



* IN THE HIGH COURT OF DELHI AT NEW DELHI

%

Judgment reserved on: 27th March, 2026

Judgment pronounced on: 24th July, 2026

I.A. 45300/2024

IN

+

CS(COMM) 1028/2024, I.A. 45301/2024 & I.A.26192/2025

ANI MEDIA PVT. LTD.

.....Plaintiff

Through: Mr. Sidhant Kumar, Mr. Akshit Mago,
Ms. Manyaa Chandok, Ms. Anshika
Saxena and Ms. Lahar Jain, Advocates

Versus

OPEN AI OPCO LLC

.....Defendant

Through: Mr. Amit Sibal, Senior Advocate with
Mr. Sanjeev Kapoor, Mr. Nirupam
Lodha, Mr. Madhav Khosla, Mr.
Gautam Wadhwa, Ms. Moha Paranjpe,
Mr. Abhi Udai Singh Gautam, Ms.
Rebecca Cardoso, Mr. Hardik Malik,
Ms. Vanshika Thapliyal, Mr. Rajat
Bector, Advocates
Mr. Akhil Sibal, Senior Advocate with
Mr. Aditya Gupta, Ms. Asavari Jain,
Mr. Shuvam Bhattacharya, Ms. Vani
Kaushik, Ms. Riddhie Bajaj, Ms.
Jahnvi Siddhu, Ms. Aishwarya Kane
and Mr. Sauhard Alung, Advocates for
IGAP/Intervenor in I.A. 4616/2025
Mr. Kapil Sibal, Senior Advocate and
Mr. Arvind P. Datar, Senior Advocate
with Mr. Shashank Mishra, Ms. Akshi
Rastogi, Mr. Parv Kaushik and Mr.
Suvaroop Saha Roy, Advocates for
Broadband India Forum



2026:DHC:5900



(BIF)/Intervenor in I.A. 10766/2025
Ms. Haripriya Padmanabhan, Senior Advocate with Mr. Shrutanjaya Bhardwaj, Mr. Akshat Agrawal, Mr. Tushar Srivastava, Mr. Shourya Das Gupta, Ms. Siddhi Nagwekar, Ms. Yashi Bajpai and Mr. Yash Tayal, Advocates for Flux AI Labs/
Intervenor in I.A. 4615/2025
Mr. Ankit Sahni, Ms. Kritika Sahni, Mr. Chirag Ahluwalia, Mr. Mohit Maru, Ms. Tanisha Sharma, Advocates for the Federation of Indian Publishers, Intervenor in I.A. No. 470 of 2025
Mr. Rajshekhar Rao, Senior Advocate and Mr. Chander M. Lall, Senior Advocate with Mr. Ameet Datta, Mr. Harsh Kaushik, Ms. Riddima Sharma, Mr. Akshay Nagarajan, Ms. Rishikaa, Ms. Gauri Khanna and Ms. Annanya Mehan, Advocates for DNPA/Intervenors in I.A. 2199/2025 & IMI/Intervenors in I.A. 4027/2025
Mr. Adarsh Ramanujan with Mr. Parth Singh, Advocate, *Amicus Curiae*
Professor Arul George Scaria, *Amicus Curiae*.

CORAM:
HON'BLE MR. JUSTICE AMIT BANSAL

JUDGMENT

AMIT BANSAL, J.

I.A. 45300/2024 (under Order XXXIX Rules 1 and 2 CPC)



INDEX

PREFACE -----	4
WORKING OF LLMS-----	8
A. TRAINING DATA -----	8
B. DATA PREPARATION AND MODEL TRAINING-----	9
C. POST-TRAINING-----	10
D. USER INTERACTION-----	10
E. RETRIEVAL-AUGMENTED GENERATION (RAG)-----	11
PROCEEDINGS IN THE SUIT-----	11
ISSUE NO.4	
WHETHER THE COURTS IN INDIA HAVE JURISDICTION TO ENTERTAIN THE PRESENT LAW SUIT CONSIDERING THAT THE SERVERS OF THE DEFENDANTS ARE LOCATED IN THE UNITED STATES OF AMERICA -----	13
SUBMISSIONS -----	13
DISCUSSION AND ANALYSIS -----	19
i. <i>Whether this court has territorial jurisdiction to entertain the present suit. -----</i>	<i>19</i>
ii. <i>Whether the Copyright Act, 1957 would apply in relation to training claim since according to Open AI training takes place on servers located outside India -----</i>	<i>20</i>
FINDINGS-----	22
ISSUE NO.2	
WHETHER THE USE BY THE DEFENDANTS OF PLAINTIFF'S COPYRIGHTED DATA IN ORDER TO GENERATE RESPONSES FOR ITS USERS, WOULD AMOUNT TO INFRINGEMENT OF THE PLAINTIFF'S COPYRIGHT. -----	23
SUBMISSIONS -----	23
DISCUSSION AND ANALYSIS -----	34
i. <i>Whether Open AI memorizes and regurgitates ANI's copyrighted literary works in the form of responses? -----</i>	<i>37</i>
ii. <i>Whether ChatGPT's responses are substantial reproduction of ANI's copyrighted literary works?-----</i>	<i>42</i>
FINDINGS-----	60
ISSUE NO.1	
WHETHER THE STORAGE BY THE DEFENDANTS OF PLAINTIFF'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-----	61
SUBMISSIONS -----	61
ISSUE NO.3	
WHETHER THE DEFENDANTS' USE OF PLAINTIFF'S COPYRIGHTED DATA QUALIFIES AS 'FAIR USE' IN TERMS OF SECTION 52 OF THE COPYRIGHT ACT, 1957. -----	68
SUBMISSIONS -----	68
DISCUSSION AND ANALYSIS OF ISSUE NOS. 1 AND 3 -----	86



2026:DHC:5900



A. Scope of Section 14(a)(i) of the Copyright Act, 1957	86
B. Legislative History of Section 52(1)(a)	92
C. Interpretation of Section 52(1)(a) of the Copyright Act	93
D. Scope of Section 52(1)(a)	95
I. PURPOSE TEST	96
a. Whether commercial use is entitled to protection under Section 52(1)(a)(i)	97
b. Requirement of “non-infringing copy” under Section 52(1)(a) of the Copyright Act.	104
c. Whether use by Open AI amounts to “private or personal use, including research”	108
II. FAIRNESS TEST	115
a. Whether Open AI’s use of ANI’s original literary works is limited to training its LLMs underlying ChatGPT?	123
b. Whether the usage of ANI’s literary works by Open AI would result in economic competition and would prejudice the legitimate interests of ANI, thereby causing actual or potential damage to ANI?	124
c. Whether the functions performed by Open AI through ChatGPT serve the overall public interest?	128
FINDINGS	129
BALANCE OF CONVENIENCE AND IRREPARABLE INJURY	129
SUBMISSIONS	129
DISCUSSION AND ANALYSIS	132
FINDINGS	134
CONCLUSION	134



PREFACE

1. Innovations affecting the works protected under the Copyright Act, 1957 (hereinafter the ‘Copyright Act’), namely literary, dramatic, musical and artistic works, cinematograph films, and sound recordings, have consistently given rise to interpretive challenges in the application of traditional legal provisions.
2. While tracing the history of copyright law, it is noteworthy that globally, copyright law and technological innovation have long existed in a dynamic relationship, traceable at least to the advent of the Gutenberg Press in the fifteenth century, well before the enactment of the Statute of Anne in 1710¹.
3. With time, successive technologies have evolved, enabling new forms of creative expression and widening avenues for dissemination, disrupting settled copyright arrangements. Some of these technologies worth referring to are player pianos, radio broadcasting, cable television, photocopying, home video cassette recorders, digital downloading and streaming, and, more recently, artificial intelligence (hereinafter ‘AI’). Therefore, technological advancement continues to generate novel questions that necessitate periodic re-examination of established copyright doctrine.
4. It is the judiciary, as the primary forum for adjudication, that is invariably called upon to address the nuanced legal questions arising from such disruptive technologies, often resulting in conflicts, as in the present

¹ In 1709 British parliament enacted the Statute of Anne; short title: Copyright Act 1709 8 Anne c.21; long title: *An Act for the Encouragement of Learning, by vesting the Copies of Printed Books in the Authors or purchasers of such Copies, during the Times therein mentioned*. Named after Anne, Queen of Great Britain, this was the first copyright statute in the Kingdom of Great Britain, and the first full-fledged copyright statute in the world. It was enacted in the regnal year 1709 to 1710, and entered into force on April 10, 1710.



case. Therefore, the Courts are required to step in to fill the gap between advancing technologies and existing laws.

5. AI is currently a technology of widespread interest, and India is considered a forerunner in this field. In response to international AI competition, Sarvam AI of India launched its first domestic Large Language Model (hereinafter ‘LLM’) on 18th February 2026². Research and development of similar LLMs is also underway in India to make India an AI hub. Data is the oil for LLMs to work efficiently.

6. Procuring publicly available data by AI innovators for training LLMs without the permission of the respective copyright holders potentially will raise legal issues, as arising in the present suit. The present suit has been filed by the plaintiff (hereinafter ‘ANI’) against the defendant (hereinafter ‘Open AI’) alleging unauthorised use of ANI’s copyrighted works by Open AI. ANI claims copyright infringement on two counts:

- i. Training claim - copying and storage of ANI’s data to train Open AI’s LLM; and
- ii. Reproduction or Output claim - reproduction of ANI’s works in the output produced by Open AI’s LLM through its application ChatGPT.

7. This case is a perfect illustration of how the law must continually catch up with technology. The case has to be adjudicated on the anvil of the Copyright Act, originally enacted in 1957, though amended from time to time. Clearly, when the Copyright Act or any of its amendments were enacted, the legislature could not have imagined the vast technological advances in the

²Aroon Deep, *Bengaluru firm unveils two AI language models*, The Hindu (Feb. 18, 2026), <https://www.thehindu.com/sci-tech/technology/bengaluru-firm-unveils-two-ai-language-models/article70648438.ece>, accessed on 20th February 2026.



field of AI.

8. In these uncharted areas, Courts often take aid from judgments of foreign Courts on similar issues. On various occasions, while dealing with Intellectual Property laws, the Indian Courts have recognized the importance of global doctrinal interdependence. The Courts laid emphasis on harmonization of basic principles of laws in foreign jurisprudence with the national laws, especially when the national laws do not indicate a contrary intent³. However, while doing so, the Courts must be mindful that the underlying legislation in foreign jurisdictions may be quite different from that prevailing in India.

9. For the purposes of adjudication of the present application for interim injunction filed by ANI, this Court had framed issues on 19th November, 2024, which are set out below:

- I. *Whether the storage by the defendants of plaintiff'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.*
- II. *Whether the use by the defendants of plaintiff's copyrighted data in order to generate responses for its users, would amount to infringement of the plaintiff's copyright.*
- III. *Whether the defendants' use of plaintiff's copyrighted data qualifies as 'fair use' in terms of Section 52 of the Copyright Act, 1957.*
- IV. *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."*

Even though Issue no. 3 refers to '*fair use*' [as it was framed on the first date of hearing] it should be read as "*fair dealing*" which is the expression used in Section 52(1)(a) of the Copyright Act.

³ Intex Technologies v. Telefonaktiebolaget LM Ericsson, 2023:DHC:2243-DB.



10. Since the question of this Court's jurisdiction to adjudicate the present application is foundational to the determination of the remaining issues, the Court shall first consider Issue No. 4. Thereafter, in view of its broader bearing on Issue Nos. 1 and 3, the Court shall examine Issue No. 2. As the *Amici Curiae* and counsel for the parties were *ad idem* that Issue Nos. 1 and 3 are intertwined, particularly in light of Open AI's plea of fair dealing in response to the allegations relating to storage, the said issues shall be considered together, subsequently.

11. With the aforesaid backdrop, I proceed to decide the present application.

WORKING OF LLMs

12. It is necessary to understand the functioning of LLMs for the purpose of adjudicating the present application, as the issues identified herein arise from, and are intrinsically connected with, the manner in which LLMs operate. Accordingly, it is apposite to first examine the technical functioning of LLMs.

13. Based on the submissions of the parties and learned *Amici Curiae* and also from the common general knowledge available, the steps involved in the working of LLMs, may be summarised in the following paragraphs.

A. Training Data

14. A large language model is a computational system that can predict the next word from previous words. That is, given a context or prefix of words, a language model assigns a probability distribution over the possible next words.

15. For enabling a model to predict subsequent words, the model is first



trained on an enormous text corpus with datasets of hundreds of billions of words, generally scraped from the web, which is referred to as the ‘*raw data*’.

16. The ‘*raw data*’ used for training LLMs would include automatically scraped data from publicly available sources on the internet, as well as licensed materials. The web text is usually taken from corpora of automatically crawled web pages, such as the Common Crawl. The data comprises materials, including but not limited to books, articles, blog posts, etc., so that, based on the aforesaid data, the model can make predictions.

17. Pre-training data collected from the web is filtered for both quality and safety. Quality filtering uses automated tools to score documents based on their usefulness and reliability, often by comparing them with trusted sources such as books and reputable websites, while avoiding content containing personal information or adult material. This generally improves language model performance. Safety filtering aims to remove harmful or toxic content, typically using existing toxicity-detection tools.⁴

18. Website owners can indicate that they do not want their sites crawled by web crawlers.⁵

B. Data preparation and Model training

19. The raw data is divided into smaller pieces called tokens, each assigned a ‘*Token ID*’, which is just a number. This is called tokenisation. This process is the first step in training most of the LLMs.⁶

⁴ Matthew Sag and Peter K. Yu, ‘The Globalization of Copyright Exceptions for AI Training’ (Emory Law Journal, 2025, Forthcoming) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4976393.

⁵ Daniel Jurafsky and James H. Martin. 2026. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models, 3rd edition. Online manuscript released January 6, 2026. <https://web.stanford.edu/~jurafsky/slp3>.

⁶ Daniel J. Gervais et al., ‘The Heart of the Matter: Copyright, AI Training, and LLMs’ (2024) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4963711.



20. Thereafter, each token is converted to a vector *via* an ‘*embedding*’ process. ‘*Vectors*’ are a numerical representation of semantic information about a ‘*token*’.

21. In the training exercise, the LLM is repeatedly fed parts of the training data where a word or a sequence of words is masked, and it is made to predict the next words or sentences, which are then compared with the original ‘*raw text*’ to verify correctness. The initial predictions may not be accurate, but with each repetition, the model readjusts and refines its prediction to reach the most probable statistical output in each scenario.

22. The choice of which word to generate in LLMs is done by sampling from the distribution of possible next words.

C. Post-training

23. Thereafter, the model is again fine-tuned to reduce prediction errors and to follow instructions better, for example, to answer questions, give summaries, write code, translate sentences, and so on. It does this by being trained on a special corpus containing many texts with both instructions and their correct responses.

24. The model is then adjusted to give more useful responses and to avoid producing harmful or offensive content.

D. User interaction

25. A user instructs LLM *via* a prompt, a text string that a user issues to a language model to get the model to do something useful by iteratively generating tokens conditioned on the prompt. The process of finding effective prompts for a task is known as prompt engineering.⁷

⁷ Mark Lemley, ‘How Generative AI Turns Copyright Law Upside Down’ (2024) 25(2) Sci. & Tech. LR 36.



E. Retrieval-Augmented Generation (RAG)

26. Retrieval-Augmented Generation ('RAG') is an innovative feature of LLMs that optimises their output. This feature references an authoritative knowledge base outside of its training data sources before generating a response. This feature does not predominantly rely on training data to generate output. Instead, it uses an information retrieval component that utilises the user prompt to pull information from external data storage.⁸ The LLM uses both training data and external information to generate better responses *via* RAG.

27. For the sake of convenience, '*LLM(s)*' and '*ChatGPT*' has been used interchangeably in the judgment and must be read likewise. When reference is made to '*ChatGPT search*' function, it has been specifically mentioned in the judgment as such.

PROCEEDINGS IN THE SUIT

28. The present suit, accompanied by an application seeking interim injunction and other ancillary applications, was first listed before this Court on 19th November 2024. On the same date, four issues were framed and the Court appointed two *Amici Curiae* to assist the Court.

29. Thereafter, multiple intervention applications were filed in the present suit, the details of which are given below:

29.1. **I.A. 470/2025** has been filed on behalf of **Federation of Indian Publishers**, a national representative body of the Indian publishing industry.

⁸ *What is RAG? - Retrieval-Augmented Generation AI Explained - AWS*, <https://aws.amazon.com/what-is/retrieval-augmented-generation/>.



- 29.2. **I.A. 2199/2025** has been filed on behalf of **Digital News Publishers Association** and its members, which are digital news platforms.
- 29.3. **I.A. 4027/2025** has been filed on behalf of the **Indian Music Industry**, which is an organization established to defend, preserve and develop the rights of the Phonogram Producers.
- 29.4. **I.A. 4615/2025** has been filed on behalf of **Flux Labs AI Private Ltd.**, a Generative Artificial Intelligence enabled start-up company.
- 29.5. **I.A. 4616/2025** has been filed on behalf of **IGAP Project LLP**, a think tank focusing on the intersection of law, policy, and emerging technologies, including AI.
- 29.6. **I.A. 10766/2025** has been filed on behalf of **Broadband India Forum**, an independent policy forum and think tank.
30. The intervenors in I.A. 470/2025, I.A. 2199/2025 and I.A. 4027/2025 are broadly supporting the case of ANI and the intervenors in I.A. 4615/2025, I.A. 4616/2025 and I.A. 10766/2025 are broadly supporting the case of Open AI.
31. On 4th July 2025, this Court allowed the aforesaid intervention applications and permitted the intervenors to make submissions limited to the legal issues framed on 19th November 2024.
32. ANI, Open AI as well as the intervenors have filed their written submissions along with judgments in support.
33. Submissions on behalf of the parties were heard on 21st February 2025, 10th March 2025, 18th March 2025, 28th March 2025, 2nd April 2025, 22nd April 2025, 29th April 2025, 4th July 2025, 23rd July 2025, 30th July 2025, 5th August



2025, 18th August 2025, 12th September 2025, 23rd September 2025, 17th October 2025, 7th November 2025, 21st November 2025, 23rd January 2026, 20th February 2026, 20th March 2026 and 27th March 2026, when the judgment was reserved and the parties were given liberty to file their closing submissions.

34. ANI, Open AI and the two *Amici Curiae* have filed their concluding written submissions.

34.1 The concluding written submissions dated 2nd April 2026 sent on behalf of Professor Arul George Scaria *via* email to the Court Master, have also been taken on record.

35. With the aforesaid backdrop, I proceed with my issues-wise analysis.

Issue No.4 ***Whether the Courts in India have jurisdiction to entertain the present law suit considering that the servers of the defendants are located in the United States of America***

SUBMISSIONS

Submissions on behalf of the Defendant (OPEN AI)⁹

36. Mr. Amit Sibal, senior counsel appearing on behalf of Open AI made the following submissions:

36.1. ANI has joined two separate causes of action in the present suit i.e. - (i) use of publicly available data by Open AI on servers located outside India to train LLMs underlying ChatGPT (*'training claim'*) and, (ii) the responses received by ANI from ChatGPT allegedly similar to content over which ANI holds copyright (*'reproduction claim'*). Therefore, ANI has to establish

⁹ Since the issue of jurisdiction has been raised by the defendant, it is deemed appropriate that the submissions of the defendant are recorded first.



jurisdiction *qua* both causes of action.

36.2. Insofar as the training claim is concerned, it is an undisputed position that the training takes place outside India on the basis of data that is stored in servers outside India. Therefore, merely because the jurisdiction can be established in respect of the reproduction claim, this Court would not have jurisdiction in respect of training claim.

36.3. The Copyright Act and in particular Section 62 thereof read with Section 1(2) of the Copyright Act, confers jurisdiction only within the territory of India. The Copyright Act was not intended to have extra-territorial jurisdiction.

36.4. The reliance placed by Professor Scaria on the judgment of a Coordinate Bench in *Neetu Singh v. Telegram*¹⁰ is misplaced insofar as the jurisdiction in respect of training claim is concerned. In *Neetu Singh* (supra), the infringement occurred within the jurisdiction of this Court and in that context, the Court observed that it would be immaterial if telegram servers are located outside India. Similarly, reliance placed on the judgment of the Supreme Court in *Exphar SA v. Eupharma Laboratories*¹¹ and the judgment of Division Bench in *Blueberry Books v. Google India*¹² is misplaced.

36.5. The test laid down by the judgment of the Division Bench in *Banyan Tree Holding v. A. Murali Krishna Reddy*¹³ for the purposes of jurisdiction is limited. The purposeful availment test as laid down in *Banyan Tree* (supra) cannot be used to confer jurisdiction in respect of causes of action occurring outside India.

¹⁰ 2022 SCC OnLine Del 2637.

¹¹ (2004) 3 SCC 688.

¹² 2016 SCC Online Del 3338.

¹³ 2009 SCC OnLine Del 3780.



36.6. Insofar as the reproduction claim is concerned, at the time of using the ChatGPT service, the user agrees to '*terms of use*' of ChatGPT, which provides jurisdiction of the Courts in San Francisco, California.

Submissions on behalf of the Plaintiff (ANI)

37. Mr. Sidhant Kumar, counsel appearing on behalf of ANI made the following submissions:

37.1. In terms of Section 62(2) of the Copyright Act, copyright owners can file suit for infringement in the jurisdiction where they reside or conduct business. Therefore, the physical location of Open AI's servers is of no consequence. ANI's principal and registered office is located within the jurisdiction of this Court.

37.2. Even in terms of Section 20(b) of the Code of Civil Procedure, 1908 (hereinafter '*CPC*'), since Open AI operates an interactive website that enables commercial transactions within the jurisdiction of this Court, i.e. Open AI collects subscription fee from users located in this jurisdiction, Open AI would be amenable to jurisdiction of this Court.

37.3. Open AI actively attracts the consumers through advertising and public outreach efforts in India, that are within the jurisdiction of this Court. Purposeful engagement of ANI's consumers and a targeted marketing sufficiently demonstrate compliance of Section 20(c) of the CPC.

37.4. The various infringement activities alleged in the plaint have taken place within the jurisdiction of this Court. The disputes raised in the present suit are independent of '*terms of use*' of ChatGPT. The acts that constitute copyright infringement are independent of any contractual agreements between the users and Open AI. The statutory remedies such as those provided



under Section 62 of the Copyright Act cannot be waived through contractual stipulations.

Intervenors supporting ANI

i. Submissions on behalf of the Intervenor – DNPA (I.A. 2199/2025)

38. Mr. Rajshekhar Rao, senior counsel appearing on behalf of intervenors in I.A. 2199/2025 made the following submissions:

38.1. The copies/reproduction of Indian works which are first published in India and thereafter made available by Open AI to attract Indian users, would imply that at the very least part of the cause of action arises within the jurisdiction of this Court. Therefore, in terms of Section 62 of the Copyright Act, read with judgment of the Supreme Court in **Indian Performing Rights Society v. Sanjay Dalia**¹⁴, this Court would have the jurisdiction to entertain the present suit.

38.2. Open AI is entering into commercial transactions within the jurisdiction of this Court and specifically targeting the customers. Reliance is placed on **Banyan Tree** (supra).

38.3. Neither the ‘terms of use’ of Open AI’s ChatGPT, nor the fact that the servers located outside India, would oust the jurisdiction of this Court under Section 62 of the Copyright Act as well as Section 20 of the CPC.

ii. Submissions on behalf of the Intervenor - Indian Music Industry (I.A. 4027/2025)

39. Mr. C.M. Lall, senior counsel appearing on behalf of intervenors in I.A. 4027/2025 made the following submissions:

¹⁴ (2015) 10 SCC 161.



39.1. Jurisdiction is also made out on the parameters of ‘*purposeful availment*’ as laid down in ***Banyan Tree*** (supra). In the present case, availment can be demonstrated since Open AI has ‘*paid subscribers*’ of its services within the jurisdiction of this Court and the said subscribers and users interact with the service within the jurisdiction.

iii. Written Submissions filed by Mr. Ankit Sahni, Advocate on behalf of the Intervenor – Federation of Indian Publishers (FIP) (I.A. 470/2025)

40. This Court would have jurisdiction under Section 62 of the Copyright Act. In this regard, reliance has been placed on the judgment of the Supreme Court in ***Exphar***¹⁵ (supra) and ***Dabur India Ltd. v. K.R. Industries***¹⁶ to submit that Section 62 of the Copyright Act was inserted by the Parliament to prescribe an additional ground for attracting jurisdiction over and above Section 20 of CPC.

40.1 Since Open AI’s services are accessible within the jurisdiction of this Court and the outputs generated by Open AI’s ChatGPT application infringes the copyright of ANI, a substantial part of cause of action arises within the jurisdiction of this Court to invoke Section 20 of CPC.

Submissions of Mr. Adarsh Ramanujan, Amicus Curiae

41. Mr. Ramanujan made the following submissions:

41.1. The location of Open AI’s servers outside India is irrelevant as ANI has invoked Section 62 of the Copyright Act. There is no challenge by Open AI to the assertion of ANI that its principal place of business is within the jurisdiction of this Court.

¹⁵ ¶ 13.

¹⁶ (2008) 10 SCC 595, ¶ 32.



41.2. Insofar as the submission of Open AI is concerned that the infringement relating to the training claim occurs outside the jurisdiction of this Court, what has to be seen is that the training process involves access of data from India and transmission of the said data from India. Therefore, it cannot be said that just because the training occurs outside the jurisdiction of this Court, this Court would not have jurisdiction. The training claim has to be seen in a holistic manner which would necessarily involve access and transmission of data from India before the same is stored overseas.

Submissions of Professor Arul George Scaria, *Amicus Curiae*

42. Prof. Scaria made the following submissions:

42.1. There is no dispute raised to the fact that ANI has its principal office in New Delhi. Therefore, this Court would have jurisdiction under Section 62(2) of the Copyright Act.

42.2. Since Open AI is making its active services available to the users in India, including those in Delhi, the Courts in India would have jurisdiction in terms of Section 20 of the CPC.

42.3. Reliance is placed on the judgment of the Supreme Court in *Exphar* (supra) and the judgment passed by the Division Bench of this Court in *Blueberry Books* (supra) and the judgment of a Coordinate Bench in *Neetu Singh* (supra).

42.4. Insofar as aspect of the territoriality of Indian Copyright Act is concerned, he supports the contention of Open AI that since the entire training process occurs outside India merely because Open AI accesses data from India would not make the infringement subject to Indian Copyright Act.



42.5. In *Getty Images v. Stability AI*¹⁷ the plaintiff in the United Kingdom ('UK') did not press the training claim as the training took place outside UK.

DISCUSSION AND ANALYSIS

43. As discussed earlier, since the issue of jurisdiction goes to the root of the matter, it would be taken for consideration at the beginning.

44. On the aspect of jurisdiction, based on the objections raised by Open AI, essentially the following two issues arise for consideration:

- i. Whether this Court has territorial jurisdiction to entertain the present suit.
 - ii. Whether the Copyright Act would apply in relation to training claim since according to Open AI training takes place on servers located outside India.
- i. ***Whether this Court has territorial jurisdiction to entertain the present suit.***

45. Insofar as the territorial jurisdiction of this Court to entertain the present suit is concerned, it is an admitted position that ANI has its principal place of business as well as its registered office within the jurisdiction of this Court. Therefore, in terms of Section 62(2) of the Copyright Act, this Court would have the territorial jurisdiction to entertain the present suit.

46. It is also an admitted position that Open AI specifically targets/offers its services to users/subscribers across India, including those located within the jurisdiction of this Court.

47. Even alleged infringing activities mentioned in the plaint, i.e. responses

¹⁷ [2025] EWHC 38 (Ch).



generated by Open AI in India based on the prompts by ANI, took place within the jurisdiction of this Court.

48. Accordingly, at this *prima facie* stage, I hold that this Court would have the territorial jurisdiction to entertain the present suit in terms of Section 20 of the CPC as well as Section 62(2) of the Copyright Act.

ii. ***Whether the Copyright Act would apply in relation to training claim since according to Open AI training takes place on servers located outside India***

49. It has been argued on behalf of Open AI that its LLMs are trained outside India and the training data is stored and processed in the servers located in the United States of America (hereinafter ‘US’). On behalf of Open AI, it is contended that the Indian Copyright Act would not apply outside the territory of India. In support of its contention, reliance is placed on the judgment of the Division Bench in ***Blueberry Books*** (supra). In the said case, Amazon, which was a defendant in the suit took the defence that the infringing act took place exclusively in the US. Though Amazon admitted that it has business/commercial activities in India, the alleged infringing activity did not occur within the jurisdiction of this Court.

50. The Division Bench, in the said case, upheld the territorial jurisdiction of this Court. Insofar as the argument of Amazon with regard to the infringing activity taking place in the US was concerned, the Division Bench did not uphold the contention of Amazon. Paragraph 27 of ***Blueberry Books*** (supra) is set out below:

“27. As far as inapplicability of the provisions of the Act to infringements within the USA are concerned, we note that the submission is based on Amazon’s arguments to that effect. Its unilateral



assertion that there are downloading restrictions, based on a bare reading of the documents - or that only 33 downloads took place, for a small amount, are not relevant; they are contentious and accepting them would be giving disproportionate weight to the pleadings of one party and precluding the rights of the plaintiff, finally.”

51. In my view, the storing of ANI’s works on the US servers is a terminal step in the chain of events which begin from access of copyrighted works from India and transmission of the same abroad. There can be no doubt that Copyright Act cannot have jurisdiction outside the territory of India. However, the Copyright Act does not require this Court to sever the chain of events and examine only the last step. On a *prima facie* view, Open AI’s argument is untenable as it would lead to evasion of Indian copyright law by infringers who shift the terminal link to servers abroad.

52. In *Neetu Singh* (supra), a Coordinate Bench of this Court has observed that merely because Telegram chooses to locate its servers in Singapore, that would not affect the right of the copyright owners, to sue Telegram in India. The relevant observations from the aforesaid judgment are set out below:

“46. In view of the above factual and legal position, in the opinion of this Court, merely because Telegram chooses to locate its server in Singapore, the same cannot result in the Plaintiffs’ - who are copyright owners of course materials - being left completely remediless against the actual infringers, especially in order to claim damages and avail of other legal remedies in accordance with law. If such an argument is accepted, in the current world where most dissemination happens through online messaging services and platforms, IP violations would go completely unchecked. This cannot be the intention of law. The provisions of the IT Act and the Rules made therein have to be construed harmoniously with the rights and remedies provided to the copyright owners under the Copyright Act. Indian Courts are competent to decide issues relating to infringement of copyright and the mere fact that Telegram is operating a messaging service in India which chooses not to locate its servers in India cannot



divest the Indian Courts from dealing with copyright disputes or divest copyright owners from availing their remedies in Indian Courts. In the present age of cloud computing and diminishing national boundaries in data storage, conventional concepts of territoriality cannot be strictly applied. The dynamic evolution of law is essential to ensure appropriate remedies in case of violation of copyright and other IP laws.”

[Emphasis supplied]

53. The aforesaid reasoning would be fully applicable in the present case. In *Neetu Singh* (supra), even though the data was stored in servers outside India, the same was being circulated within the jurisdiction of this Court.

54. In my view, the act of Open AI in scraping of ANI’s copyrighted material and training of its LLM cannot be entirely divorced from the output claim. Since the output claim is based on training of the LLM, and the output is reproduced within the jurisdiction of this Court, I am not inclined to separate the two causes of action.

FINDINGS

55. On a *prima facie* view, it cannot be said at this stage that this Court does not have the territorial jurisdiction to entertain the present suit or that the alleged infringement on account of training would involve extra-territorial application of the Indian Copyright Act.

56. As discussed earlier, next, I would take up Issue No. 2 for consideration.



Issue No.2

Whether the use by the defendants of Plaintiff's copyrighted data in order to generate responses for its users, would amount to infringement of the Plaintiff's copyright.

SUBMISSIONS

Submissions on behalf of the Plaintiff (ANI)

57. Mr. Sidhant Kumar, counsel appearing on behalf of ANI made the following submissions:

57.1. LLMs generate responses by matching user input with the vector database and predicting the token sequences. This is the process of 'de-tokenization'. Once a response is formed, detokenization converts tokens into readable text by reassembling words, phrases, and sentences. Detokenization reconstructs raw data into its original textual form, converting training data (including ANI's works) back into a format readable by humans. Therefore, the whole process amounts to reproduction of raw data.

57.2. ANI has provided instances of ChatGPT's unauthorized reproduction of its works while responding to the ChatGPT's users, in paragraph 15 of the interim application. ANI claims these examples are illustrative and not exhaustive.

57.3. ChatGPT's output retains the original and creative elements of the ANI's text and thereby violates ANI's proprietary rights. Some of the responses demonstrate exact or nearly exact copies of ANI's data. This occurs due to memorization of the raw data/ training data.

57.4. The outputs generated by ChatGPT are not derivative literary works either, as it requires the author to exercise independent skill, labour, and



resources [Ref. *Eastern Book Company v. D.B. Modak*¹⁸]. In the present case, the user generated responses show that there is no independent application of mind. At best, the user generated responses can be considered adaptations of ANI's works as these responses are merely a re-arrangement of ANI's works.

57.5. Open AI itself admits that verbatim reproduction of content from the raw data occurs in cases where the same content occurs repeatedly in the raw data, through multiple sources on the internet. The very nature of ANI's business model makes its copyrighted works available on various platforms including its subscribers' websites, which are leading news publications.

57.6. The defence of Open AI that ANI's works are available in the public domain and can be used for training Open AI's LLM without obtaining a license from ANI, is not valid and strikes at the root of copyright law. Public availability does not nullify copyright or grant a universal license for unrestricted use. ANI is not divested of its exclusive rights in its works only because it has licensed its content to its subscribers.

57.7. Open AI has recognized similar proprietary rights of other news organizations and entered into license agreements with other news organizations such as Financial Times, Associated Press and Condé Nast, and has acknowledged the copyright of these entities.

57.8. Open AI has also infringed ANI's exclusive right to translate the works, for which ANI invests significant time and effort. In this regard, reliance is placed on the English version of the interview given by Mr. Neeraj Chopra's mother, which was originally in Hindi and was translated into English

¹⁸ (2008) 1 SCC 1.



by ANI.

57.9. In his rejoinder submissions, counsel for ANI placed reliance on the following foreign judgments:

- (i) In *The Associated Press v. Meltwater U.S. Holdings, Inc.*¹⁹, the US District Court held that summarizing plaintiff's works using computer programme without adding any commentary is not transformative and hence, it does not constitute fair use under the US Copyright Act.
- (ii) In *GEMA v. Open AI*²⁰, the Munich Regional Court granted an injunctive relief in favour of GEMA and against Open AI holding that there is memorization of training data and consequential reproduction.
- (iii) In *The Advance Local Media LLC et al. v. Cohere Inc.*²¹, the US District Court denied Cohere's partial motion to dismiss, holding that substitutive summaries could constitute copyright infringement, and that the generation of hallucinated articles falsely attributed to publishers/plaintiffs constituted a classic passing-off claim.

Intervenors supporting ANI

i. Submissions on behalf of the Intervenor – DNPA (I.A. 2199/2025)

58. Mr. Rajshekhar Rao, senior counsel appearing on behalf of intervenors in I.A. 2199/2025 made the following submissions:

58.1. The LLM models inherently memorize content. This can be implied

¹⁹ 931 F. Supp. 2d 537 (S.D.N.Y. 2013).

²⁰ case no. 42 O 14139/24, decided on 11 November 2025.

²¹ Advance Local Media LLC v. Cohere Inc., 1:25-cv-01305, (S.D.N.Y.).



from the fact that LLMs can reproduce verbatim copies of unique forms of content on which they are trained.

58.2. The reproduction of copies of training data implies that LLMs retain such data. The submission made on behalf of Open AI that LLMs only keep the non-expressive components of copyrighted works is incorrect. It is also incorrect to suggest that LLM models only learn rules of grammar and syntax.

58.3. The Copyright Act does not distinguish between ‘*expressive*’ and ‘*non-expressive*’ uses for the purpose of copyright infringement.

ii. *Submissions on behalf of the Intervenor - Indian Music Industry (I.A. 4027/2025)*

59. Mr. C.M. Lall, senior counsel appearing on behalf of intervenors in I.A. 4027/2025 made the following submissions:

59.1. Where the output amounts to substantial reproduction of original work, the same would amount to copyright infringement. It is well-established that LLM models retain copies of their training data and this is apparent from the admission made by Open AI with regard to ‘*regurgitation*’.

59.2. Accordingly, regurgitation or extraction of training data from the LLM model, which results in a response/ output to a user prompt being a substantial reproduction of the original work, will constitute infringement.

iii. *Written Submissions filed by Mr. Ankit Sahni, Advocate on behalf of the Intervenor – Federation of Indian Publishers (FIP) (I.A. 470/2025)*

60. ChatGPT regurgitates works on which it has been trained in the outputs/responses generated for its users and this would amount to infringement under the Copyright Act as the outputs/responses are a substantial reproduction of the copyrighted works.



Submissions on behalf of the Defendant (OPEN AI)

61. Mr. Amit Sibal, senior counsel appearing on behalf of Open AI made the following submissions:

61.1. The cut-off date for training an LLM is several months prior to the launch of a model. In respect of the models available to ANI in the lead up to the present suit, the cut-off dates for training were either April 2022 (for GPT 4) or April 2024 (for GPT 4o). Each of the 14 (fourteen) works that ANI alleges have been infringed in the plaint were admittedly authored and published subsequent to April 2024 (in either August or September 2024). It is therefore undisputed that the training data on which Open AI's LLMs were trained did not contain the 14 (fourteen) works that ANI alleges were infringed.

61.2. ANI has not identified any other works that it claims copyright over. Therefore, ANI cannot seek a speculative finding of infringement over unidentified works.

61.3. There is no exclusive right to 'use' any works under Section 14 of the Copyright Act. Therefore, ANI's allegation regarding 'use' of ANI's work by Open AI cannot amount to infringement. The LLMs underlying ChatGPT do not store or reproduce training data. Nor do they access data on which they were trained once the pre-training phase is complete. The only 'use' for which this data is deployed is during the training of Open AI's LLMs underlying ChatGPT.

61.4. ChatGPT is not designed to reproduce the extracts of content forming a part of its training data, but may use information or facts that it has '*learnt*' from such content in the course of its training towards responses to the user



prompts. Therefore, ChatGPT cannot be said to ‘*duplicate*’ or ‘*reproduce*’ the ‘*manner of presentation*’ of content on which the models underlying ChatGPT have been trained.

61.5. It is the settled position of law that copyright holder’s rights are limited to the protection of the manner in which an idea or a fact is expressed. However, there is no such protection on an underlying idea or fact. This is especially true in the context of news, where the underlying fact is incapable of copyright protection. In fact, there is a predominant public interest in the wide dissemination of such facts. Hence, ANI cannot claim monopoly of use over the said facts.

61.6. Copyright in news and other factual literary content is limited only to the expression thereof. In the case of ‘*news*’, the threshold for establishing substantial similarity in the expression is higher. [Ref: ***Feist Publications v. Rural Telephone Service***²²].

61.7. Grammatical rules and syntax by themselves cannot be protected by copyright because they are basic rules for sentence formation.

61.8. Indian copyright law explicitly rejects the ‘*sweat of the brow*’ doctrine and has adopted the ‘*skill and judgment test*’ to determine whether a work would be entitled to protection under Indian copyright law. [Ref. ***Eastern Book Company v. D.B. Modak*** (supra)]

61.9. ANI has sought to present limited extracts of a particular work rather than the whole work to allege infringement. For the purpose of determining copyright infringement, it is not permissible to split or compartmentalize or

²² 499 U.S. 340 (1991).



dissect ANI's work into parts, then seek to compare such parts against the alleged infringed work. [Ref. ***Star India v. Leo Burnett***²³]

61.10. ANI has not even been able to demonstrate substantial similarity, let alone verbatim reproduction in respect of the limited instances that have been cited in the plaint. This is despite the fact that ANI made extreme efforts to try and obtain a verbatim/ substantially similar response from ChatGPT through a series of adversarial prompts.

61.11. An infringing copy must be a substantial or material reproduction. Where similarities in the two works exist alongside broad dissimilarities, which negate the intention to copy the original and the coincidences appearing in the two works are clearly incidental, no infringement of the copyright comes into existence. [Ref. ***R.G. Anand v. Deluxe Films***²⁴; ***Bikramjeet Singh Bhullar v. Yash Raj Films***²⁵]

61.12. Different ChatGPT users receive different responses based on the same prompts.

61.13. Even in respect of ANI's allegations that the ChatGPT search feature reproduces results from ANI's website, it is submitted that a mere perusal of the alleged response(s) relied upon by ANI makes it apparent that the same are neither identical nor substantially similar to the corresponding articles relied upon by ANI. Further, in such a case, ChatGPT provides a reference/ link to the source but does not reproduce content therefrom.

²³ 2002 SCC OnLine Bom 942.

²⁴ (1978) 4 SCC 118.

²⁵ (2023) SCC OnLine Del 8212.



61.14. ‘*Regurgitation*’ of data on which the LLM models are trained, is a rare glitch that is being actively addressed by Open AI. ANI has not been able to identify any specific instance where ‘*regurgitation*’ of their works occurred.

Intervenors supporting OPEN AI

i. *Submissions on behalf of the Intervenor - Indian Governance and Policy Project (IGAP) (I.A. 4616/2025)*

62. Mr. Akhil Sibal, senior counsel appearing on behalf of the intervenor in I.A. 4616/2025 made the following submissions:

62.1. In ***Andrea Bartz v. Anthropic***²⁶, the US District Court emphasized that what is relevant is not the amount of data used for making a copy, but the amount and substantiality of what is made accessible to the public by the secondary use. Since LLMs do not regurgitate an author’s work or produce substantially similar works, it would not amount to copyright infringement.

62.2. An LLM model does not store the training data and memorization of training data is a rare occurrence. Hence, the display at the output creation stage is a result of manipulation of the system by the user alone.

ii. *Submissions on behalf of the Intervenor - Broadband India Forum (BIF) (I.A. 10766/2025)*

63. Mr. Kapil Sibal and Mr. Arvind P. Datar, senior counsel appearing on behalf of the intervenor in I.A. 10766/2025 made the following submissions:

63.1. Newspaper reports, including articles which are in the public domain are not reproduced in the output in any material form by LLMs. If there is a prompt on a particular event, the LLM is able to examine the voluminous

²⁶ Bartz v. Anthropic PBC, 3:24-cv-05417, (N.D. Cal.).



information available in the public domain and prepare a unique response. Therefore, there is no reproduction in any material form. The LLMs often provide links to the sources that have been accessed.

63.2. Several newspapers, magazines, etc. have created ‘paywalls’ or ‘crawler blockers’ to prevent access to their contents. The data from these websites is completely blocked and cannot be accessed. Preparing the responses based on publicly available websites will not amount to violation of Section 14(a)(i) of the Copyright Act, so long as there is no reproduction in material form. LLMs do not reproduce the literary work in any material form.

63.3. LLMs are not designed to materially reproduce any work or a substantial part of it in their response to prompts. They merely understand the ideas, concepts and the facts in those works, which can be used to generate output. It does not amount to adaptation of copyrighted works.

iii. Submissions on behalf of the Intervenor – Flux Labs Pvt. Ltd. (I.A. 4615/2025)

64. Ms. Haripriya Padmanabhan, senior counsel appearing on behalf of intervenor in I.A. 4615/2025 made the following submissions:

64.1. The examples given by ANI in the plaint do not show substantial reproduction of copyrighted works of ANI.

64.2. Unlike any other customer of an AI platform, ANI gave a ‘direct prompt’ to ChatGPT in the said examples. Yet, the responses given by ChatGPT only summarised the interviews in its own words and referred to the author in third person.

64.3. The AI platforms also provide a reference and do not claim that the



information is theirs. Hence, the output/responses generated by AI platforms cannot be stated to be a violation of copyright under Section 14(a)(i) of the Copyright Act.

Submissions of Mr. Adarsh Ramanujan, *Amicus Curiae*

65. Mr. Ramanujan made the following submissions:

65.1. The issue whether Open AI's LLMs underlying ChatGPT retain the full 'raw data' and reproduce duplicate or near duplicate extracts of the 'raw data' as alleged by ANI can be determined only after a full trial. However, ANI's submission in this regard does not appear to be *prima facie* tenable as the function of an LLM is based on patterns and therefore, it would not make sense for an LLM to store raw data on a permanent basis to generate responses for a query.

65.2. In some instances, training of an LLM leads to memorization. This tendency of verbatim memorization occurs when an LLM is trained on basic facts or unique pieces of information, since the training data may not be sufficiently large for the model to build a generalization. Hence, it is possible that training of ChatGPT leads the underlying LLMs to memorize parts of ANI's works.

65.3. In the event memorization occurs, even if inadvertently, the communication of the same to the public at the output stage would amount to infringement.

65.4. The decision of the Supreme Court in ***R.G. Anand*** (supra), is the leading judgment on the aspect of copyright infringement. In terms of paragraph 46 of ***R.G. Anand*** (supra), the Supreme Court has clearly stated



that “*form, manner and arrangement and expression of the idea*” would only be entitled to copyright protection.

65.5. Retrieval Augmented Generation (RAG)²⁷ is an aspect which has not been pleaded in the plaint and has been raised only in the course of arguments. Therefore, whether RAG involves infringement of ANI’s works should also be a subject matter of a trial.

Submissions of Professor Arul George Scaria, *Amicus Curiae*

66. Prof. Scaria made the following submissions:

66.1. A reading of the judgment by the Supreme Court in ***Eastern Book Company v. D.B. Modak*** (supra), highlights that India has explicitly discarded the ‘*sweat of the brow*’ approach and adopted the ‘*skill and judgment*’ standard. Therefore, the current legal standard of originality under Indian copyright law is of a standard higher than the ‘*sweat of the brow*’ standard.

66.2. Taking into account the current standard of originality, substantial chunk of news related materials might be outside copyright protection in India. Therefore, issues such as existence of a copyright in favour of ANI and the need for use of news material by LLMs to provide accurate responses can only be determined in a trial.

66.3. The non-expressive use of copyrighted works by LLMs to provide more optimal responses to any user prompts may not constitute copyright infringement.

66.4. Since news is based on facts and when there are limited ways in which

²⁷ As discussed in the Section: Working of LLMs.



a fact can be expressed, in terms of ‘*merger doctrine*’ as discussed in *Syndicate of the Press of the University of Cambridge v. B.D. Bhandari*²⁸, news should not get copyright protection or should get only a thin layer of protection.

DISCUSSION AND ANALYSIS

67. Under the Copyright Act, copyright subsists only for the classes of works enlisted under Section 13 of the Copyright Act. ANI claims that news articles and interviews published on its digital platforms are “*original literary works*” under Section 13(1)(a)²⁹ of the Copyright Act. This position has not been disputed by Open AI. However, Open AI submits that ANI is not the owner of works published on its website in terms of Section 17³⁰ of the Copyright Act.

68. To establish ownership of the works published on its website, ANI has placed on record a sample ‘*Professional Services Agreement*’ in terms of which copyright would vest with ANI in respect of original works, such as images, video clippings, articles, reports, interviews, etc. created/authored/obtained by personnel on behalf of ANI.

²⁸ 2011 SCC OnLine Del 3215.

²⁹ **13. Works in which copyright subsists.**—(1) Subject to the provisions of this section and the other provisions of this Act, copyright shall subsist throughout India in the following classes of works, that is to say,—

(a) original literary, dramatic, musical and artistic works;

³⁰ **17. First owner of copyright.**—Subject to the provisions of this Act, the author of a work shall be the first owner of the copyright therein:

Provided that—

(a) in the case of a literary, dramatic or artistic work made by the author in the course of his employment by the proprietor of a newspaper, magazine or similar periodical under a contract of service or apprenticeship, for the purpose of publication in a newspaper, magazine or similar periodical, the said proprietor shall, in the absence of any agreement to the contrary, be the first owner of the copyright in the work, in so far as the copyright relates to the publication of the work in any newspaper, magazine or similar periodical, or to the reproduction of the work for the purpose of its being so published, but in all other respects the author shall be the first owner of the copyright in the work;



69. On a *prima facie* view, this Court is satisfied that there is an explicit agreement between ANI and its professionals which recognises ANI as the owner of the copyright subsisting in “*original literary works*” published on its website in terms of Section 17 of the Copyright Act.

70. It is not disputed by Open AI that “*original literary works*” that are available free of cost and publicly accessible are nevertheless entitled to copyright protection. Therefore, for adjudicating present application this Court is of the *prima facie* view that ANI, as an owner, would have copyright over the “*original literary works*”, even if they are freely and publicly available on its website. Any challenge to the ownership of individual works which are subject matter of the present suit can be adjudicated at the stage of trial.

71. As a copyright owner of the original literary works, ANI enjoys certain exclusive rights under Section 14 of the Copyright Act, which include the following:

- (i) the exclusive right to reproduction of the work in terms of Section (14)(a)(i) of the Copyright Act and,
- (ii) the right to “*communicate it to the public*” in terms of Section 14(a)(iii) of the Copyright Act.

For the ease of reference, the relevant extracts from Section 14 are set out below:

“14. Meaning of copyright.—For the purposes of this Act, “copyright” means the exclusive right subject to the provisions of this Act, to do or authorise the doing of any of the following acts in respect of a work or any substantial part thereof, namely:— (a) in the case of a literary, dramatic or musical work, not being a computer programme,—



(i) to reproduce the work in any material form including the storing of it in any medium by electronic means;

(iii) to perform the work in public, or communicate it to the public;”

72. Section 2 (ff)³¹ of the Copyright Act defines “*communication to the public*”.

73. Thus, the act of reproduction and communicating the work of ANI to the public by a third party without license would amount to infringement as per Section 51³² of the Copyright Act

74. It is the case of ANI that Open AI’s LLM model memorizes the content of ANI’s articles and specific responses are generated by ChatGPT through memorisation and the original literary works of ANI are communicated to the public. This amounts to infringement under Section 51 of the Copyright Act.

³¹ (ff) “**communication to the public**” means making any work or performance available for being seen or heard or otherwise enjoyed by the public directly or by any means of display or diffusion other than by issuing physical copies of it, whether simultaneously or at places and times chosen individually, regardless of whether any member of the public actually sees, hears or otherwise enjoys the work or performance so made available.

³² 51. When copyright infringed.—Copyright in a work shall be deemed to be infringed—

(a) when any person, without a licence granted by the owner of the Copyright or the Registrar of Copyrights under this Act or in contravention of the conditions of a licence so granted or of any condition imposed by a competent authority under this Act—

(i) does anything, the exclusive right to do which is by this Act conferred upon the owner of the copyright, or

(ii) permits for profit, any place to be used for the communication of the work to the public where such communication constitutes an infringement of the copyright in the work, unless he was not aware and had no reasonable ground for believing that such communication to the public would be an infringement of copyright; or

(b) when any person—

(i) makes for sale or hire, or sells or lets for hire, or by way of trade displays or offers for sale or hire, or

(ii) distributes either for the purpose of trade or to such an extent as to affect prejudicially the owner of the copyright, or

(iii) by way of trade exhibits in public, or

(iv) imports into India,

any infringing copies of the work:

Provided that nothing in sub-clause (iv) shall apply to the import of one copy of any work for the private and domestic use of the importer.

Explanation.—For the purposes of this section, the reproduction of a literary, dramatic, musical or artistic work in the form of a cinematograph film shall be deemed to be an “infringing copy”.



75. Therefore, this Court would have to examine whether the responses generated by ChatGPT would result in reproduction and communication of the original literary works of ANI to the public without authorisation from ANI and hence, would amount to infringement. More particularly, the Court would have to examine the following two issues:

- i. whether Open AI memorizes and regurgitates ANI's copyrighted literary works in the form of responses and,
- ii. whether the ChatGPT's responses are substantial reproduction of ANI's copyrighted literary works.

i. ***Whether Open AI memorizes and regurgitates ANI's copyrighted literary works in the form of responses?***

76. It is contended on behalf of ANI and the intervenors supporting ANI that LLM models underlying ChatGPT inherently memorize content and retain copies of their training data which results in regurgitation. According to ANI and intervenors, LLMs are capable of providing verbatim or near-identical reproductions of the training data in response to user prompts, thereby demonstrating that expressive elements of the copyrighted works are also retained and reproduced.

77. *Per contra*, it has been highlighted on behalf of Open AI as well as intervenors supporting Open AI that an LLM model is not designed to reproduce extracts of the contents that form part of their training data. It was also argued that the LLMs underlying ChatGPT are not trained to duplicate or regurgitate the exact manner of presentation of content on which these models have been trained. Based on a particular prompt, the LLM is able to prepare a unique response based on the voluminous data on which it has been trained.



Further, generative AI models also provide links to the sources that have been accessed by them. In this regard, Open AI has highlighted that the same prompts by different users would result in different responses from the ChatGPT (*document 10 in the defendant's documents filed along with the written statement*).

78. Professor Scaria, one of the *Amici Curiae*, in his written submissions has argued that in vast majority of cases there is minimal memorisation in the working of LLMs and mostly memorisation occurs when specific attempts are made to generate potential copyright infringing outputs. The relevant extracts from the written submissions of Prof. Scaria are set out below:

“58. A couple of general observations may be added here. First, as the technical literature suggests, though LLMs may be using copyrighted training data to generate outputs, in most instances, they may not be generating potentially copyright infringing responses. Scholars like Prof. Matthew Sag points out that this is due to the different steps taken during model training like the process of decomposition and abstraction, blending of latent concepts, and noise injection [Sag, 2023]. However, in some instances, particularly due to the memorisation problem discussed earlier as part of Issue II, there may be outputs which are identical or similar to the input (copyrighted) training data. This is seen mostly during the so-called “extraction attacks”, wherein specific attempts are made to generate potentially copyright infringing outputs. Some of the empirical studies suggest that memorisation was seen in around 0.03% of sample images during extraction attacks and they were successful mostly in instances wherein the images were duplicated more than 100 times [Sag, 2023]. This also suggests that in the vast majority of use scenarios, the use of copyrighted materials for generating responses are non-infringing, because only non-expressive uses are made of the copyrighted materials.”

[Emphasis supplied]

79. Mr. Adarsh Ramanujan in his written submissions has argued that



ANI's contention that Open AI retains the full raw data is *prima facie* untenable as an LLM is merely a set of mathematical functions to connect input to output. An LLM's function is to predict based on patterns. The chances of memorization of the raw data may increase when an LLM Model is trained on basic facts or unique pieces of information. This tendency occurs because there may be absence of general patterns leading to direct reproduction at the time of output generation.

80. The US District Court for Northern District of California in *Bartz v. Anthropic* (supra) examined the issue of memorization in LLMs and regurgitation of responses from its training data and made the following observations:

"Either way Meta added the books it downloaded to the datasets it used to train the Llama models. It is also post-trained its models to prevent them from "memorizing and outputting certain text from their training data, including copyrighted material. These training efforts, which Meta calls "mitigations," appear to have been successful. Meta's expert witness tested them using a method designed to get LLMs to regurgitate material from its training data (which Meta calls "adversarial prompting"). Even using that method, the expert could get no model to generate more than 50 words and punctuation marks (that is, "tokens") from the plaintiffs' books. And the plaintiffs' expert could only get the Llama model best at regurgitation to generate 50 words and punctuation marks from the plaintiffs' books in 60% of tests. She also testified that Llama was not able to reproduce "any significant percentage" of them. Meta MSJ Ex.24 at 237:16-19; see also Pls Ex. 79, 70-72, 79, 82-83, 92; Meta MSJ Ex.23 at 179:22-25, 180:17-181:16. In short, Llama cannot currently be used to read or other wise meaningfully access the plaintiffs' books."

[Emphasis supplied]

81. From the aforesaid observations, it appears that even the plaintiffs' expert in the field of AI could only make the LLM Model regurgitate



insignificant amounts from the copyrighted works of the plaintiffs.

82. ANI's case is based on '*memorisation*', whereby the LLM memorises the training data verbatim or substantially. In the plaint, ANI has given nine illustrations to demonstrate substantial reproduction. It is an undisputed position that the training of the LLMs underlying ChatGPT ended in April 2022 (for GPT 4) or April 2024 (for GPT 4o), which was much prior to the publication of the illustrative articles used in the plaint, to allege reproduction of ANI's copyrighted works. The table below illustrates this aspect:

TITLE(S) OF THE ARTICLE(S)	DATE OF PUBLICATION
"My father was forced...": Omar Abdullah opens up on IC 814 hijack incident, reveals difficult decisions faced by his father	6 th September 2024
T20 World Cup: "Pakistan never a challenge in Women's cricket": Ex-player Reema Malhotra backs India	30 th September 2024
"Lord Balaji will not forgive him even if he climbs up the seven hills on his knees": TDP spokesperson slams Jagan Mohan Reddy	26 th September 2024
"Connecting this to UP not right," Himachal AICC in-charge Rajeev Shukla defends state govt on ID cards for vendors	26 th September 2024
"It's gone underground...": JKPC chairman Sajad Gani Lone on 'Azadi' sentiments in Kashmir	10 th September 2024



“Jinnah like mentality,” Minister Hardeep Puri accuses Rahul Gandhi of crossing limit for “political interest”	17 th September 2024
Neeraj Chopra’s mother celebrates Olympic Silver; praises Pakistan’s Arshad Nadeem for bagging Gold	9 th August 2024
“Ashwin best in India, Lyon better overall”: Monty Panesar sets-up best Test spinner debate	23 rd September 2024

83. Therefore, it is apparent that the illustrations of ChatGPT responses relied upon in the plaint could not have been produced on account of training of Open AI’s LLMs on the ANI’s literary works, as the literary works in question were published only after the completion of the training process.

84. As discussed earlier under the Section ‘*Working of LLMs*’, it could be inferred that when an LLM refers to information on which it was not trained, it is using RAG technique. Using the RAG technique, LLMs retrieve relevant information from external sources, in this case ANI’s website, before generating a response. Relevant data retrieved from the website, which is contextually aligned with the user’s prompt, is then generated as responses grounded in the website’s data rather than relying solely on its pre-trained knowledge. Therefore, responses that are not based on the trained data cannot be used against Open AI in respect of claims based on memorisation. Such allegations must be tested based on the identical or substantially similar reproduction of the trained data. Therefore, this Court is of the *prima facie* view that the instances given in the plaint alleging infringement are not a



result of memorisation, rather they are in the nature of live links, perhaps reflecting RAG technique. Whether outputs produced using RAG would amount to copyright infringement is an aspect which has not been pleaded in the plaint, though it was referred during the course of submissions.

ii. ***Whether ChatGPT's responses are substantial reproduction of ANI's copyrighted literary works?***

85. To begin with, it may be relevant to discuss legal principles regarding copyright infringement elucidated in some of the instructive judgments of the Supreme Court that have been relied upon by both sides.

86. In ***R.G. Anand*** (supra), the plaintiff sued the defendants for damages claiming infringement of the copyright of the plaintiff in a dramatic play, which was made into a motion picture by the defendants. The Supreme Court held that an idea, principle, theme or facts cannot be subject matter of copyright protection. Where two writers write on the same subject, similarities are bound to occur because the central idea of both are the same. Therefore, the Court has to determine whether the defendant has only adopted the idea of the copyrighted work or has also adopted the form, manner, arrangement and expression of an idea with minor changes. If the defendant's work appears to be a transparent rephrasing or a copy of the substantial and material part of the original, it would amount to copyright infringement. The Supreme Court in paragraph 46 of the judgment has laid down the following principles to determine infringement of copyright, which are set out below:

“46. Thus, on a careful consideration and elucidation of the various authorities and the case law on the subject discussed above, the following propositions emerge:



1. There can be no copyright in an idea, subject-matter, themes, plots or historical or legendary facts and violation of the copyright in such cases is confined to the form, manner and arrangement and expression of the idea by the author of the copyrighted work.
2. Where the same idea is being developed in a different manner, it is manifest that the source being common, similarities are bound to occur. In such a case the courts should determine whether or not the similarities are on fundamental or substantial aspects of the mode of expression adopted in the copyrighted work. If the defendant's work is nothing but a literal imitation of the copyrighted work with some variations here and there it would amount to violation of the copyright. In other words, in order to be actionable the copy must be a substantial and material one which at once leads to the conclusion that the defendant is guilty of an act of piracy.
3. One of the surest and the safest test to determine whether or not there has been a violation of copyright is to see if the reader, spectator or the viewer after having read or seen both the works is clearly of the opinion and gets an unmistakable impression that the subsequent work appears to be a copy of the original.
4. Where the theme is the same but is presented and treated differently so that the subsequent work becomes a completely new work, no question of violation of copyright arises.
5. Where however apart from the similarities appearing in the two works there are also material and broad dissimilarities which negative the intention to copy the original and the coincidences appearing in the two works are clearly incidental no infringement of the copyright comes into existence.
6. As a violation of copyright amounts to an act of piracy it must be proved by clear and cogent evidence after applying the various tests laid down by the case-law discussed above.
7. Where however the question is of the violation of the copyright of stage play by a film producer or a director the task of the plaintiff becomes more difficult to prove piracy. It is manifest that unlike a stage play a film has a much broader prospective, wider field and a bigger background where the defendants can by introducing a variety of incidents give a colour and complexion different from the manner in which the copyrighted work has expressed the idea. Even so, if the viewer after seeing the film gets a totality of impression that the film is by and large a copy of the original play, violation of the copyright may be said to be proved."

[Emphasis supplied]



87. In *Eastern Book Company v. D.B. Modak* (supra), the appellants/plaintiffs filed a suit along with an application for temporary injunction seeking to restrain the respondents/defendants from reproducing the plaintiffs' version of copy-edited judgments and orders of the Supreme Court on CD-ROMs. The defendants were selling software packages on CD-ROMs with the text of the copy-edited judgments of the Supreme Court along with headnotes prepared by the plaintiffs for its own law report, 'SCC'. The Supreme Court, relying upon the judgment of the Supreme Court of Canada in *CCH Canadian v. Law Society of Upper Canada*³³ held that 'sweat of the brow' approach to originality is not the correct standard as it fails to protect the public interest in production and dissemination of works. On the other hand, the creativity standard of originality as applied by the US Courts, is a standard of originality with a higher threshold. The underlying impetus for this modification was the need to balance the rights of the copyright owner with the public interest.

88. In paragraph 15 of the *Eastern Book Company v. D.B. Modak* (supra), it was specifically observed that for a literary work to be original and to claim copyright, it should possess some amount of creativity in the work, more than mere facts. For ease of reference, the aforesaid paragraph is set out below:

*"15. Copyright is a right to stop others from exploiting the work without the consent or assent of the owner of the copyright. A copyright law presents a balance between the interests and rights of the author and that of the public in protecting the public domain, or to claim the copyright and protect it under the copyright statute. **One of the key requirements is that of originality which contributes, and has a direct nexus, in maintaining the interests of the author as well as that of***

³³ (2004) 1 R.C.S. 339.



public in protecting the matters in public domain. It is a well-accepted principle of copyright law that there is no copyright in the facts per se, as the facts are not created nor have they originated with the author of any work which embodies these facts. The issue of copyright is closely connected to that of commercial viability, and commercial consequences and implications.”

[Emphasis supplied]

89. Accordingly, the Supreme Court laid down the ‘*skill and judgment test*’ as opposed to ‘*sweat of the brow*’ test that was applied earlier in India. In terms of the ‘*skill and judgment test*’, the author of a derivative work has to satisfy that the work has been produced from exercise of skill and judgment and, the skill and judgment should not be of a trivial nature. It was observed that the ‘*sweat of the brow*’ doctrine creates a monopoly in materials available in public domain without encouraging the creation of writings by the authors. In this regard, the relevant paragraph of the *Eastern Book Company v. D.B. Modak* (supra) is set out below:

“57. The Copyright Act is not concerned with the original idea but with the expression of thought. Copyright has nothing to do with originality or literary merit. Copyrighted material is that what is created by the author by his own skill, labour and investment of capital, maybe it is a derivative work which gives a flavour of creativity. The copyright work which comes into being should be original in the sense that by virtue of selection, coordination or arrangement of pre-existing data contained in the work, a work somewhat different in character is produced by the author. On the face of the provisions of the Copyright Act, 1957, we think that the principle laid down by the Canadian Court would be applicable in copyright of the judgments of the Apex Court. We make it clear that the decision of ours would be confined to the judgments of the courts which are in the public domain as by virtue of Section 52 of the Act there is no copyright in the original text of the judgments. To claim copyright in a compilation, the author must produce the material with exercise of



his skill and judgment which may not be creativity in the sense that it is novel or non-obvious, but at the same time it is not a product of merely labour and capital. The derivative work produced by the author must have some distinguishable features and flavour to raw text of the judgments delivered by the court. The trivial variation or inputs put in the judgment would not satisfy the test of copyright of an author.”

[Emphasis supplied]

90. The legal position that emerges from *Eastern Book Company v. D.B. Modak* (supra) is that the work which is produced by the author by using his skill and judgement with a flavour of creativity is copyrightable. The protection would be only in respect of creative and original works. There cannot be any copyright protection in respect of facts even if someone has put in effort to obtain those facts.

91. In *Feist Publications* (supra), the US Supreme Court laid down the test of ‘*modicum of creativity*’. In this case, a telephone utility company brought a copyright infringement action against the publisher of a telephone directory. The US Supreme Court held that no one can claim originality as to the facts. A person who reports a particular fact has not created the said fact but has only discovered its existence. The relevant paragraphs of *Feist Publications* (supra) are set out below:

“[15] It is this bedrock principle of copyright that mandates the law’s seemingly disparate treatment of facts and factual compilations. “No one may claim originality as to facts.” *Id.*, § 2.11[A], p. 2-157. This is because facts do not owe their origin to an act of authorship. The distinction is one between creation and discovery : the first person to find and report a particular fact has not created the fact; he or she has merely discovered its existence. To borrow from Burrow-Giles, one who discovers a fact is not its “maker” or “originator.” 111 U.S., at 58. “The discoverer merely finds and records.” *Nimmer* § 2.03[E]. *Census-*



takers, for example, do not “create” the population figures that emerge from their efforts; in a sense, they copy these figures from the world around them. Denicola, *Copyright in Collections of Facts : A Theory for the Protection of Nonfiction Literary Works*, 81 Colum. L. Rev. 516, 525 (1981) (hereinafter Denicola). Census data therefore do not trigger copyright because these data are not “original” in the constitutional sense. Nimmer [p*348] § 2.03 [E]. The same is true of all facts -- scientific, historical, biographical, and news of the day. “They may not be copyrighted and are part of the public domain available to every person.” Miller, supra, at 1369.

[18] This inevitably means that the copyright in a factual compilation is thin. Notwithstanding a valid copyright, a subsequent compiler remains free to use the facts contained in an another's publication to aid in preparing a competing work, so long as the competing work does not feature the same selection and arrangement. As one commentator explains it: “No matter how much original authorship the work displays, the facts and ideas it exposes are free for the taking.... The very same facts and ideas may be divorced from the context imposed by the author, and restated or reshuffled by second comers, even if the author was the first to discover the facts or to propose the ideas.” “Ginsburg 1868.”

[Emphasis supplied]

92. ANI has relied upon judgment of the House of Lords in **Walter v. Lane**³⁴ delivered in 1900, to show that it has copyright in respect of the news circulated by its platform. In **Walter v. Lane** (supra), the issue before the House of Lords was whether a person who makes notes of speech delivered in public, transcribes the same and publishes in the newspaper would be entitled to copyright protection. The House of Lords held that each of the reporters reporting the speech would have a copyright in his own published report. The House of Lords came to this conclusion based on the doctrine of

³⁴ [1900] AC 539.



‘*sweat of the brow*’ and it was premised on the principle that an author should be able to reap the benefits of its labour.

93. On behalf of Open AI, the aforesaid judgment is sought to be distinguished as it was in the context of copyright laws that existed in England in 1842. The then copyright laws in England did not require originality.

94. **Walter v. Lane** (supra) was considered by the Supreme Court in **Eastern Book Company v. D.B. Modak** (supra), as the appellant therein relied upon the same, and it was held that the doctrine of ‘*sweat of the brow*’ does not hold good anymore as it gives more weight to the copyright owner’s right as opposed to the public interest.

95. In **B.D. Bhandari** (supra), the Division Bench delved on the ‘*merger doctrine*’ and noted that when there is limited manner in which a fact can be expressed, the facts would merge with the expression. The facts and the expression would become so intrinsically connected that it would be difficult to distinguish one from another, making the expression itself non-copyrightable. For ease of reference, the relevant paragraphs of **B.D. Bhandari** (supra) are set out below:

“106. While answering the first question, the learned Single Judge expanded the “the Idea-Expression Dichotomy and the Doctrine of Merger” and noted that there is no copyright protection on ideas and facts and it is only the creative expression of such ideas and facts that is rewarded by law. In the process, expression “originality” in artistic, literary, dramatic and musical works were also explained holding that with the passage of time, doctrine of “sweat of the brow” had given way to the “modicum of creativity” and therefore, in order to have a copyright in a work, it was necessary to show that it involves a certain degree of creativity.

107. The learned Single Judge also deemed it proper to refer to the “doctrine of merger”, peculiar to copyright law in the following manner:



“18. An aspect peculiar to copyright law, is the doctrine of merger, is involved in this case. This doctrine posits that where the idea and expression are intrinsically connected, and that the expression is indistinguishable from the idea, copyright protection cannot be granted. Applying this doctrine courts have refused to protect the expression of an idea that can be expressed only one manner, or in a very restricted manner, because doing so would confer monopoly on the idea itself. The decision in Herbert Rosenthal Jewelry Corporation v. Kalpakian, 446 F.2d 738 (1971) is illustrative in this regard. In that case the plaintiffs sued the defendants asking them to refrain from manufacturing bee shaped jewel pins. The Court held that the jewel shaped bee pin was a an idea that anyone was free to copy, the expression of which could be possible only in a few ways, therefore, no copyright could subsist in it.”

[Emphasis supplied]

96. In *Akuate Internet Services v. Star India*³⁵, Star India was granted exclusive broadcasting rights for a cricket match organised by BCCI including mobile distribution rights. The defendants/appellants were disseminating match information/score alerts without obtaining a license from the Star India (plaintiff/respondent). In the said case, the plaintiff was seeking copyright protection over facts that were available in public domain. The Division Bench of this Court held that there cannot be any copyright protection over facts and information that is available in public.

97. From a reading of the aforesaid judicial precedents, it is clear that there cannot be any copyright in facts. Therefore, in the context of news, copyright would subsist only in the form and manner of expression of news and not on the underlying facts. What has to be seen is whether the expression used by

³⁵ 2013 SCC OnLine Del 3344.



Open AI in its ChatGPT responses is materially similar to the expression used by ANI in its news articles resulting in reproduction and making the work available to the public in terms of Sections 14(a)(i) and 14(a)(iii) of the Copyright Act. ANI cannot claim any copyright in the facts underlying the news articles.

98. ANI has given nine instances of substantial reproduction of ANI's works by Open AI (*paragraph 33.1 to 33.9 of the plaint*). ANI claims that the aforesaid nine instances have been given only as an illustration and there are many more instances of copyright infringement by Open AI. All the nine instances given in the plaint are in respect of news/interviews which were originally published by ANI.

99. During oral submissions, ANI had highlighted the similarities between the works of ANI and the response produced by ChatGPT in paragraph 33.7 of the plaint, which dealt with an interview of the mother of Neeraj Chopra who had secured a silver medal at the Olympics, 2024 in Men's Javelin Throw. Admittedly, the said interview was given in Hindi and was translated into English by ANI for reporting. A comparison given by ANI in paragraph 33.7 of the plaint is set out below:



Plaintiff's Work	Response produced by ChatGPT
Panipat (Haryana), August 09 (ANI): "Neeraj Chopra's mother, Saroj Devi elated as her son secured a silver medal in the Paris Olympics men's javelin throw, equating the achievement to gold. She also praised Arshad Nadeem's performance, calling him "her child" and acknowledging the hard work of all the athletes. Speaking to ANI, Saroj Devi said, "We are very happy, for us silver is also equal to gold...he was injured, so we are happy with his performance. The one who won the gold [Nadeem] is also my child. They all work hard to get there."	Prompt: "What did Neeraj Chopra's mother tell the media about the Pakistani gold medallist?" Response: "Neeraj Chopra's mother, Saroj Devi, made a heartfelt statement to the media after the 2024 Paris Olympics, where her son won the silver medal in the men's javelin throw. Despite her son finishing second, she expressed immense pride in his performance and graciously acknowledged the achievement of Pakistan's Arshad Nadeem, who won the gold medal. She remarked that "the one who got the gold is also like our son, highlighting the bond between the two athletes and showing a sense of sportsmanship that transcends national rivalry. Her comments were widely praised in both India and Pakistan as a beautiful gesture of unity and respect." Prompt: "Please try to tell me exactly what she said" Response: "Neeraj Chopra's mother, Saroj Devi, expressed her feelings after the 2024 Paris Olympics by saying that "Arshad bhi mera beta hai" (Arshad is also like my son). She further added "We are very happy. For us, silver is also equal to gold. He [Neeraj] was injured, so we are happy with his performance." These words reflect her warmth and the sportsmanship that she showed towards Arshad Nadeem, the gold medalist from Pakistan, acknowledging him with a motherly affection despite the intense competition."

*The colour highlighting has been done for ease of comparison

100. As per ANI, the response produced by ChatGPT shows that ChatGPT



has produced exact or nearly exact copy of ANI's news article. The original interview and responses were in Hindi which were translated into English by ANI. Open AI has also infringed ANI's exclusive rights to translate the works for which ANI invested significant time and effort.

101. What emerges from a reading of the aforesaid extracts is that ChatGPT generated this output in response to two different prompts. The first prompt was:

***“Prompt:** What did Neeraj Chopra's mother tell the media about the Pakistani gold medallist?”*

102. A comparison of the response given by ChatGPT to the first prompt as highlighted above (in blue) with what was reported by ANI as highlighted above (in orange) shows that the response given by ChatGPT was not a substantial reproduction or nearly exact copy, though the essence of the response is similar as it based on facts i.e. a particular interview of Neeraj Chopra's mother. ChatGPT has added its own commentary to the news article published by ANI, which bears its own expression and is quite dissimilar to the expression used in ANI's article.

103. Realising that the response was not a substantial reproduction, ANI gave another prompt in the following terms:

***“Prompt :** Please try to tell me exactly what she said”*

104. To be noted, in the second prompt, ANI used the word ‘*exactly*’, which clearly shows that ANI was seeking to elicit a response from ChatGPT in exact terms i.e. an adversarial prompt (“*carefully designed inputs that*



manipulate model outputs”³⁶). In other words, deliberately a prompt is being put in a manner so as to elicit a specific response to show that ChatGPT has substantially reproduced ANI’s works.

105. Upon the said prompt, ChatGPT came out with a quote of what was stated by Neeraj Chopra’s mother as highlighted above (in orange). The only similarity which emerges from the aforesaid comparison is relating to a part of a quote from Neeraj Chopra’s mother in the interview. However, ChatGPT’s response has given its own context to the aforesaid quotes from Neeraj Chopra’s mother’s interview. In addition, a subsequent paragraph was added (as highlighted in green) completely on its own by ChatGPT to explain the quoted text which was not a part of ANI’s news report.

106. In the present case, ANI has only demonstrated that Open AI has extracted a quote from an interview accompanied by its own commentary. According to Section 17(cc)³⁷ of the Copyright Act, the author of an address or speech made in public is the person who made it. The news article gives the quote of what was said by the Neeraj Chopra’s mother in the video. Applying the underlying principle of Section 17(cc) of the Copyright Act to the facts of the present case, in the *prima facie* view of this Court, Neeraj

³⁶ Ahmet Emre Ergün, Aytuğ Onan, Adversarial Prompt Detection in Large Language Models: A Classification-Driven Approach, Computers, Materials and Continua, Volume 83, Issue 3, 2025, Pages 4855-4877 ISSN 1546-2218, <https://doi.org/10.32604/cmc.2025.063826>. (<https://www.sciencedirect.com/science/article/pii/S1546221825004898>).

³⁷ **17. First owner of copyright.**—Subject to the provisions of this Act, the author of a work shall be the first owner of the copyright therein:
Provided that—

...
(cc) in the case of any address or speech delivered in public, the person who has delivered such address or speech or if such person has delivered such address or speech on behalf of any other person, such other person shall be the first owner of the copyright therein **notwithstanding that the person who delivers such address or speech, or, as the case may be, the person on whose behalf such address or speech is delivered, is employed by any other person who arranges such address or speech or on whose behalf or premises such address or speech is delivered;**



Chopra's mother would be first owner of the copyright in the said quotation. Similarly, in terms of Section 14(a)(vii) of the Copyright Act, the right to translate would also vest with Neeraj Chopra's mother. There is nothing produced by ANI to show that Neeraj Chopra's mother has assigned this ownership in favour of ANI. Therefore, in the *prima facie* opinion of this Court, the reproduction of quotes of Neeraj Chopra's mother by ChatGPT in its responses would not amount to infringement of ANI's copyright.

107. Pertinently, the original article of ANI (*page 54, document no.11 in the plaintiff's documents filed along with the plaint*) with which the comparison has been made, was much longer than the text which was quoted in paragraph 33.7 of the plaint for the purposes of comparison. The response given by ChatGPT does not reproduce the remaining text which is a part of the original article.

108. In ***R.G. Anand*** (supra), the Supreme Court has held that if a comparison has to be made between two works it should be made for the works as a whole and not selected parts of the works.

109. In this regard, Open AI has relied upon the judgment of Bombay High Court in ***Leo Burnett*** (supra). In the said judgment, the Bombay High Court held that a work in which the plaintiff's claim copyright has to be compared as a whole and not in parts in order to determine substantial copying. It is not permissible to split or dissect the plaintiff's work into parts and then compare the same with the allegedly infringing work. Paragraph 26 of the said judgment is set out below:

"26. The issue then is whether if the whole of the respective works is compared and not only parts, factually or materially as it stands, is there copying or substantial copying. I am not at this stage also not considering whether the "work" as an episode or all the episodes which



have to be considered in totality. Even for arguments sake if an episode is considered as a 'work', still what is to be compared is the whole work as a single entity. It is not permissible to split or compartmentalize or dissect the plaintiffs' work into parts, then seek to compare such parts against the alleged infringed work. What is to be seen is the entirety of each work. It may be made clear that I am not considering an issue whether a particular part of the film or a scene in the film which may have a striking resemblance, if copied, would amount to a substantial copying of the film. For this purpose reference may be made to the decision in (Francis Day and Hunter Limited v. Twentieth Century Fox Corporation Limited), 1940 A.C. 112 (Privy Council); (Ladbroke (Football) Limited v. William Hill (Football) Limited), 1964 (1) W.L.R. 273; (Merchandising Corporation of America Inc. v. Harpbond Limited), 1983 F.S.R. 32; R.G. Anand v. Delux Films supra and (Fritco Lay Limited v. Uncle Chips Private Limited), 2000 P.T.C. 341. Applying the tests on the material as it stands there is no substantial copying on facts."

[Emphasis supplied]

110. The aforesaid findings have been followed by Coordinate Bench of this Court in **Bikramjeet Singh** (supra), wherein it was observed that rival works have to be compared as a whole and as a single entity. It is not possible to split or dissect the works into parts and then compare them.

111. During the course of oral submissions, ANI has not addressed on any of the other illustrations in paragraph 33 of the plaint. Hence, I do not propose to examine the same. Counsel on behalf of Open AI during the course of submissions has handed over a note to highlight that in none of the instances cited by ANI in paragraph 33 of the plaint, ChatGPT has given responses identical or substantially similar to ANI's articles.

111.1. The said note has been taken on record.

111.2. The contents of the said note have not been rebutted on behalf of ANI.

112. In paragraph 33.9 of the plaint, ANI has stated that on a specific prompt



2026:DHC:5900



about the latest news published by ANI, ChatGPT has provided updates from ANI's website. For the sake of convenience, the screenshot of the specific prompt which has been given in paragraph 33.9 of the plaint has been transcribed below:

“Prompt - “Give me the latest news updates from ANI News.in

Response - Here are some of the latest news updates from ANI News:

- 1. In a legal development, AAP leader Durgesh Pathak has been granted bail by the Rouse Avenue court (ANI News).*
- 2. Indian stock markets saw a positive opening, buoyed by a global rally after U.S. inflation data indicated easing in August (ANI News).*
- 3. The Union Cabinet has approved Phase 3 of the Bangalore Metro Rail Project, which will include two elevated corridors covering 44.65 km with 31 stations (ANI News).*
- 4. Wrestler Bajrang Punia has moved a petition to the Delhi High Court, leading to a notice being issued to the National Anti-doping Agency (ANI News).”*

For more detailed updates, you can visit ANI News.”

113. ANI has also filed the corresponding news stories from its website (pages 32 to 46 in the documents filed by the plaintiff). The table below gives a comparison of the title of ANI's articles and the news updates generated by ChatGPT based on the prompt set out above.

TITLE(S) OF ARTICLE(S) PUBLISHED BY ANI	NEWS UPDATES GENERATED BY CHATGPT
Excise Police case: Delhi Court grants bail to Durgesh Pathak, Arvind Kejriwal's custody extended till September 25	In a legal development, AAP leader Durgesh Pathak has been granted bail by the Rouse Avenue Court (ANI News).



Indian indices opened in green after ease in US inflation	Indian stock markets saw a positive opening, buoyed by a global rally after U.S. inflation data indicated easing in August (ANI News).
Cabinet approves two corridors of Bangalore Metro Rail Project Phase-3 project for 44.65 km with 31 stations	The Union Cabinet has approved Phase 3 of the Bangalore Metro Rail Project, which will include two elevated corridors covering 44.65 km with 31 stations (ANI News).
Delhi HC issues notice to NADA on Bajrang Punia's plea challenging his suspension	Wrestler Bajrang Punia has moved a petition to the Delhi High Court leading to a notice being issued to the National Anti-Doping Agency (ANI News).

114. A comparison above shows that even the titles in the response generated by ChatGPT are substantially different from the titles of the articles on ANI's website.

115. Applying the test for infringement laid down in **R.G. Anand** (supra), in my *prima facie* view, when the articles of ANI are compared with the responses generated by ChatGPT as a whole, there is no substantial similarity.

116. Now, I shall proceed to deal with the judgments relied upon by ANI.

117. In **GEMA v. Open AI** (supra), the plaintiff claimed copyright infringement in respect of German song lyrics that were memorised and reproduced by the LLMs underlying ChatGPT. The Munich Regional Court observed that the plaintiff therein had expressly excluded the '*online search function*' and in those circumstances a finding of fact was given by the Court that memorisation had occurred as exact song lyrics were reproduced on the basis of non-adversarial prompts. The defendant therein (Open AI) was held to be liable for generating outputs that are identical/substantially similar to the



copyrighted lyrics.

118. In ***GEMA v. Open AI*** (supra), the Court found that there were repeated verbatim reproduction of lyrics by ChatGPT despite non-adversarial prompts. In that context, the Court held that there was memorisation.

119. As noted above, in the present case, the training was completed before the illustrative articles in the plaint were published by ANI and therefore, there cannot be any question of memorization. Secondly, the Munich Court noted that the responses generated were based on non-adversarial prompts given by the user. In the present case, ANI has given fairly detailed, repeated prompts and has even used the term ‘*exactly*’, to extract copyrighted contents from the LLM models underlying ChatGPT, which cannot be termed as non-adversarial. Even after giving such adversarial prompts, ANI was not able to eke out a response which can be termed as ‘*substantial*’ reproduction of its copyrighted material.

120. In ***Associated Press v. Meltwater*** (supra), the defendant-Meltwater was a news-monitoring agency which scraped the articles published by the plaintiff-Associated Press using web crawlers and provided excerpts from these articles to its subscribers through emails or through online account on Meltwater’s website. This was not a case of data being scraped by an LLM. The Court in ***Associated Press v. Meltwater*** (supra) specifically held that there were 33 articles which were reproduced verbatim by the defendant. In the present case, as noted above, there is no substantial reproduction of ANI’s works.

121. In ***Cohere*** (supra), the plaintiffs, publishers of news and magazines in the USA, claimed that the defendant’s LLM Models were using the copyrighted works of the plaintiffs for training and were



reproducing/delivering verbatim copies of the plaintiffs' works using the RAG function. This judgment was in the context of motion to dismiss, which is similar to Order VII Rule 11 of the CPC in India. In the said case, the plaintiffs had given 75 instances of copying, out of which, 50 were alleged to be verbatim. Hence, the motion to dismiss was denied.

122. ANI has also placed reliance on the judgment of the United States Court of Appeal for the Fifth Circuit in ***Positive Black Talk v. Cash Money Records***³⁸ in support of its submission that copyright infringement is made out if there is '*probative similarity*' between the plaintiff's work and the work of the defendant. However, this decision was in the context of copyright infringement in respect of a song and hence, the observations made by the Court were specific to the said factual background. The application of a standard for determining whether there is copyright infringement in respect of a song would be quite different from its application in the context of news. The creativity involved in writing lyrics of a song would ordinarily not be present to the same extent in a news article as the fundamental purpose of a news article is to report events that have actually occurred. Therefore, threshold of establishing substantial similarity in expression would be higher in the case of news. Therefore, this judgment would not be of much assistance to ANI in the present case.

123. The judgments cited on behalf of ANI and the intervenor-DNPA, ***Infopaq International v. Danske Dagblades Forening***³⁹ and ***Newspaper Licensing Agency v. Meltwater Holding***⁴⁰, would have little relevance in the

³⁸ 394 F.3d 357, 368 (5th Cir. 2004).

³⁹ ECLI:EU:C:2009:465.

⁴⁰ [2010] EWHC 3099 (Ch).



present case. In both the aforesaid judgments, it was held that the defendants were providing verbatim extracts of the plaintiff's works. In *Meltwater* (supra), it was held that the defendants have produced extracts and headlines of news articles of the plaintiff therein and had accordingly infringed the copyright in the news articles. Same was the position in *Infopaq* (supra) where the defendant had copied extracts of news articles of the plaintiff. In the present case, as noted above, there is no substantial reproduction of ANI's works.

FINDINGS

124. As highlighted above, the illustrations given in the plaint are post the training of Open AI's LLMs and a case for memorization of ANI's works on the basis of the said illustrations cannot be made out. Therefore, at this *prima facie* stage, the contention of ANI that Open AI permanently stores the training data in order to memorize and regurgitate ANI's works cannot be accepted. At best, these are disputed questions which can only be determined during trial upon parties leading evidence.

125. Based on discussion above it cannot be concluded that the responses produced by ChatGPT are a substantial reproduction of the news articles of ANI. Hence, at a *prima facie* stage, ANI has failed to make out a case of copyright infringement based on the responses generated by ChatGPT i.e. the output claim.

126. As noted earlier, Issue nos. 1 and 3 are intertwined and have to be considered together. Hence, I now proceed to deal with Issue nos. 1 and 3.



Issue No.1

Whether the storage by the defendants of plaintiff'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

SUBMISSIONS

Submissions on behalf of the Plaintiff (ANI)

127. Mr. Sidhant Kumar, counsel appearing on behalf of ANI made the following submissions:

127.1. Open AI collects and stores the 'raw data' as a repository of ANI's works in their original form. The 'raw data' is processed further and filtered. The process of filtration is impossible without storage of the copyrighted works of ANI.

127.2. Tokenisation of raw data is nothing but numerical representation of raw data which belongs to ANI. A token may constitute an individual character, whole words or even parts of words. The LLM decodes these tokens and reconstructs the raw data to generate responses.

127.3. In the second step, the tokens are encoded into vectors, which are thereafter stored in the memory of the LLM to create a vector database. Even if it is assumed that the vector database is an adaptation, such reproduction in adapted form also results in infringement of copyrighted works in terms of Section 2(a)(v) of the Copyright Act.

127.4. Training of LLM inherently involves storage and reproduction of original works. The entire processes undertaken by Open AI including storage



of data, tokenisation, vectorisation amounts to copyright violation in terms of Section 51 read with Section 14 of the Copyright Act.

127.5. What Open AI uses in training is only the expression used by ANI and not the actual facts that occurred. Therefore, training occurs in respect of expression of language.

127.6. Once the training data is fed into Open AI's LLM during the training process, it cannot be removed and continues to be reflected in the outputs generated by the LLM.

Intervenors supporting ANI

- i. Submission on behalf of Intervenor – Indian Music Industry (I.A. 4027/2025)

128. Mr. C.M. Lall, senior counsel appearing on behalf of intervenors in I.A. 4027/2025 made the following submissions:

128.1. The right to reproduce a work under Section 14(a)(i) of the Copyright Act includes the right to store the work. Therefore, scraping, collection, coalition and processing of data for training purposes constitutes infringement.

128.2. Just because copyrighted works are available on online resources/websites would not mean that these works can be copied/reproduced without a license or authorization.

128.3. Regardless of the purpose of the activity, storage of substantial part of copyrighted work amounts to infringement.



ii. Written Submissions filed by Mr. Ankit Sahni, Advocate on behalf of Intervenor – Federation of Indian Publishers (I.A. 470/2025)

129. Open AI admits that collection and compilation of raw data into training data sets involves creation of copies of the training data. The various activities carried out by Open AI makes it clear that Open AI not only collects data, it also stores the same for training its LLM.

129.1. Open AI's act of copying/storing copyrighted works for purposes of text and data mining/text and data analysis amounts to infringement since there is no specific provision in the Copyright Act creating an exception for the same.

iii. Submissions on behalf of Intervenor – Digital News Publishers Association & Ors. (I.A. 2199/2025)

130. Mr. Rajshekhar Rao, senior counsel appearing on behalf of intervenors in I.A. 2199/2025 made the following submissions:

130.1. By using the words “*any medium*” and “*any material form*”, Parliament has made it clear that Section 14(a)(i) of the Copyright Act must not be diluted. Merely because protected content is not perceivable by naked eye or meant for human consumption would not take it outside the ambit of Section 14 of the Copyright Act. There is no requirement in Indian copyright law for the copyrighted works to be stored in a tangible medium. Therefore, storage of literary works in digital formats and their reproduction in the shape of vectors/tokens would fall within the sweep of the expression “*reproduction/storage*” in “*any medium*” or in “*any material form*”.

130.2. The scheme of the Copyright Act does not distinguish between expressive and non-expressive uses. The very same argument raised by the



defendant in ***Richard Kadrey v. Meta Platforms***⁴¹ was rejected by the United States District Court. Even in ***Bartz v. Anthropic*** (supra), it was noted that “*literary works were chosen for their expressive qualities in building a central library and then in training specific LLMs*”.

130.3. Even the intermediate step of introducing and storing data for training cannot be said to be non-infringing as substantial portions of copyrighted works are extracted. Once again, this view finds support from the judgment in ***Richard Kadrey v. Meta Platforms*** (supra).

Submissions of the Defendant (OPEN AI)

131. Mr. Amit Sibal, senior counsel appearing on behalf of Open AI made the following submissions:

131.1. It is an admitted position that storage of data takes place only for the purposes of training. Creation of copy during the training process is an intermediate step. Once the training process is over, an LLM does not have access to data.

131.2. Purposive interpretation should be given by the Courts to the provisions of Copyright Act. Accordingly, it is submitted that Section 14 (a)(i) of the Copyright Act should not be interpreted in a literal manner. In this regard, reliance is placed on:

- i. ***Shailesh Dhairyawan v. Mohan Balkrishna Lulla***⁴²;
- ii. ***State of Maharashtra v. Praful B. Desai***⁴³.

⁴¹ Kadrey v. Meta Platforms, Inc., 3:23-cv-03417, (N.D. Cal.).

⁴² (2016) 3 SCC 619.

⁴³ (2003) 4 SCC 601.



131.3. Training is not based on ANI's data alone. ANI's data along with other data is put in database. ANI's data would only be a miniscule amount of the total data on which training takes place.

131.4. For the purposes of training, Open AI uses the non-expressive elements of literary works of ANI such as grammar, syntax or linguistic patterns.

131.5. The training process is unconcerned with each work's unique expression. The LLMs underlying ChatGPT encode the learnt meanings in the form of a '*vector database*' during the pre-training process, which is equivalent to a dictionary and does not amount to additional copies of raw data. Hence, it is incorrect to state that '*vectorization*' is an '*adaptation*' or '*translation*' of ANI's works.

131.6. The storage is not intended for human consumption or comprehension. Storage is only an intermediate step to derive the non-expressive element.

Intervenors supporting OPEN AI

- i. *Submissions on behalf of Intervenor – Broadband India Forum (I.A. 10766/2025)*

132. Mr. Kapil Sibal and Mr. Arvind P. Datar, senior counsel appearing on behalf of the intervenor in I.A. 10766/2025 made the following submissions:

132.1. Any storage of copyrighted works for training LLMs should be protected in line with the statutory intent of making and enabling new technologies. The storage of works is incidental i.e. training of LLMs and LLMs themselves do not store the works.



ii. Submission on behalf of Intervenor – Flux Labs AI Private Ltd. (I.A. 4615/2025)

133. Ms. Haripriya Padmanabhan, senior counsel appearing on behalf of intervenor in I.A. 4615/2025 made the following submissions:

133.1. The purpose of copyright law is that the work ought not to be circulated in public. Taking this purpose into consideration, the reference to “*storage*” in Section 14(a)(i) of the Copyright Act has to be in the context of reproduction of the work. Mere storage of the work would not amount to copyright infringement.

Submissions of Professor Arul George Scaria, Amicus Curiae

134. Prof. Scaria made the following submissions:

134.1. In the present case, the use of copyrighted works by Open AI is ‘*non-expressive use*’, hence, it would not amount to infringement. The training data is used by Open AI only to extract vital information such as patterns, trend and correlations in phrases and sentences.

134.2. The Court has to consider whether it is feasible to develop an LLM without giving access to copyrighted material as the strength of the LLM depends upon the extent to which materials are available for training.

134.3. The Court must consider if all companies developing LLMs have to obtain licenses from every copyright holder, it will cause practical difficulties. It will also hinder the development of LLMs by small domestic entities.

134.4. The expression “*storage*” used in Section 14(a)(i) of the Copyright Act is only in the context of “*reproduction*” and not storage in an independent manner. Therefore, if there is no reproduction, there cannot be any



infringement on the basis of storage alone. There is no independent right of storage available to the copyright owner.

Submissions of Mr. Adarsh Ramanujan, *Amicus Curiae*

135. Mr. Adarsh Ramanujan made the following submissions:

135.1. In terms of Section 14 (a)(i) of the Copyright Act, collection of data and storage of the same would amount to reproduction. Therefore, infringement occurs when Open AI stores the data of ANI. Even if there is a one-time storage, it would result in infringement.

135.2. It is likely that while building its corpus of ‘raw data’, Open AI has used and copied ANI’s publicly available works. Even if stored ‘raw data’ is deleted after the training is complete, it does not change the fact that storage has occurred in the first place.

135.3. Reproduction of ANI’s works during the training process either in the original text form or as a collection of numerical tokens, on a *prima facie* view, constitutes copyright infringement. Such temporary copies of copyrighted works do not qualify as “transient” or “incidental” copies protected under Section 51(1)(b) of the Copyright Act.

135.4. Even if the purpose of storage is to mine data, the purpose of storage or the intention behind it is not relevant for the purposes of infringement under Section 51 of the Copyright Act.

135.5. The subsequent processes of vectorization/tokenization raise technical issues and therefore, whether these acts amount to infringement can be decided only in a trial.

135.6. Whether the training process involves further reproduction of ANI’s data, can also be determined only in a trial.



Issue No.3

Whether the defendants' use of plaintiff's copyrighted data qualifies as 'fair use'⁴⁴ in terms of Section 52 of the Copyright Act, 1957.

SUBMISSIONS

Submissions of the Defendant (OPEN AI)⁴⁵

136. Mr. Amit Sibal, senior counsel appearing on behalf of Open AI made the following submissions:

136.1. Section 14 of the Copyright Act has to be read with Section 52 of the Copyright Act as it uses the words “*subject to provisions of this Act*”. Accordingly, Section 52 is not an exception, but a part of what is infringement as defined in Section 14 of the Copyright Act. Therefore, the submission on behalf of ANI that Section 52 should be interpreted in a narrow manner, cannot be sustained. On the contrary, Section 52 of the Copyright Act should be liberally construed. Reliance in this regard is made on ***B.D. Bhandari*** (supra) and ***Chancellor Masters & Scholars of The University of Oxford v. Narendera Publishing***⁴⁶.

136.2. Under Section 2(m) of the Copyright Act, a copy of the work would be “*infringing copy*” only if the requirements of Sections 14, 51 and 52 of the Copyright Act are fulfilled, as Section 2(m) uses the expression “*in contravention of the provisions of this Act*”. Reliance in this regard is placed on ***B.D. Bhandari*** (supra) and ***Narendera Publishing*** (supra).

⁴⁴ To be read as ‘fair dealing’.

⁴⁵ Since this defence was raised on behalf of the defendant, it is deemed appropriate that the submissions of the defendant are recorded first.

⁴⁶ 2008 SCC OnLine Del 1058.



136.3. The Copyright Act as originally enacted in 1957 used the words “*research or private study*” in its fair dealing clause. The Notes on clauses for Copyright Amendment Bill, 1994 (which was enacted as Copyright Amendment Act, 1994) makes it clear that the amendment was proposed in 1992 to expand the “*unduly narrow interpretation of the words “private study”*”. The latest amendment in the Copyright Act, 1957 included the words “*personal use*”, in addition to the already existing words “*private use*” and “*research*”.

136.4. There is no restriction on “*private use*” being only for non-commercial purposes. This view is reinforced from a perusal of sub-clauses (ii) and (iii) of Section 52 (1)(a) of the Copyright Act, which do not make the distinction between commercial and non-commercial. Wherever the Copyright Act makes a distinction between commercial and non-commercial uses, it has been stipulated therein. Reference in this regard may be made to Section 52(1)(ad) and Section 32(6)(d) of the Copyright Act. Merely because the use by Open AI is commercial, would not make it unfair. Reference in this regard is made to the judgment in ***Super Cassettes Industries v. Hamar Television Network***⁴⁷.

136.5. ANI has wrongfully placed reliance on the judgment of ***Rupendra Kashyap v. Jiwan Publishing House***⁴⁸. In the said case, the defendant no.3/CBSE had outsourced printing of its previous years’ question papers to third parties/defendants no.1 and 2, hence it was held that the defence under Section 52 would not be available as it is available only to a defendant conducting private study or research and not to a third party. ***B.D. Bhandari***

⁴⁷ 2010 SCC OnLine Del 2086.

⁴⁸ 1996 SCC OnLine Del 466.



(supra) holds that even a body corporate is entitled to the defence under Section 52(1)(a) of the Copyright Act.

136.6. The Explanation to Section 52(1)(a) was also added in 2012. In terms of the Explanation to Section 52(1)(a) of the Copyright Act, the expression “*not itself being an infringing copy*”, is restricted to “*computer programme*” and not to the other forms of storage. The Explanation’s language and intent is clear that the protection of Section 52(1)(a) extends to works stored in an electronic medium for the purposes mentioned under the said provision. The limiting language of the Explanation within the commas only applies to computer programmes.

136.7. If ANI’s argument is to be accepted, no electronic storage of a copy of a work published in an electronic form (such as an online article which is freely accessible over the internet) would ever be eligible for fair dealing protection under Section 52(1)(a). This would mean that making physical copies (for instance, by hand, a typewriter, or a mechanical medium) of a work would be capable of protection under Section 52(1)(a), while the same act done in an electronic medium would be barred from a fair dealing defence.

136.8. ANI wrongly contends that the defence under Section 52(1)(a) of the Copyright Act would not be available to storage as storage is covered under Section 52(1)(n) and 52(1)(p) of the Copyright Act. The aforesaid Sections cover specific types of storage and on that basis, it cannot be said that storage is not covered under Section 52(1)(a) of the Copyright Act.

136.9. Unlike ANI’s works, which are meant for human consumption for its expressive elements, the use of works for training models by Open AI is non-expressive.



136.10. The use of Open AI would amount to “*private use*” as mentioned in sub-clause (i) of Section 52(1)(a) of the Copyright Act, as it is only for the purposes of training of the Open AI’s LLM. If Section 52(1)(a)(i) of the Copyright Act is interpreted to mean that the exception is available only for personal use, the word “*private*” would be rendered otiose.

136.11. The word “*research*” must be given a broad and liberal interpretation so that it is not limited in any manner to a private context. Reliance in this regard is placed on the judgment of Canadian Court in ***CCH Canadian*** (supra).

136.12. Open AI’s storage and use of training data falls within the ambit of “*private use*”, including “*research*”, which is protected under Section 52(1)(a)(i) of the Copyright Act and does not amount to infringement of copyright.

136.13. Indian Courts have adapted the US four-factor test in India. By various judgments of the Indian Court, the US four-factor test has been adapted under Indian copyright law for the purposes of Section 52(1)(a) of the Copyright Act. The Indian judgments which have relied upon the four-factor test are as follows:

- i. ***India TV Independent News Service v. Yashraj Films***⁴⁹
- ii. ***Super Cassettes Industries Limited v. Chintamani Rao***⁵⁰
- iii. ***B.D. Bhandari*** (supra)
- iv. ***Narendera Publishing*** (supra)

Indian Courts have also recognized that the most crucial test out of the four-factor test is transformative character of the use. The use of publicly available

⁴⁹ (2012) SCC OnLine Del 4298.

⁵⁰ (2011) SCC OnLine Del 4712.



works forming part of the training data for training the LLMs is highly transformative. [Ref. *Narendera Publishing* (supra) and *B.D. Bhandari* (supra)].

136.14. Open AI's ChatGPT is not a substitute for subscription to news agencies, such as ANI.

Intervenors supporting OPEN AI

- i. *Submissions on behalf of Intervenor - IGAP Project LLP (I.A. 4616/2025)*

137. Mr. Akhil Sibal, senior counsel appearing on behalf of the intervenor in I.A. 4616/2025 made the following submissions:

137.1. When originally enacted, Section 52(1)(a)(i) of the Copyright Act limited the purposes to “*research or private study*”. The provision was amended in 1994 to “*private use including research*”.

137.2. The use of copyrighted data to train LLMs meets both, the purpose and the fairness limitations as envisaged under Section 52(1)(a) of the Copyright Act.

137.3. Section 52(1)(a) of the Copyright Act does not provide that commercial uses would not be covered therein. In contradistinction, Section 52(1)(k)(ii), Section 52(1)(l) and Section 52(1)(o) of the Copyright Act have excluded commercial uses. Hence, the legislature has deliberately chosen not to put an element of commerciality in Section 52(1)(a) of the Copyright Act.

137.4. Reliance placed by ANI on *Jiwan Publishing House* (supra) and *B.D. Bhandari* (supra) to submit that commerciality defeats fair dealing is misplaced. *Jiwan Publishing House* (supra) was a case where a publisher of infringing books claimed that since its customers are using the books for



personal use, the publisher's use should also be considered private or personal, which was rejected by the Court.

137.5. **B.D. Bhandari** (supra) holds that a commercial use is not fair in the context of Section 52(1)(h) of the Copyright Act and not in the context of Section 52(1)(a) of the Copyright Act. There is no general proposition laid down in the aforesaid judgment that commercial use defeats fair dealing.

137.6. The purposes mentioned in Section 52(1)(a) of the Copyright Act, such as “*research*”, “*review*” and “*reporting of current news and current affairs*” are well recognized as permissible uses of copyright, which are often carried out for commercial purposes.

137.7. While interpreting terms such as “*research*”, “*review*” and “*private use*”, the Court should apply the doctrine of updating construction i.e. the Parliament, while enacting a law is expected to anticipate future developments. Accordingly, the terms “*research*”, “*private use*” and “*review*” may not be limited to acts of human beings alone but to machine learning as well. Therefore, these terms should be interpreted liberally.

137.8. The term “*private use*” must be contradistinguished from the term “*personal use*”. Supreme Court in **Academy of General Education, Manipal v. B. Malini Mallya**⁵¹ has adopted the expansive meaning of the word “*private*” by holding that a dance performance at an educational institution can also be within the meaning of “*private use*” for research, review or criticism. ‘*Pre-training dataset*’ obtained by LLMs is private in nature i.e. accessible only to LLM models themselves. In other words, this data set is not

⁵¹ (2009) 4 SCC 256.



publicly available for either access or download. Therefore, ‘use’ of any copyrighted material is totally private.

137.9. In ***B.D. Bhandari*** (supra), the Division Bench adopted the transformative test adopted by the American Courts for determining what constitutes fair use. The Court considered whether contribution in writing a ‘*guide book*’ is such that a nature of guide book is different than the textbook. It was held that if the purpose served by the guide book is substantially different from the purpose served by the text book, it would be treated as transformative.

137.10. In ***The Chancellor, Masters & Scholars of University of Oxford v. Rameshwari Photocopy Services (DB)***⁵², it was held that the purposes of use would determine whether a particular use qualifies as being fair. The fairness is determined by whether the extent is justified by its purpose. In the context of reproduction of literary works for creating course packs, the Court held that copying of entire articles was justified for the purposes of education.

137.11. The DNPA and IMI have wrongfully relied upon ***Bartz v. Anthropic*** (supra) to submit that the fair use defence for training is unavailable if the first copy is not authorized by the copyright holder. In ***Bartz v. Anthropic*** (supra), the defendant had used ‘*pirated books*’ for creating a permanent general-purpose library, which was held to be not fair use. In the present case, Open AI does not create any library and the data is used only for training purpose.

137.12. The US Courts in ***Bartz v. Anthropic*** (supra) and ***Kadrey v. Meta Platforms*** (supra) have come to the conclusion that the use of copyrighted materials for training LLM is transformative. In ***Bartz v. Anthropic*** (supra),

⁵² 2016 SCC OnLine Del 6229.



the Court drew parallels between the training process and a human being reading books and generating new works. It was observed that the entire purpose of training was not to replicate or supplant the original words on which the LLM was trained but to create something different.

137.13. Since the purpose and character of the use of ANI's work is only for training the data sets along with millions of other works being used for the same purpose, the same would amount to transformative purpose. Further, since the format of storage on the input side is totally different from that of the ANI's data, i.e. the data stored in a machine-readable algorithm which cannot be read by humans, the same would amount to transformative use.

137.14. In the present case, the objective of ANI and Open AI are fundamentally different. ANI is a news agency which is mainly focused on creating, verifying and distributing news which involved a process of fact-gathering, investigative journalism and contextualization of stories. On the other hand, LLMs serve as tools such as content creation, researching, brainstorming and idea generation, language translation, learning and education, creative writing, summarizing or paraphrasing. Therefore, activities of Open AI cannot substitute the activities carried out by ANI. No evidence has been provided by ANI that Open AI's use of AI generative use had led to a decrease in their market share.

Submissions of the Plaintiff (ANI)

138. Mr. Sidhant Kumar, counsel appearing on behalf of ANI made the following submissions:

138.1. Open AI's actions do not fall under "*private or personal use, including research*" as it is used purely for commercial purposes and for making profit.



It is settled law that the defence under Section 52(1)(a)(i) is unavailable to person who uses the copyrighted works for commercial exploitation [Ref. ***Jiwan Publishing House*** (supra)]. Consequently, Open AI's claim of use for personal and private research is also unsustainable as it uses ANI's works for a manifestly commercial purpose, that is, to earn revenues from enhanced functioning of the ChatGPT model.

138.2. The contention of Open AI that the condition of “*non-infringing copy*” is only for a computer programme is misconceived. The Explanation appended to Section 52(1)(a) of the Copyright Act extends the specified exemption to an electronically stored copy of a work subject to the condition that the work itself is lawfully stored. The Explanation is centred around the term “*any work*”. Consequently, the condition that the work must not “*itself*” be an infringing copy, has to be a reference to the “*work*” itself and not to a computer programme that may be incidentally involved. The word “*itself*” emphasizes that the concern is about the copy stored. It does not refer to some other copy, not the programme in the abstract, but the particular copy. This is significant because it echoes the concerns about the distinction between the lawfulness of the act (storage for fair dealing) and the lawfulness of the stored copy.

138.3. The opening line of Section 52(1)(a) expressly excludes computer programmes from its ambit. It is settled law that Explanation cannot enlarge the scope of the parent provision. Consequently, the Explanation cannot be read to provide a fair dealing right with respect to computer programmes that the parent clause itself has denied.



138.4. The Explanation to Section 52(1)(a) of the Copyright Act only provides protection to incidental storage of non-infringing copies. Therefore, permanent storage of ANI's work would amount to infringement of copyright.

138.5. Open AI's use of ANI's works would not qualify as "research" under Section 52(1)(a)(i) of the Copyright Act. Open AI just copies ANI's creative expression and generates responses without any creative input from Open AI.

138.6. Without prejudice to the aforesaid submissions, the four-factor test under the US Law has no application in the Indian copyright Law.

138.7. Open AI's claim of transformative use is flawed. The expressive and non-expressive element of raw data are not segregated before feeding the information in the LLM.

138.8. Open AI's ChatGPT works as an alternative to the ANI's publications and diminishes the ANI's market.

Intervenors supporting ANI

- i. Submission on behalf Intervenor – Digital News Publishers Association and Ors. (I.A. 2199/2025)

139. Mr. Rajshekhar Rao, senior counsel appearing on behalf of intervenors in I.A. 2199/2025 made the following submissions:

139.1. Under Section 52(1)(a) of the Copyright Act, Open AI has to show the following to be eligible to claim the defence to copyright infringement:

- (a) The works utilized were non-infringing copies as mandated under the Explanation to Section 52(1)(a) of the Copyright Act that have been lawfully accessed.



(b) The use of works was for any one (or more) of the purposes enumerated at Section 52(1)(a)(i), (ii) and (iii) of the Copyright Act.

(c) The use of the works amounted to fair dealing.

139.2. Storage of copyright materials which are publicly available by a commercial entity for disseminating information to the public can never be “*private or personal use*”. Under Section 52(1)(a) of the Copyright Act, Open AI’s LLM is only available to an individual researcher or student, this cannot be claimed by a commercial organization. Since the elemental purpose is dissemination of commercial platform to subscribers and users, the same cannot amount to “*private use*”.

139.3. Data sets containing copyrighted information are accessed and extracted unlawfully from sources without a license. Therefore, the defence of fair dealing cannot be invoked.

139.4. As per Indian judgments, the American four-factor test has only been applied as a guiding principle to determine fairness of a dealing and only has a persuasive value. There are significant structural and textual differences between the Indian concept of fair dealing and American test of fair use. To reiterate, right to ‘*transform*’ a work is explicitly recognized under the US Copyright Act, whereas no such recognition exists under the Indian Copyright Act.

139.5. The submission of Open AI that storage of works for extracting information in the process of training an AI model is a transformative purpose, is misplaced. The US Supreme Court in *Andy Warhol v. Lynn Goldsmith*⁵³

⁵³ 2023 SCC OnLine US SC 19.



held that merely because the secondary work adds new expression to the source material, it will not be transformative.

139.6. The ultimate purpose of training an LLM and that of a news platform is the same i.e. to disseminate accurate information to the public. When a secondary work shares a common purpose with the original, work ceases to be transformative. Reliance is placed on the observations made by the US District Court in *Kadrey v. Meta Platforms* (supra), where the Court has noted the unlawful advantage that training an AI model makes of creative expression and the effect it has on the market share for news and other fact-based expression.

ii. *Submission on behalf of Intervenor – Indian Music Industry and Ors. (I.A. 4027/2025)*

140. Mr. C.M. Lall, senior counsel appearing on behalf of intervenors in I.A. 4027/2025 made the following submissions:

140.1. The copies of works used by AI models are obtained from unlawful sources/works used without a license from the copyright owner.

140.2. Use of any copyright protected content for the purposes of training AI models does not *ex facie* amount to fair dealing of the works under Section 52(1)(a) of the Indian Copyright Act.

140.3. The Explanation, which was introduced by the 2012 Amendment to the Copyright Act, effectively incorporated an additional condition to Section 52(1)(a) that the storage of any work for the “*purposes*” under Section 51(1)(a) must be of a non-infringing copy of the work.

140.4. The word “*including*” occurring between “*storing of any work in any electronic medium for the purposes mentioned in this clause*” and “*the*



incidental storage of any computer programme” is the crucial connecting link. It is settled law that the usage of the word “*including*” in statutory provisions is conjunctive. The placement of commas does not control the meaning of the Explanation to Section 52(1)(a). The opening part of the Explanation also refers to “*any work*”.

140.5. The ultimate purpose of the Open AI is not research or private or personal use, but to develop and make available a commercial AI model to its subscribers/users. The word “*private*” would have to be viewed in the context of the word “*research*”. Open AI’s activities are not entitled to any defences available under Section 52 of the Copyright Act.

140.6. The wholesale reproduction of entire copyrighted content would not amount to fair dealing. The use of the word “*including*” in Section 52(1)(a) is in fact a limitation and not an expansion. The expression “*private use*” thus effectively ousts ‘commercial use’. The use by Open AI does not amount to “*transient or incidental storage*” which is exempted under Section 52(1)(b) of the Copyright Act.

140.7. The open-ended US Copyright Law Doctrine of ‘Fair Use’ is not sanctioned by the statute in India. The limitations under Section 52(1)(a) of the Copyright Act would have to be satisfied in order to invoke the defence specified therein. The concept of ‘transformative use’ applied by American Courts cannot be adopted in the Indian context. The Courts in UK and Australia have desisted from incorporating the ‘*fair use*’ principle into their domestic law.



140.8. The judgments of the US Courts in *Thomson Reuters v. Ross Intelligence*⁵⁴, *Bartz v. Anthropic* (supra) and *Kadrey v. Meta Platforms* (supra), support the view that use of the copyrighted material for training an LLM cannot be considered fair use.

iii. Written Submissions filed by Mr. Ankit Sahni, Advocate on behalf of the Intervenors – Federation of Indian Publishers (I.A. 470/2025)

141. The use of the expression “*namely*” in Section 52 makes it clear that only the acts that are specifically mentioned in Section 52 are exempted from being construed as infringement.

141.1. The defence under Section 52(1)(a)(i) of the Copyright Act is available to an individual user and not to a defendant who is exploiting copyrighted works for its commercial activities. [Ref. *TIPS Industries v. Wynn Music*⁵⁵]

141.2. The *de-minimis* defence taken by Open AI that the copyrighted work used by Open AI forms very miniscule part of the news related work being used by Open AI is misplaced as *de-minimis* is to be seen *qua* the extent of the work of the copyrighted owner that has been used by the defendant and not *qua* the entire copyrighted work of the various copyrighted owners being used by the defendant.

141.3. The four-factor test adopted in the US in respect of fair use would not be applicable in India.

⁵⁴ Case No. 1:20-cv-613-SB.

⁵⁵ 2019 SCC OnLine Bom 13087.



Submissions of Professor Arul George Scaria, *Amicus Curiae*

142. Prof. Scaria made the following submissions:

142.1. The fair dealing analysis requires a two-step assessment:

- i. The first step, whether the use falls under one of the specific purpose categories mentioned in Section 52 of the Copyright Act.
- ii. The second step, whether the use can be considered as “*fair*” or not.

142.2. The terms “*private use*” and “*research*” mentioned in Section 52(1)(a)(i) of the Copyright Act and the words “*criticism or review*” in Section 52 (1)(a)(ii) of the Copyright Act and “*reporting of current news events*” in Section 52(1)(a)(iii) of the Copyright Act must be interpreted in a manner which is consistent with the broader purpose of copyright law i.e. to promote creation and dissemination of new works.

142.3. As the learning process of the LLM does not involve giving access to any external human being and such learning is primarily done with the objective of promoting research, such storage would be covered within the broad ambit of Section 52(1)(a)(i) of the Copyright Act.

142.4. Insofar as interplay between Section 51 and Section 52 of the Copyright Act is concerned, the Court should take into account the opening words of Section 52 which provide that the “*following acts shall not constitute an infringement*”. Therefore, Section 52 should be interpreted by a Court in a liberal manner. Reference is made to Single Bench judgment in ***Chancellor, Masters & Scholars of the University of Oxford v. Rameshwari Photocopy***



*Services*⁵⁶ and the judgment of the Supreme Court in *B. Malini Mallya*⁵⁷ (supra).

142.5. There is no requirement in Section 52(1)(a) of the Copyright Act for Open AI to acquire lawful access to ANI's copyrighted works. Wherever a statute requires lawful access the same has been specifically provided in the statute [Ref. Section 52(1)(ab) and Section 52(1)(ad)].

142.6. Section 52(1)(a) of the Copyright Act specifically excludes computer programme as a computer programme is specially covered under other provisions of Section 52(1) of the Copyright Act. The Explanation to Section 52(1)(a) is in respect of incidental use of computer programme for the purposes of storage of a work.

142.7. There is no prohibition for '*commercial use*' under Section 52(1)(a)(i) of the Copyright Act. Therefore, the same is not a relevant factor for the purpose of Section 52(1)(a)(i) of the Copyright Act.

142.8. Section 29 of the Canadian copyright law also provides a similar fair dealing exception as provided under the Indian Copyright Act. In *CCH Canadian* (supra), Canadian Supreme Court held that the expression "*research*" used in Section 52(1)(a)(i) of the Copyright Act should be given a broad and liberal meaning. Reliance in this regard is also placed on judgment of the Canadian Supreme Court in *Society of Composers, Authors and Music Publishers of Canada v. Bell Canada*⁵⁸.

142.9. The four-factor test adopted in United States is not applicable in India. Therefore, the Courts in India are free to adopt an appropriate test to determine

⁵⁶ 2016 SCC OnLine Del 6713, ¶ 28.

⁵⁷ ¶¶ 38-39.

⁵⁸ [2012] 2 S.C.R. 326.



fair dealing. As is evident from the judgments passed by Indian Courts, there is no single approach that the Courts have used to determine fair dealing.

Submissions of Mr. Ramanujan, *Amicus Curiae*

143. Mr. Ramanujan made the following submissions:

143.1. Section 52(1)(a) of the Copyright Act should not be expanded beyond its expressed terms is apparent from the use of the term “*namely*”, which suggests that the list which follows is exhaustive.

143.2. The use of the words “*private or personal*” in Section 52(1)(a)(i) of the Copyright Act indicate that the exemption is intended to apply at a personal level to an individual and not to a defendant with a commercial motive. [Ref. *TIPS Industries v. Wynk Music* (supra)].

143.3. The four-factor test adopted by the US Courts cannot be adopted in India in the absence of any statutory framework. The same can be used only to supplement and not to supplant the statutory mandate of Section 52 of the Copyright Act. The Indian fair dealing exception is purpose limited, whereas the US ‘*fair use*’ exception is purpose agnostic. [Ref. *Super Cassettes v. Hamar Television* (supra); *Jiwan Publishing House* (supra) and *Syndicate of Press of University of Cambridge v. Kasturi Lal*⁵⁹].

143.4. The judgment of Division Bench in *B.D. Bhandari* (supra), clearly lays down that Section 52 is in the nature of an exception to Section 51 of the Copyright Act. Therefore, there should be a two-step analysis –

- (i) whether there is infringement and,

⁵⁹ 2005 SCC OnLine Del 1448.



- (ii) whether the infringing act falls in any of the exceptions provided under Section 52 of the Copyright Act.

143.5. The expression “*private use*” in Section 52(1)(a)(i) of the Copyright Act would not apply to Open AI in the present case. Even though the process adopted by Open AI for training is private, the use is not private. As a result of the training, the output which is produced is for the use by the public.

143.6. Reliance placed by Open AI on the judgment in ***B. Malini Mallya*** (supra) may not be apposite as the observations made in paragraph 38 of the said judgment that institutions can take benefit of Section 52 of the Copyright Act were ‘*sub silentio*’ observations.

143.7. The expressions “*private*” or “*personal*” used in Section 52(1)(a)(i) of the Copyright Act have been used in an overlapping manner and would apply only to use by individuals and not by corporations like Open AI. The judgment in ***Narendera Publishing*** (supra), was in the context of Section 52(1)(a)(ii) and not Section 52(1)(a)(i) of the Copyright Act. Sub-clauses (ii) and (iii) of Section 52(1)(a) do not use the word “*private*” or “*personal*”. The Single Bench applied the transformative test in the context of “*review*” under Section 52(1)(a)(ii) of the Copyright Act and not in the context of fair dealing.

143.8. In India, the test to be applied to determine fair dealing would be ‘*substantiality of copying*’ and ‘*motive to compete*’ as stated in ***Blackwood v. A.N. Parasuraman***⁶⁰, which was cited in ***Rameshwari Photocopy Services (DB)*** (supra).

143.9. In ***B.D. Bhandari*** (supra), the test that was adopted by the Division Bench in the context of Section 52(1)(h) was whether a guide book is a

⁶⁰ 1958 SCC OnLine Mad 62.



‘*derivative work*’, and the test of what would amount to derivative work was laid down in ***Eastern Book Company v. D.B. Modak*** (supra). In paragraph 119 of the said judgment, the said test of derivative work has also been extended to Section 52(1)(a)(ii) of the Copyright Act. It appears the expression ‘*derivative work*’ has been used synonymously with the expression ‘*transformative work*’. Therefore, both the judgments of the Single Bench and the Division Bench in ***B.D. Bhandari*** (supra) do not provide a definitive ruling that ‘*transformative work*’ would apply to Section 52(1)(a)(i) of the Copyright Act.

143.10. The exception of ‘*transformative use*’ is a part of the first factor of the four-factor test, which is applied in the United States of America and the same would not be applicable in India.

DISCUSSION AND ANALYSIS OF ISSUE NOS. 1 AND 3

144. Now, this Court shall examine the Issue Nos. 1 and 3, as follows:

A. Scope of Section 14(a)(i) of the Copyright Act, 1957

145. As observed in Issue No. 2, this Court is of *prima facie* view that the copyright in respect of original literary works available on the website of ANI vests with ANI. Therefore, ANI enjoys certain exclusive rights under Section 14 of the Copyright Act, which includes the right of “*reproduction*” under Section 14(a)(i) of the Copyright Act. The right of reproduction includes the right of storing the literary works owned by ANI “*in any medium by electronic means*”. For ease of reference, Section 14(a)(i) of the Copyright Act is set out below:



“14. Meaning of Copyright—

For the purposes of this Act, “copyright” means the exclusive right subject to the provisions of this Act, to do or authorise the doing of any of the following acts in respect of a work or any substantial part thereof, namely:—

(a) in the case of a literary, dramatic or musical work, not being a computer programme,—

(i) to reproduce the work in any material form including the storing of it in any medium by electronic means;”

[Emphasis supplied]

146. The original Section 14(1)(a)(i) as enacted on 4th June, 1957 only gave the owner of the literary work, the “*right to reproduce it in any material form*”. The phrase “*including the storing of it in any medium by electronic means*” was incorporated in Section 14(a)(i) through the Copyright (Amendment) Act, 1994 enacted on 9th June, 1994. A comparison between the original text of Section 14 (relevant part) and the amended text is tabulated below:

<u>THE COPYRIGHT ACT, 1957</u> [4 th June 1957]	<u>THE COPYRIGHT (AMENDMENT) ACT, 1994</u> [9 th June 1994]
14. (1) For the purposes of this Act, “copyright” means the exclusive right, by virtue of, and subject to the provisions of, this Act, — (a) in the case of a literary, dramatic or musical work, to do and authorise the doing of any of the following acts, namely: — (i) to reproduce the work in any material form,	14. For the purposes of this Act, “copyright” means the exclusive right subject to the provisions of this Act, to do or authorise the doing of any of the following acts in respect of a work or any substantial part thereof, namely:— (a) in the case of a literary, dramatic or musical work, not being a computer programme, (i) reproduce the work in any material form <u>including the storing of it in any medium by electronic means;</u>



147. By way of the said amendment, the right of reproduction was expanded to incorporate the digital storage of a literary work. This amendment was enacted to keep pace with the evolving technological realities.

148. The scope and ambit of Section 14(a)(i) of the Copyright Act insofar as storage of a literary work is concerned, is very wide as it uses the expression “storing of it in any medium by electronic means” and the expression “in any material form”. In terms of Section 14(a)(i) of the Copyright Act, the storage of a copyrighted literary work in any medium by electronic means, would therefore, amount to reproduction of the work.

149. Section 14 (a)(i) of the Copyright Act, when it speaks of storage, does not make a distinction between temporary or permanent storage. Similarly, under Section 51⁶¹ of the Copyright Act, the intention/purpose of storage is not relevant when determining an infringement claim. Thus, if a literary work is stored temporarily or permanently in any medium using electronic means, the same would violate the exclusive right of the copyright holder as per Section 14(a)(i) of the Copyright Act and would be an act of infringement under Section 51 of the Copyright Act.

150. However, it is pertinent to note that Section 14 of the Copyright Act uses the expression, “*subject to the provisions of this Act*”, thus making it clear

⁶¹ **51. When copyright infringed.**—Copyright in a work shall be deemed to be infringed—

(a) when any person, without a licence granted by the owner of the Copyright or the Registrar of Copyrights under this Act or in contravention of the conditions of a licence so granted or of any condition imposed by a competent authority under this Act—

(i) does anything, the exclusive right to do which is by this Act conferred upon the owner of the copyright, or

(ii) permits for profit, any place to be used for the communication of the work to the public where such communication constitutes an infringement of the copyright in the work, unless he was not aware and had no reasonable ground for believing that such communication to the public would be an infringement of copyright; or



that the exclusive rights therein would be subject to other provisions of the Copyright Act. The other provisions would also necessarily include Section 52 of the Copyright Act which provides for various acts that do not constitute infringement under the Copyright Act. In this regard, reference may be made to opening sentence of Section 52 of the Copyright Act:

“52. Certain acts not to be infringement of copyright.—(1) The following acts shall not constitute an infringement of copyright, namely—”

Thus, specific acts enlisted under Section 52 of the Copyright Act cannot be considered to be falling under Section 14 and 51 of the Copyright Act.

151. Therefore, while analysing the rights of copyright holder under Section 14 of the Copyright Act, the Court would also have to bear in mind Section 52 of the Copyright Act as that would make the ‘act’ which would have been otherwise infringing under Section 51, to be non-infringing. The expression “*infringing copy*” as defined under Section 2(m)⁶² of the Copyright Act also uses the words “*in contravention of the provisions of this Act*”, therefore, if the work/act is in accordance with provisions of Section 52 then it would not be an infringing copy.

152. This view also finds support from the observations made by the Division Bench in **B.D. Bhandari** (supra). In the said decision, the Division Bench observed that when a defendant raises a defence under Section 52 of

⁶² **2. Interpretation.**—In this Act, unless the context otherwise requires, —

...
(m) “infringing copy” means,—

(i) in relation to a literary, dramatic, musical or artistic work, a reproduction thereof otherwise than in the form of a cinematographic film;

...
if such reproduction, copy or sound recording is made or imported in contravention of the provisions of this Act;



the Copyright Act, the Courts are expected to give a finding whether the act of the defendant falls within the categories enlisted under the said section before giving a finding of infringement. The relevant paragraphs from **B.D. Bhandari** (supra) are set out below:

“38. It is, thus, clear that when a work is in public domain, the natural consequence is that it does not enjoy copy right and question of violation thereof would not arise at all inasmuch as anybody and everybody has right to use and/or copy the same. On the other hand, when we talk of “fair use”, it would in the context that there is someone enjoys copyright in that work, but the user thereof comes within the domain of eventualities provide under Section 52 of the Copyright Act. Thus, all the cases of fair dealing are infringements, but the law examines the same based on the principles of “modicum of creativity”, access and affordability based on social, economic, educational, industrial, etc. considerations of society.

39. What follows form the above is that the two concepts, viz., ‘public domain’ and ‘fair use’ work in altogether different fields. When the defence is taken that a particular work is in public domain, the question for consideration would arise as to whether the appellants enjoy copyright therein or not. On the other hands, when plea of fair use under Section 52 of the Copyright Act is adopted by the respondents what is expected from the Court on the application of the provisions of Section 52 of the Act is not to examine whether the activity is infringement is not, but to examine whether the conditions stated in a particular clause of Section 52 of the Act invoked by the respondents is satisfied or not.”

[Emphasis supplied]

153. The Single Bench in **Rameshwari Photocopy Services** (supra) also observed that if an act, which is an exclusive right of the copyright holder, is provided in Section 52 of the Copyright Act, it would not amount to copyright infringement. The relevant paragraph of **Rameshwari Photocopy Services**



(Single Bench) (supra) is set out below:

“28. Section 51 prescribes that copyright is infringed *inter alia* when any person does anything exclusive right to do which has been conferred by the Act on the owner of copyright. It follows, if there is no exclusive right, there is no infringement. Section 52 lists the acts which do not constitute infringement. Thus, even if exclusive right to do something constitutes copyright, if it finds mention in Section 52, doing thereof will still not constitute infringement and the outcome thereof will not be infringing copy within the meaning of Section 2(m). Section 55 also, entitles the owner of copyright to remedies by way of injunction, damages, accounts or otherwise as are conferred by law for infringement of a right, only when copyright is infringed and except as otherwise provided by the Copyright Act. Thus unless there is infringement of copyright within the meaning of the Act, owner of copyright is not entitled to sue.”

[Emphasis supplied]

An appeal was preferred against the Single Bench judgment. The Division Bench⁶³ did not interfere with the finding of the Single Bench with respect to interpretation of Section 52. In fact, the Division Bench noted that the counsel for the parties were in agreement that Section 52 is not a proviso or an exception to Section 51 of the Copyright Act.

154. Therefore, in my opinion, Issue No.1, as framed by this Court dealing with the storage of ANI’s copyrighted works by Open AI and Issue No.3 on the aspect of “*fair dealing*” would have to be considered together and cannot be considered independently.

155. In the present case, the defence taken by Open AI is under Section 52(1)(a)(i) of the Copyright Act.

⁶³ The Chancellor, Masters & Scholars of University of Oxford v. Rameshwari Photocopy Services (DB), 2016 SCC OnLine Del 6229, ¶ 25.



B. Legislative History of Section 52(1)(a)

156. Section 52 forms part of the Copyright Act from its inception in 1957. Thereafter, it has undergone multiple amendments, with the last amendment in 2012.

157. In the Copyright Act of 1957, Section 52(1)(a) as originally enacted, read as under:

“(a) *a fair dealing with a literary, dramatic, musical or artistic work for the purposes of –*

(i) ***research or private study;***

(ii) *criticism or review, whether of that work or of any other work.”*

[Emphasis supplied]

158. Section 52(1)(a) was amended in 1994 to read as under:

“(a) *a fair dealing with a literary, dramatic, musical or artistic work, not being a computer programme, for the purposes of –*

(i) ***private use, including research;***

(ii) *criticism or review, whether of that work or of any other work.”*

[Emphasis supplied]

159. In the ‘**Notes on clauses**’ in the amendment bill of 1994, the rationale given for the amendment was that “*an unduly narrow interpretation of words “private study” (as existed) may result in harassment to public.*”

160. Once again, Section 52(1)(a) was amended by Copyright (Amendment) Act, 2012. The amended provision which exists as on date, reads as under:

“(a) *a fair dealing with any work, not being a computer programme, for the purposes of –*

(i) ***private or personal use, including research;***

(ii) *criticism or review, whether of that work or of any other work*



- (iii) *the reporting of current events and current affairs, including the reporting of a lecture delivered to public.*

Explanation.—The storing of any work in any electronic medium for the purposes mentioned in this clause, including the incidental storage of any computer programme which is not itself an infringing copy for the said purposes, shall not constitute infringement of copyright.”

[Emphasis supplied]

161. The purpose of various amendments carried out in Section 52(1)(a) has been to widen the scope of Section 52(1)(a). More specifically in the Copyright (Amendment) Act, 2012, an explanatory provision to Section 52(1)(a) of the Copyright Act was incorporated, which carves out an exception in respect of ‘*digital reproduction rights*’ to align it with the amendment in Section 14(a)(i), recognizing digital storage as part of reproduction.

C. Interpretation of Section 52(1)(a) of the Copyright Act

162. It has been vehemently contended on behalf of ANI and intervenors supporting ANI that since Section 52 is in the nature of an exception to copyright infringement the same should be interpreted narrowly. However, Open AI contends that Section 52 of the Copyright Act has to be construed broadly, fulfilling the objective of the Copyright Act as it has been interpreted by various Courts.

163. While interpreting Section 52(1) of the Copyright Act, a Coordinate Bench of this Court in ***Narendera Publishing*** (supra) held that Section 52 of the Copyright Act has to be construed liberally. The relevant paragraph of the



said judgment is set out below:

“33. The doctrine of fair use then, legitimizes the reproduction of a copyrightable work. Coupled with a limited copyright term, it guarantees not only a public pool of ideas and information, but also a vibrant public domain in expression, from which an individual can draw as well as replenish. Fair use provisions, then must be interpreted so as to strike a balance between the exclusive rights granted to the copyright holder, and the often competing interest of enriching the public domain. Section 52 therefore cannot be interpreted to stifle creativity, and the same time must discourage blatant plagiarism. It, therefore, must receive a liberal construction in harmony with the objectives of copyright law. Section 52 of the Act only details the broad heads, use under which would not amount to infringement. Resort, must, therefore be made to the principles enunciated by the courts to identify fair use.”

[Emphasis supplied]

164. The matter went up in appeal before the Division Bench. However, the Division bench⁶⁴ did not disturb the aforesaid finding of Single Bench. In fact, a holistic reading of the judgment of the Division Bench indicates that Section 52 should not be read as a proviso or an exception to Section 51. Rather, Section 52 has been construed liberally by the Division Bench.

165. At this juncture, it would be appropriate to examine how similar exceptions to copyright infringement are interpreted internationally. In **CCH Canadian** (supra), the Supreme Court of Canada held that fair dealing exception is an integral part of the Copyright Act (of Canada) rather than simply a defence. Therefore, in order to maintain a proper balance between the rights of copyright owners and users, it should not be interpreted

⁶⁴ Syndicate of the Press of the University of Cambridge v. B.D. Bhandari, 2011 SCC OnLine Del 3215.



restrictively. The relevant paragraph from the said judgment is set out below:

*“Before reviewing the scope of the fair dealing exception under the Copyright Act, it is important to clarify some general considerations about exceptions to copyright infringement. **Procedurally, a defendant is required to prove that his or her dealing with a work has been fair; however, the fair dealing exception is perhaps more properly understood as an integral part of the Copyright Act than simply a defence. Any act falling within the fair dealing exception will not be an infringement of copyright. The fair dealing exception, like other exceptions in the Copyright Act, is a user’s right. In order to maintain the proper balance between the rights of a copyright owner and users’ interests, it must not be interpreted restrictively.** As Professor Vaver, supra, has explained, at p. 171: “User rights are not just loopholes. Both owner rights and user rights should therefore be given the fair and balanced reading that befits remedial legislation.””*

[Emphasis supplied]

166. The legal principle that emerges from a reading of the aforesaid judgments is that Section 52 of the Copyright Act strikes a balance between the exclusive rights granted to the owner of the copyright and the competing public interest of encouraging creativity and disseminating knowledge. Hence, Section 52 is not in the nature of a proviso or an exception to the Section 51 of the Copyright Act.

167. In my respectful view, Section 52 independently defines the rights/privileges of a user in respect of copyrighted works and must not be interpreted in a restrictive manner. It is an integral part of the Copyright Act and therefore, it has to be given broad and liberal interpretation.

D. Scope of Section 52(1)(a)

168. To determine whether the storage by Open AI falls under Section



52(1)(a) of the Copyright Act, a two-step examination is required:

- i. Whether the storage by Open AI is for one of the specific purposes mentioned in Section 52(1)(a) of the Copyright Act. (Purpose Test)
- ii. Whether the storage can be considered as “*fair dealing*” or not. (Fairness Test)

I. PURPOSE TEST

169. In the present case, it is an admitted position that Open AI electronically stores ANI’s original literary works, at least on a temporary basis. Even if literary works of ANI collected and stored by Open AI are deleted after the completion of training, it would not alter the position that temporary storage occurred in the first place.

170. Hence, to satisfy the first requirement under Section 52(1)(a) of the Copyright Act, it has to be determined whether the purpose for which Open AI has stored ANI’s literary works falls under Section 52(1)(a) of the Copyright Act. The relevant extracts from Section 52(1)(a) providing the purposes mentioned therein are set out below:

“52. Certain acts not to be infringement of copyright.—(1) The following acts shall not constitute an infringement of copyright, namely—

(a) a fair dealing with any work, not being a computer programme, for the purposes of—

- (i) private or personal use, including research;**
- (ii) criticism or review, whether of that work or of any other work;**
- (iii) the reporting of current events and current affairs, including the reporting of a lecture delivered to public.**

Explanation.—*The storing of any work in any electronic medium for the purposes mentioned in this clause, including the incidental storage of*



any computer programme which is not itself an infringing copy for the said purposes, shall not constitute infringement of copyright.”

[Emphasis supplied]

171. On a plain reading of Section 52(1)(a), it is manifest that sub-clauses (i), (ii) and (iii) of Section 52(1)(a) are disjunctive in nature. Therefore, if the purpose falls under any of the three sub-clauses of Section 52(1)(a), the use would be entitled to protection against an infringement action. In the present case, Open AI claims protection under Section 52(1)(a)(i) of the Copyright Act, i.e. *“private or personal use, including research”*.

172. In support of its contention that Open AI does not fulfil the requirements of the purpose test under Section 52(1)(a)(i), ANI has made the following submissions:

- a. The private use/research cannot be for commercial purposes.
- b. The first copy of the work has to be a non-infringing copy.
- c. The use by Open AI does not amount to *“private or personal use, including research”*

I will deal with each of these separately.

a. Whether commercial use is entitled to protection under Section 52(1)(a)(i)

173. On behalf of ANI, it is contended that the phrase *“private or personal use, including research”* under Section 52(1)(a)(i) of the Copyright Act would not be applicable in the present case since ANI’s works are being used by Open AI for commercial purposes and to earn profits. Moreover, Open AI is a business venture which profits from large scale reproduction, adaptation and distribution of copyrighted content.



174. *Per Contra*, it is contended by Open AI that limiting the expression “*private*” use including “*research*” only for non-commercial purposes and to individuals would amount to a narrow interpretation of Section 52 (1)(a) of the Copyright Act. According to Open AI, wherever non-commercial use is intended, it has been specifically provided in the statute.

175. It has been argued on behalf of the intervenors IMI and DNPA that since Section 52(1)(ad)⁶⁵ of the Copyright Act makes a reference to “*non-commercial personal use*”, the said limitation should also be read in Section 52(1)(a) of the Act.

176. I am unable to agree with the said submission. A perusal of various provisions of Section 52(1) of the Copyright Act clearly demonstrate that the legislature was aware of the distinction between ‘*commercial*’ and ‘*non-commercial*’ use and thought it fit to exclude commercial use only in specific cases.

177. Wherever the protection under Section 52(1) of the Copyright Act is provided only for non-commercial use, it has been specifically provided. For instance, Section 52(1)(ad) makes it clear that the protection therein is available only for “*non-commercial personal use*”. Similarly, Section

⁶⁵ (ad) the making of copies or adaptation of the computer programme from a personally legally obtained copy for **non-commercial personal use**;



52(1)(k)(ii)⁶⁶, Section 52(1)(l)⁶⁷, Section 52(1)(n)⁶⁸ and Section 52(1)(o)⁶⁹ of the Copyright Act, specifically, stipulate that the benefit of the exceptions contained therein would be available for non-commercial purposes only. The requirement of non-commercial use is conspicuously absent in Section 52(1)(a).

178. The defence under Section 52(1)(a) of the Copyright Act is not restricted to non-commercial use can also be gauged from the fact that a person may be writing a book review that could be done on a commercial basis, and yet be covered under Section 52(1)(a)(ii) of Copyright Act. In *Super Cassettes v. Hamar Televisions* (supra), while dealing with the defence under Section 52(1)(a)(ii) of the Copyright Act, it was held that commercial use of a copyrighted work would not make it unfair or take it away from the ambit of fair dealing.

179. Similarly, even “*reporting of current events and current affairs*” can be for commercial purposes. For example, a journalist or a news agency reporting a current event, may be a paid activity and yet the journalist or the news agency would be entitled to take the defence under Section 52(1)(a)(iii) of the Copyright Act. Therefore, in my view, the limitation of non-commercial use cannot be read into Section 52(1)(a)(i). Merely because the use of a defendant is commercial would not mean that the defendant would not be

⁶⁶ (k) the causing of a recording to be heard in public by utilising it, —

(i) ...

(ii) as part of the activities of a club or similar organisation which is not established or **conducted for profit**;

⁶⁷ (l) the performance of a literary, dramatic or musical work by an amateur club or society, if the performance is given to a **non-paying audience**, or **for the benefit of a religious institution**;

⁶⁸ (n) the storing of a work in any medium by electronic means by a **non-commercial public library**, for preservation if the library already possesses a non-digital copy of the work;

⁶⁹ (o) the making of not more than three copies of a book (including a pamphlet, sheet of music, map, chart or plan) by or under the direction of the person in charge of a **non-commercial public library** for the use of the library if such book is not available for sale in India;



entitled to take the defences provided under Section 52 of the Copyright Act unless there is a specific bar on commercial use as stipulated in some of the provisions of Section 52 of the Copyright Act.

180. In this regard, a reference may be made to the judgment of the Canadian Supreme Court in ***CCH Canadian*** (supra). In the said case, the Canadian Supreme Court was dealing with Section 29⁷⁰ of Copyright Act of Canada (R.S.C., 1985, c. C-42) which is *pari materia* with Section 52(1)(a) of the Copyright Act. It was held that “*research*” should be given a liberal interpretation and that research should not be confined to non-commercial use. The relevant paragraph of the said judgment is set out below:

*“51. The fair dealing exception under s. 29 is open to those who can show that their dealings with a copyrighted work were for the purpose of research or private study. **“Research” must be given a large and liberal interpretation in order to ensure that users’ rights are not unduly constrained. I agree with the Court of Appeal that research is not limited to non-commercial or private contexts. The Court of Appeal correctly noted, at para. 128, that “[r]esearch for the purpose of advising clients, giving opinions, arguing cases, preparing briefs and factums is nonetheless research.” Lawyers carrying on the business of law for profit are conducting research within the meaning of s. 29 of the Copyright Act.**”*

[Emphasis supplied]

181. Other common law jurisdictions have also taken a similar view that merely because the use is a commercial one, may not take it outside the defence of “*fair dealing*”⁷¹.

⁷⁰ **Research, private study, etc.**

29. Fair dealing for the purpose of research, private study, education, parody or satire does not infringe copyright.

⁷¹ *Society of Composers, Authors and Music Publishers of Canada v. Bell Canada*, 2012 SCC 36; *Television New Zealand Ltd v. Newsmonitor Services Ltd*, [1994] 2 NZLR 91.



182. ANI has relied upon paragraph 65 of **B.D. Bhandari** (supra) to argue that the defences under Section 52 of the Copyright Act would not be available for commercial use. The relevant paragraph of **B.D. Bhandari** (supra) relied upon by ANI is set out below:

“65. It is clear from the reading of the aforesaid provision that this would not come to the aid of the respondents who have published the guide for commercial exploitation. In view of the law laid down in various cases, commercial exploitation will take the matter outside the ambit of ‘fair use’. The defence of ‘fair use’ as provided under Section 52(1)(h) of the Act is only available to the teacher and pupil to reproduce the literary work in the course of instructions or examination paper setter to reproduce the literary work as part of the questions or to the pupil to reproduce the literary work as answers to such questions. The respondents are neither teachers nor students nor a person giving or receiving instruction. The defence of ‘fair use’ under Section 52(1)(h) of the Act only provided to a teacher and pupil and not otherwise. If this defence is allowed to a publisher, then it would result in a situation where every publisher, without permission from the owner of copyright, would reproduce the ad verbatim literary text from the educational textbooks prescribed by the University into their books and exploit the same for commercial gains and benefits and that too without giving any royalty or payment to the right owners towards such reproduction. This practice would obviously discourage creativity of authors who put their skill, labour, years of knowledge, expertise, etc. into these educational books and encourage infringers like the respondents to reproduce ad verbatim literary text from original publications into their infringing books on the pretext of ‘fair use’. Obviously, therefore, Section 52(1)(h) of the Act has no application to the facts of this case.”

[Emphasis supplied]

183. The aforesaid observations made in **B.D. Bhandari** (supra) were in the context of Section 52(1)(h) of the Copyright Act [now Section 52(1)(i) of the Copyright Act]. For ease of reference, Section 52(1)(i) of the Copyright Act



is set out below:

- “(i) *the reproduction of any work—*
 (i) *by a teacher or a pupil in the course of instruction; or*
 (ii) *as part of the questions to be answered in an examination; or*
 (iii) *in answers to such questions;”*

The scope of Section 52(1)(i) is entirely different from Section 52(1)(a) as Section 52(1)(i) deals with reproduction of any work in the “*course of instruction*” or “*as part of the questions to be answered in an examination*” or “*in answers to such questions*”. The Division Bench held that the defence under Section 52(1)(h) is only applicable to a teacher and pupil and not to a publisher.

184. Ultimately, the Division Bench held that the act of the defendants was non-infringing and amounted to creating a ‘*derivative work*’ of distinctive character. In ***B.D. Bhandari*** (supra), the guidebooks were being sold by defendants for commercial purposes and yet the act of the defendants was held to be non-infringing.

185. In ***Rameshwari Photocopy Services (DB)*** (supra), while dealing with the defence under Section 52(1)(h) [now section 52(1)(i) of the Copyright Act], the Division Bench gave the benefit of Section 52(1)(h) to the defendant, even though the use of the defendant was for commercial purposes. It was held that a teacher or a pupil cannot be expected to buy photocopiers and the argument that there cannot be an intermediary for photocopying, was rejected.

186. The plaintiff has also placed reliance on paragraph 21 of the judgment of this Court in ***Jiwan Publishing House*** (supra), in support of its contention that the defence under Section 52(1)(a) of the Copyright Act would not be available to a publisher who publishes a book for commercial exploitation.



The relevant paragraph of the said judgment is set out below:

*“21. In the written statement of defendants 1 and 2 vide para 4 of the preliminary objections, as also in the reply to the application, reference has been made to Section 52(1)(a)(i) and S 52(1)(b)(iii) of the Copyright Act as providing the defences available to defendants 1 and 2. In the Copyright Act, there is nothing like sub clause (iii) to be found under clause (b) of sub-section (1) of Section 52 of the Act. That reference appears to be mistaken. It would, therefore, suffice to deal with the plea raised by reference to S. 52(1)(a)(i) of the Act, which provides inter alia, that a fair dealing with the literary work for the purpose of research or private study does not constitute infringement of copyright. The words ‘research or private study’ have been substituted by the words ‘private use including research’ by the Copyright (Amendment) Act, 1994 (Act 381 of 1994). What is contemplated is a defence to the person conducting research or private study who while doing so, if dealing fairly with a literary work, may not incur wrath of the copyright having been infringed. **But, if a publisher publishes a book for commercial exploitation and in doing so infringes a Copyright, the defence under section 52(1)(a)(i) would not be available to such a publisher though the book published by him may be used or be meant for use in research or private study. The defence raised by defendants 1 and 2 based on Section 52(1) (a) (i) is not available to them and the plea so raised has to be rejected.**”*

[Emphasis supplied]

187. In ***Jiwan Publishing House*** (supra), the defendants were publishing previous years’ question papers of the examinations conducted by Central Board of Secondary Education (‘CBSE’), without taking any license from CBSE. The plaintiff contended that it was the exclusive licensee of CBSE in respect of reprinting and reproduction of the question papers and actions of the defendants were in breach of plaintiff’s rights. In those circumstances, it was held that the defence under Section 52(1)(a)(i) of the Copyright Act is available to a person conducting “research” or “private study”. The said defence would not be available to a publisher whose book may be used for



“research” or “private study” by third parties such as students.

188. ANI has also placed reliance on the judgment of the Bombay High Court in *TIPS Industries v. Wynk Music* (supra) in support of its contention that where the defendants commercially exploit copyrighted material of the plaintiff, it would not be entitled to the defence under Section 52(1)(a) of the Copyright Act. Once again, in this case, there was no “private use” of the copyrighted material by the defendants. On the contrary, the defendants were making available the plaintiff’s music to third party users on a commercial basis through online streaming services and the defence taken was that the said users were using the content in a private or personal capacity.

189. In the present case, as is explained hereinafter⁷², the use by Open AI is only for the purposes of training its LLM, which is completely an internal process and does not involve providing the copyrighted material of ANI to any third party. Even after the training is complete, the training data is never made available either in its natural language or tokenized form to any person. Therefore, the judgments in *Jiwan Publishing House* (supra) and *TIPS Industries v. Wynk* (supra) are of no assistance to ANI in the present case.

190. In light of the discussion above, in my view, merely because the use of Open AI is for a commercial purpose would not mean that Open AI is not entitled to take the defence under Section 52(1)(a) of the Copyright Act.

b. Requirement of “non-infringing copy” under Section 52(1)(a) of the Copyright Act.

191. The Explanation to Section 52(1)(a) of the Copyright Act makes it clear

⁷² Refer to sub-heading ‘Whether use by Open AI amounts to “private or personal use, including research”’.



that storage (temporary/permanent) of a work in electronic means for purposes mentioned in Section 52(1)(a) of the Copyright Act shall not constitute infringement of copyright subject to the test of “*fair dealing*”. The Explanation to Section 52(1)(a) of the Copyright Act has been the subject matter of intense debate between the counsel for the parties. For the sake of convenience, Section 52(1)(a) is set out below:

“52. Certain acts not to be infringement of copyright.—(1) The following acts shall not constitute an infringement of copyright, namely—

(a) a fair dealing with any work, not being a computer programme, for the purposes of—

- (i) private or personal use, including research;**
- (ii) criticism or review, whether of that work or of any other work’**
- (iii) the reporting of current events and current affairs, including the reporting of a lecture delivered to public.**

Explanation.—The storing of any work in any electronic medium for the purposes mentioned in this clause, including the incidental storage of any computer programme which is not itself an infringing copy for the said purposes, shall not constitute infringement of copyright.”

[Emphasis supplied]

192. The submission made on behalf of Open AI is that the use of ‘,’ (commas) before and after the expression “*including the incidental storage of any computer programme which is not itself an infringing copy for the said purposes*” makes it clear that the limitation of not being an infringing copy is applicable only in respect of a computer programme.

193. It is the submission of ANI as well as the intervenors, DNPA and IMI that Open AI stores a copy of ANI’s works that has been obtained in an unauthorised manner. The storage of an infringing copy of a work would not



qualify as a defence in light of Explanation to Section 52(1)(a) of the Copyright Act. Any other interpretation of the Explanation would expand the scope of Section 52(1)(a) of the Copyright Act as “*computer programmes*” have already been excluded from the main text of Section 52(1)(a).

194. No doubt “*computer programmes*” have been excluded from the main text of Section 52(1)(a) of the Copyright Act. What the Explanation refers to is the “*incidental storage of any computer programme*” i.e. computer programme is incidental to the electronic storage of a work. The limitation of not being an infringing copy is, therefore, only applicable to the incidental storage of computer programmes. This position is quite clear from the positioning of the ‘,’ (commas) in the Explanation.

195. The Explanation to Section 52(1)(a) does not expand the scope of protection with respect to “*computer programmes*” and merely clarifies that storage of an infringing copy of a computer programme would not be protected even if such storage is incidental. Therefore, ANI’s submission with respect to expansion of scope of Section 52(1)(a) cannot be accepted.

196. In Sections 52(1)(aa), 52(1)(ab) and 52(1)(ad), the words “*lawful*”/ “*legally obtained copy*” has been used only in respect of computer programmes. Therefore, it cannot be said that “*non-infringing copy*” would apply to storage of other works unless it has been specifically mentioned. For instance, in Section 52(1)(c)⁷³ it has been specifically mentioned that for the

⁷³ (c) transient or incidental storage of a work or performance for the purpose of providing electronic links, access or integration, where such links, access or integration has not been expressly prohibited by the right holder, **unless the person responsible is aware or has reasonable grounds for believing that such storage is of an infringing copy:**

Provided that if the person responsible for the storage of the copy has received a written complaint from the owner of copyright in the work, complaining that such transient or incidental storage is an infringement, such person responsible for the storage shall refrain from facilitating such access for a period of twenty-one days or till he receives an order from the competent court refraining from facilitating access and in case no such



purposes of providing electronic links, the person responsible must have reasonable grounds to believe that an infringing copy is not being stored.

197. If the interpretation given by ANI is accepted, it would imply that a researcher going to a public library and photocopying a page from a book for the purposes mentioned under Section 52(1)(a) of the Copyright Act would be entitled to defence of fair dealing. However, if the same researcher scans the said page electronically and stores the same in its device for the same purpose, the protection would not be available.

198. In my considered view, this distinction between storage in electronic medium and storage in the physical medium cannot be the intent of Section 52(1)(a) of the Copyright Act. Therefore, I do not find any merit in the contention that since Open AI stores an infringing copy of ANI's work, it would not be entitled to defence under Section 52(1)(a).

199. In *Bartz v. Anthropic* (supra), the US District Court clearly made a distinction between acquiring plaintiff's books in a lawful manner and obtaining copies from 'Shadow Libraries'. The US Court held that scanning of lawfully acquired books amounted to fair use, whereas obtaining copies from shadow libraries was not fair use.

200. In the present case, it is not the case of ANI that Open AI obtains copies of ANI's works from unauthorized sources or by breaking through the paywall of ANI. It is not disputed that Open AI obtains copies from information/data that is freely available on ANI's website. Therefore, in any event, it cannot be said that Open AI has obtained an infringing copy of ANI's works.

201. Therefore, in my view, under the Explanation to Section 52(1)(a) the

order is received before the expiry of such period of twenty-one days, he may continue to provide the facility of such access;



limitation of not being an infringing copy is applicable only in respect of a ‘computer programme’ and not in respect of the works stored in electronic mode.

c. Whether use by Open AI amounts to “*private or personal use, including research*”

202. To begin with, as noted in my findings in Issue no. 2, the cut-off dates for storage and training of the LLM models underlying ChatGPT were April 2022 (*for GPT 4*) and April 2024 (*for GPT 4o*), which is prior to the publication of articles used by ANI as instances of reproduction. During the course of hearing, ANI has not provided any further instances to show the extracted data from ANI’s website has been used for training. Therefore, there is no factual foundation in the plaint to reach a finding of infringement in respect of the training claim. However, since it is an admitted position that the original literary works of ANI are stored by Open AI at least temporarily during the course of training⁷⁴, this Court shall examine whether the storage of literary works by Open AI for training LLMs underlying ChatGPT qualifies as “*private or personal use, including research*” under Section 52(1)(a) of the Copyright Act.

203. Open AI submits that the purpose of the storage of ANI’s original literary works falls within of Section 52(1)(a)(i) i.e. “*private or personal use, including research*”. More particularly, Open AI’s electronic storage of ANI’s original literary works for training LLMs underlying ChatGPT, would fall

⁷⁴ Paragraph 11.3 of Response on behalf of the Defendant to the written submissions on behalf of Prof. Arul George Scaria, *Amicus Curiae*; Paragraphs 7, 13.1, 58.3 of Response on behalf of the Defendant to the written submissions on behalf of Mr. Adarsh Ramanujan, *Amicus Curiae*; Paragraphs 63, 76 of the Written Submissions on behalf of the Defendant dated 28th March, 2025.



within the ambit of “*private*” use, including “*research*”.

204. It is the submission of ANI duly supported by Mr. Adarsh Ramanujan that “*private use*” under Section 52(1)(a)(i) of the Copyright Act is applicable only to an individual/human being.

205. The interpretation advanced by Mr. Ramanujan is that the words “*private*” and “*personal*” are overlapping and they draw meaning from each other. Therefore, the interpretative rule of *noscitur a sociis* would require that the meaning of “*private*” and “*personal*” be interpreted in an analogous and cognate sense.

206. According to Prof. Scaria, the Copyright Act does not define the term “*private*”. He refers to the definition of “*private*” given in Meriam Webster Dictionary and Cambridge Dictionary to submit that the term “*private*” cannot be confined to individual persons.

207. The **Merriam Webster Dictionary** gives the meaning of “*private*” as - “*intended for or restricted to the use of a particular person, group, or class*” and “*belonging to or concerning an individual person, company, or interest*”. The **Cambridge Dictionary** defines “*private*” as “*only for one person or group and not for everyone*”.

208. From the aforesaid definitions, it transpires that the term “*private*” cannot be confined to an individual and can apply to a closed group or a company. The expression “*personal*” may be confined to individual persons, however, the term “*private*” would include other private entities, including private companies.

209. As can be discerned from the amendments carried out in Section 52(1)(a) from time-to-time, the legislature has throughout been expanding the scope of Section 52(1)(a) of the Copyright Act. The latest amendment of 2012



provides that both “*private*” as well as “*personal*” use are eligible for protection under Section 52(1)(a) of the Copyright Act. If the submission of Mr. Ramanujan that “*private*” and “*personal*” are to be used interchangeably, is accepted, then the word “*private*” would be rendered otiose.

210. The view that the term “*private*” should be given an expansive meaning also finds support from the judgment of the Supreme Court in ***B. Malini Mallya*** (supra). The respondent-plaintiff therein claimed copyright ownership over a ballet dance form which was performed by the performers at the appellant/defendant’s institute in New Delhi without any permission. In this case, the Supreme Court gave the benefit of Section 52 to the defendant institute by holding that a dance performance at an educational institution would be covered under Section 52(1)(a)(i) of the Copyright Act as it amounted to “*private use*”. The relevant extracts from ***B. Malini Mallya*** (supra) are set out below:

“38. The High Court, in our opinion, should have clarified that the appellants can also take the statutory benefit of the provisions contained in clauses (a), (i) and (l) of sub-section (1) of Section 52 of the Act.

39. Section 52 of the Act provides for certain acts which would not constitute an infringement of copyright. When a fair dealing is made, inter alia, of a literary or dramatic work for the purpose of private use including research and criticism or review, whether of that work or of any other work, the right in terms of the provisions of the said Act cannot be claimed. Thus, if some performance or dance is carried out within the purview of the said clause, the order of injunction shall not be applicable.

40. Similarly, the appellant being an educational institution, if the dance is performed within the meaning of provisions of clause (i) of sub-section (1) of Section 52 of the Act strictly, the order of injunction shall not apply thereto also. Yet again, if such performance is conducted before a non-paying audience by the appellant, which is an



institution if it comes within the purview of amateur club or society, the same would not constitute any violation of the said order of injunction.”

[Emphasis supplied]

211. In the present case, Open AI stores the literary works in a closed space without access to the public. The data obtained by the LLMs for training purposes is used for private purposes. The said data is accessible only to the LLM models themselves. The said data is not publicly available to any human entity either for access or for download. Therefore, in my opinion, the use amounts to being purely private.

212. The term “*research*” is not defined under the Copyright Act. However, Merriam-Webster Dictionary gives the meaning of “*research*” as “*to study or investigate carefully*”.

213. The Madras High Court in *Blackwood v. A.N. Parasuraman* (supra), used the dictionary meaning of “*research*” for the purpose of interpreting Section 52(1)(a) of the Copyright Act. The relevant paragraphs of the said judgment are reproduced here below:

“154. There is no definition of the word “research” in the Act, and, therefore its meaning has to be ascertained from ordinary English, usage. In Funk and Wagnalls English dictionary “research” is said to mean, I am extracting that meaning which appears to be most relevant to the present context, “Diligent protracted investigation especially for the purpose of adding to human knowledge, studious enquiry.”
155. In “The Shorter Oxford Dictionary” the meaning attributed is “An investigation directed to the discovery of some fact by careful study of a subject; investigation, inquiry into things.”...”

[Emphasis supplied]

214. The process of “*research*” is generally an intermediary process in all cases. It is undertaken before an output is generated. Research happens before



an academician writes a book or a research paper, or before a judge writes a judgment, or before a medicinal chemist develops a drug, or before any such developmental process. It is normally a closed activity and is not disclosed to the general public. It is always the output of the research that is communicated to the public.⁷⁵

215. From the submissions made in the present case, it transpires that the process of training LLMs underlying ChatGPT involves machine learning of the stored literary works by screening and organising them. In this process, the data is analysed by making extractions from the literary works and converting them into machine-readable training inputs. This data is used to train the LLMs iteratively to rectify errors and generate the most probable statistical output. Accordingly, the process of training of LLMs underlying ChatGPT can be considered as a form of research for generating new knowledge and advancing artificial intelligence systems.

216. Clearly, when Section 52(1)(a) was last amended in 2012, the legislature could not have imagined the advent of artificial intelligence (AI) and/or LLMs. However, when the Court interprets the aforesaid provisions in the light of modern-day technological developments, the Court has to give a liberal and purposive interpretation using the ‘*doctrine of updating construction*’.

217. The doctrine of updating construction gives effect to the true intention of the legislature by interpreting words in a statute in accordance with changing social conditions and developments in science and technology. This doctrine was applied by a Constitutional Bench of Supreme Court in ***State***

⁷⁵ Merriam Webster Dictionary defines the verb “research” as “to study or investigate carefully” <<https://www.merriam-webster.com/dictionary/research>>.



*(Through CBI/ New Delhi) v. S.J. Choudhary*⁷⁶. In the said case, the Supreme Court was dealing with a question of law i.e. “*whether the opinion of a typewriter expert is admissible in evidence under Section 45 of the Indian Evidence Act, 1872*”. The Supreme Court allowed the appeal and observed that at the time of enactment of the Indian Evidence Act in 1872, the typewriters were “*practically unknown*”. Hence, typewriting was not specifically mentioned in Section 45 of Indian Evidence Act. The Supreme Court relied upon Francis Bennion’s *Statutory Interpretation*, 2nd Edn., to give a wider meaning to the expression “*science*” used in Section 45⁷⁷ of the Indian Evidence Act, to hold that the opinion of an expert in typewriting is admissible. The relevant paragraph of *S.J. Choudhary* (supra) which quotes Bennion’s *Statutory Interpretation* is reproduced below:

“10. *Statutory Interpretation by Francis Bennion, 2nd Edn., Section 288 with the heading “Presumption that updating construction to be given” states one of the rules thus: (p. 617)*

“****

(2) It is presumed that Parliament intends the court to apply to an ongoing Act a construction that continuously updates its wording to allow for changes since the Act was initially framed (an updating construction). While it remains law, it is to be treated as always speaking. This means that in its application on any date, the language of the Act, though necessarily embedded in its own time, is nevertheless to be construed in accordance with the need to treat it as current law.

***”

⁷⁶ (1996) 2 SCC 428.

⁷⁷ **45. Opinions of experts.** - When the Court has to form an opinion upon a point of foreign law or of science, or art, or as to identity of handwriting, or finger impressions, the opinions upon that point of persons specially skilled in such foreign law, science or art, or in questions as to identity of handwriting or finger impressions are relevant facts.

Such persons are called experts.



In the comments that follow it is pointed out that an ongoing Act is taken to be always speaking. It is also, further, stated thus: (pp. 618-19)

“In construing an ongoing Act, the interpreter is to presume that Parliament intended the Act to be applied at any future time in such a way as to give effect to the true original intention. Accordingly the interpreter is to make allowances for any relevant changes that have occurred, since the Act’s passing, in law, social conditions, technology, the meaning of words, and other matters. Just as the US Constitution is regarded as ‘a living Constitution’, so an ongoing British Act is regarded as ‘a living Act’. That today’s construction involves the supposition that Parliament was catering long ago for a state of affairs that did not then exist is no argument against that construction. Parliament, in the wording of an enactment, is expected to anticipate temporal developments. The drafter will try to foresee the future, and allow for it in the wording.

An enactment of former days is thus to be read today, in the light of dynamic processing received over the years, with such modification of the current meaning of its language as will now give effect to the original legislative intention. The reality and effect of dynamic processing provides the gradual adjustment. It is constituted by judicial interpretation, year in and year out. It also comprises processing by executive officials.”

[Emphasis supplied]

218. In my opinion, the expression “research” should be given an updating construction by taking into account the modern-day technical advancements. With the advent of these technologies, research/ learning is no longer confined to humans. It is now being done through Artificial Intelligence. However, ultimately the research is at the behest of humans and for the benefit of humans. For example, Section 52(1)(i) of the Copyright Act exempts the act of reproduction by a teacher “in the course of instruction” from the ambit of



infringement. If tomorrow a human teacher is replaced by an AI bot or Robot⁷⁸ to say that the said exception could only be used by a human would be a regressive view. Such an approach would limit societal progress. Therefore, the acts of further research cannot be confined to acts of human being alone and the same would extend to machine learning as well.

219. Therefore, on a *prima facie* view, from the above analysis, this Court is of the view that the process of training LLMs underlying ChatGPT undertaken by Open AI using stored literary work of ANI falls under “*private or personal use, including research*” as provided in Section 52(1)(a) of the Copyright Act and fulfils the purpose test.

II. FAIRNESS TEST

220. Now, I will examine the second aspect, i.e. whether the acts of Open AI can be considered “*fair dealing*” under Section 52 of the Copyright Act.

221. The term “*fair dealing*” has not been defined in the Copyright Act. Therefore, it would be useful to examine how Courts in India have interpreted the term “*fair dealing*”.

222. Open AI has placed reliance on ***Narendera Publishing*** (supra). In the said judgment, the plaintiff was a well-known publisher of academic books, who filed a suit against the defendants, who were selling guide books, which reproduced excerpts from the plaintiff’s books.

223. In ***Narendera Publishing*** (supra), the Court placed partial reliance on

⁷⁸ Economic Times, ‘I am Sophie’: 17-year-old from a small UP city’s Hindi-medium school builds AI robot teacher that can teach any subject, The Economic Times (Nov. 29, 2025), <https://economictimes.indiatimes.com/news/india/i-am-sophie-class-12-student-from-a-small-up-citys-hindi-medium-school-builds-ai-robot-teacher-that-can-teach-any-subject/articleshow/125653059.cms?from=mdr>.



the four-factor test⁷⁹ which is a part of the Copyright Act of 1976 of the United States of America. The Court placed emphasis on the transformative character of the use i.e. whether the purpose of the use is substantially different from the purpose served by the prior work. It was also observed that if the use is transformative, it will not serve as a market substitute of the original. Hence, such use will not affect the market share of the original work. Based on the above, the Court held that the defendant's work can be said to be 'transformative' and hence, entitled to defence under Section 52(1)(a)(ii) of the Copyright Act. The relevant paragraphs dealing with this aspect from *Narendera Publishing* (supra) are set out below:

*“34. One crucial test, of the four-factor test, as developed by the American courts, is the transformative character of the use. **The Courts should in cases like the present ask whether the purpose served by the subsequent (or infringing) work is substantially different (or is the same) from the purpose served by the prior work. The subsequent work must be different in character; it must not be a mere substitute, in that, it not sufficient that only superficial changes are made, the basic character remaining the same, to be called transformative.** This determination, according to the Court is closely knit with the other three factors, and therefore, central to the determination of fair use. **If the work is transformative, then it might not matter that the copying is whole or substantial. Again, if it is transformative, it may not act as a***

⁷⁹ Section 107 - Notwithstanding the provisions of sections 106 and 106A, the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include—

- (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
 - (2) the nature of the copyrighted work;
 - (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
 - (4) the effect of the use upon the potential market for or value of the copyrighted work.
- The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors.



market substitute and consequently, will not affect the market share of the prior work.”

[Emphasis supplied]

224. The judgment in *Narendera Publishing* (supra) was taken up in appeal by the plaintiffs before the Division Bench. The Division Bench⁸⁰ decided the said appeal along with another appeal dealing with Section 52(1)(h) [now Section 52(1)(i)]. The Division Bench re-emphasized that the purpose of the Copyright Act was to balance the rights of the author on one hand and the society on the other. After referring to the judgments of US Courts dealing with the aspect of fair use, the Division Bench held that the guidebooks used by the defendant in *B.D. Bhandari* (supra), are in the nature of ‘*derivative works*’. The guide books have a different purpose or character than the original work of the plaintiffs/appellants. Hence, it was held that there was no infringement. On this aspect, the observations of the Division Bench are set out below:

*“78. It follows from the above discussion that when a published work is prescribed as text book, a ‘guide’ can be published provided it fulfils the test of a ‘derivative work’. At the same time, this guide book has to be materially different from the textbook. Guide book should not be verbatim reproduction of the text book. If a guide book reproduces the original work to a great measure without its contribution, it may amount to plagiarism and may infringe the copyright law of the person in original work. **On the other hand, if guide book written in a different format, albeit, dealing with the subject matter which is contained in the original book, with a purpose to help, assist and support the students, the problems given in the text book, then it would be a work different from the original work. The guiding principle is to see as to whether such guide book provides explanation and/or step by step process for reaching the answer, detailed analysis of any problem with***

⁸⁰ Syndicate of the Press of the University of Cambridge v. B.D. Bhandari, 2011 SCC OnLine Del 3215.



the objective of making the task simpler in understanding the subject provided in the textbook. In writing this guide book, if the contribution of aforesaid nature is made, it may be treated as a work which is somewhat different in nature. This is the crucial test which is to be applied and has been characterized as the transformative character of the use by the American Courts. Therefore, the task would be to ascertain as to whether the purpose served by the guide book is substantially different from the purpose served by the textbook. If the guide book is different in character and not a mere substitute of the original work/textbook, it would be treated as transformative. However, the character must be substantially different and it is not sufficient that superficial changes are made with basic character of the textbook creeping in the guide book. In that eventuality, it would not qualify as a work of transformative character. If this derivative (sic) work in the guidebook has assumed different character, it would not amount to infringement of the original work, though it would have been drawn, to certain extent, from the original work. And, in that sense, it would amount to ‘fair use’ as described in Amazon. Com. Inc. (supra).

119. We have already formulated the test for ascertaining as to whether in a given situation publication of a guide book would be violative of a copyright in a textbook in detail while dealing with RFA (OS) No. 21 of 2009. It is not necessary to repeat the same. Suffice is to state that it will be permissible to publish a guide book if it qualifies the proper description of a guide book and becomes a derivative work of the author who has produced such guide book.”

[Emphasis supplied]

225. Based on the aforesaid findings, the judgment of the Single Bench in *Narendera Publishing* (supra) was upheld by the Division Bench. The Division Bench in *B.D. Bhandari* (supra) used the expression ‘*derivative work*’ as opposed to the ‘*transformative work*’ used by Single Bench.

226. Mr. Ramanujan, *Amicus Curiae*, submits that the judgment in *Narendera Publishing* (supra) applied the ‘*transformative*’ test in the context of “review” under Section 52(1)(a)(ii) of the Copyright Act and not in the



context of Section 52(1)(a)(i). In ***B.D. Bhandari*** (supra), the test finally adopted under Section 52(1)(h) was whether a guidebook is a ‘*derivative work*’ (used synonymously for ‘*transformative work*’). In paragraph 119 of the Division Bench judgment in ***B.D. Bhandari*** (supra), the test of derivative work was extended to Section 52(1)(a)(ii) also.

227. Even though Division Bench judgment in ***B.D. Bhandari*** (supra) is in the context of Section 52(1)(a)(ii) of the Copyright Act, there is no reason why the fair dealing principles discussed therein cannot be applied in the context of Section 52(1)(a)(i).

228. In ***Rameshwari Photocopy Services (DB)*** (supra), the plaintiffs, who were major publishing houses, claimed that the defendant, a photocopy services shop situated within the precincts of Delhi School of Economics (University of Delhi), was photocopying pages from the books published by the plaintiffs. In the facts of the said case, the Court held that the defendant was entitled to defence under Section 52(1)(h) as it existed then [now Section 52(1)(i)]. The Division Bench in ***Rameshwari Photocopy Services*** (supra) also held that the concept of “*fair dealing*” would not be applicable in cases involving Section 52(1)(i). In the Indian context, the concept of “*fair dealing*” is only a part of Section 52(1)(a) of the Copyright Act and not the remaining provisions of Section 52.

229. The Division Bench specifically noted that the four-factor test is a part of the Copyright Act in United States of America (Section 107 of the US Copyright Act). Therefore, it was held that the decisions of the Courts in US, applying the aforesaid test, would have no persuasive value in the Indian context.

230. In ***Blackwood v. A.N. Parasuraman*** (supra), cited by Mr. Adarsh



Ramanujan, Madras High Court adopted a test for determining fair dealing based on quantity and quality of the work that has been reproduced.

231. In **ESPN Star Sports v. Global Broadcast News**⁸¹, a Coordinate Bench of this Court took note of the test laid down by the English Courts. In paragraph 34, the Court observed that the following three factors have to be considered for fair dealing:

“34. The test indicated in Ashdown was based on the Textbook “The Modern Law of Copyright and Designs (3rd edn, 2000)” by Laddie, Prescott and Vitoria. The court held that:

“the authors suggest that the success or failure of the defence depends on three factors : (1) whether the alleged fair dealing is in commercial competition with the owner's exploitation of work, (2) whether the work has already been published or otherwise exposed to the public and (3) the amount and importance of the work which has been taken.” ”

232. In **Super Cassettes v. Chintamani Rao** (supra), Coordinate Bench of this Court observed that when assessing fair dealing, the Court must consider proportionality and whether there is any intellectual input or any original mental exercise undertaken by the defendant.

233. The Division Bench in **India TV Independent News Service v. Yashraj Films** (supra), made an observation to the effect that “*even in India, Fair Use is determined on the same four factors*” i.e. the four-factor test used in US. However, in the said judgment, the Division Bench did not apply the four-factor test but denied interim injunction to the plaintiff by applying the ‘*de-minimis*’ principle.

234. From the abovesaid analysis of judgments of the Indian Courts on the

⁸¹ 2008 SCC OnLine Del 1766.



aspect of fair dealing, the position that emerges is that different Courts have adopted different tests according to the facts and circumstances of the case before them. There is no single test which has been adopted uniformly or consistently by the Indian Courts. There is broad consensus amongst the counsel in the present case also that the US four-factor test is not applicable in India and there is no single uniform test for examining whether an act under Section 52(1)(a) is fair dealing or not.

235. In light of the aforesaid, this Court shall attempt to formulate factors to determine “*fair dealing*” which are suitable in the facts and circumstances of the present case. Whether use by Open AI amounts to fair dealing or not is ultimately a question of fact, degree and the overall impression carried by the Court.

236. As discussed above, Section 52 of the Copyright Act balances the author’s right and the public interest. Therefore, Court has to consider if the training of LLMs by using the original literary works of ANI violates ANI’s rights. To determine fairness, it has to be examined whether the aforesaid use by Open AI of ANI’s works prejudice interests of ANI, resulting in competition, actual or potential damage to ANI’s legitimate commercial interests. Simultaneously, consideration must be given to the extent to which the activities of Open AI serve broader public interests, including scientific research, technological innovation, access to information, and the dissemination of knowledge. Thus, in the facts and circumstances of the case, in my opinion, the factors relevant for determining fair dealing can be summarised as under:

- a. Whether Open AI’s use of ANI’s original literary works is limited to training its LLMs underlying ChatGPT?



- b. Whether the usage of ANI's literary works by Open AI would result in economic competition and would prejudice the legitimate interests of ANI, thereby causing actual or potential damage to ANI?
- c. Whether the functions performed by Open AI through ChatGPT serve the overall public interest?

237. The first two factors examine how Open AI's use of ANI's works does not affect the author's rights, and the third factor examines whether functions performed by Open AI serve public interest.

238. The aforesaid factors are broadly in line with the principles emerging from Article 9 of **Berne Convention for the Protection of Literary and Artistic Works (1886)** (hereinafter 'Berne Convention'). Article 9 of the Berne Convention which discusses the rights available to a copyright holder and its possible exceptions is reproduced below:

Article 9 Right of Reproduction:

1. Generally; 2. Possible exceptions; 3. Sound and visual recordings

(1) Authors of literary and artistic works protected by this Convention shall have the exclusive right of authorizing the reproduction of these works, in any manner or form.

(2) It shall be a matter for legislation in the countries of the Union to permit the reproduction of such works in certain special cases, provided that such reproduction does not conflict with a normal exploitation of the work and does not unreasonably prejudice the legitimate interests of the author.

(3) Any sound or visual recording shall be considered as a reproduction for the purposes of this Convention.

[Emphasis supplied]

239. The Statements of Objects of the Copyright (Amendment) Act, 1983⁸²

⁸² **Statement of Objects and Reasons of Amendment Act 23 of 1983.**—India is a member of the two International conventions on copyright, namely the Berne Convention for the Protection of Literary and



provides that the Copyright Act is in conformity with the Berne Convention.
240. Now, I proceed to examine the facts and circumstances of the present case in light of the aforesaid factors.

a. Whether Open AI's use of ANI's original literary works is limited to training its LLMs underlying ChatGPT?

241. It has already been observed in the findings in respect of Issue No.2 that ANI has not been able to establish memorization of data or its regurgitation, or communication of the exact copies of ANI's works to the public.

242. The documents filed by Open AI (*document nos. 8 and 9 of the documents filed by defendant along with the reply to I.A. 43500 of 2024*) demonstrate that the LLMs underlying ChatGPT are not trained to reproduce or communicate the training material to the public. Instead, the training material enables it to generate new responses and perform novel tasks when presented with novel prompts.

243. This Court has not been given any instance where Open AI has used the literary works of ANI for any purposes other than for training. Hence, in my opinion, it would be safe to conclude that the storage by Open AI is only for the purposes of training Open AI's LLMs underlying ChatGPT.

Artistic Works (1948 Brussels Text) and the Universal Copyright Convention (1952). The Copyright Act, 1957, conforms to the provisions of these two conventions.



b. Whether the usage of ANI's literary works by Open AI would result in economic competition and would prejudice the legitimate interests of ANI, thereby causing actual or potential damage to ANI?

244. ANI is a news agency that is in the business of collecting and distributing news. ANI has engaged a team of professionals to collectively produce original news content, which is published on ANI's syndicated news feed. The subscribers of ANI, which include news organisations, can access the content posted on ANI's syndicated news feed. A non-exclusive license is granted to ANI's subscribers, who can use ANI's works for commercial purposes, specifically excluding any right of further syndication of ANI's works. Apart from the revenue generated from these subscription and licensing agreements, ANI earns by publishing the news articles on its website, YouTube channel and other social media accounts. The user traffic on its websites and social media channels is directly correlated to revenue raised through advertising by ANI.⁸³

245. On the other hand, Open AI's ChatGPT is an application, which allows the users to interact with LLMs. LLMs serve as tools for content creation, research, language translation, learning and education, summarising, paraphrasing, creation of images based on text description. ChatGPT is available for free to all users. There is also an option for users to sign up for a premium monthly subscription that makes additional features and capabilities available. Open AI also offers an enterprise version for businesses to purchase. The API (Application Programming Interface) version of Open AI allows the developers to integrate the benefits of Open AI's models into their

⁸³ ¶¶ 9-15 of the plaint.



own application. The developers pay to access the API.⁸⁴

246. As is evident from aforesaid, the functions performed by ChatGPT are quite distinct from the functions performed by ANI. ChatGPT has multifarious functions, whereas business of ANI is limited to news reporting and syndication of news. In the context of news, ChatGPT only provides a summary or snippets of the news articles or reports carried by ANI along with a reference to its website. As held in Issue No.2, the expressions used are not the same as used by ANI. Even when a user searches for daily news, the responses given by ChatGPT are similar to that of an AI-enabled search engine, using RAG technique to provide multiple sources for the snippets of the news content. In sum, the purpose and character of use of copyrighted works by Open AI is fundamentally different from that of the news articles of ANI. Hence, it cannot be said that responses produced by ChatGPT are substitutes for news articles published by ANI.

247. This view also finds support from the judgment of the United States District Court for Northern District of California in **Bartz v. Anthropic** (supra), wherein it has been observed that use of copyright to train LLMs was essentially transformative:

“ **The copies used to train specific LLMs did not and will not displace demand for copies of Authors’ works, or not in the way that counts under the Copyright Act.**”

*Instead, Authors contend generically that training LLMs will result in an explosion of works competing with their works - such as by creating alternative summaries of factual events, alternative examples of compelling writing about fictional events, and so on. This order assumes that is so (Opp. 22-23 (citing, e.g., Opp. Exh. 38)). **But***

⁸⁴ Document 8 of the documents filed by the defendant along with reply to I.A. No. 45300 of 2024.



Authors' complaint is no different than it would be if they complained that training schoolchildren to write well would result in an explosion of competing works. This is not the kind of competitive or creative displacement that concerns the Copyright Act. The Act seeks to advance original works of authorship, not to protect authors against competition.

In short, the purpose and character of using copyrighted works to train LLMs to generate new text was quintessentially transformative. Like any reader aspiring to be a writer, Anthropic's LLMs trained upon works not to race ahead and replicate or supplant them — but to turn a hard corner and create something different. If this training process reasonably required making copies within the LLM or otherwise, those copies were engaged in a transformative use."

[Emphasis supplied]

248. Similar observations were also made by the United States District Court for Northern District of California in *Kadrey v. Meta Platforms*, (supra) i.e. the purpose should be transformative:

"This factor favors Meta. There is no serious question that Meta's use of the plaintiffs' books had a "further purpose" and "different character" than the books that it was highly transformative. The purpose of Meta's copying was to train its LLMs, which are innovative tools that can be used to generate diverse text and perform a wide range of functions. Cf. Oracle, 593 US. at 30 (transformative to use copyrighted computer code "to create a new platform that could be readily used by programmers") Users can ask Llama to edit an email they have written, translate an excerpt from or into a foreign language, write a skit based on a hypothetical scenario, or do any number of other tasks. The purpose of the plaintiffs books, by contrast, is to be read for entertainment or education.

The plaintiffs do not meaningfully disagree about Llama's purpose. To the contrary, they acknowledge that LLMs have "end uses" including serving "as a personal tutor," assisting "with creative ideation," and helping users "generate business reports." And several of the plaintiffs testified to using LLMs for various purposes, all distinct



from creating or reading in expressive work like a novel or biography for instance, to find recipes, get tax or medical advice, translate documents, or conduct research. All of these functions are different from the use to which the plaintiffs books are generally put. So copying the books to develop a tool that can perform those functions is a use with a different purpose and character than the books themselves.

[Emphasis supplied]

249. In *Authors Guild v. Google*⁸⁵, the authors of published books filed a copyright infringement action against Google claiming that Google made digital copies of the books and allowed the public to search the texts of the digitally copied books and see displays of snippets of the text, which amounted to infringement of the copyright of the authors. The United States Court of Appeals held that the more transformative the new work, the lesser will be the significance of commercialism. In the said case, Google was digitizing copyright protected works of authors. The Court held that the purpose of copying is highly transformative as the public display of text is limited and the activities of Google are not in the nature of market substitution of the original books. The mere fact that the activities of Google are commercial in nature and profit-motivated would not deny Google the defence of fair use.

250. Except for bare averments, nothing has been placed on record on behalf of ANI to show that they have lost any market share or there has been a reduction in subscription revenues due to operations of Open AI.

251. Hence, in my *prima facie* opinion, the use of ANI's works by Open AI would not result in market substitution of ANI's works and consequently, will

⁸⁵ 804 F.3d 202 (2nd Cir. 2015).



not affect its market share or cause actual or potential damage to ANI.

c. Whether the functions performed by Open AI through ChatGPT serve the overall public interest?

252. The public benefits flowing from such trained LLMs underlying ChatGPT are considerable. They are designed to assist users in analysing and generating text, improving access to information, supporting education, assisting scientific research, facilitating software development, enabling translation and communication, and creating tools for persons with disabilities. The technology is also capable of supporting research, learning, and knowledge dissemination at scale, thereby advancing science and the useful arts in a broader societal sense.

253. Training LLMs underlying ChatGPT therefore contributes to advancing scientific knowledge, developing innovative computational tools, disseminating information, promoting education, enhancing accessibility, and advancing artificial intelligence research. These objectives resonate with the broader constitutional and statutory commitment to advancing knowledge and learning. While the rights of copyright owners remain important, the societal benefits arising from scientific and technological research constitute a relevant consideration in assessing the fairness of a dealing.

254. Therefore, the factor of public interest also stands established.

255. Thus, on a *prima facie* view, all the factors for establishing the aspect of fair dealing stand satisfied in the present case and the fairness test stands fulfilled.



FINDINGS

256. In light of the discussion above, both the purpose test as well as the fairness test under Section 52(1)(a) stand fulfilled. Hence, in my *prima facie* view, Open AI's acts of storage of the literary works of ANI for the training of its LLMs would fall under Section 52(1)(a) of the Copyright Act and hence, would not amount to infringement.

BALANCE OF CONVENIENCE AND IRREPARABLE INJURY

SUBMISSIONS

Submissions on behalf of the Plaintiff (ANI)

257. Mr. Sidhant Kumar, counsel appearing on behalf of ANI made the following submissions:

257.1. Open AI, by offering direct quotes and summaries of ANI's work, is diverting traffic from ANI's website to ChatGPT thereby reducing the revenue generated by ANI through advertisements. Open AI's unlawful use and storage of ANI's works curtails ANI's commercial avenues including the ability to use AI to provide services including high level summaries of its works.

257.2. Open AI has entered into licensing arrangements and strategic partnerships with similarly-placed news agencies with articles, news and other works in the training and response generation process. However, it has consciously opted not to compensate ANI for using its data. Therefore, Open AI is unjustly enriching itself at the expense of ANI.

Submissions on behalf of the Defendant (OPEN AI)

258. Mr. Amit Sibal, senior counsel appearing on behalf of Open AI made



the following submissions:

258.1. The interim relief sought by ANI to delete the “*plaintiff’s works held and stored in whatever form*” by Open AI would amount to grant of the final relief sought in the suit. This would also be inconsistent with Open AI’s obligation under the laws of United States of America to preserve such training data.

258.2. ANI itself had offered Open AI a license to all its digital media video, imagery, photographic and/or news content owned or within the control of ANI for a fee of USD 7.5 million in terms of ANI’s letter dated 3rd October 2024. Therefore, it is clear that the claim of ANI is quantifiable and ANI can be compensated in monetary terms if ANI succeeds in the suit. On the other hand, Open AI cannot be compensated in monetary terms if an interim injunction is granted against Open AI, even if Open AI succeeds in the suit.

258.3. The AI platform of Open AI, i.e. ChatGPT, has transformed information in new ways and is being used all around the world for a variety of beneficial purposes. Any injunction granted would be detrimental to public interest.

258.4. Without prejudice to their rights, Open AI has stopped ANI’s website both for the purposes of training as well as RAG.

Intervenors supporting OPEN AI

- i. *Submissions on behalf of the Intervenor - Broadband India Forum (BIF) (I.A. 10766/2025)*

259. Mr. Arvind P. Datar, senior counsel appearing on behalf of the intervenor in I.A. 10766/2025 made the following submissions:

259.1. Any injunction granted will have cascading effects for several LLM



platforms as it would curtail access to information for LLM developers and the public. It would also require an LLM to take multiple licenses with each news agency, individual newspaper, magazine etc. Hence, when ANI has chosen not to engage ‘paywalls’ or ‘crawler blockers’, no injunction can be granted to access information that is freely available on the web.

259.2. Hindering access to publicly available data would degrade LLM’s performance and stifle information as developers would struggle to improve their AI models.

259.3. ANI has claimed damages in the suit and hence, no interlocutory injunction should be granted in favour of ANI. The issues raised by ANI would require large amount of evidence in the trial. Therefore, no injunction should be granted in the absence of a trial.

Submissions of Professor Arul George Scaria, *Amicus Curiae*

260. Prof. Scaria made the following submissions:

260.1. In the event the Court finds that the Open AI is required to take a license from ANI, it would be detrimental to the growth of LLMs and more particularly, LLMs being sought to be developed in India.

260.2. While granting an interim injunction, the Court should consider the two competing interests, i.e. protecting the interest of the authors and the broader public interest in developing LLMs in India.

260.3. It is an undisputed fact that at least some users use ChatGPT for gathering information about events and facts. Therefore, the Courts should take a broader view so as to enable fulfilment of this broader purpose.



Submissions of Mr. Adarsh Ramanujan, *Amicus Curiae*

261. Mr. Ramanujan made the following submissions:

261.1. ChatGPT is currently being used by millions of users in India and a blanket injunction order passed by this Court would adversely affect the said users.

261.2. Insofar as ANI seeks a direction that Open AI should obtain license from ANI in respect of its works, it is demonstrative of the fact that ANI can be compensated in monetary terms.

261.3. ANI is in a position to block its websites to prevent Open AI from scraping data therefrom for training of its LLM as also for the purposes of its search functionality.

DISCUSSION AND ANALYSIS

262. It is an admitted position that ANI has the ability to block its website vis-à-vis any third-party including Open AI. The opting-out option is available to ANI for blocking the third-party web crawlers from copying their data as well as from scraping their website for the search function/RAG. Despite having an option of opt-out, evidently ANI has not exercised the same. In fact, it has been stated on behalf of Open AI that it has internally blocked ANI's website from its web crawlers or bots for the purposes of scraping of data.⁸⁶ During the course of oral submissions Open AI has also submitted that Open AI itself has blocked ANI's website from '*ChatGPT search function/RAG*'.

263. ANI contends that there are various paid subscribers of ANI, whose websites carry the copyrighted data of ANI and Open AI can still access and copy ANI's copyrighted data from the aforesaid websites. Even if that be so,

⁸⁶ Document 21 of the plaintiff's documents filed along with the plaint.



it would be open for the subscribers of ANI also to block their websites to Open AI. Clearly, if they have chosen not to do so, it could be possibly for their own commercial reasons.

264. There is no averment in the plaint nor has any material been placed on record to show that the activities of Open AI have resulted in ANI losing subscribers or suffering a loss from its news syndication business.

265. ANI itself has offered a license to Open AI for its content for a fee of USD 7.5 million *vide* its communication dated 3rd October 2024.⁸⁷ Therefore, the claim of ANI is quantifiable and ANI can be compensated in monetary terms, if ANI succeeds in the present case. On the other hand, if there is an injunction granted against Open AI, it would have a significant effect on its functioning, which may not be possible to be compensated in monetary terms.

266. There is yet another aspect of public interest for not granting an interim injunction in favour of ANI and against Open AI. As discussed earlier, Artificial Intelligence (AI) including generative AI has brought about a transformational change in the lives of the people all over the world, including India. It has completely revolutionized the manner in which people seek and obtain information.

267. As noted earlier, AI and its applications are being used beneficially in several sectors, such as education, healthcare, financial support sector, agriculture and for providing other skill development resources. [Ref. *AI for Inclusive Societal Development*, October 2025 - Niti Aayog paper at page 7389 of Pleadings Vol. 6]

268. The key to success of an AI model is to access the information in public

⁸⁷ Document 24 of the plaintiff's documents filed along with the plaint.



domain. The development of LLMs and their success depends on availability of data. It would be economically unviable to develop an LLM if training of an LLM would require licenses from multiple sources.

269. Any interim injunction granted at this stage would, in my opinion, be detrimental to the growth of AI and more particularly, to the LLMs being developed in India. It would also have adverse impact on public interest, including millions of users of ChatGPT in India, many of whom would not be paid subscribers. Public interest is an important aspect or the fourth factor that the Courts have to consider while granting interim injunction, more particularly in intellectual property cases. [Pls. see. **Zydus Lifesciences v. E.R. Squibb**⁸⁸; **F-Hoffman-La Roche AG v. NATCO Pharma**⁸⁹; **Astrazeneca AB and Ors. v. Intas Pharmaceuticals Limited**⁹⁰].

FINDINGS

270. In light of the aforesaid analysis, on the touchstone of balance of convenience as well as irreparable loss and injury, this Court is not inclined to grant an interim injunction this stage.

CONCLUSION

271. Based on the discussion above, I am of the *prima facie* view that Open AI's act of storing ANI's original literary works for training LLMs underlying ChatGPT falls under Section 52(1)(a) of the Copyright Act and therefore, does not amount to infringement under Section 51 of the Copyright Act. I am also of the *prima facie* view that the outputs generated by ChatGPT using RAG

⁸⁸ FAO(OS) (COMM) 120/2025, decided on 12th January 2026.

⁸⁹ CS(COMM) 567/2024, decided on 24th March, 2025.

⁹⁰ 2020:DHC:3125.



technique does not amount to infringement under Section 51 of the Copyright Act since the outputs generated by Open AI were not substantially similar to ANI's original literary works. Further, ANI has failed to satisfy this Court that any memorization and regurgitation of ANI's original literary works has happened through the responses generated by ChatGPT.

272. In view of the discussion above, ANI has failed to make out a *prima facie* case for grant of interim injunction. Balance of convenience is also against grant of interim injunction. Irreparable injury would be caused not only to Open AI but also to the public at large, if an interim injunction is granted in favour of ANI at this stage.

273. Accordingly, I.A. 45300/2024 is dismissed.

274. Needless to say, any observations made herein are only for the purpose of adjudication of the aforesaid application and would have no bearing on the final outcome of the suit.

275. I would like to express my sincere appreciation to all the learned counsel who appeared in the matter as well as the learned *Amici Curiae* who rendered invaluable assistance to the Court in rendering this judgment.

**AMIT BANSAL
(JUDGE)**

JULY 24, 2026

Vivek/-