USE OF COPYRIGHTED MATERIAL FOR TRAINING LARGE LANGUAGE MODELS: INDIAN LAW AFTER *ANI V. OPENAI*

Vatsal Sutaria, NMIMS' Kirit P Mehta School of Law, Mumbai

Archie Kotwani, NMIMS' Kirit P Mehta School of Law, Mumbai

ABSTRACT

Large language models are trained on very large collections of text, and much of that text is protected by copyright. This paper asks whether Indian law allows that use without a licence. It explains which steps in training involve copying, reads Sections 14 and 52 of the Copyright Act, 1957 against those steps, and examines the Delhi High Court's interim order of July 24, 2026 in *ANI Media v. OpenAI*. It then compares other jurisdictions and assesses the licensing model proposed by the Indian government in December 2025. The paper argues that the *ANI* order leaves two questions open, the source of training data and memorised output, and that reform in India should address those questions directly.

I. Introduction

A large language model (LLM) learns to predict the next word in a sequence by processing billions of words of text. Much of that text is under copyright, and its owners were rarely asked. Since 2023, owners have sued developers over this practice in the United States, the United Kingdom, Germany and India.

NASSCOM and Boston Consulting Group estimated in 2024 that the Indian AI market would reach about USD 17 billion by 2027.[1] The legal question became concrete on July 24, 2026, when Justice Amit Bansal of the Delhi High Court refused an interim injunction sought by the news agency ANI against OpenAI. He held, prima facie, that storing ANI's articles to train ChatGPT fell within the fair dealing exception in Section 52(1)(a) of the Copyright Act, 1957.[2]

[1]Nat'l Ass'n of Software & Serv. Cos. & Boston Consulting Grp., AI-Powered Tech Services: A Roadmap for Future-Ready Firms (2024).

[2]ANI Media Pvt. Ltd. v. Open AI OpCo LLC, I.A. 45300/2024 in CS(COMM) 1028/2024, 2026:DHC:5900 (Del. H.C. July 24, 2026) (India); *see also* Cyril Amarchand Mangaldas, *ANI v. Open AI: Delhi High Court Refuses Interim Injunction in Landmark AI Copyright Dispute*, Cyril Amarchand Mangaldas Blog (July 2026), https://corporate.cyrilamarchandblogs.com/2026/07/ani-v-open-ai-delhi-high-court-refuses-interim-injunction-

ANI has appealed to a Division Bench.[3]

Part II describes where copying happens during training. Part III applies the Indian statute and the *ANI* order to those facts. Part IV sets out how other jurisdictions have dealt with the same question. Part V evaluates the working paper issued by the Department for Promotion of Industry and Internal Trade (DPIIT), and Part VI makes proposals for Indian law.

II. Where copying happens in training

Training involves four stages, and copyright law cares about each of them. The developer first gathers works, usually by running a web crawler or downloading existing datasets, and stores them. The stored text is then cleaned, deduplicated and converted into tokens, which are numerical codes for words or parts of words. Next, the model reads the tokens many times and adjusts its internal parameters, called weights, to reduce its prediction errors. Finally, the trained model is deployed and produces text in answer to user prompts. Some systems also fetch live web content at the time of a query, which is called retrieval-augmented generation.

The first two stages plainly produce copies. A training corpus is a set of complete stored copies, even if they are later tokenised. The third stage is harder. Developers say the weights record statistical relations between tokens and do not contain the works. Grimmelmann describes this kind of machine reading as non-expressive use, because the machine does not take from the work anything a human reader would enjoy.[4] Lemley and Casey argue along similar lines that copying in order to learn unprotectable patterns should generally be permitted.[5] Sobel had earlier pointed out that the argument is weaker for what he called expressive machine learning, where a model is trained to produce works of the same kind as those it learns from.[6]

There is also a factual limit to the developers' argument. Carlini and his co-authors showed that a language model can memorise parts of its training data and reproduce them word for word when prompted in certain ways.[7] The U.S. Copyright Office took the view in 2025 that where

in-landmark-ai-copyright-dispute/.

[3] *ANI Challenges Delhi High Court Order Denying Interim Relief Against OpenAI*, Adgully (Sept. 2026), https://www.adgully.com/post/20258/ani-challenges-delhi-high-court-order-denying-interim-relief-against-openai.

[4] James Grimmelmann, *Copyright for Literate Robots*, 101 Iowa L. Rev. 657 (2016).

[5] Mark A. Lemley & Bryan Casey, *Fair Learning*, 99 Tex. L. Rev. 743 (2021).

[6] Benjamin L.W. Sobel, *Artificial Intelligence's Fair Use Crisis*, 41 Colum. J.L. & Arts 45 (2017).

[7] Nicholas Carlini et al., *Extracting Training Data from Large Language Models*, *in* Proceedings of the 30th USENIX Security Symposium 2633 (2021).

a model can output substantial protected expression, copying or distributing its weights may itself engage the reproduction right.[8] Whether a given model has memorised a given work is a question of evidence.

III. The Indian position

A. The reproduction right

Section 14(a)(i) gives the owner of a literary work the exclusive right to reproduce it "in any material form including the storing of it in any medium by electronic means."[9] Downloading an article into a training corpus stores it by electronic means, so it is a reproduction unless an exception applies. In *ANI*, the court accepted that OpenAI's storage of ANI's works during training engaged the reproduction right.[10] The contest was therefore over the exceptions.

Copyright subsists only in original works.[11] In *Eastern Book Co. v. D.B. Modak*, the Supreme Court held that originality requires a minimal degree of creativity together with skill and judgment, and it rejected the pure "sweat of the brow" test.[12] For news, this means the facts in a report are free to use and only the reporter's expression is protected. The court in *ANI* relied on this distinction when it rejected ANI's complaint about ChatGPT's answers.[13]

B. Fair dealing and storage under Section 52

Section 52 lists acts that do not infringe. Unlike American fair use, it is a closed list tied to named purposes. Section 52(1)(a) permits fair dealing with any work other than a computer program for private or personal use including research, for criticism or review, and for reporting current events. An Explanation added in 2012 provides that storing a work in any electronic medium for those purposes does not infringe.[14] Section 52(1)(b) covers transient or incidental storage purely in the technical process of electronic transmission. Section 52(1)(c)

[8]U.S. Copyright Off., Copyright and Artificial Intelligence, Part 3: Generative AI Training (pre-publication version, May 2025).
[9]The Copyright Act, 1957, No. 14, Acts of Parliament, 1957, § 14(a)(i) (India).
[10]*ANI Media*, 2026:DHC:5900; *see* Intepat, *ANI vs OpenAI Judgment: What the Delhi High Court Actually Held*, Intepat Blog (2026), https://www.intepat.com/blog/ani-vs-openai-judgment-delhi-high-court.
[11]The Copyright Act, 1957 § 13(1)(a).
[12]Eastern Book Co. v. D.B. Modak, (2008) 1 SCC 1 (India).
[13]*ANI Media*, 2026:DHC:5900; LexCounsel, *India's First AI Copyright Ruling: The Delhi High Court in ANI v. OpenAI*, Legal 500 (2026), https://www.legal500.com/intelligence/india/intellectual-property/indias-first-ai-copyright-ruling-the-delhi-high-court-in-ani-v-openai.
[14]The Copyright Act, 1957 § 52(1)(a) & expl.; The Copyright (Amendment) Act, 2012, No. 27, Acts of Parliament, 2012 (India).

covers transient or incidental storage for providing electronic links, access or integration, unless the right holder has expressly prohibited it or the person responsible knows the stored copy is infringing.[15]

Clauses (b) and (c) do not fit training well. A training corpus is not stored in the course of transmitting anything, and developers often keep it for months and reuse it across training runs. That is difficult to call transient. The developer's case therefore rests on clause (a), and in particular on reading "research" in sub-clause (i) to include the building of a commercial model.

The text of sub-clause (i) supports two readings. On one reading, "research" takes its meaning from the words "private or personal" before it, and does not reach research carried out by a company to build a product used by millions of people. On the other, Parliament added "including research" in 2012 without any non-commercial condition, and a court should not supply one. Indian courts have not read Section 52 grudgingly. In the *Rameshwari Photocopy* case, the Delhi High Court declined to give the educational exception in Section 52(1)(i) a narrow reading.[16] Fairness is still a separate inquiry. In *Civic Chandran v. Ammini Amma*, the Kerala High Court looked at the quantity taken, the purpose of the use and whether the use competed with the original.[17]

C. The ANI order

ANI sued in November 2024. It made two claims: that OpenAI had scraped and stored its news content to train ChatGPT, and that ChatGPT reproduced that content in its answers.[18] The court rejected OpenAI's objection that it lacked jurisdiction because training took place on servers in the United States. It treated storage abroad as the final step in a process that began with access to works in India, and it noted that OpenAI offers its services to users in India.[19]

On the training claim, the court held prima facie that OpenAI's storage of ANI's works fell within Section 52(1)(a). ANI had argued that the Explanation to that clause requires every stored work to be a non-infringing copy. The court rejected that argument, reading the non-infringing copy condition as part of the words dealing with incidental storage of computer

[15]The Copyright Act, 1957 § 52(1)(b)–(c).
[16]Chancellor, Masters & Scholars of the Univ. of Oxford v. Rameshwari Photocopy Servs., CS(OS) 2439/2012 (Del. H.C. Sept. 16, 2016) (India).
[17]Civic Chandran v. Ammini Amma, 1996 PTR 142 (Ker. H.C.) (India).
[18]LexCounsel, *supra* note 13.
[19]*ANI Media*, 2026:DHC:5900; Cyril Amarchand Mangaldas, *supra* note 2; LexCounsel, *supra* note 13.

programs.[20] On the output claim, the court found that ChatGPT's answers were not substantially similar to ANI's articles and that ANI had not proved systematic memorisation and reproduction of its text.[21] On the balance of convenience, the court noted that ANI had not used the opt-out mechanisms available to it to stop OpenAI's crawler.[22]

The findings are prima facie, and the order decides less than some reports suggest. The defence succeeded on material that was freely available online, and ANI did not allege that OpenAI had obtained it unlawfully. Pirated or paywalled sources were not before the court. Retrieval-augmented output was not pleaded as a separate category and was not decided.[23] The leading American cases turn on these very points.

IV. Other jurisdictions

A. United States

Section 107 of the U.S. Copyright Act directs courts to four factors: the purpose and character of the use, the nature of the work, the amount used, and the effect on the market for the work.[24] Before generative AI, the Second Circuit held that scanning entire books to build a searchable index was fair use, because the index gave users information about the books without supplying their expression.[25] In *Andy Warhol Foundation v. Goldsmith*, the Supreme Court narrowed the transformative use inquiry. It asked whether the new use served substantially the same purpose as the original, and treated commercial substitution as weighing against fair use.[26]

Three district courts applied these rules to AI training in 2025. In *Thomson Reuters v. Ross Intelligence*, Judge Bibas held that Ross's use of Westlaw headnotes to train a legal research tool was not fair use. Ross had built a competing product, and its use harmed Thomson Reuters's existing market and a potential market for licensing content as AI training data. He limited the ruling to non-generative AI.[27] The Third Circuit heard argument on the interlocutory

[20]Cyril Amarchand Mangaldas, *supra* note 2.
[21]*ANI Challenges Delhi High Court Order*, *supra* note 3.
[22]LexCounsel, *supra* note 13.
[23]Intepat, *supra* note 10.
[24]17 U.S.C. § 107.
[25]Authors Guild v. Google, Inc., 804 F.3d 202, 217 (2d Cir. 2015); *see also* Authors Guild, Inc. v. HathiTrust, 755 F.3d 87, 97 (2d Cir. 2014).
[26]Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith, 598 U.S. 508 (2023); *see* Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 579 (1994).
[27]Thomson Reuters Enter. Ctr. GmbH v. Ross Intel. Inc., 765 F. Supp. 3d 382 (D. Del. 2025).

appeal on June 11, 2026, and its decision was pending at the time of writing.[28]

In *Bartz v. Anthropic*, Judge Alsup held that using books to train Anthropic's Claude models was fair use, and that scanning lawfully purchased print books into digital form was also fair use. He reached a different result on the millions of pirated books Anthropic had downloaded to build a central library, holding that the later training use did not excuse that copying.[29] The parties settled. On July 20, 2026, the court gave final approval to a USD 1.5 billion settlement covering about 482,460 works. Anthropic must destroy the pirated files, and the release covers only past conduct, leaving claims about outputs untouched.[30]

Two days after *Bartz*, Judge Chhabria granted summary judgment to Meta in *Kadrey v. Meta*, but only because the authors had not produced evidence of market harm. He said that training on copyrighted works without permission would be unlawful in many cases, and he treated seriously the theory that a model can dilute the market by producing large numbers of works that compete with the originals.[31] Claims centred on output, such as those brought by *The New York Times* against Microsoft and OpenAI, remain pending.[32] So far, American courts have accepted that learning from works can be transformative, but they look closely at how the works were obtained and whether the product competes with them.

B. European Union and United Kingdom

Article 3 of the 2019 Directive on Copyright in the Digital Single Market permits text and data mining by research organisations and cultural heritage institutions for scientific research. Article 4 permits text and data mining by anyone, including commercial developers, where the user has lawful access and the right holder has not reserved its rights. For content made publicly available online, the reservation must be made in an appropriate manner, such as machine-readable means.[33] The AI Act adds two duties for providers of general-purpose AI models.

[28]Thomson Reuters Enter. Ctr. GmbH v. Ross Intel. Inc., No. 25-2153 (3d Cir. argued June 11, 2026); *see* Stevens & Lee, *Third Circuit Poised to Decide Whether Training an AI Model on Copyrighted Content Is Fair Use* (June 29, 2026), https://www.stevenslee.com/appellate/third-circuit-poised-to-decide-whether-training-an-ai-model-on-copyrighted-content-is-fair-use/.
[29]Bartz v. Anthropic PBC, No. C 24-05417 WHA (N.D. Cal. June 23, 2025) (order on fair use).
[30]Bartz v. Anthropic PBC, No. 3:24-cv-05417-AMO (N.D. Cal. July 20, 2026) (order granting final approval of class settlement); *see* Haim Ravia & Dotan Hammer, *Federal Court Approves $1.5 Billion Anthropic Copyright Settlement*, Pearl Cohen (July 30, 2026), https://pearlcohen.com/federal-court-approves-1-5-billion-anthropic-copyright-settlement-largest-in-history.
[31]Kadrey v. Meta Platforms, Inc., No. 23-cv-03417-VC (N.D. Cal. June 25, 2025) (order on cross-motions for partial summary judgment).
[32]*See* Complaint, N.Y. Times Co. v. Microsoft Corp., No. 1:23-cv-11195 (S.D.N.Y. Dec. 27, 2023).
[33]Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on Copyright and

They must have a policy to comply with EU copyright law, including rights reservations under Article 4(3), and they must publish a sufficiently detailed summary of the content used for training.[34]

German courts have applied these rules. In *Kneschke v. LAION*, the Hamburg Regional Court held that a non-profit association's creation of an image and text dataset fell within the scientific research exception.[35] In *GEMA v. OpenAI*, the Munich Regional Court held that song lyrics memorised by OpenAI's models were reproduced in the models themselves, and that the text and data mining exception did not cover that reproduction.[36]

The United Kingdom permits text and data mining only for non-commercial research, under Section 29A of the Copyright, Designs and Patents Act 1988.[37] In *Getty Images v. Stability AI*, Getty dropped its main training claim because it could not show that training had taken place in the United Kingdom.[38] On the secondary infringement claim that remained, Joanna Smith J held that the Stable Diffusion model weights did not store Getty's works and so were not an infringing copy.[39]

C. Japan and Singapore

Article 30-4 of Japan's Copyright Act permits use of a work for data analysis and other purposes that do not involve enjoying the thoughts or sentiments expressed in it, except where the use would unreasonably prejudice the owner's interests.[40] Singapore's Copyright Act 2021 permits copying for computational data analysis where the user has lawful access to the work.[41] Neither exception is limited to non-commercial use.

D. The three-step test

India is bound by Article 9(2) of the Berne Convention and Article 13 of the TRIPS Agreement.

Related Rights in the Digital Single Market, arts. 3–4, 2019 O.J. (L 130) 92.
[34]Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence, art. 53(1)(c)–(d), 2024 O.J. (L 1689) 1.
[35]Landgericht Hamburg [LG] [Regional Court of Hamburg] Sept. 27, 2024, 310 O 227/23 (Ger.) (Kneschke v. LAION e.V.).
[36]Landgericht München I [LG] [Regional Court of Munich I] Nov. 11, 2025, 42 O 14139/24 (Ger.) (GEMA v. OpenAI).
[37]Copyright, Designs and Patents Act 1988, c. 48, § 29A (U.K.).
[38]Getty Images (US) Inc. v. Stability AI Ltd. [2025] EWHC 2863 (Ch) (Eng.).
[39]*Id.* at [600].
[40]Chosakukenhō [Copyright Act], Law No. 48 of 1970, art. 30-4 (Japan).
[41]Copyright Act 2021, § 244 (Sing.).

Under both, exceptions to exclusive rights must be confined to certain special cases, must not conflict with normal exploitation of the work, and must not unreasonably prejudice the legitimate interests of the right holder.[42] A broad exception for commercial training has to meet this test. Judge Bibas treated the licensing of training data as a relevant market.[43] As that market grows, it becomes harder to say that unlicensed training does not conflict with normal exploitation.

V. The DPIIT working paper

In April 2025, DPIIT set up an eight-member committee to examine whether the Copyright Act can deal with generative AI. On December 8, 2025, it published Part I of a working paper.[44] The committee found existing models, including a text and data mining exception with opt-out and extended collective licensing, unsuited to India, and rejected a zero-price licence as harmful to the incentive to create. It proposed a "hybrid model" in which AI developers receive a mandatory blanket licence to train on all lawfully accessed works, and right holders cannot withhold their works. Royalties become payable only when the AI system is commercialised. A government-appointed committee would set the rates, subject to judicial review, and a new central body would collect them.[45] According to one law firm's summary, the royalty would be a share of the system's global gross revenue, and the scheme would also apply to models already trained.[46] NASSCOM, a member of the committee, dissented and supported a text and data mining exception with opt-out.[47]

The proposal has real strengths. It removes the problem of individual owners holding out, gives developers legal certainty, and pays creators, which a fair dealing victory for developers in *ANI* would not do. Its lawful access condition also excludes pirated collections.

It also has four difficulties. The first is consent. Copyright is an exclusive right, and replacing the owner's choice with a compulsory licence for every work is a large step that must itself

[42]Berne Convention for the Protection of Literary and Artistic Works art. 9(2), Sept. 9, 1886, as revised July 24, 1971, 1161 U.N.T.S. 3; Agreement on Trade-Related Aspects of Intellectual Property Rights art. 13, Apr. 15, 1994, 1869 U.N.T.S. 299.
[43]*Thomson Reuters*, 765 F. Supp. 3d 382.
[44]Dep't for Promotion of Indus. & Internal Trade, Generative AI and Copyright, Part I: One Nation, One Licence, One Payment: Balancing AI and Innovation (Working Paper, Dec. 8, 2025) (India).
[45]*Id.*
[46]Ikigai Law, *Exploring the DPIIT's Working Paper on Generative AI and Copyright*, Ikigai Law (Dec. 2025), https://www.ikigailaw.com/article/655/exploring-the-dpiits-working-paper-on-generative-ai-and-copyright.
[47]Intepat, *AI Training and Copyright in India: What the DPIIT Hybrid Model Means for Builders*, Intepat Blog (2026), https://www.intepat.com/blog/ai-training-copyright-india-dpiit-hybrid-model.

satisfy the three-step test. The second is reach. A royalty on global revenue captures income that has nothing to do with Indian works. The third is the conflict with *ANI*. If training on freely available content is fair dealing, as the Delhi High Court held prima facie, there is nothing to license. A statutory licence would first need an amendment narrowing Section 52(1)(a). The fourth is distribution. Paying millions of authors from one pool requires data on which works were used, and that data will not exist unless developers are required to disclose it.

VI. Proposals

Indian law treats "training" as one question. It is several, and Parliament need not wait for the *ANI* appeal to answer them.

The first concerns where the data came from. The Explanation to Section 52(1)(a) should state that storage for training is protected only where the developer had lawful access to the work, and not where the copy came from a pirate library or was obtained by getting around a paywall. Section 65A does not do this work today. It punishes circumvention only when done with the intention of infringing, and it expressly permits circumvention for purposes the Act does not prohibit.[48] *Bartz* shows why the distinction matters. The same judge who upheld training as fair use treated the pirated library separately, and that part of the case led to a USD 1.5 billion settlement.

The second concerns opting out. India has no statutory method by which an owner can reserve works from text and data mining. The *ANI* court nonetheless counted ANI's failure to block the crawler against it. A statutory rule modelled on Article 4(3) of the EU Directive would give this practice a legal footing and tell owners exactly what they must do.

The third concerns memorisation and output. Where a model reproduces a substantial part of a work, as the Munich court found for song lyrics, that reproduction should be treated as infringing however the training copy was made. Developers should be required to test for memorisation and to filter outputs, measures Sag has proposed as part of a set of copyright safety practices.[49] ANI lost its output claim partly for want of proof, and most of that proof sits with the developer.

[48]The Copyright Act, 1957 § 65A(1), (2)(a).
[49]Matthew Sag, *Copyright Safety for Generative AI*, 61 Hous. L. Rev. 295 (2023).

The fourth concerns disclosure. Neither a fair dealing inquiry nor a royalty scheme works well without knowing what was used. A duty to publish a summary of training content, similar to the one in the AI Act, would help courts and would be a precondition for any scheme of the kind DPIIT has proposed.[50]

If Parliament does adopt a remuneration scheme, it should apply only to commercial models above a defined size, leave research and non-commercial models within the existing research exception, and let owners opt out. That design sits more easily with the three-step test.

VII. Conclusion

The Delhi High Court has given developers a prima facie answer for training on freely available content. It has not said what happens when the data is pirated, when the model memorises, or when the output competes with the source. Courts in the United States and Germany have decided cases on exactly those points.[51] Indian law will have to answer them next, whether through the *ANI* appeal, the DPIIT process or an amendment to Section 52.

[50] *See* Regulation (EU) 2024/1689, *supra* note 34, art. 53(1)(c); Directive (EU) 2019/790, *supra* note 33, art. 4(3).
[51] *See, e.g.*, *Bartz*, No. C 24-05417 WHA; *Kadrey*, No. 23-cv-03417-VC; *GEMA v. OpenAI*, 42 O 14139/24.

Bibliography

No.	Source	Type
1	Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith, 598 U.S. 508 (2023).	Cases
2	ANI Media Pvt. Ltd. v. Open AI OpCo LLC, I.A. 45300/2024 in CS(COMM) 1028/2024, 2026:DHC:5900 (Del. H.C. July 24, 2026) (India).	Cases
3	Authors Guild v. Google, Inc., 804 F.3d 202 (2d Cir. 2015).	Cases
4	Authors Guild, Inc. v. HathiTrust, 755 F.3d 87 (2d Cir. 2014).	Cases
5	Bartz v. Anthropic PBC, No. C 24-05417 WHA (N.D. Cal. June 23, 2025) (order on fair use).	Cases
6	Bartz v. Anthropic PBC, No. 3:24-cv-05417-AMO (N.D. Cal. July 20, 2026) (order granting final approval of class settlement).	Cases
7	Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569 (1994).	Cases
8	Chancellor, Masters & Scholars of the Univ. of Oxford v. Rameshwari Photocopy Servs., CS(OS) 2439/2012 (Del. H.C. Sept. 16, 2016) (India).	Cases
9	Civic Chandran v. Ammini Amma, 1996 PTR 142 (Ker. H.C.) (India).	Cases
10	Eastern Book Co. v. D.B. Modak, (2008) 1 SCC 1 (India).	Cases
11	Getty Images (US) Inc. v. Stability AI Ltd. [2025] EWHC 2863 (Ch) (Eng.).	Cases
12	Kadrey v. Meta Platforms, Inc., No. 23-cv-03417-VC (N.D. Cal. June 25, 2025).	Cases
13	Landgericht Hamburg [LG] [Regional Court of Hamburg] Sept. 27, 2024, 310 O 227/23 (Ger.) (Kneschke v. LAION e.V.).	Cases
14	Landgericht München I [LG] [Regional Court of Munich I] Nov. 11, 2025, 42 O 14139/24 (Ger.) (GEMA v. OpenAI).	Cases
15	N.Y. Times Co. v. Microsoft Corp., No. 1:23-cv-11195 (S.D.N.Y. filed Dec. 27, 2023).	Cases
16	Thomson Reuters Enter. Ctr. GmbH v. Ross Intel. Inc., 765 F. Supp. 3d 382 (D. Del. 2025), *appeal docketed*, No. 25-2153 (3d Cir.).	Cases
17	17 U.S.C. § 107.	Statutes, regulations and treaties

No.	Source	Type
18	Agreement on Trade-Related Aspects of Intellectual Property Rights, Apr. 15, 1994, 1869 U.N.T.S. 299.	Statutes, regulations and treaties
19	Berne Convention for the Protection of Literary and Artistic Works, Sept. 9, 1886, as revised July 24, 1971, 1161 U.N.T.S. 3.	Statutes, regulations and treaties
20	Chosakukenhō [Copyright Act], Law No. 48 of 1970 (Japan).	Statutes, regulations and treaties
21	Copyright Act 2021 (Sing.).	Statutes, regulations and treaties
22	Copyright, Designs and Patents Act 1988, c. 48 (U.K.).	Statutes, regulations and treaties
23	Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on Copyright and Related Rights in the Digital Single Market, 2019 O.J. (L 130) 92.	Statutes, regulations and treaties
24	Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence, 2024 O.J. (L 1689) 1.	Statutes, regulations and treaties
25	The Copyright Act, 1957, No. 14, Acts of Parliament, 1957 (India).	Statutes, regulations and treaties
26	The Copyright (Amendment) Act, 2012, No. 27, Acts of Parliament, 2012 (India).	Statutes, regulations and treaties
27	DEP'T FOR PROMOTION OF INDUS. & INTERNAL TRADE, GENERATIVE AI AND COPYRIGHT, PART I: ONE NATION, ONE LICENCE, ONE PAYMENT: BALANCING AI AND INNOVATION (Working Paper, Dec. 8, 2025) (India).	Government and institutional reports
28	NAT'L ASS'N OF SOFTWARE & SERV. COS. & BOSTON CONSULTING GRP., AI-POWERED TECH SERVICES: A ROADMAP FOR FUTURE-READY FIRMS (2024).	Government and institutional reports
29	U.S. COPYRIGHT OFF., COPYRIGHT AND ARTIFICIAL INTELLIGENCE, PART 3: GENERATIVE AI TRAINING (pre-publication version, May 2025).	Government and institutional reports
30	Benjamin L.W. Sobel, *Artificial Intelligence's Fair Use Crisis*, 41 COLUM. J.L. & ARTS 45 (2017).	Articles and conference papers

No.	Source	Type
31	James Grimmelmann, *Copyright for Literate Robots*, 101 IOWA L. REV. 657 (2016).	Articles and conference papers
32	Mark A. Lemley & Bryan Casey, *Fair Learning*, 99 TEX. L. REV. 743 (2021).	Articles and conference papers
33	Matthew Sag, *Copyright Safety for Generative AI*, 61 HOUS. L. REV. 295 (2023).	Articles and conference papers
34	Nicholas Carlini et al., *Extracting Training Data from Large Language Models*, *in* PROCEEDINGS OF THE 30TH USENIX SECURITY SYMPOSIUM 2633 (2021).	Articles and conference papers
35	*ANI Challenges Delhi High Court Order Denying Interim Relief Against OpenAI*, ADGULLY (Sept. 2026), https://www.adgully.com/post/20258/ani-challenges-delhi-high-court-order-denying-interim-relief-against-openai.	Online sources
36	Cyril Amarchand Mangaldas, *ANI v. Open AI: Delhi High Court Refuses Interim Injunction in Landmark AI Copyright Dispute*, CYRIL AMARCHAND MANGALDAS BLOG (July 2026), https://corporate.cyrilamarchandblogs.com/2026/07/ani-v-open-ai-delhi-high-court-refuses-interim-injunction-in-landmark-ai-copyright-dispute/.	Online sources
37	Haim Ravia & Dotan Hammer, *Federal Court Approves $1.5 Billion Anthropic Copyright Settlement*, PEARL COHEN (July 30, 2026), https://pearlcohen.com/federal-court-approves-1-5-billion-anthropic-copyright-settlement-largest-in-history.	Online sources
38	Ikigai Law, *Exploring the DPIIT's Working Paper on Generative AI and Copyright*, IKIGAI LAW (Dec. 2025), https://www.ikigailaw.com/article/655/exploring-the-dpiits-working-paper-on-generative-ai-and-copyright.	Online sources
39	Intepat, *AI Training and Copyright in India: What the DPIIT Hybrid Model Means for Builders*, INTEPAT BLOG (2026), https://www.intepat.com/blog/ai-training-copyright-india-dpiit-hybrid-model.	Online sources
40	Intepat, *ANI vs OpenAI Judgment: What the Delhi High Court Actually Held*, INTEPAT BLOG (2026), https://www.intepat.com/blog/ani-vs-openai-judgment-delhi-high-court.	Online sources

No.	Source	Type
41	LexCounsel, *India's First AI Copyright Ruling: The Delhi High Court in ANI v. OpenAI*, LEGAL 500 (2026), https://www.legal500.com/intelligence/india/intellectual-property/indias-first-ai-copyright-ruling-the-delhi-high-court-in-ani-v-openai.	Online sources
42	Stevens & Lee, *Third Circuit Poised to Decide Whether Training an AI Model on Copyrighted Content Is Fair Use* (June 29, 2026), https://www.stevenslee.com/appellate/third-circuit-poised-to-decide-whether-training-an-ai-model-on-copyrighted-content-is-fair-use/.	Online sources