OWNING THE FUTURE: AI, COPYRIGHT AND THE GLOBAL STRUGGLE FOR INTELLECTUAL PROPERTY IN 2025

by Lim Zhi Jian & Evien See

On 24 June 2025, Judge William Alsup delivered a pivotal ruling in *Bartz et al. v. Anthropic PBC*. This case has since become the largest certified copyright class action involving artificial intelligence to date, with the court allowing millions of U.S. authors to proceed together. At the heart of the dispute was whether AI developers may train their systems with copyrighted works without obtaining any licences.

The plaintiffs alleged that Anthropic ingested thousands of books into its Claude training datasets. Some were legitimately acquired, others pirated. The court drew a clear line: where the books were lawfully acquired, training on them amounted to fair use. The judge described the process as "transformative", similar to how a student studies and synthesises different works and then creates something new without copying any one source verbatim. However, where pirated works were retained and indexed, the court refused to extend the protection of fair use.

Instead of proceeding to trial, Anthropic raised settlement with the plaintiffs on 26 August 2025. While the terms remain confidential, the court filing noted that preliminary approval must be sought in early September. This marked the first major resolution of an AI copyright class action, with authors' counsel calling it "historic." The settlement also underscored the scale of potential liability, as judges had previously warned Anthropic could face billions in statutory damages over an alleged ingestion of up to seven million pirated books.

This ruling is significant because it gave the AI industry some breathing space: training on lawfully acquired works could fall within fair use. At the same time, the combination of class certification and settlement sent a warning that "data hygiene" matters. Using or failing to filter out pirated materials not only leaves companies exposed to infringement claims but now carries the real prospect of costly settlements or damages awards.

The Anthropic decision is not an isolated case. It sits within a larger global debate on three key questions: Who owns AI-generated works? Can training on copyrighted content without permission ever be justified? And what role should intellectual property law play when machines are increasingly involved in creative expression?

Generative AI And Ownership

Generative AI systems differ from traditional software in that they produce new text, images, music, video, and even product designs based on patterns learned from vast datasets. In simple words, the user provides a prompt, and the system generates content in response.

This gives rise to an immediate ownership dilemma. Is the AI company, having built and trained the model, the true creator? Is it the user, whose prompts guide the output? Should the output be considered a collaboration? Or should the law accept that these works cannot be owned at all?

There is still no consistent answer. Courts, regulators and creators are taking different positions depending on their jurisdiction and perspective.

United States: Human Authorship Remains Non-Negotiable

The U.S. has consistently rejected the idea that machines themselves can be authors.

In *Thaler v. Perlmutter*, Dr. Stephen Thaler tried to register a work generated by his "Creativity Machine," naming the AI as the author. Both the Copyright Office and the D.C. Circuit refused. The court held that the Copyright Act requires human authorship, and that listing only the AI meant the application had to be denied.

Similarly, in *Allen v. Perlmutter*, Jason Allen sought to register "Théâtre d'opéra Spatial," a work produced using Midjourney and later refined with software editing. He argued that his hundreds of prompt refinements amounted to authorship. The Copyright Office disagreed. It maintained that prompts alone are insufficient, since the "traditional elements of authorship" are executed by the AI, not the human. As Allen refused to disclaim the AI-generated portions, his application was rejected.

The *Bartz v. Anthropic* case shifts the focus from outputs to inputs. In this case, Judge Alsup gave a measure of clarity: lawfully acquired data may be used to train AI under fair use, but pirated data remains outside that protection. This ruling has made clear to the industry — data provenance is as important as the use of the final output.

Other U.S. disputes add to this picture. In *Thomson Reuters v. ROSS Intelligence,* the Delaware court found that ROSS had infringed copyright by using Westlaw's headnotes to build a competing legal research tool. The use was not fair use because it substituted directly for Westlaw's market. The ongoing New York Times v. Microsoft/OpenAI litigation raises similar issues in the context of news archives, while *Disney v. Midjourney* goes to the heart of whether training on visual works like Disney characters can be justified without licences.

There is a perceivable trend in the U.S.. Courts have firmly rejected the notion of machine authorship, insisting that only works created by humans can qualify

for protection. At the same time, judges are willing to accept that training on copyrighted works may fall within the scope of fair use, but only where the material has been lawfully obtained and the use is genuinely transformative. What U.S. courts will not accept is the use of pirated datasets or outputs that operate as a direct substitute for the original work. Through class certification, that principle was weaponised: infringement was no longer a one-to-one dispute, but a potential multibillion-dollar class liability.

China: Recognising The Prompt Engineer

China has shown greater openness to treating AI as a tool and crediting the human who guides it.

In *Li v. Liu*, the court held that Mr. Li's detailed prompts and technical settings in Stable Diffusion amounted to sufficient intellectual input. The image he generated was therefore protected, and the defendant's use without permission was infringing.

This reasoning was reinforced in the "Half Heart" case, *Chen (Lin) v. Changshu Qin Hong Real Estate*. Mr. Lin tested multiple prompts in Midjourney, selected the best results, and refined them in Photoshop. The court held that this reflected original intellectual input, sufficient for copyright. However, the ruling was limited to his 2D image; a 3D installation inspired by the same work was held not to infringe, because copyright protects expressions, not ideas.

By contrast, in *Feng Runjuan v Kuashi Plastic,* the court dismissed the plaintiff's claim when she could not demonstrate sufficient creative input. Her use of Midjourney prompts was deemed too vague and unoriginal, and she admitted she could not reproduce the work due to AI randomness. The court found that the AI had done the substantive creative work, leaving her with no valid copyright claim.

These cases show that Chinese courts are drawing a clear distinction. Where the human input is meaningful, demonstrable, and creative, the courts have been prepared to recognise copyright protection and treat the AI as no more than a tool. By contrast, where the involvement is minimal, mechanical, or incapable of being evidenced, the claim will fail.

Training Data: The Real Battleground

While questions of authorship remain important, the bigger disputes are now about training data. Can AI companies ingest copyrighted works without permission?

The U.S. Copyright Office's ongoing reports have addressed this head-on. The report in May 2025 reaffirmed that purely AI-generated works are not protectable, and that creators must disclose any AI involvement when registering works. The Office concluded that while fair use remains the applicable framework, there is no blanket safe harbour. Wholesale ingestion of creative works is risky, particularly where the outputs substitute for originals or undermine existing licensing markets. It also encouraged AI developers to adopt safeguards such as licensing agreements, filtering systems, and provenance tracking.

These principles are now playing out in courtrooms and negotiations. The New York Times case is testing whether mass ingestion of articles without permission can be justified as fair use. The BBC's dispute with Perplexity AI, which it accused of reproducing its articles verbatim, illustrates how news organisations are beginning to push back. Perplexity's response was to roll out a "Publishers' Program", offering to share advertising revenue whenever publishers' content surfaces in AI answers. Disney's suit against Midjourney raises similar concerns in the entertainment industry, particularly over the scraping of visual characters.

At the same time, OpenAI's accusations against DeepSeek highlight the industry paradox. OpenAI argues that its own use of public works in training is protected as fair use, but also objects to DeepSeek training its rival model on ChatGPT outputs. If it pursues litigation, it risks undermining the very defence it relies upon elsewhere.

All of this is unfolding against a political backdrop that is shaping how the law develops. The removal of U.S. Register of Copyrights Shira Perlmutter in May 2025 was a sharp reminder that copyright policy does not evolve in isolation. Perlmutter had warned against assuming fair use for wholesale ingestion of expressive works such as music. Her dismissal prompted questions about whether industry pressure played a role and about the independence of the Copyright Office itself. At the same time, public figures like Jack Dorsey and Elon Musk have fanned the flames of debate by questioning the value of intellectual property altogether. Their calls to "delete all IP law" are unrealistic, but they reflect a growing dissatisfaction with a system many view as ill-suited to an AI-driven economy.

What emerges is a fragmented and unstable landscape: companies are litigating, regulators are warning, and politicians and industry voices are pulling the debate in different directions. In practice, wholesale abolition of IP is off the table, but reform is inevitable. More likely than not, this reform will take the form of incremental steps—collective licensing arrangements, clearer exceptions for AI training, and stricter transparency requirements.

Looking Forward

As the law continues to develop, a few trends are already taking shape. This divergence creates a patchwork of rules that companies operating globally must carefully navigate.

Second, data provenance has become critical. Judge Alsup's ruling in the *Anthropic* case makes clear that fair use may protect training on lawfully acquired works, but the use of pirated or unlawfully obtained materials falls outside any safe harbour. For AI developers, this means that recordkeeping, dataset auditing, and provenance tracking are no longer optional but essential compliance measures.

Third, licensing and revenue-sharing mechanisms are beginning to take root. The "Publishers' Program" introduced by Perplexity AI is one example of how the industry may move towards a collective licensing model, similar to the way the music sector addressed digital disruption through blanket licences. While not a

complete solution, these frameworks offer a practical compromise, compensating rights holders while giving AI firms continued access to the data they need.

Fourth, transparency obligations are expanding. The U.S. Copyright Office has already introduced requirements for applicants to disclose AI involvement when registering works. Broader adoption of provenance tracking, labelling, and disclosure mechanisms is likely. These measures serve two purposes: they provide clarity for users and rights holders, and they help AI companies defend themselves against claims of hidden infringement.

Finally, policy remains heavily shaped by politics. The dismissal of Register Shira Perlmutter highlighted the degree to which copyright guidance can be influenced by industry lobbying and shifting political priorities. This means that the trajectory of copyright law in AI will not be determined by courts alone but will also depend on how governments balance innovation with protection of creative industries.

Conclusion: 2025 As A Turning Point

The *Bartz v. Anthropic* decision reflects the tension at the heart of today's copyright debates. By affirming that fair use can apply to training on lawfully acquired works while excluding pirated data, the court provided both reassurance and caution to AI developers.

The ripple effects are being felt globally. In the United States, the decision will influence ongoing disputes such as the New York Times case. In China, courts continue to carve out space for prompt engineers as authors. Meanwhile, companies are experimenting with licensing and revenue-sharing to reduce risk.

2025 may well be remembered as the year copyright law faced its most serious test since the digital revolution. Whether the law adapts by extending existing doctrines, developing new licensing models, or embracing collective solutions will determine not only who owns AI-generated works, but who owns the creative future itself.

Lim Zhi Jian | *Partner*
Intellectual Property
jian@rdslawpartners.com

Evien See | *Associate*
Intellectual Property
evien@rdslawpartners.com