



The Indian Journal for Research in Law and Management

Open Access Law Journal – Copyright © 2026

Editor-in-Chief – Dr. Muktai Deb Chavan; Publisher – Alden Vas; ISSN: 2583-9896

This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike 4.0 International (CC-BY-NC-SA 4.0) License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium provided the original work is properly cited.

IS TRAINING AN AI MODEL "FAIR DEALING" UNDER INDIAN LAW, OR INFRINGEMENT?

Saad Mulani

ABSTRACT

This article examines whether training generative artificial intelligence models on copyrighted works constitutes fair dealing under Section 52(1)(a) of the Copyright Act, 1957. It analyses the Delhi High Court's interim observations in ANI Media Pvt. Ltd. v. OpenAI OpCo LLC and compares Indian law with United States decisions concerning AI training, transformative use, and market harm. The article argues that neither technological innovation nor commercial research automatically justifies copying protected works. It recommends a clearer Indian legal framework addressing lawful data acquisition, transparency, anti-piracy safeguards, and market substitution while balancing creators' rights with technological development.

Keywords: Artificial Intelligence, Copyright, Fair Dealing, AI Training, Section 52, ANI v. OpenAI.

THE QUESTION IS NO LONGER HYPOTHETICAL

A generative AI system can process millions of books, articles, images, recordings and other works before producing a single answer. The copyright difficulty therefore begins before the output. AI training ordinarily requires works to be copied, stored, processed and repeatedly used by computational systems. The central question is consequently not merely whether an AI output resembles a protected work, but whether the upstream copying required for training is itself lawful.

India now has an important judicial starting point. In ANI Media Pvt. Ltd. v. OpenAI OpCo LLC, the Delhi High Court considered whether OpenAI's storage of ANI's literary works for training its

large language models (“LLMs”) could fall within Section 52(1)(a) of the Copyright Act, 1957. Justice Amit Bansal held, at the interim stage, that the training-related storage was prima facie covered by the fair-dealing exception and therefore did not amount to infringement. The Court, however, expressly limited its observations to the interim application and clarified that they would not determine the final outcome of the suit.¹

ANI therefore does not create a blanket licence to scrape copyrighted material. It instead exposes the central doctrinal problem: Section 52 was drafted long before machine learning, while AI can convert millions of individual acts of copying into one industrial-scale process.

SECTION 52: FAIR DEALING, NOT AN OPEN LICENCE

Section 51 identifies copyright infringement, while Section 52 specifies acts that do not constitute infringement. Section 52(1)(a)² permits “fair dealing” with a work, other than a computer programme, for purposes including private or personal use and research, criticism or review, and reporting current events. The accompanying Explanation also recognises electronic storage for the purposes covered by the provision.

Two questions therefore arise. First, does AI training fall within a recognised statutory purpose, particularly “research”? Second, if it does, is the dealing actually fair? A commercial developer cannot simply label machine learning “research” and end the inquiry. Purpose is only the first stage.

Indian jurisprudence has treated fairness as a contextual inquiry rather than a purely quantitative one. In *Super Cassettes Industries Ltd. v. Chintamani Rao*, the Delhi High Court considered the extent and qualitative significance of copying, its purpose, commercial character and the circumstances of the use. It also observed that commercial use does not automatically make dealing unfair and recognised that transformative use may be relevant in appropriate circumstances.³

This creates both an opportunity and a limitation for AI. A model does not ordinarily consume a novel for its expressive value; it processes the work to identify patterns from which a

¹ *ANI Media Pvt. Ltd. v. OpenAI OpCo LLC*, CS(COMM) 1028/2024, I.A. 45300/2024 (Delhi High Court July 24, 2026) (Bansal, J.)

² Copyright Act, 1957, No. 14 of 1957, § 52(1)(a) (India).

³ *Super Cassettes Industries Ltd. v. Chintamani Rao*, 2012 (49) PTC 1 (Del.), ¶¶ 27-28, 81-82 (Del. 2011).

computational system is developed. But the scale of copying is unprecedented. The fact that the ultimate function is different does not, by itself, answer whether copying millions of works is fair.

WHAT ANI V. OPENAI ACTUALLY DECIDED

ANI is currently India's most direct judicial engagement with AI training. The Court distinguished between the alleged reproduction of ANI material in ChatGPT responses and the storage and use of ANI material during training. On the record before it, ANI did not establish sufficient evidence of memorisation, regurgitation or substantial reproduction in the challenged outputs.⁴

On training, the Court adopted a two-stage analysis: a **purpose test** followed by a **fairness test**. It held prima facie that training LLMs could fall within “private or personal use, including research” under Section 52(1)(a). Importantly, the Court reasoned that research could not be confined to activity physically undertaken by human beings; machine learning could also facilitate research.⁵

After that, the question of fairness came up before the Court. In lieu of simply transplanting the concept of “fair use” as applied in the U.S., the Court developed criteria that fit the context of India’s copyright legislation. These included the extent of use being restricted to training the LLM only, whether it created any economic rivalry or harm to the interests of the copyright owner, and whether the AI system helped achieve other public objectives like scientific research, technology development, education, and spreading of information.⁶

The Court ruled in favour of fulfilling both the criteria, prima facie. It also took into consideration the practical problem involved in obtaining licenses from different sources for developing LLMs and said that granting an interim injunction would be detrimental to the development of AI systems and public interest. More importantly, however, the Court made it explicitly clear that these observations were not binding in any way on the determination of the case.⁷

WHAT THE U.S. CASES ADD

The United States may serve as a good comparison to India; however, its fair-use policy cannot be directly applied here. For example, in the case of *Authors Guild v. Google*⁸, the Second Circuit

⁴ *ANI Media Pvt. Ltd. v. OpenAI OpCo LLC*, CS(COMM) 1028/2024, I.A. 45300/2024 (Delhi High Court July 24, 2026), ¶¶ 271-74.

⁵ *Id.* ¶¶ 202-19.

⁶ *Id.* ¶ 220-38.

⁷ *Id.* ¶ 274.

⁸ *Authors Guild, Inc. v. Google, Inc.*, 804 F.3d 202, ¶¶ 214-25 (2d Cir. 2015).

held that Google's scanning of millions of books to create a searchable database was a separate and distinct informational purpose, and it did not substitute the books.

However, *Andy Warhol Foundation v. Goldsmith* shows why "transformative" must not become a magic word. The Supreme Court of the United States said that the use of a different expression or message is not necessarily fair use, especially if the secondary use performs the same commercial function.⁹

The recent AI cases make the picture more complicated. In the *Bartz v. Anthropic PBC* case, the Northern District of California ruled that the use of books to train LLMs is fair use in that case, but did not consider the unlawful acquisition and possession of pirated copies in the same light.¹⁰ In *Kadrey v. Meta Platforms, Inc.*, Meta was similarly granted summary judgment on its training claim, owing to the highly transformative use and the lack of evidence for any market dilution by the plaintiffs.¹¹

A more recent appeal decision serves as a good example in this regard. In the *Thomson Reuters Enterprise Centre GmbH v. Ross Intelligence Inc.* case, the Third Circuit ruled that the Ross copying of Westlaw headnotes to build an AI-based competitor product was not fair use. The case involved a non-generative system and a directly competing product, but it demonstrates that "AI training" is not itself a defence. Purpose, transformation, commercial competition and market effects remain decisive.¹²

THE INDIAN PROBLEM: SCALE, SOURCE AND MARKET EFFECT

The difficult question for India is therefore one of scale. Traditional fair-dealing disputes involved identifiable acts of copying. AI training can involve scraping websites, creating multiple copies during preprocessing, tokenising and storing datasets, training successive models and retaining material for future development.

Scale may also affect market harm. Copyright law protects creative incentives partly by preserving the economic value of creative works. If AI developers can acquire the same corpus without

⁹ *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, ¶¶ 528-39 (2023).

¹⁰ *Bartz v. Anthropic PBC*, No. 3:24-cv-05417-WHA, 2025 WL 1741691, ¶¶ 1-31 (N.D. Cal. June 23, 2025).

¹¹ *Kadrey v. Meta Platforms, Inc.*, 788 F. Supp. 3d 1026, ¶¶ 1033-54 (N.D. Cal. 2025).

¹² *Thomson Reuters Enter. Centre GmbH v. Ross Intelligence Inc.*, No. 25-2153, slip op. ¶¶ 1-31 (3d Cir. Sept. 29, 2026).

payment and then compete with publishers, journalists, authors or other creators in markets built around those works, the economic consequences may become relevant to fairness.

At the same time, requiring a separate licence for every work in a massive training corpus could make large-scale AI research economically impracticable. *ANI* expressly recognised this concern.¹³ The law therefore has to avoid both extremes: treating every form of training as infringement and treating everything available online as automatically free for computational use.

WHAT SHOULD INDIAN LAW DO NEXT?

India should resist the immediate temptation to create a blanket AI-training exception. *ANI* reveals that Section 52 possesses enough flexibility to cope with technological advancements, yet it also illustrates the need for fact-based analysis.

Future regulations will need to differentiate between legitimately acquired materials and pirated content, between training copies and archives, and between research-led development and technology which is designed to replace the preexisting creative markets. Disclosure of the sources of the training would also help to establish whether those uses were unlawful.

Thus, the regulatory approach can take into consideration a narrow exception from copyright for the sake of training with certain requirements of transparency, anti-piracy measures, opt-out procedures, and licensing/remuneration where the market substitution takes place.

The response to the question is neither the training is always fair dealing nor that each act of training constitutes an infringement. In accordance with Indian law, whether the use was lawful will depend on the purpose of the copyright exception and the fairness of the particular use. *ANI* vs. *OpenAI* case serves as a good foundation but not the rule.

Endnotes:

1. *ANI Media Pvt. Ltd. v. OpenAI OpCo LLC*, CS(COMM) 1028/2024, I.A. 45300/2024 (Delhi High Court July 24, 2026) (Bansal, J.).

¹³ *ANI Media Pvt. Ltd. v. OpenAI OpCo LLC*, CS(COMM) 1028/2024, I.A. 45300/2024 (Delhi High Court July 24, 2026), ¶¶ 265–70.

2. Copyright Act, 1957, No. 14 of 1957, § 52(1)(a) (India).
3. *Super Cassettes Industries Ltd. v. Chintamani Rao*, 2012 (49) PTC 1 (Del.), ¶¶ 27-28, 81-82 (Del. 2011).
4. *ANI Media Pvt. Ltd.*, CS(COMM) 1028/2024, ¶¶ 271-74.
5. *Id.* ¶¶ 202-19.
6. *Id.* ¶¶ 220-38.
7. *Id.* ¶ 274.
8. *Authors Guild, Inc. v. Google, Inc.*, 804 F.3d 202, 214-25 (2d Cir. 2015).
9. *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 528-39 (2023).
10. *Bartz v. Anthropic PBC*, No. 3:24-cv-05417-WHA, 2025 WL 1741691, at 1-31 (N.D. Cal. June 23, 2025).
11. *Kadrey v. Meta Platforms, Inc.*, 788 F. Supp. 3d 1026, 1033-54 (N.D. Cal. 2025).
12. *Thomson Reuters Enter. Centre GmbH v. Ross Intelligence Inc.*, No. 25-2153, slip op. at 1–31 (3d Cir. Sept. 29, 2026).
13. *ANI Media Pvt. Ltd.*, CS(COMM) 1028/2024, ¶¶ 265–70.