected by copyright since the firm supplied the parameters and rules for the AI's output. This ruling demonstrates a complex understanding of copyright law by implying that, under certain circumstances, works produced by AI under the supervision of a person or organisation may be eligible for copyright protection.

In the UK and some Commonwealth countries, copyright laws recognize "computer-generated works" under a distinct category, attributing copyright to the individual who planned for the work's creation. This approach assigns ownership to the programmer or user who directed the AI rather than to the AI itself, effectively treating AI as a tool rather than a creator. Although works produced solely by an autonomous AI without human input are not covered under the UK Copyright, Designs, and Patents Act, it does contain clauses that provide authorship to the person who starts the creative process. As AI technology develops, these varying interpretations of copyright law underscore the lack of agreement on AI authorship and point to the need for clarification.

Economic impact on creative industries

The creative industries have been significantly impacted by the rise of AI-generated material, which has raised questions about fair competition and market value. Artificial intelligence programs like DALL-E and Midjourney may quickly and affordably produce high-quality writing, music, and graphics, competing with human producers and possibly reducing the demand for human-made goods. Writers, musicians, and artists worry that AI might oversupply the market with low-cost, mass-produced content, undermining the value of original works and affecting the earnings of professionals who depend on their creative abilities.

Expressionism created by DALL-E 2.

For instance, by analysing existing songs, AI-generated music might create melodies, harmonies, and entire compositions that might be identical to human-composed music. Because AI models trained on copyrighted songs may produce works that inadvertently resemble current ones, this raises worries about copyright infringement in the music industry. The legal ambiguity around the subject of whether AI-generated music violates copyright when it mimics particular musical styles or elements could result in conflicts between musicians and AI developers.

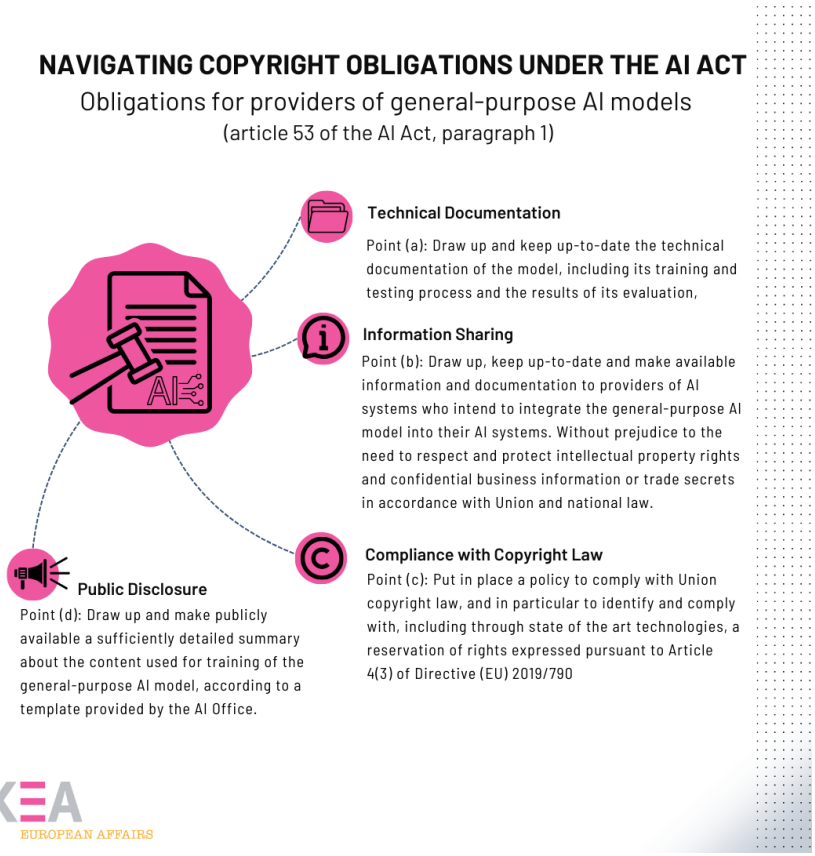
Similar issues have surfaced in literature and visual art, where AI-generated work can imitate well-known authors' and painters' subjects, styles, and approaches. Particularly for commercial applications like illustration, content generation, and backdrop art, writers and artists fear that generative AI will replace some of their labour. Concerns over creative integrity have also arisen because of the quick rise of AI-generated material, since consumers would find it difficult to tell the difference between works created by humans and those created by AI, which could reduce the cultural value of uniqueness and individual expression.

Beyond just individual artists, AI has a significant economic impact on broader sectors of the creative economy, such as publishing, advertising, and entertainment. There are concerns about job displacement and market saturation because of the growing usage of AI systems in certain industries to create content that would otherwise require qualified experts. The broad use of AI creates ethical and economic issues that current copyright rules were not intended to address, even though technology has enormous potential to support the creative industries by automating monotonous activities.

Global regulatory challenges and ethical implications

The necessity for uniform copyright laws is highlighted by the global scope of AI development and data utilisation, since disparities among jurisdictions pose difficulties for businesses doing business internationally. Developers in areas with more lenient copyright rules can acquire greater datasets and build their AI skills more quickly than those in areas with more stringent regulations due to differences in copyright laws, TDM exceptions, and fair use principles.

In the EU, the DSM Directive's TDM exemptions are designed to protect creators while allowing for AI innovation. However, the Directive's opt-out approach contrasts with the US's fair use policy, which favours access to data for commercial and non-commercial AI development without requiring rights holders' permission. This discrepancy has implications for competition, as companies in more restrictive jurisdictions may struggle to keep pace with their counterparts in the US and other regions with lenient copyright policies.



Furthermore, since businesses increasingly operate across borders, AI copyright challenges in one jurisdiction may have worldwide repercussions. The legal concerns that developers confront are highlighted by well-known cases involving AI-generated content, such as Stability AI's legal objections over its usage of copyrighted photos in training. The conflict between copyright protection and AI advancement is highlighted by Getty Images' lawsuit against Stability AI, alleging the business exploited millions of its copyrighted photos for training purposes without authorisation.

The necessity of international cooperation on AI copyright norms is highlighted by this litigation. However, because different nations have different perspectives on intellectual property, fair use, and artists' rights, it is difficult to harmonise copyright rules worldwide. The necessity to modify copyright laws to account for the distinct features of AI-generated output is becoming more widely acknowledged as AI technology advances.

The use of copyrighted content in AI training presents ethical concerns around respect for authors' rights and the integrity of creative work, in addition to legal and financial ones. Concerns regarding the unapproved use of copyrighted works have been raised by AI's reliance on massive datasets that are scraped from the internet, leading to requests for increased openness in AI development. In addition to increased accountability for businesses that make money off AI-generated

material, artists and rights holders are calling for more clear controls over how their creations are utilised in AI training.

Though progress has been slow because of the complexity of the issue and varying priorities among member nations, international organisations such as the World Intellectual Property Organisation (WIPO) and the United Nations (UN) have begun to address these ethical concerns. To ensure that AI can advance while honouring the contributions of artists, writers, musicians, and other creators whose works inspire AI's skills, policymakers must strike a balance between innovation and creators' rights.

Some stakeholders have asked for a brief halt in AI development to give time for the creation of suitable legislation as knowledge of these ethical issues emerges. Elon Musk and other prominent figures in the technology industry wrote an open letter in 2023 calling for a halt to the creation of sophisticated AI models. The signatories emphasised the need for careful regulation to safeguard creators' rights and maintain fair competition, arguing that unrestrained AI development might cause major societal changes.

Major Parties Involved

The European Union (EU)

In an attempt to strike a balance between innovation and intellectual property rights, the European Union has taken the lead in regulating AI and digital copyright. The EU created exclusions for text and data mining (TDM) in the 2019 Copyright in the Digital Single Market (DSM) Directive. While Article 4 expands TDM to commercial purposes if rights holders haven't opted out, Article 3 permits non-commercial TDM for scientific study. In contrast to the US fair use theory, this "opt-out" paradigm requires content creators to actively prevent unauthorised usage of their work. The EU is a key participant in worldwide copyright negotiations because of its policies, which emphasise conflicts between stringent copyright rules and the need for global AI development as it seeks to establish ethical norms.

The United States (US)

The "fair use" concept, which permits limited, unlicensed use of copyrighted material for research and criticism, is the cornerstone of the United States' more lenient approach to copyright issues pertaining to AI. The permissive framework of fair use fosters innovation by lowering obstacles to data access for AI developers, allowing for the wider use of copyrighted content without express consent. This position has raised discussions about possible weaknesses in creator protections, particularly regarding generative AI models that use datasets that are protected by copyright. As it promotes data access to

facilitate AI advancement, the US has a significant influence in international copyright negotiations because of its policy, which permits US AI developers to operate in a more liberal environment than EU norms.

China

Thanks to government funding and policies that prioritise technological development, China is a key worldwide player in AI. Recent court decisions have recognised AI-generated works under certain conditions, such as Tencent's Dreamwriter AI, whose news stories acquired copyright protection based on human monitoring, even though traditional Chinese copyright law emphasises human authorship. This transparency raises the possibility that, under the supervision of people or organisations, China may grant copyright rights to content produced by AI. China's approach has a big impact on other countries as it improves its standards, particularly in balancing copyright protection with the quick development of its AI capabilities. Because China is a major hub for AI technology, it plays a crucial role in talks about data governance and AI worldwide.

World Intellectual Property Organization (WIPO)

Established in 1967, the World Intellectual Property Organisation (WIPO) is a United Nations body that works to advance and defend intellectual property (IP) rights around the world. WIPO oversees international agreements that assist innovators and creators and provide a platform for discussions on IP policy. Given the speed at which AI is developing, WIPO has been actively involved in discussions about copyright and AI, with a particular emphasis on how to safeguard rights holders as generative AI advances. To promote AI development and protect creators' rights, WIPO's AI and IP policy discussions aim to standardise international norms. WIPO's function as an impartial forum is crucial for promoting global collaboration and creating cross-border intellectual property regulations that tackle the intricacies of content produced by artificial intelligence.

United Kingdom (UK)

AI-generated works are treated differently under UK copyright law. The UK treats AI as a tool by recognising "computer-generated works" under its Copyright, Designs, and Patents Act (CDPA) and granting copyright to people who arrange for the development of the AI. This enables the protection of AI-generated work without completely circumventing the conventional human authorship criterion. The UK has taken the lead in talks on striking a balance between creator rights and AI innovation by investigating independent digital copyright policy since Brexit. The UK's inclusion of computer-generated works highlights its position as a leading voice in AI and copyright legislation and provides a potential model for other nations with comparable legal systems.

Japan

Leading exporter of culture and technology, Japan actively encourages AI development while resolving copyright issues. Although human authorship is required under Japanese copyright law to be eligible for copyright, Japan has explored potential TDM exceptions to encourage AI-driven research in copyright-sensitive fields including publishing and music. Japan's position on global copyright concerns is influenced by its national AI policy, which places a strong emphasis on openness, moral AI development, and equitable data use. Japan is a key player in international copyright negotiations because of its emphasis on striking a balance between intellectual property rights and technical advancement, promoting a solution that upholds artists' rights while fostering technological innovation.

Timeline of Events

April 5th, 2016	The Next Rembrandt Project Revealed: A generative AI produces a fresh piece of art influenced by Rembrandt, igniting debate about the creative possibilities of AI and its consequences for copyright around the world.
May 18th, 2018	Japan adapts Copyright Law for TDM: Under certain research and non-commercial criteria, Japan permits data mining for AI training by introducing TDM exclusions for AI research.
March 26th, 2019	EU Approves DSM Directive: The DSM Directive, which emphasises authors' ownership over their content in AI training and establishes TDM exceptions and a "opt-out" model for copyright, is formally approved by the European Parliament.
September 12th, 2019	*Cofemel v. G-Star Raw CJEU Ruling*: The Court of Justice of the European Union (CJEU) clarifies the definition of "original work," requiring human originality, reinforcing the human authorship standard across the EU.
September 22nd, 2021	UK Releases National AI Strategy: The UK publishes its AI and IP policy, establishing rules for AI innovation after Brexit and reaffirming copyright for "computer-generated works."
October 12th, 2022	US Copyright Office Refuses to Register AI-Generated Work: The US Copyright Office insists that human authorship is required for copyright eligibility and refuses to register an artwork that was produced solely by AI.

January, 2023	Getty Images Files Suit Against Stability AI: Getty Images sues Stability AI in both the US and UK for unauthorised use of millions of copyrighted images for AI training, highlighting copyright concerns in AI training practices.
August 18th, 2023	US District Court Ruling (Thaler v. Perlmutter): A US federal court upholds the "human authorship" paradigm by confirming that AI-generated works devoid of human authorship are not entitled to copyright protection.
November 27th, 2023	*Bejing Internet Court Rules in Dreamwriter Case*: Tencent's AI-generated news stories are granted copyright by China's Beijing Internet Court, which is a step towards recognising AI-works under copyright where only human supervision is involved.
May 21st, 2024	*EU AI Act Approved by European Parliament*: The European Parliament formally adopts the AI Act, introducing stringent requirements for transparency in AI training data to protect copyright and foster responsible AI use.

Possible Solutions

Establishing an International AI Copyright framework

The UN could help to establish a uniform, global copyright framework for AI as one potential remedy. This approach would resolve discrepancies between areas such as the US and the EU, where disparities in copyright laws—for example, the US's "fair use" doctrine vs the EU's "opt-out" TDM model—create obstacles for AI developers working internationally. An international framework should provide balanced protections for artists' rights while fostering AI innovation, create common criteria for human vs AI creation, and specify rules for exploiting protected information in AI training. This strategy will encourage ethical AI practices, foster fair competition, and lessen the regulatory burden for international AI enterprises by standardising copyright rules. With required public disclosures of dataset sources, the framework might incorporate rules for openness in AI training data, giving authors the ability to monitor and manage the usage of their creations.

Sui Generis Rights for AI-Generated Content with Human Input

Sui generis rights, a unique form of intellectual property reserved only for AI-generated content in which human input has been incorporated, form a second option. Under this arrangement, works produced with some human assistance—for example, when a user selects prompts or makes

substantial adjustments to AI-generated content—would be given copyright-like protections. This strategy will acknowledge the human contributions to AI-assisted works while putting fully autonomous AI products in the public domain. Sui generis rights might provide a brief protection period (ten to fifteen years, for example) to strike a compromise between encouraging creators and allowing public access. While preventing a copyright monopoly on completely autonomous AI-generated content, such rights would promote investment in AI. By reducing copyright conflicts, preserving access to AI-generated resources, and assisting in the clarification of copyright ownership for works created by AI under human supervision, this approach could promote innovation without jeopardising the rights of authors.

Sources

Bird & Bird, & Hembt, S. (2023, September 1). *The debate on AI-created art and copyright protection*. Bird & Bird. https://www.twobirds.com/en/insights/2023/global/the-debate-on-ai-created-art-and-copyright-protection

European Parliament, & Madiega, T. (2024, September 2). *Artificial intelligence act | Think Tank | European Parliament*. https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2021)698792

Guadamuz, A. & WIPO. (2017, October). *Artificial intelligence and copyright*. WIPO. https://www.wipo.int/wipo_magazine/en/2017/05/article_0003.html#:~:text=The%20copyright%20lies%20with%20the,machine%20can%20do%20its%20thing.

Hill, M. & Charles Russel Speechlys. (2024, May 13). *Copyright in the Age of AI*. Charles Russel Speechlys. https://www.charlesrussellspeechlys.com/en/insights/quick-reads/102j7d1-copyright-in-the-age-of-ai/

KEA. (2024, July 25). *EU AI Act: shaping Copyright compliance in the age of AI Innovation*. https://keanet.eu/eu-ai-act-shaping-copyright-compliance-in-the-age-of-ai-innovation/

Sun, H. (2022). Redesigning copyright protection in the era of artificial intelligence. In *IOWA LAW REVIEW* (Vol. 107, pp. 1213–1251). https://ilr.law.uiowa.edu/sites/ilr.law.uiowa.edu/files/2023-02/A6_Sun.pdf

Texts adopted - Wednesday, 13 March 2024. (2024, March 13). © European Union, 2024 - Source: European Parliament. https://www.europarl.europa.eu/doceo/document/TA-9-2024-03-13_EN.html#top

Wikborg Rein, Grøner, O., & Ersvaer, E. I. (2023, April 5). *Kampen om opphavsrett i AI-æraen*. Wikborg Rein. https://www.wr.no/en/news/the-battle-of-copyright-in-the-ai-era