

JURISPRUDENCE ON ARTIFICIAL INTELLIGENCE IN INTELLECTUAL PROPERTY DISPUTES

by Lim Zhi Jian & Wong Qi Shu

There is little that Artificial Intelligence ("AI") does not have the potential to revolutionise. Inevitably disrupting creative and technical industries, Intellectual Property ("IP") disputes naturally arise. The essence of these disputes lies in how AI-generated works are treated under IP laws.

AI, through machine learning algorithms, processes vast amounts of data to "learn" and create new content. This involves data ingestion where AI systems ingest various datasets, which could be copyrighted, raising concerns about the legality of using such material without explicit permission. The AI then uses this data to generate derivative or original works or even generate inventions, which challenges conventional notions of authorship in Copyright Law and inventorship in Patent Law.

This article examines developing cases from various jurisdictions.

Issue 1: Infringement of copyright by Training AI

Exclusive Rights

- **Right of Reproduction**

Copyright Law grants authors exclusive rights to reproduce their works in different material forms. When AI algorithms engage with copyrighted material, reproduction may occur in several stages: from the work being copied into the AI's database to when the AI produces an output based on the same. The stages may constitute separate acts of reproduction, potentially acts of infringement.

- **Right of Communication to the Public**

AI platforms may violate this right if they distribute generated content that infringes copyrighted material. Even when the output is modified or altered, infringement may still occur if essential, recognisable components of the original work are present.

- **Right of Distribution**

The inclusion of unpublished works in AI training databases without consent can raise serious infringement issues. Making content derived from these works available, even if not identical, could breach the author's rights. The concern is heightened when AI-generated content circumvents the author's control over the first publication and distribution of their original work.

Case Commentaries

- ***HiQ Labs, Inc. v. LinkedIn Corp*, No. 17-16783 (9th Cir. 2022)**

HiQ Labs, a data analytics company, used automated tools to scrape publicly available data from LinkedIn user profiles. It then uses that information, along with a proprietary predictive algorithm, to yield "people analytics," which it sells to business clients. LinkedIn attempted to block hiQ's data scraping and sent a cease-and-desist letter, arguing that hiQ's practices violated LinkedIn's terms of service and constituted unauthorized access under the United States of America's ("U.S.") Computer Fraud and Abuse Act ("CFAA"). HiQ sought an injunction to stop LinkedIn from blocking its access to the data.

The 9th Circuit Court ruled in favour of hiQ, affirming that scraping publicly accessible data does not necessarily violate the CFAA. The Court's decision emphasised that, even though LinkedIn's terms of service prohibited such scraping, accessing data that was already public did not equate to breaching protected computer information. Moreover, the Court emphasised the public's interest in open access to data and balanced it against concerns over privacy and contractual terms.

The ruling did not definitively address broader IP concerns.

- ***Andersen v. Stability AI Ltd*, 23-cv-00201-WHO (N.D. Cal. Aug. 12, 2024)**

A group of artists, led by Andersen, filed a class-action lawsuit against Stability AI (creator of Stable Diffusion) and other AI companies. They claimed that Stability AI unlawfully downloaded billions of copyrighted images from the internet to train generative AI models, violating the artists' copyright.

The court has allowed Andersen's copyright infringement claims to proceed against Stability AI, representing a significant challenge to the AI community. Andersen specifically asserted that her copyright-protected artworks were included in the training datasets used to develop Stable Diffusion.

Since the AI model was trained to analyse and learn from these images, the act of downloading and processing could be considered an infringing reproduction, even if the outputs do not identically match the original works. Andersen argued that this use amounted to unauthorised reproduction of copyrighted content, violating their exclusive rights as copyright holders.

The Court deemed Andersen's claim of ownership and registration, along with her attestations and general description of how her registered works were exploited, sufficient to move forward with the case.

- ***Thomson Reuters v. Ross Intelligence*, No. 1:20-cv-613-SB.**

Ross Intelligence developed an AI-driven legal research platform by training its model on data from Westlaw, a legal service owned by Thomson Reuters. Ross's AI was designed to analyse and learn language patterns from Westlaw's legal summaries, specifically headnotes, to enhance its ability to retrieve and analyse judicial opinions. Thomson Reuters accused Ross of copyright infringement, alleging that Ross copied Westlaw's headnotes to train its AI without authorisation.

The Court denied motions for summary judgment on copyright infringement and fair use claims, meaning these issues will be decided by a jury. The case centres on whether Ross's use of the headnotes was "transformative," with the key issue being whether the AI learned language patterns to generate unprotected quotes or whether it reproduced the creative elements of Westlaw's headnotes.

- ***J.L. v. Alphabet Inc*, U.S. District Court for the Northern District of California, No. 3:23-cv-03440.**

A class action lawsuit was filed against Google, alleging that the company used web scraping to train its AI models, including Google Bard, without consent from copyright holders. The Plaintiffs claim that their copyrighted content was scraped from various websites and used to develop AI systems, constitutes an unauthorised reproduction, as well as an infringement of their exclusive rights under Copyright Law.

The class action centres on issues related to unauthorised reproduction and communication of protected works through AI-generated outputs. The case highlights two key legal questions to be decided: (1) whether the training of Google's large language model on publicly available online resources constitutes fair use under U.S. Copyright Law, and (2) whether shifting the liability for potential copyright infringement—related to both the input data and AI-generated output—to users through contractual agreements is enforceable.

- ***Getty Images v Stability AI, (1:23-cv-00135)***

Getty Images filed a lawsuit against Stability AI, claiming the company unlawfully used millions of copyrighted images to train its AI model, Stable Diffusion. The images were scraped from Getty's website without permission, and Getty argues that these images were reproduced or transformed without authorisation and are thus in violation of Copyright and database rights.

The court denied Stability AI's attempt to dismiss the copyright infringement claims, allowing the case to proceed. The judge determined that there was sufficient evidence to suggest that the training of Stable Diffusion involved the use of Getty's images in a manner that could be considered an infringement. Despite Stability AI's arguments that much of the AI training occurred outside of the United Kingdom ("UK"), the Court emphasised that parts of the process took place in the UK, which made the claim valid under UK Copyright, Designs and Patents Act 1988 ("CDPA").

- ***Doe 1 v. GitHub, Inc., No. 22-CV-06823-JST***

The Plaintiffs allege that GitHub's AI-driven code suggestion tool, Copilot, infringes on their copyrights by using copyrighted code repositories for AI training without proper authorisation. They claim that GitHub scraped public repositories to build its AI model, which resulted in unauthorised reproductions of their work, thus violating provisions of the U.S.'s Digital Millennium Copyright Act ("DMCA"). Specifically, the Plaintiffs argue that Copilot violated Section 1202(b) of the DMCA by removing copyright management information ("CMI") from the code when generating suggestions, stripping away any attribution that would identify the original authors.

The court analysed whether Copilot's output constituted a direct replication of copyrighted material and whether the removal of CMI was sufficient to suggest an intention to conceal infringement. The court concluded that for the Section 1202(b) claims to succeed, the Plaintiffs must demonstrate exact copying of protected works, not just similarities or functional patterns. Without clear evidence of such identical copying, the court dismissed the claims, highlighting the significant burden of proof required under the DMCA.

- ***Authors Guild v. Google No. 13-4829-cv***

Google digitised over 20 million books for its Google Books project, scanning copyrighted works without permission and allowing users to view snippets. The Plaintiffs, including the Authors Guild, claimed that this process violated copyright by reproducing their works without consent.

In determining whether Google's use qualified as fair use, the court applied the four statutory factors. It found Google's use to be transformative, as it created a new, non-substituting function—enabling users to search and access books in a way that was previously unavailable. Despite the works being creative, the court noted that Google's educational and non-commercial purposes supported the fair use claim. The court also ruled that the amount of text copied (snippets) was reasonable, as it didn't undermine the original works. Additionally, the court found no market harm, as Google's service didn't replace the books but rather increased access to them.

Ultimately, the court ruled in favour of fair use, emphasising the transformative nature of Google's use as a key factor. This case highlights how transformative use can be a decisive factor in fair use claims, even if other statutory factors are less clear-cut. In the context of AI training, the focus on transformation could set a precedent for future cases where AI uses copyrighted content to generate new outputs.

Issue 2: Human Element

Originality In Copyright

In Malaysia, Section 7(3)(a) of the CA 1987 requires that a work be original, meaning it must be a result of the author's skill and labour. This was emphasised in *University of London Press Ltd v University Tutorial Press Ltd* [1916] 2 Ch 601, which stressed that originality necessitates human involvement.

In the UK, the CDPA has a specific provision for computer-generated works, where the work is created by a computer without a human author. Under Section 9(3) of the CDPA, the person responsible for the arrangements necessary to create the work is considered the author. This acknowledges the potential for copyright protection even where there is no direct human creation involved. However, in *Infopaq International A/S v. Danske Dagblades Forening* (C-5/08) the European Court of Justice emphasised that for a work to be original, it must reflect intellectual effort and creativity by the author, which is typically human.

In this context, the current interpretation of Copyright Law suggests that while AI can generate works, it does not fulfil the criteria for originality under both Malaysian and UK law, as AI lacks the human element required for intellectual creation or the skill and labour typically associated with authorship. Thus, for a work to qualify for copyright protection, it must involve human effort, which AI inherently lacks.

Case Commentaries



• Jason M. Allen's "Théâtre D'opéra Spatial"

¹Jason M. Allen's AI-generated artwork *Théâtre D'opéra Spatial* won first place in the digital art category at the Colorado State Fair in 2022. The artwork was created using Midjourney, an artificial intelligence

¹ United States Copyright Office. (2023, September 5). Re: Second Request for Reconsideration for Refusal to Register *Théâtre D'opéra Spatial* (SR # 1-11743923581; Correspondence ID: 1-ST5320R).

tool that translates textual descriptions into visual art. Allen has input numerous revisions and text prompts at least 624 times to generate the final piece. Following his success in the competition, Allen sought to secure copyright protection for his creation.

The U.S. Copyright Office ("Office") stated that contains Théâtre D'opéra Spatial more than a de minimis amount of content generated by AI, and this content must therefore be disclaimed in an application for registration. However, Mr. Allen is unwilling to disclaim the AI-generated material, the Work cannot be registered as submitted.

The above incident is one of the factors which led to the Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence issued by the Office on 16 March 2023, the Office emphasised the need of human authorship for the work to be registered. The office stated that when an AI technology receives solely a prompt from a human and produces complex written, visual, or musical works in response, the 'traditional elements of authorship' are determined and executed by the technology—not the human user. It suggests that works produced merely on prompts are not registrable in the US.

- **Cf. Li v. Liu (2023)**



²In this case, the Plaintiff, Li, utilised the AI model Stable Diffusion to create an image using an iterative and meticulous process. Li generated the image through over 150 detailed prompts, carefully arranging and refining parameters to achieve a final result that expressed his artistic vision. The image was subsequently shared on Xiaohongshu, a Chinese social media platform. Liu, the defendant, used this image to accompany an article on a separate website without obtaining Li's permission and even removed the watermark that identified Li as the creator.

The Beijing Internet Court ruled that the AI-generated image qualified as a "work" under Chinese Copyright Law, as it exhibited originality and resulted from human intellectual effort. The court emphasised that the intellectual input from Li (from selecting prompts to refining outputs) met the criteria for authorship. It ruled that the authorship could not be attributed to the AI model, as Chinese law only recognises natural persons or legal entities as authors. Consequently, Liu's use of the image, without proper attribution or authorisation, constituted copyright infringement.

The court recognised that although the work was done by AI software, the person who gave the prompts demonstrated sufficient choice and arrangement to be considered the author. This case illustrates that AI is merely a tool and that a strong degree of human intervention makes the work original.

This shows different approaches taken internationally regarding AI-generated works.

² Beijing Internet Court Civil Judgment (2023) Jing 0491 Min Chu No. 11279.
<https://english.bjinternetcourt.gov.cn/pdf/>

Issue 3: Human Authors and Inventors

Stephen Thaler



³The Stephen Thaler cases centre around the legal recognition of AI as an inventor and author, raising fundamental questions about human involvement in creative and inventive processes under Copyright and Patent Law. Thaler, the creator of the DABUS AI system, filed patent applications for inventions generated by AI, such as a "Food Container" and "Devices and Methods for Attracting Enhanced Attention." Additionally, he sought copyright protection for a two-dimensional artwork titled "A Recent Entrance to Paradise" created by the AI algorithm, Creativity Machine. In both cases, Thaler asserted that AI was capable of independently generating inventive concepts and creative works

In the realm of copyright, Thaler's attempt to have Creativity Machine recognised as the author of creative works faced significant opposition. In most jurisdictions, Copyright Law explicitly requires a human author, emphasising the need for human intellectual creation. Courts rejected Thaler's argument that AI could be considered an author, reinforcing the idea that copyright protections are reserved for works that are the result of human intellectual effort. The UK Intellectual Property Office and the U.S. Copyright Office similarly ruled that AI could not be designated as the author of works, underlining that human authorship is a bedrock requirement of copyright.

In patent applications for inventions generated by DABUS, Thaler argued that the AI system should be listed as the inventor, as it had independently developed the inventions without human intervention. However, patent offices, including in the UK, the U.S., and Europe, rejected this claim. The UK Supreme Court and the European Patent Office (EPO) ruled that Patent Law requires a human inventor based on the interpretation that an inventor must be a person who has the legal capacity to own rights, and AI does not qualify as a legal entity with rights or duties.

Conclusion

The jurisprudence surrounding AI in IP disputes highlights the ongoing tension between technological advancements and traditional IP laws. AI systems, capable of generating creative works or inventions, challenge the foundational concepts of authorship and inventorship, which have traditionally required a human.

The issues currently faced underscore the need for legal reforms to adapt to AI's growing influence in creative and inventive fields.

Lim Zhi Jian | Partner
Intellectual Property
jian@rdsllawpartners.com

Wong Qi Shu | Paralegal
Intellectual Property
qishu@rdsllawpartners.co

³ United States Copyright Office. (2022, February 14). Re: Second Request for Reconsideration for Refusal to Register A Recent Entrance to Paradise (Correspondence ID 1-3ZPC6C3; SR # 1-7100387071). <https://www.copyright.gov/rulings-filings/review-board/docs/a-recent-entrance-to-paradise.pdf>