

Kluwer Copyright Blog

Does Human Learning equal Machine Learning? High Court of Delhi to rule on lawfulness of TDM for Machine Learning

Pragya Jha, Bernd Justin Jütte (University College Dublin) · Monday, May 19th, 2025

The debate on whether works protected by copyright can be used for the training of artificial intelligence (AI) has reached India. While dozens of US District Courts are currently grappling with the question of whether AI training with protected works constitutes fair use, the UK High Court is largely grappling with jurisdictional questions, and EU courts are mainly concerned with the modalities of rights reservations (see for overviews [here](#), as well as [here](#) and [here](#)), it is now the High Court of Delhi's turn. The essential question in this first Indian AI case is whether the use of works for training purposes is covered by an exception, or whether AI developers must obtain authorisation for the works used to train their AI systems.

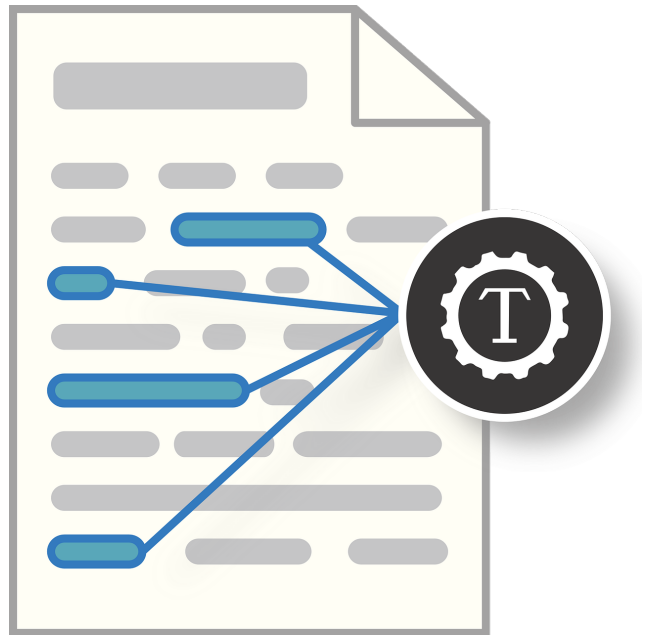


Image by [mcmurryjulie](#) from Pixabay

Background:

In November 2024, the news agency Asian News International Media Private Limited (ANI) filed a case before the High Court of Delhi, India (Ani Media Pvt Ltd vs Open AI Inc & Anr. [CS(COMM) 1028/2024]). ANI alleged that Open AI had used ANI's content to train its Large Language Model (LLM), particularly Open AI's ChatGPT, without obtaining adequate permission from ANI for such usage. ANI contends that [some of its material was accessible only to its subscribers and that Open AI has no authorisation to use the openly available and paywalled materials](#) which were [republished by ANI's subscribers](#). Furthermore, ANI claims that Open AI had falsely attributed false news to the agency, damaging its reputation/ spreading misinformation.

In its application, ANI sought an ex parte and interim injunction on two matters. First, that Open AI or any person acting on Open AI's behalf be restricted from 'storing, publishing, reproducing or in any manner using, including through the ChatGPT model, the copyrighted work of ANI or any

other original works of ANI.’ And, second, that ‘Open AI be directed to disable access of ChatGPT to ANI’s works published anywhere by ANI or its subscribers.’

Open AI submitted that content accessible on ‘www.aninews.in’ had already been blocklisted in October 2024 and that the domain will be excluded from any future training of Open AI. In its order dated 19 November 2024, the High Court of Delhi has set up the following questions for consideration:

1. Whether the storage by Open AI of ANI’s data (*which is in the nature of news and is claimed to be protected under the Copyright Act, 1957*) for training its software *i.e.*, ChatGPT, would amount to infringement of plaintiff’s copyright?
2. Whether the use by Open AI of ANI’s copyrighted data in order to generate responses for its users, would amount to infringement of ANI’s copyright?
3. Whether Open AI’s use of ANI’s copyrighted data qualifies as ‘fair use’ in terms of Section 52 of the Copyright Act, 1957?
4. Whether the Courts in India have jurisdiction to entertain the present lawsuit considering that the servers of the defendants are located in the United States of America?

Arguments of Amici Curiae

To help answer these questions, the Court invited submissions from two *Amici Curiae* (Prof. Dr. Arul Scaria and Advocate Adarsh Ramanujan). The *Amici* made oral submissions during two hearings on 21 February and 10 March. Both argued that ANI must establish that their content is protected by copyright and that it is the lawful owner of that content, and neither *Amicus* seems to contest that the acts of OpenAI engaged the reproduction right under section 14(a)(i) of the [Indian Copyright Act, 1957](#). However, their appreciation of the applicability of the statutory exceptions to the various stages of AI training differ significantly.

As a preliminary point, apart from the questions on the interpretation of substantive Indian copyright law, OpenAI challenges the High Court of Delhi’s jurisdiction to decide on the matter. It advances the argument that none of the relevant acts have been performed in India, a strategy that has also been adopted by Open AI in the *Getty Images v Stability AI* litigation in the UK. Neither *Amici* seems to agree with that contestation. Although the relevant acts of alleged infringement took place outside India, both argue that, according to Section 62 of the [Indian Copyright Act, 1957](#), a suit concerning copyright infringement can be instituted in the court where the plaintiff resides or carries on business. Since ANI has its place of business in New Delhi (which is so far undisputed in these proceedings), the High Court of Delhi would have jurisdiction to hear this matter. The [Judge](#) on the matter mentioned that he would not deal with the jurisdiction issue as a preliminary issue and would hear arguments on both merits and jurisdiction.

On the substance, it seems that the *Amici* did not engage in a detailed analysis of the restricted acts under Section 14 (‘Exclusive rights of reproduction vested with the Copyright owner’) in connection with Section 51 of the Act (‘Acts which amount to the infringement of copyright’). The majority of the arguments advanced focus on the question of whether such acts can be justified based on an exception under Section 52 of the Act.

Section 52 of the [Indian Copyright Act, 1957](#) provides for certain exceptions to the exclusive rights and follows a ‘hybrid’ system of exceptions. Section 52(1)(a) provides three ‘*fair dealing*’ exceptions for private or personal use, which expressly include uses for research, for criticism or review, and for the reporting of current events and current affairs. Section 52 further contains a series of other specific statutory exemptions. However, none of the exceptions listed in section 52 expressly provide for the use of works for TDM, similar to articles 3 and 4 of the EU [CDSM Directive](#), or as in the UK for computational analysis under s. 29A of the [Copyright, Designs and Patents Act 1988](#).

In the absence of an express exception, the *Amici* discuss whether and how dealing with protected subject matter can be accommodated within Indian copyright law. More concretely, both submissions discuss whether reproduction for the creation of training datasets and the training itself fall under the exception for private and personal use, including for research purposes, under Section 52 (1)(a)(i) of the Act.

The [submissions](#) made by the first *Amicus*, Arul Scaria, suggests that the extraction of information for purposes of AI training constitutes a non-expressive use of copyrighted works. In his oral submissions he suggests that a machine learning process is similar to the human learning process and that therefore the relevant exception under Section 52 would apply to human as well as machine learning. He advances the argument that learning is permissible under the current framework of Indian Copyright law because the AI system is trained by ‘learning’ the ingested materials. In addition, AI applications assist individuals with learning and research and storage for such purposes is also permissible under the Indian Copyright Law. Finally, Scaria proposes that exceptions under Section 52 apply to all types of use, including uses by commercial providers of AI systems.

The second *Amicus*, Advocate Adarsh Ramanujan, [argues](#) that LLM training can be divided into three parts: collection of raw data, tokenisation of the collected data, and training of the model, a distinction the first *Amicus* had not made. He agreed with the first *Amicus* only to the extent that tokenising and vectorising of the collected data constitutes a non-expressive use which does not reproduce the original expression. Therefore, this stage would not constitute copyright infringement. However, the other stages (collection of raw data and training of the model) involved expressive use, which amounted to infringement. He stated that collecting and storing publicly accessible data amounted to reproduction under Section 14(a)(i) of the Act and therefore comes within the scope of infringement prescribed under Section 51 of the Act. Ramanujan seemed sceptical that any of the narrowly formulated specific exceptions listed under Section 52(1) apply to machine learning, but it would eventually be Open AI’s onus to demonstrate that the relevant acts are covered under Section 52(1).

-

ANI’s arguments

ANI’s lawyer [argued](#) (in part) before the High Court of Delhi on 10 March and 18 March. Building on Ramanujan’s [argument](#) which separates the training process into three stages he stated that infringement occurred at all stages of the training process as the vectorisation process resulted in an adaptation (Section 14(a)(vi) of the Act) of ANI’s work. In addition to the infringements at the three stages of the training process, further infringement occurred at the output stage. Furthermore, ANI, as the copyright owner, had an exclusive right to use the work and any breach of that

exclusive right amounts to infringement under Section 51 of the Act. These infringements cannot be justified, since Section 52 provides for an exhaustive list of instances in which *prima facie* infringing uses do not require authorisation, and no further permitted uses could be read into the statute apart from the ones that are expressly listed.

Comment

The outcome of the pending case before the High Court of Delhi will carry a certain significance. Whilst the written submissions of the *Amicus* remain unpublished, the reports of the hearings foreshadow an intense proceeding with high stakes. Beyond the issues discussed in this post, the *Amici* have also alluded to the question of opt-outs and filtering of generated outputs, neither of which have a statutory basis in the Indian Copyright Act. Therefore, it can reasonably be expected that the High Court of Delhi will focus on the interplay of exclusive rights and permitted uses.

In the absence of a clearly applicable exception, the answer to the question of whether the use of works for AI training purposes is lawful will determine whether India offers a tech-friendly legal copyright framework. A negative answer might induce the government to take legislative action to address an obvious *lacuna* in Indian copyright law. The arguably required overhaul of India's copyright exceptions will have to address similar policy questions that are currently being debated in the UK.

Substantively, questions that are equally debated in the EU and the US have surfaced against a much more rudimentary statutory background: whether commercial uses of protected subject matter require authorization. Here, the *Amici* are in stark disagreement, which also seems to reflect the respective normative preferences of the *Amici*.

While Arul Scaria's arguments are suggestive of *how the law should be read*, i.e. equating the machine learning with human learning in the light of the broader implications of AI on the Indian economic and innovation ecosystem, the arguments advanced by Adarsh Ramanujan seem to highlight the current position of law i.e. *what the law is* and how the acts of Open AI are infringing copyright unless it is demonstrated that they are exempted under Section 52.

Ramanujan's approach aligns with the written response submitted in the Upper House of Parliament in 2024 by the Union Minister of State for Commerce and Industry (subsequently [published](#) by the Press Information Bureau), which stated that the existing legislation obligates the user of generative AI to obtain permission to use the copyrighted works from the owner of such work if the use was intended for commercial purposes i.e. if the use of such copyrighted work was not exempted under Section 52 of the Act.

To make sure you do not miss out on regular updates from the Kluwer Copyright Blog, please [subscribe here](#).

2024 Future Ready Lawyer Survey Report

Legal innovation: Seizing the future or falling behind?

[Download your free copy →](#)

 Wolters Kluwer



This entry was posted on Monday, May 19th, 2025 at 10:22 am and is filed under [Artificial Intelligence \(AI\)](#), [Case Law](#), [Exceptions and Limitations](#), [India](#), [Infringement](#), [Text and Data Mining \(TDM\)](#)

You can follow any responses to this entry through the [Comments \(RSS\)](#) feed. You can leave a response, or [trackback](#) from your own site.