

# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

Impact Factor: 8.3 [www.ijesh.com](http://www.ijesh.com) ISSN: 2250 3552

## **Fair Dealing at the Frontier: Generative Artificial Intelligence, Section 52(1)(a) of the Copyright Act, 1957, and the Case for a Calibrated Text and Data Mining Regime in India**

**Parul**

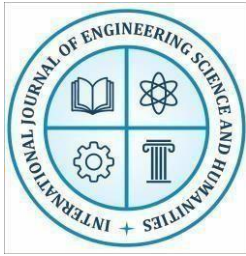
Assistant Professor, Chaudhary Ranbir Singh University, Jind

### **Abstract**

The training of generative artificial intelligence systems on copyright-protected material has become the defining doctrinal controversy of contemporary intellectual property law. Large language models are built by copying, tokenising and statistically encoding vast corpora of protected expression, an activity that engages the reproduction right in Section 14 of the Copyright Act, 1957, yet fits uneasily within any exception the Indian legislature has enacted. This paper examines how Indian copyright law has begun to answer that question and argues that the answer presently emerging is doctrinally fragile and institutionally misplaced. The analysis proceeds from the Delhi High Court's decision of 24 July 2026 in ANI Media Pvt. Ltd. v. Open AI OpCo LLC, the first substantive judicial engagement in India with model training, in which the Court held on a prima facie view that the storage of scraped news content for training falls within the fair dealing exception for private or personal use including research under Section 52(1)(a)(i), and refused an interim injunction. The paper reconstructs the two-limb purpose-and-fairness test applied in that ruling and evaluates it against the closed-list architecture of Section 52, the three-step test in Article 9(2) of the Berne Convention and Article 13 of TRIPS, and the Indian jurisprudence on fair dealing running from Civic Chandran through the Delhi University photocopying litigation.

It then situates the Indian position within a comparative field that has moved sharply in the opposite direction: the Munich Regional Court in GEMA v. OpenAI treating memorisation within model parameters as an act of reproduction outside the European text and data mining exceptions, the United Kingdom abandoning its proposed opt-out exception in March 2026, and the United States generating contradictory first-instance fair use outcomes with no appellate resolution. Against that background the paper assesses the Department for Promotion of Industry and Internal Trade Working Paper of December 2025, which proposes a statutory blanket licence administered through a Copyright Royalties Collective for AI Training. The paper concludes that neither an expansive judicial reading of Section 52(1)(a) nor an unconditional statutory licence is defensible, and proposes a calibrated legislative model built on a lawful-access-conditioned text and data mining exception, an enforceable machine-readable reservation of rights, mandatory training-data disclosure, and a separately governed remuneration mechanism for commercially deployed models.

**Keywords:** Generative Artificial Intelligence; Copyright Act, 1957; Fair Dealing; Section 52(1)(a); Text and Data Mining; ANI Media v. Open AI



# International Journal of Engineering, Science and Humanities

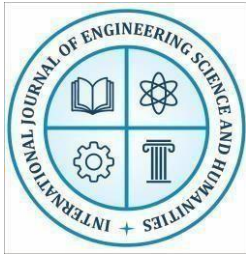
An international peer reviewed, refereed, open access journal

**Impact Factor: 8.3** [www.ijesh.com](http://www.ijesh.com) **ISSN: 2250 3552**

## 1. Introduction

Copyright law has repeatedly been asked to absorb technologies its drafters could not have anticipated. The photocopier, the videocassette recorder, the peer-to-peer network and the search engine each provoked a period of doctrinal instability that was eventually resolved, more or less satisfactorily, through a combination of judicial interpretation and legislative adjustment. Generative artificial intelligence presents a harder version of the same problem, because it does not merely reproduce or distribute protected works but consumes them at industrial scale in order to produce a system capable of generating expression that competes with the very works consumed. The economic and expressive stakes are therefore of a different order, and the doctrinal categories available for resolving them are correspondingly strained. The technical process is well understood even if its legal characterisation is not. A large language model is developed by assembling a corpus, frequently through automated scraping of publicly accessible websites, converting that corpus into numerical tokens, and adjusting billions of model parameters so that the system learns the statistical relationships between those tokens. At no stage is the model instructed to memorise any particular work. Yet the assembly of the corpus necessarily involves the making of copies, the retention of those copies for the duration of training involves storage, and the resulting parameters may in some circumstances encode fragments of protected expression sufficiently faithfully that the model can be induced to reproduce them. Each of these three moments engages a different aspect of the exclusive rights conferred by Section 14 of the Copyright Act, 1957, and each attracts a different analysis. India entered this debate later than the principal common law jurisdictions but has arrived at a position of unusual significance. In November 2024, Asian News International instituted proceedings in the Delhi High Court against OpenAI, alleging both that its news reports had been copied and stored to train the models underlying ChatGPT and that the chatbot reproduced and misattributed its content in responses to users. After thirty-two hearings, the appointment of two amici curiae and interventions by six industry bodies representing publishers, the music industry, broadband providers and artificial intelligence developers, Justice Amit Bansal delivered judgment on the interim application on 24 July 2026. The Court refused an injunction and held, on a prima facie view, that the storage of the plaintiff's literary works for the purpose of training large language models falls within the fair dealing exception in Section 52(1)(a) of the Copyright Act, 1957. It is the first considered engagement by an Indian court with the training question and the first anywhere in the common law world to construe an enumerated fair dealing provision, rather than an open-ended fair use standard, as accommodating machine learning.

That ruling did not emerge in a policy vacuum. On 28 April 2025 the Department for Promotion of Industry and Internal Trade constituted an expert committee to examine whether the Indian copyright regime adequately addresses the questions raised by artificial intelligence. Its first Working Paper, published on 8 December 2025 and open to public comment until 7 January 2026, addressed the input side of the problem and proposed a hybrid model under which



# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

**Impact Factor: 8.3** [www.ijesh.com](http://www.ijesh.com) **ISSN: 2250 3552**

artificial intelligence developers would obtain a statutory blanket licence to use lawfully accessed works, subject to compulsory remuneration administered by a new Copyright Royalties Collective for AI Training. The Working Paper proceeded from the express premise that Section 52 is narrowly drawn, purpose-specific and inadequate to the scale of contemporary machine learning. The Delhi High Court, some seven months later, reached a materially different conclusion about the reach of the same provision. The Indian position is therefore internally divided between a judicial reading that treats existing law as sufficient and an executive assessment that treats it as deficient.

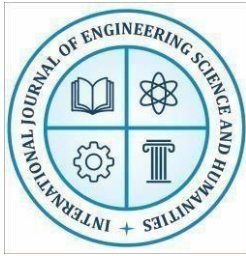
This paper argues that both positions are unsatisfactory, though for different reasons, and that the resolution lies in neither. The judicial reading strains the closed enumeration of Section 52 beyond what its structure will bear and rests on a conception of research that would, if applied consistently, hollow out the exception. The executive proposal, by conferring an unconditional and non-waivable licence on developers, subordinates the exclusive rights of authors to a compensation mechanism whose administrative viability is unproven and which denies rightsholders any meaningful control over the use of their work. What is required is a legislative intervention that is narrower than the Working Paper proposes and more principled than the Court's interpretation permits.

## **2. Statement of the Problem, Objectives and Research Design**

The problem this study addresses is the absence of any provision in Indian copyright law that was drafted with machine learning in view, combined with the emergence of a judicial interpretation that supplies that absence by expanding an exception designed for individual scholarly use into a licence for industrial-scale commercial data ingestion. The consequence is legal uncertainty for creators, developers and investors alike, and a growing divergence between the Indian position and that of the jurisdictions with which India competes for both creative and technological investment.

The study pursues four objectives. First, to establish with precision which acts within the artificial intelligence development pipeline engage the exclusive rights conferred by the Copyright Act, 1957, and to distinguish them from acts that do not. Second, to subject the reasoning in *ANI Media Pvt. Ltd. v. Open AI OpCo LLC* to doctrinal scrutiny against the structure of Section 52, the settled Indian authority on fair dealing, and India's obligations under the Berne Convention and the Agreement on Trade-Related Aspects of Intellectual Property Rights. Third, to situate the Indian approach within a comparative framework encompassing the United States, the European Union, the United Kingdom, Japan and Singapore, with attention to why jurisdictions with similar economic interests have diverged so sharply. Fourth, to evaluate the hybrid statutory licensing model proposed by the Department for Promotion of Industry and Internal Trade and to develop an alternative legislative framework.

Four research questions follow. Does the storage and processing of copyright-protected works for the purpose of training a large language model constitute infringement under Sections 14



# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

**Impact Factor: 8.3** [www.ijesh.com](http://www.ijesh.com) **ISSN: 2250 3552**

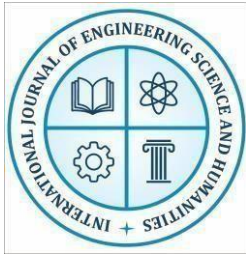
and 51 of the Copyright Act, 1957? Can that activity be accommodated within Section 52(1)(a) consistently with the closed-list character of the provision and with the three-step test? What lessons does comparative practice offer for a jurisdiction that has neither an open-ended fair use standard nor a dedicated text and data mining exception? What legislative architecture would best reconcile the interests of authors, developers and the public?

The methodology is doctrinal and comparative. It relies on primary sources comprising the Copyright Act, 1957, the judgments of Indian and foreign courts, the Berne Convention, the Agreement on Trade-Related Aspects of Intellectual Property Rights, Directive (EU) 2019/790, Regulation (EU) 2024/1689, the Copyright, Designs and Patents Act 1988 and the Data (Use and Access) Act 2025 of the United Kingdom, and on secondary sources comprising the Working Paper of the Department for Promotion of Industry and Internal Trade, the reports of the United States Copyright Office, the United Kingdom Report on Copyright and Artificial Intelligence of 18 March 2026, and the scholarly literature. No empirical or survey component is undertaken. The study is confined to copyright and does not address the distinct questions raised by patents over artificial intelligence inventions, the protection of trade secrets in model weights, personality and publicity rights in digital replicas, or the data protection dimensions of training corpora, save where these intersect with the copyright analysis.

### 3. Review of Literature

Scholarly engagement with machine learning and copyright predates the generative turn by roughly a decade and was initially framed around the copying involved in automated text analysis. Grimmelmann (2016) advanced the influential claim that copying which serves only to enable a machine to extract non-expressive information should not be treated as an infringement of a right designed to protect human expressive enjoyment, an argument developed at greater length by Sag (2019) under the rubric of non-expressive use. Lemley and Casey (2021) extended this position to machine learning generally, contending that a fair learning principle is necessary if the law is not to obstruct the development of systems whose social value depends on training breadth. Sobel (2017) offered the most sustained early challenge, arguing that where the machine learning output competes in the market for the works ingested, the non-expressive characterisation conceals rather than resolves the distributive question.

The arrival of generative systems capable of producing substitutable expressive output has shifted the centre of gravity of this literature. Samuelson (2023) identified the critical distinction between training-stage copying and output-stage reproduction and cautioned that conflating them produces unstable doctrine. Sag (2023) revised his earlier position to accept that memorisation and regurgitation present a genuine infringement risk requiring engineering as well as legal responses. Senftleben (2023) reframed the debate around remuneration rather than prohibition, arguing that the appropriate response to unavoidable large-scale ingestion is a levy or licensing mechanism directed at sustaining authorship rather than an injunction that will not in practice be granted.



# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

**Impact Factor: 8.3** [www.ijesh.com](http://www.ijesh.com) **ISSN: 2250 3552**

European scholarship has concentrated on the adequacy of the text and data mining exceptions introduced by Directive (EU) 2019/790. Geiger, Frosio and Bulayenko (2018) argued during the legislative process that Articles 3 and 4 were drawn too narrowly to serve their stated innovation purpose, a criticism that has acquired renewed force as courts have been asked to determine whether generative training falls within them at all. Hugenholtz and Quintais (2021) examined the authorship consequences of machine creation and concluded that European law's insistence on the author's own intellectual creation excludes wholly machine-produced output. Quintais (2025) has analysed the interaction between the copyright *acquis* and the transparency obligations imposed on general-purpose model providers by Regulation (EU) 2024/1689, observing that the Regulation assumes an underlying copyright entitlement whose scope remains unsettled. Ginsburg (2018) and Gervais (2019) supply the doctrinal foundation for the human authorship requirement that has since been applied by the United States Copyright Office and affirmed in *Thaler v. Perlmutter*.

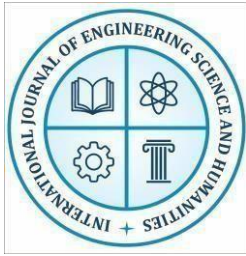
Indian scholarship on the subject remains comparatively thin and has developed principally through policy commentary rather than sustained academic treatment. Reddy and Chandrashekar (2017) provide the essential historical account of how Indian copyright exceptions were shaped by access considerations rather than by any theory of technological accommodation, a lineage that bears directly on how Section 52 should be construed today. Scaria (2014) documents the gap between formal copyright entitlements and enforcement realities in Indian creative industries. Ahuja (2023) and Narayanan (2017) supply the standard doctrinal treatment of Section 52 as a closed enumeration. What the Indian literature has not yet produced is a rigorous assessment of whether the enumerated fair dealing model can bear the interpretive weight that has now been placed upon it, or a worked proposal for statutory reform calibrated to Indian conditions. This paper addresses that gap.

## **4. The Anatomy of Model Training and the Rights It Engages**

Legal analysis in this field has been impeded by imprecision about what is actually being done. It is therefore necessary to disaggregate the development pipeline into its constituent acts and to identify, for each, the right in Section 14 that is engaged.

### **4.1 Corpus assembly and retention**

The first stage involves the acquisition of material, typically by automated crawling of publicly accessible web resources, and its retention in a training corpus. This is the clearest infringement point. Section 14(a)(i) confers on the owner of copyright in a literary work the exclusive right to reproduce the work in any material form, including the storing of it in any medium by electronic means. The words were introduced by the Copyright (Amendment) Act, 1994 precisely to capture digital storage, and they leave little room for argument. A crawler that downloads and retains a news report makes a reproduction. That the copy is transient, or that it is retained only for the duration of a training run, does not alter the position under Indian law, which contains no analogue to the temporary copying exception in Article 5(1) of Directive



# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

Impact Factor: 8.3 [www.ijesh.com](http://www.ijesh.com) ISSN: 2250 3552

2001/29/EC. It was for this reason that the defendant in ANI Media did not dispute that storage had occurred and rested its case entirely on the exception.

## **4.2 Tokenisation, vectorisation and parameter adjustment**

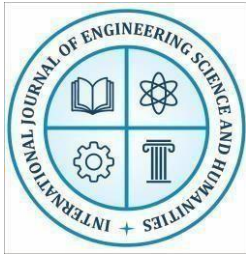
The second stage converts retained text into tokens and then into numerical vectors capturing statistical relationships, and adjusts model parameters accordingly. Whether this constitutes an adaptation within the meaning of Section 2(a) and Section 14(a)(vi) is more difficult. The better view is that it does not. Adaptation under the Act contemplates the conversion of a work into another form in which the expression of the original survives in recognisable shape. A vector representation does not preserve expression; it preserves statistical relationships abstracted from expression. The idea-expression dichotomy affirmed in *R.G. Anand v. Delux Films* supports this conclusion, as does the principle that copyright does not subsist in facts or in the information a work conveys. To this extent the non-expressive use scholarship is doctrinally sound.

## **4.3 Memorisation within model parameters**

The third stage is the most contested. Where a work appears many times in a corpus, or where its expression is highly distinctive, the trained parameters may encode it sufficiently precisely that the model can reproduce substantial portions on an appropriate prompt. The Munich Regional Court in *GEMA v. OpenAI* held on 11 November 2025 that such memorisation is itself an act of reproduction within Sections 15 and 16 of the German Copyright Act, on the reasoning that reproduction covers any material fixation from which the work can be perceived, and that the text and data mining exceptions in Sections 44b and 60d do not extend to it because mining is directed at the evaluation of abstract patterns rather than the retention of the works themselves. Whether Indian law would reach the same conclusion is untested. Section 14(a)(i) is at least as broad as the German provision, and there is no principled reason why a fixation in model weights should be treated differently from a fixation in any other medium if the work can in fact be extracted from it. The Delhi High Court did not decide the point, holding instead that the plaintiff had not established memorisation on the evidence before it.

## **4.4 Output generation and retrieval-augmented generation**

The final stage concerns what the deployed system produces. Two distinct situations must be separated. Where the model generates text from its learned parameters, infringement depends on whether the output is substantially similar to a protected work, applying the ordinary observer standard familiar from *R.G. Anand*. Where the system employs retrieval-augmented generation, fetching and summarising live web content at the moment of answering, the analysis is entirely different: the system is making a contemporaneous copy of an identified work for the purpose of producing a derivative summary, which raises questions closer to those in traditional news aggregation disputes than to those raised by training. The Delhi High Court recognised this distinction but found that the illustrative outputs relied upon by the plaintiff had been generated after the relevant training was complete and could not therefore establish that the model had been trained to memorise and later regurgitate those specific articles. This



# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

Impact Factor: 8.3 [www.ijesh.com](http://www.ijesh.com) ISSN: 2250 3552

is a proposition about evidence rather than about law, and it identifies the central practical difficulty facing rightsholders: memorisation must be proved with material that predates the training run, which is ordinarily in the exclusive possession of the developer.

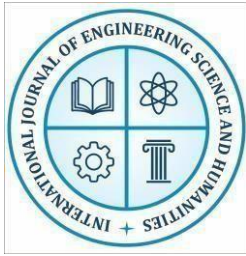
## 5. The Indian Statutory Framework and the Architecture of Section 52

India follows the enumerated exception model inherited from the Imperial Copyright Act, 1911 and preserved through the Copyright Act, 1957. Section 52 does not confer a general defence of fairness; it lists specified acts which shall not constitute infringement. The distinction from the United States model is fundamental. Section 107 of Title 17 of the United States Code directs a court to weigh four factors in relation to any use whatever, whereas Section 52 requires the defendant first to bring the impugned act within an enumerated head and only then to demonstrate that the dealing was fair. A use may be entirely fair in the abstract and nonetheless infringe if it does not correspond to a listed purpose.

Section 52(1)(a) permits a fair dealing with any work, other than 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; and (iii) the reporting of current events and current affairs. The present wording derives from the Copyright (Amendment) Act, 2012, which replaced the earlier formulation covering research or private study. That amendment is significant in a direction the Delhi High Court did not fully explore. By subordinating research to the governing phrase private or personal use, the legislature arguably narrowed rather than widened the head, indicating that the research contemplated is research undertaken by or for a person in a personal capacity rather than research as an institutional or industrial activity.

The Indian case law on fair dealing has developed a set of criteria that operate within, not in substitution for, this enumerated structure. In *Civic Chandran v. Ammini Amma* the Kerala High Court identified the quantum and value of the matter taken, the purpose of the taking and the likelihood of competition between the two works as the governing considerations. In *Super Cassettes Industries Ltd. v. Hamar Television Network Pvt. Ltd.* the Delhi High Court held that the commercial character of a use does not by itself defeat fair dealing, since the statute imposes no non-commercial condition on Section 52(1)(a). In *India TV Independent News Service Pvt. Ltd. v. Yashraj Films Pvt. Ltd.* the same Court emphasised that the purpose and character of the use, rather than the mere fact of copying, is central. The Division Bench in *The Chancellor, Masters and Scholars of the University of Oxford v. Rameshwari Photocopy Services* held that the reproduction of substantial extracts in university course packs fell within Section 52(1)(i), and articulated the underlying principle that copyright is not a natural or divine right but a statutory instrument for stimulating creativity, to be construed in light of that purpose.

Two constraints on this jurisprudence must be kept in view. The first is that the *Rameshwari Photocopy* decision rested on the distinct head relating to instruction, and its reasoning about the primacy of educational access does not transfer automatically to commercial technology development. The second is the three-step test. Article 9(2) of the Berne Convention, to which India has been a party since 1928, and Article 13 of the Agreement on Trade-Related Aspects



# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

Impact Factor: 8.3 [www.ijesh.com](http://www.ijesh.com) ISSN: 2250 3552

of Intellectual Property Rights confine permitted exceptions to certain special cases which do not conflict with a normal exploitation of the work and do not unreasonably prejudice the legitimate interests of the rightholder. An interpretation of Section 52(1)(a) broad enough to authorise the wholesale ingestion of a publisher's entire output for commercial model development is difficult to describe as a certain special case, and becomes harder still to reconcile with the second and third steps as a licensing market for training data develops.

It is also relevant that the Act already contains a considered legislative response to situations in which transactional licensing is impracticable. Sections 31, 31A and 31D provide for compulsory and statutory licensing subject to rates fixed by the Appellate Board, and the Supreme Court in *Entertainment Network (India) Ltd. v. Super Cassettes Industries Ltd.* explained that these provisions exist precisely to reconcile the exclusive right with the public interest in dissemination where refusal to license would be unreasonable. In *Tips Industries Ltd. v. Wynk Music Ltd.* the Bombay High Court held that Section 31D is to be construed strictly and does not extend to services the legislature did not contemplate. The existence of this machinery is a strong indication that the Act treats the problem of unavoidable large-scale use as a licensing question to be resolved by the legislature, not as an interpretive question to be resolved by expanding an exception.

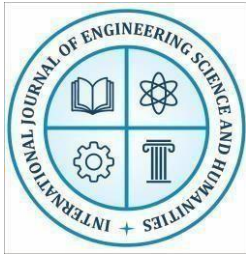
## **6. ANI Media Pvt. Ltd. v. Open AI OpCo LLC: Analysis and Critique**

### **6.1 The proceedings and the issues**

The suit, numbered CS(COMM) 1028/2024, was instituted in November 2024 by an Indian news agency alleging that OpenAI had scraped and stored its copyrighted news reports to train the models underlying ChatGPT, and that the chatbot reproduced its content and attributed fabricated statements to it. Four issues were framed by order of 19 November 2024: whether storage of the plaintiff's works for training amounts to infringement; whether the use of those works to generate responses amounts to infringement; whether any such use is protected by Section 52; and whether an Indian court has territorial jurisdiction over a foreign defendant whose training occurs on servers abroad. Two amici curiae were appointed and six industry bodies intervened. Judgment on the interim application was reserved on 27 March 2026 and delivered on 24 July 2026.

### **6.2 Territorial jurisdiction**

The Court rejected the contention that training conducted on servers in the United States lies beyond the reach of the Copyright Act, 1957. Reading Section 62 of the Act with Section 20 of the Code of Civil Procedure, 1908, it held that storage abroad is the terminal step in a chain of events beginning with access to the works from India, and that severing the two would permit evasion of Indian copyright law by the simple expedient of relocating the final link. It followed the reasoning of a coordinate bench in *Neetu Singh v. Telegram FZ LLC*, which had held that the location of servers in Singapore did not deprive an Indian rightholder of a remedy. The purposeful availment test in *Banyan Tree Holding (P) Ltd. v. A. Murali Krishna Reddy* was distinguished as inapposite to a training claim.



# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

**Impact Factor: 8.3** [www.ijesh.com](http://www.ijesh.com) **ISSN: 2250 3552**

This aspect of the judgment is the most persuasive and the most consequential. It stands in instructive contrast to the position in the United Kingdom, where Getty Images abandoned its principal training claim against Stability AI once it became clear that the training had occurred outside the jurisdiction, and to the approach of the Munich Regional Court in *GEMA v. Suno*, which resolved the same difficulty by asserting jurisdiction over training conducted in the United States and applying United States law to it. The Delhi High Court's reasoning avoids both the abdication of the first course and the awkwardness of the second, and offers a model likely to be influential in other jurisdictions confronting the territoriality problem.

### **6.3 The output claim**

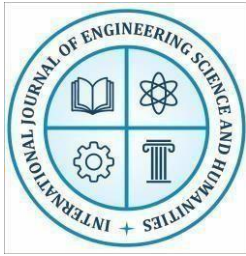
On the second issue the Court reaffirmed that copyright protects the original expression of a news report and not the underlying facts, so that the reproduction of information contained in the plaintiff's reporting is not by itself actionable. It found that the illustrative outputs annexed to the interim application had been generated after the completion of the relevant training and could not therefore demonstrate that the model had been trained to memorise and regurgitate those articles, and that on the material available it could not conclude that responses produced through retrieval-augmented generation were substantially similar to the plaintiff's works. These were, the Court held, disputed questions requiring evidence at trial.

The reasoning is defensible on its own terms but exposes a structural asymmetry. Evidence of memorisation resides in the training corpus, the model weights and the internal evaluation records of the developer. A rightsholder cannot ordinarily obtain such evidence before discovery, and if pre-training illustrations are required to establish a *prima facie* case, the interim remedy becomes practically unavailable in precisely the class of case for which it was designed. The contrast with *GEMA v. OpenAI* is stark: there, the plaintiff was able to demonstrate near-verbatim reproduction of nine identified song lyrics through simple prompts, and the court treated that demonstration as sufficient to establish fixation within the model. The difference in outcome may reflect a difference in the evidence available as much as a difference in the law.

### **6.4 The fair dealing holding**

The defendant did not dispute that copies of scraped articles were stored during training and defended solely under Section 52(1)(a)(i). Since the three limbs of the sub-clause are disjunctive, success on one sufficed. The Court applied a two-part test comprising a purpose test and a fairness test.

On purpose, three findings were made. First, commercial use is not disqualifying, because Section 52(1)(a) contains no non-commercial condition whereas other clauses of the same section, including clauses (ad), (k), (l), (n) and (o), expressly impose one; the omission was therefore treated as deliberate. Second, the training material had been obtained from publicly accessible websites rather than from an already infringing source. Third, and most significantly, research was given a wide construction. Drawing on the decision of the Supreme Court of Canada in *CCH Canadian Ltd. v. Law Society of Upper Canada*, which held that research must



# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

**Impact Factor: 8.3** [www.ijesh.com](http://www.ijesh.com) **ISSN: 2250 3552**

receive a large and liberal interpretation and extends to lawyers conducting research for profit, the Court held that the construction of a model's underlying knowledge base is itself research within Section 52(1)(a)(i).

On fairness, the Court found that the use was confined to training rather than resale or republication, that the plaintiff had placed nothing on the record beyond bare averments to show loss of market share or subscription revenue, and that the public interest favoured the defendant given the contribution of such systems to education, research and accessibility and the observation that requiring a licence from every data source would render their construction economically unviable. The Court drew on the reasoning in *Bartz v. Anthropic PBC*, *Kadrey v. Meta Platforms, Inc.* and *Authors Guild v. Google, Inc.*, treating transformative purpose rather than commercial motive as the true measure of fairness. It concluded that both limbs were satisfied and that storage for training does not, *prima facie*, infringe.

## **6.5 The refusal of interim relief**

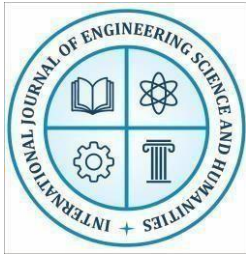
Independently of the fair dealing finding, the Court held that the balance of convenience and the risk of irreparable injury told against an injunction. Six considerations were weighed: the plaintiff had the admitted technical ability to block the defendant's crawlers and had not exercised it; the defendant had itself blocklisted the plaintiff's website in October 2024 for both training and retrieval purposes; the plaintiff had by letter of 3 October 2024 offered a licence for a stated sum, which the Court read as demonstrating that any loss was quantifiable in damages; no evidence of actual commercial loss had been adduced; an injunction would disrupt the service for a large number of Indian users, many of them non-paying; and the public interest, treated as a distinct fourth factor in intellectual property matters following *Zydyus Lifesciences Ltd. v. E.R. Squibb & Sons LLC* and *F. Hoffmann-La Roche AG v. NATCO Pharma Ltd.*, weighed against relief, the Court referring to the NITI Aayog roadmap on artificial intelligence for inclusive societal development.

## **6.6 Critique**

Four objections may be entered against the fair dealing holding, none of which detracts from the care with which the judgment was constructed.

The first concerns the construction of research. *CCH Canadian* was decided on Section 29 of the Canadian Copyright Act, which permits fair dealing for the purpose of research without qualification. Section 52(1)(a)(i) as amended in 2012 permits fair dealing for the purpose of private or personal use, including research. The word including performs work: it identifies research as a species of private or personal use rather than as a free-standing purpose. To read it as authorising research at any scale and by any actor is to treat the governing words as surplusage. A construction under which the ingestion of an entire national news archive by a commercial enterprise qualifies as private or personal use is difficult to sustain on the language the legislature chose.

The second concerns the treatment of commerciality. It is correct that Section 52(1)(a) imposes no express non-commercial condition, and *Super Cassettes* so holds. But the inference drawn



# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

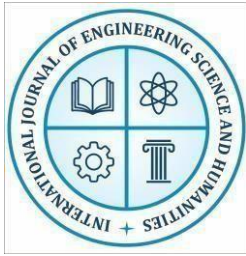
**Impact Factor: 8.3** [www.ijesh.com](http://www.ijesh.com) **ISSN: 2250 3552**

from that omission proves too much. The absence of an express condition means that commerciality is not automatically fatal; it does not mean that commerciality is irrelevant to the separate assessment of fairness, where under *Civic Chandran* the likelihood of competition between the two works is a governing criterion. The relevant competition here is not between the model and any individual news report but between the model and the market for licensing news content for training, a market whose existence the plaintiff's own licence offer demonstrated. The Court treated that offer as evidence that the claim was quantifiable; it might equally have been treated as evidence of a functioning licensing market whose displacement is a cognisable harm, which is precisely the reasoning that produced the opposite result in *Thomson Reuters Enterprise Centre GmbH v. ROSS Intelligence Inc.*

The third concerns the transplantation of transformative use. Transformativeness is a construct of the first factor of Section 107, developed in *Campbell v. Acuff-Rose Music, Inc.* and refined in *Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith*. It has no textual foothold in Section 52. Importing it as the true measure of fairness under an enumerated exception risks converting Section 52(1)(a) into a general fair use clause by judicial construction, which is a legislative act. It is also worth noting that the American authorities relied upon are first-instance decisions of divided persuasive weight: *Bartz* settled after the fair use ruling for a reported sum of USD 1.5 billion in respect of the separate holding on pirated acquisition, *Kadrey* turned expressly on the plaintiffs' failure to develop a market-harm theory rather than on the unavailability of one, and the contrary decision in *Thomson Reuters v. ROSS* was argued before the Third Circuit on 11 June 2026 without an opinion having issued. No United States appellate court has determined whether training constitutes fair use.

The fourth concerns the three-step test. If Section 52(1)(a)(i) permits an unlimited quantity of any category of work to be ingested by any commercial actor for the development of systems that generate competing expression, the exception ceases to be a certain special case within the meaning of Article 9(2) of the Berne Convention. The judgment does not address India's treaty obligations, and it is a striking feature of the reasoning that a construction of such breadth was adopted without reference to the external constraint that most directly bears upon it.

These objections should be read subject to an important qualification, which the judgment itself repeatedly makes. Every finding is expressly *prima facie* and directed to the interim application. The suit proceeds to trial, where a fuller evidentiary record on memorisation and market harm may produce a different result, and an appeal to a Division Bench remains available. The decision is a first word rather than a last one. Its immediate significance lies less in its holdings than in the fact that it has established, in a jurisdiction with a closed list of exceptions, an interpretive route by which machine learning may be accommodated without legislative amendment. Whether that route is a sound one is the question this paper answers in the negative.



# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

**Impact Factor: 8.3** [www.ijesh.com](http://www.ijesh.com) **ISSN: 2250 3552**

## 7. Comparative Perspectives

The Indian position acquires its full significance only when placed beside the approaches adopted elsewhere. What is striking about the comparative landscape of 2025 and 2026 is not that jurisdictions have diverged, but that they have diverged while confronting materially identical technology and largely similar treaty obligations.

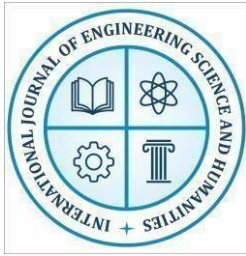
### 7.1 United States

The United States possesses the most flexible instrument and has produced the least settled outcome. Three first-instance decisions of 2025 point in different directions. In *Bartz v. Anthropic PBC* the Northern District of California held that training a language model on lawfully acquired books was transformative to a high degree and constituted fair use, while holding separately that the downloading and retention of pirated copies was not; the case subsequently settled for a reported USD 1.5 billion, with final approval granted in July 2026. In *Kadrey v. Meta Platforms, Inc.* the same district reached a similar conclusion on training but the court took care to state that the plaintiffs had lost because their market-harm arguments were underdeveloped rather than because such arguments could never succeed. In *Thomson Reuters Enterprise Centre GmbH v. ROSS Intelligence Inc.* the District of Delaware held that copying editorial headnotes to train a competing legal research tool was not fair use, the fourth factor weighing decisively against the defendant because the tool competed directly with the source and undercut the emerging market for training data licences. That decision was certified for interlocutory appeal and argued before the Third Circuit on 11 June 2026. As at the date of writing no United States court of appeals has ruled on whether training constitutes fair use. On the distinct question of output copyrightability, the human authorship requirement was affirmed in *Thaler v. Perlmutter* and certiorari was denied by the Supreme Court on 2 March 2026, consolidating the position taken by the United States Copyright Office in the second part of its report on copyright and artificial intelligence.

### 7.2 European Union

The European Union legislated for text and data mining before generative systems existed. Article 3 of Directive (EU) 2019/790 provides a mandatory and non-overridable exception for research organisations and cultural heritage institutions pursuing scientific research, and Article 4 provides a general exception for lawfully accessible works subject to the rightholder's express reservation of rights, which for online content must be expressed in machine-readable form. Article 53 of Regulation (EU) 2024/1689 requires providers of general-purpose models to implement a policy for identifying and complying with such reservations and to publish a sufficiently detailed summary of training content.

The judicial application of this framework has proved considerably less accommodating than its architecture suggested. In *GEMA v. OpenAI* the Munich Regional Court held on 11 November 2025 that the memorisation of song lyrics within model parameters constitutes reproduction under Sections 15 and 16 of the German Copyright Act; that the text and data mining exception in Section 44b does not extend to it, because mining is directed at the



# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

**Impact Factor: 8.3** [www.ijesh.com](http://www.ijesh.com) **ISSN: 2250 3552**

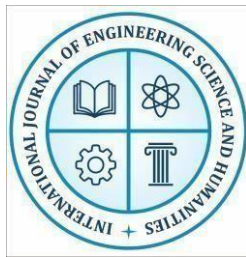
extraction of abstract patterns whereas memorisation retains the work itself; that training is not an ordinary or expected use to which authors may be taken to have impliedly consented; and that responsibility rests with the provider rather than the user. Injunctive relief, disclosure and damages were granted. The court in *GEMA v. Suno Inc.* extended this reasoning to a music generation system on 31 July 2026, additionally asserting jurisdiction over training conducted in the United States and rejecting a fair use defence under United States law. Appeals are pending and a reference to the Court of Justice remains likely. The European position, in short, is that a broadly drafted mining exception does not by itself resolve the generative training question, because the retention of expression within a model is a distinct act from the analysis the exception was designed to permit.

## **7.3 United Kingdom**

The United Kingdom offers the most instructive negative example. Section 29A of the Copyright, Designs and Patents Act 1988 permits copying for computational analysis only for non-commercial research, and therefore does not cover commercial model training. In December 2024 the Government consulted on four options and expressed a preference for a broad mining exception with a rightsholder opt-out and supporting transparency measures. That preference attracted support from approximately three per cent of more than eleven thousand respondents. Sections 135 and 136 of the Data (Use and Access) Act 2025 obliged the Government to publish an economic impact assessment and a report by 18 March 2026. Ministers had by then acknowledged that expressing the preference was a mistake, and the Report published on that date abandoned the opt-out model, committing instead to monitor developments, encourage industry-led licensing and gather further evidence. The narrower existing exception remains in force. The episode demonstrates both the political difficulty of legislating an opt-out regime in a jurisdiction with a substantial creative sector and the cost of the resulting abstention, which is that fundamental questions are left to litigation.

## **7.4 Japan and Singapore**

Article 30-4 of the Japanese Copyright Act, introduced in 2018, permits the exploitation of a work where the purpose is not to enjoy the thoughts or sentiments expressed in it, a formulation that on its face accommodates machine learning generously. The General Understanding issued by the Agency for Cultural Affairs in 2024 significantly qualified that generosity, indicating that the exception does not apply where an additional purpose of enjoyment is present, where the training is directed at generating outputs in the style of specific works, or where a licensing market for the material exists. Section 244 of the Singapore Copyright Act 2021 permits computational data analysis subject to a lawful-access condition and cannot be excluded by contract. Both models are instructive for India because both are enumerated-exception systems that chose to legislate a specific permission rather than to expand an existing head by interpretation.



# International Journal of Engineering, Science and Humanities

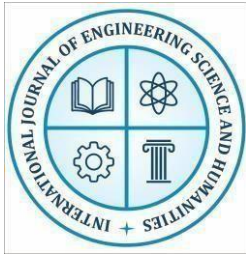
An international peer reviewed, refereed, open access journal

Impact Factor: 8.3 [www.ijesh.com](http://www.ijesh.com) ISSN: 2250 3552

The comparative position may be summarised as follows.

| Jurisdiction   | Statutory basis for training-stage copying                                       | Leading determination (2025–2026)                                                                                                | Normative orientation                              |
|----------------|----------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------|
| India          | No TDM exception; enumerated fair dealing under s. 52(1)(a), Copyright Act, 1957 | ANI Media v. Open AI (Del. HC, 24 July 2026): training storage prima facie fair dealing; interim injunction refused              | Judicially permissive; legislatively unsettled     |
| United States  | Open-ended fair use, 17 U.S.C. § 107                                             | Bartz v. Anthropic and Kadrey v. Meta (training transformative); Thomson Reuters v. ROSS (not fair use); no appellate ruling yet | Fact-sensitive and fragmented; market-harm centred |
| European Union | Arts. 3–4, Directive (EU) 2019/790; Art. 53, Regulation (EU) 2024/1689           | GEMA v. OpenAI (LG München I, 11 Nov. 2025): memorisation is reproduction; TDM defence rejected; GEMA v. Suno (31 July 2026)     | Rights-reservation and transparency centred        |
| United Kingdom | Section 29A, CDPA 1988 (non-commercial research only)                            | Getty Images v. Stability AI (2025); Report on Copyright and AI, 18 March 2026 abandoning the opt-out model                      | Deliberate legislative abstention; licensing-led   |
| Japan          | Article 30-4, Copyright Act (non-enjoyment uses)                                 | Agency for Cultural Affairs General Understanding (2024) narrowing the exception for outputs                                     | Broadest statutory permission, output-limited      |
| Singapore      | Section 244, Copyright Act 2021 (computational data analysis)                    | Lawful-access condition; non-derogable by contract                                                                               | Innovation-facing but access-conditioned           |

*Table 1: Comparative regulatory approaches to generative artificial intelligence training (position as at August 2026).*



# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

Impact Factor: 8.3 [www.ijesh.com](http://www.ijesh.com) ISSN: 2250 3552

Two conclusions follow from this survey. The first is that no jurisdiction has succeeded in resolving the training question through the interpretation of exceptions drafted for other purposes; those that have attempted it have produced either instability, as in the United States, or restriction, as in Germany. The second is that the jurisdictions that have legislated specifically, Japan and Singapore, have coupled permission with conditions relating to lawful access and to the presence of a licensing market. India, which has now attempted the interpretive route, stands alone in reading an exception for private or personal use as authorising commercial training at scale.

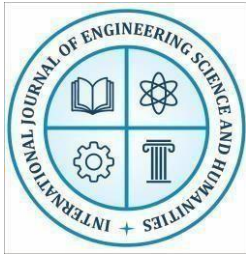
## **8. The DPIIT Working Paper and the Statutory Licensing Alternative**

The Working Paper on Generative Artificial Intelligence and Copyright published by the Department for Promotion of Industry and Internal Trade on 8 December 2025 proceeds from a diagnosis that is materially different from the one subsequently adopted by the Delhi High Court. The Committee, drawn from the Department, the Ministry of Electronics and Information Technology, industry and the academy, concluded that Section 52 is narrowly defined and purpose-specific, that India recognises no dedicated text and data mining exception, and that current training practices therefore fall outside the statutory protections available. It examined the principal regulatory models in use internationally, including mining exceptions of the European type and voluntary licensing, and concluded that none of them adequately balanced creator remuneration against innovation.

Its proposal, summarised in the formula of one nation, one licence, one payment, is a statutory blanket licence. Developers would be entitled to use lawfully accessed copyright-protected works for training without seeking individual permission; rightsholders would not be able to opt out; developers would file an artificial intelligence training data disclosure form identifying the datasets used; and once a model was commercialised, annual royalties at prescribed rates would be remitted to a new non-profit body, the Copyright Royalties Collective for AI Training, which would coordinate with existing copyright societies to distribute payments.

The proposal has genuine merits. It acknowledges, as the judgment does not, that the transaction costs of individual licensing at the scale required are prohibitive, which is the strongest argument in favour of a collective solution. It preserves a remuneration right where an exception would extinguish it. It ties permission to lawful access, thereby excluding the use of pirated corpora, a distinction that proved decisive in *Bartz*. And it imposes disclosure obligations that would go some way toward correcting the evidentiary asymmetry identified above.

Three objections nonetheless arise. The first concerns the denial of opt-out. A regime under which a rightsholder cannot decline participation converts an exclusive right into a liability rule without the rightsholder's consent. That is a substantial step, and while compulsory licensing is not unknown to Indian law, Sections 31, 31A and 31D confine it to defined categories and subject it to adjudication by a tribunal on evidence. A blanket, non-waivable licence covering the entirety of Indian creative output would be without precedent. It also sits uneasily with



# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

Impact Factor: 8.3 [www.ijesh.com](http://www.ijesh.com) ISSN: 2250 3552

Article 9(1) of the Berne Convention and with the three-step test, since the removal of any capacity to reserve rights is difficult to reconcile with the requirement that exceptions not conflict with normal exploitation, particularly where a licensing market is emerging.

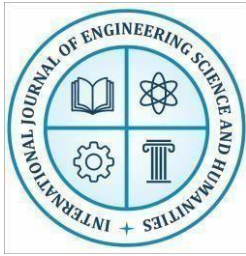
The second concerns administrative feasibility. The record of collective management in India is not encouraging. The functioning of copyright societies under Section 33 has been the subject of protracted disputes over registration, transparency and distribution, and the Copyright (Amendment) Act, 2012 was itself in large part a response to those difficulties. A new collective charged with valuing the contribution of an indeterminate number of works to models whose training data may not be fully documented, and with distributing the resulting revenue equitably, would face a task of an entirely different order. Without a workable methodology for attribution and distribution, a statutory royalty risks becoming a levy that developers pay and creators do not receive.

The third concerns sequencing. The Working Paper deals with the input side and expressly defers the output side to a second part. Yet the two are connected. If wholly machine-generated output is not copyrightable for want of a human author, as Section 2(d)(vi) read with the requirement of originality and the reasoning in *Rupendra Kashyap v. Jiwan Publishing House and Navigators Logistics Ltd. v. Kashif Qureshi* suggests, then the competitive threat to human authors takes a specific form that ought to inform the design of the input regime. Legislating the input side in isolation risks producing a framework that is internally coherent but incomplete.

## 9. Findings

The analysis yields the following findings.

- The assembly and retention of a training corpus engages the reproduction right in Section 14(a)(i) of the Copyright Act, 1957, which expressly includes electronic storage. Tokenisation and parameter adjustment, considered in isolation, do not amount to adaptation, because they preserve statistical relationships rather than expression. Memorisation within model parameters is analytically distinct from both and, on the reasoning adopted in Germany, is capable of constituting a further reproduction.
- Section 52 is a closed enumeration and not a general fairness standard. The construction of Section 52(1)(a)(i) adopted in *ANI Media* requires the phrase private or personal use to be treated as adding nothing to the word research, which is difficult to reconcile with the language introduced by the Copyright (Amendment) Act, 2012.
- The transformative use construct imported from the first factor of Section 107 of Title 17 of the United States Code has no textual basis in Indian law, and the American authorities relied upon are unsettled first-instance decisions between which no appellate court has yet chosen.
- The judgment's treatment of territorial jurisdiction is its most durable contribution and offers a coherent answer to a problem that defeated the claimant in the United Kingdom and required an unusual solution in Germany.



# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

Impact Factor: 8.3 [www.ijesh.com](http://www.ijesh.com) ISSN: 2250 3552

- The requirement of pre-training evidence of memorisation creates an evidentiary asymmetry that renders interim relief practically unavailable, since the relevant material is in the sole possession of the developer. This is an argument for mandatory disclosure rather than for a lower evidentiary threshold.
- An interpretation of Section 52(1)(a) broad enough to permit unlimited commercial ingestion is difficult to defend as a certain special case under Article 9(2) of the Berne Convention and Article 13 of the Agreement on Trade-Related Aspects of Intellectual Property Rights, and the judgment does not engage with these obligations.
- India is presently governed by two inconsistent official positions: an executive assessment that Section 52 is inadequate and a judicial holding that it suffices. This inconsistency is itself a source of the uncertainty that both institutions seek to dispel.
- The statutory blanket licence proposed by the Department for Promotion of Industry and Internal Trade correctly identifies the transaction cost problem but overcorrects by denying rightsholders any capacity to reserve rights, and depends upon a collective management infrastructure whose capacity is unproven.

## **10. Recommendations: A Calibrated Legislative Framework**

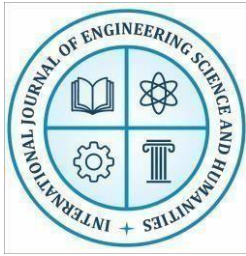
The appropriate response is legislative rather than interpretive, and should be narrower than the Working Paper proposes. Six elements are recommended.

### **10.1 A dedicated text and data mining exception**

A new Section 52(1)(zd) should be inserted permitting the reproduction and storage of a work for the purpose of computational analysis, including the training of machine learning systems, where the person carrying out the analysis has lawful access to the work. Lawful access should be defined to exclude material obtained from sources known or reasonably suspected to be infringing, and to exclude material obtained in circumvention of a technological protection measure or of terms of access applicable to the source. Placing the permission in a dedicated clause rather than leaving it to be derived from Section 52(1)(a) preserves the integrity of the enumerated model and permits the legislature to attach conditions that an existing head cannot bear.

### **10.2 An enforceable reservation of rights**

The exception should be subject to the right of the owner to reserve rights in respect of uses other than non-commercial scientific research, expressed by machine-readable means in the case of works made available online. The European experience shows that a reservation mechanism is only as effective as the technical standard supporting it, and the Central Government should accordingly be empowered to prescribe recognised formats by rule. The reservation should be non-derogable in respect of accredited research institutions, following Article 3 of Directive (EU) 2019/790 and Section 244 of the Singapore Copyright Act 2021, so that scientific research is not impaired by an instrument designed to regulate commercial development. Critically, the failure of a rightsholder to deploy a technical block should not,



# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

Impact Factor: 8.3 [www.ijesh.com](http://www.ijesh.com) ISSN: 2250 3552

absent a prescribed and functioning standard, be treated as an implied licence, as it effectively was in the interim analysis in ANI Media.

## **10.3 Mandatory training data disclosure**

Any person who commercially deploys in India a general-purpose model trained on copyright-protected works should be required to publish a summary of the categories and principal sources of the training data, in a form to be prescribed, and to retain records sufficient to establish the provenance of the corpus for a specified period. This adapts the approach of Article 53 of Regulation (EU) 2024/1689 and reflects the disclosure form proposed in the Working Paper. Its principal function is corrective: it addresses the evidentiary asymmetry that presently makes both the training claim and the memorisation claim practically unprovable at the interim stage.

## **10.4 A memorisation and output safeguard**

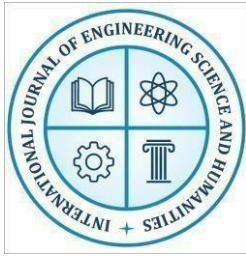
The exception should not extend to the retention within model parameters of expression from which a work or a substantial part of it can be extracted, nor to the generation of outputs substantially similar to a protected work. This preserves the distinction between analysis and retention on which the Munich Regional Court rested, and locates responsibility where it belongs, with the developer who controls model architecture and output filtering, rather than with the user who supplies a prompt. Developers should bear an obligation to implement reasonable technical measures against regurgitation, compliance with which should be a relevant consideration in assessing remedies.

## **10.5 Remuneration through the existing licensing architecture**

Where rights have been reserved and a commercial developer wishes nonetheless to use the works, recourse should be had to a licensing mechanism modelled on Sections 31 and 31D rather than to an unconditional blanket licence. A specialised licensing scheme, administered through registered copyright societies under Section 33 with rates determined by the Appellate Board on evidence, would preserve the exclusive right as the default while ensuring that unreasonable refusal to license cannot obstruct socially valuable development. This is the model the Act already adopts for analogous problems, and it avoids the constitutional and treaty difficulties of a non-waivable licence. Distribution methodology should be prescribed by rule before any such scheme is notified, so that the collection of royalties is not divorced from their equitable distribution.

## **10.6 Clarification of authorship in machine-assisted works**

Section 2(d)(vi), which designates as author of a computer-generated work the person who causes the work to be created, was drafted in 1994 for a materially different technology and should be revisited in the second phase of the reform. The statute should confirm that copyright subsists only where a human being has made an original contribution of skill and judgment satisfying the standard in *Eastern Book Company v. D.B. Modak*, and that the mere entry of a prompt does not by itself constitute such a contribution. Consideration should also be given to a separate and shorter-term neighbouring right for substantial investment in machine-generated



# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

**Impact Factor: 8.3** [www.ijesh.com](http://www.ijesh.com) **ISSN: 2250 3552**

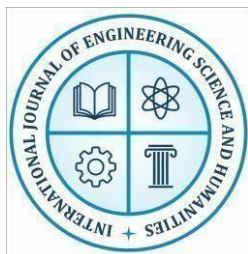
output, which would recognise commercial effort without diluting the authorship standard that anchors the entire system.

Taken together, these recommendations would place India in a position that is neither the permissive posture of the interim ruling nor the restrictive posture emerging in Germany, but a calibrated regime in which lawful access is a precondition, reservation is meaningful, disclosure is compulsory, memorisation is excluded, and remuneration operates through institutions the Act already contains.

## 11. Conclusion

The question whether a statute enacted in 1957 can accommodate a technology that did not exist until the present decade has now been answered, provisionally, in the affirmative by the Delhi High Court. This paper has argued that the answer is doctrinally insecure. The reasoning in *ANI Media Pvt. Ltd. v. Open AI OpCo LLC* is careful, well informed and alert to the practical consequences of the alternative, and its treatment of territorial jurisdiction is a genuine contribution to a problem that has embarrassed courts elsewhere. But the construction it places on Section 52(1)(a)(i) requires the phrase private or personal use to be emptied of content, imports a doctrine of transformative use for which the Indian statute makes no provision, relies on foreign first-instance authority that no appellate court has endorsed, and does not engage with the constraints that Article 9(2) of the Berne Convention and Article 13 of the Agreement on Trade-Related Aspects of Intellectual Property Rights impose on the breadth of exceptions. The comparative record supports this assessment. Every jurisdiction that has attempted to resolve the training question through the interpretation of pre-existing exceptions has produced either instability or restriction. The German courts, applying a mining exception drafted with machine analysis expressly in view, have held that it does not reach the retention of expression within a model. The United Kingdom, having proposed and then abandoned an opt-out regime, has left the question to litigation. The United States, with the most flexible standard available anywhere, has generated contradictory first-instance decisions and no appellate resolution. It is not plausible that India, with the narrowest instrument of the group, has found in an exception for private or personal use the answer that eluded all of them.

What follows is not that training should be prohibited. Prohibition is neither achievable nor desirable, and the Court was right to observe that a requirement of individual permission from every source would render model development economically unviable. What follows is that the accommodation must be legislated rather than construed, and that legislation permits the attachment of conditions that interpretation cannot supply: lawful access, an enforceable reservation of rights, compulsory disclosure of training data, an exclusion for memorisation, and a remuneration mechanism built on the licensing architecture the Copyright Act already provides. The Department for Promotion of Industry and Internal Trade has correctly identified the need for legislative intervention while proposing an instrument that goes too far in the opposite direction, converting an exclusive right into an unwaivable claim to compensation administered by an institution that does not yet exist.



# International Journal of Engineering, Science and Humanities

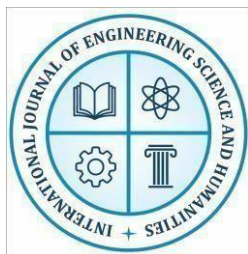
An international peer reviewed, refereed, open access journal

**Impact Factor: 8.3** [www.ijesh.com](http://www.ijesh.com) **ISSN: 2250 3552**

The Indian creative economy and the Indian technology sector are both substantial, and the country has no interest in a regime that sacrifices either to the other. The distinctive opportunity presently available is that India is legislating in the knowledge of what has failed elsewhere, and while its own jurisprudence remains at the interim stage and open to correction. That opportunity will not remain open indefinitely. Once training practices, licensing markets and judicial expectations have settled around a permissive default, the cost of legislative correction rises steeply. The case for acting now, and for acting with more precision than either the judgment or the Working Paper displays, is correspondingly strong.

## References

- Agreement on Trade-Related Aspects of Intellectual Property Rights, art. 13, 15 April 1994, 1869 U.N.T.S. 299.
- Ahuja, V. K. (2023). Law relating to intellectual property rights (4th ed.). LexisNexis.
- Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith, 598 U.S. 508 (2023).
- ANI Media Pvt. Ltd. v. Open AI OpCo LLC, CS(COMM) 1028/2024, I.A. 45300/2024 (Delhi High Court, judgment dated 24 July 2026, Bansal J.).
- Authors Guild v. Google, Inc., 804 F.3d 202 (2d Cir. 2015).
- Banyan Tree Holding (P) Ltd. v. A. Murali Krishna Reddy, 2009 SCC OnLine Del 3780.
- Bartz v. Anthropic PBC, No. 3:24-cv-05417 (N.D. Cal. 2025).
- Berne Convention for the Protection of Literary and Artistic Works, as revised at Paris on 24 July 1971, 1161 U.N.T.S. 3.
- Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569 (1994).
- CCH Canadian Ltd. v. Law Society of Upper Canada, [2004] 1 S.C.R. 339 (Supreme Court of Canada).
- Civic Chandran v. Ammini Amma, 1996 PTC (16) 670 (Kerala High Court).
- Copyright Act 2021 (Singapore), Act 22 of 2021, s. 244.
- Copyright Act, 1957 (India), Act No. 14 of 1957, as amended by the Copyright (Amendment) Act, 2012.
- Copyright, Designs and Patents Act 1988 (United Kingdom), s. 29A.
- Data (Use and Access) Act 2025 (United Kingdom), ss. 135–136.
- Department for Promotion of Industry and Internal Trade, Government of India. (2025, December 8).
- Working paper on generative artificial intelligence and copyright (Part I). Ministry of Commerce and Industry.
- Department for Science, Innovation and Technology, Intellectual Property Office & Department for Culture, Media and Sport. (2026, March 18). Report on copyright and artificial intelligence and copyright and AI impact assessment. HM Government.
- Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market, arts. 3–4, 2019 O.J. (L 130) 92.

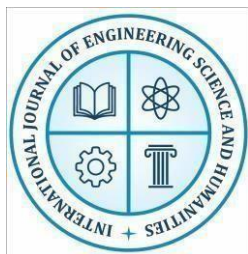


# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

**Impact Factor: 8.3** [www.ijesh.com](http://www.ijesh.com) **ISSN: 2250 3552**

- Eastern Book Company v. D.B. Modak, (2008) 1 S.C.C. 1.
- Entertainment Network (India) Ltd. v. Super Cassettes Industries Ltd., (2008) 13 S.C.C. 30.
- Expfar SA v. Eupharma Laboratories Ltd., (2004) 3 S.C.C. 688.
- F. Hoffmann-La Roche AG v. NATCO Pharma Ltd., CS(COMM) 567/2024 (Delhi High Court, 2025).
- Feist Publications, Inc. v. Rural Telephone Service Co., 499 U.S. 340 (1991).
- Geiger, C., Frosio, G., & Bulayenko, O. (2018). Text and data mining in the proposed copyright reform:
- Making the EU ready for an age of big data? IIC – International Review of Intellectual Property and Competition Law, 49(7), 814–844.
- GEMA v. OpenAI, Case No. 42 O 14139/24 (Landgericht München I, 11 November 2025).
- GEMA v. Suno Inc., Case No. 42 O 763/25 (Landgericht München I, 31 July 2026).
- Gervais, D. J. (2019). The machine as author. Iowa Law Review, 105(5), 2053–2106.
- Getty Images (US) Inc. v. Stability AI Ltd (High Court of Justice of England and Wales, Chancery Division, 2025).
- Ginsburg, J. C. (2018). People not machines: Authorship and what it means in the Berne Convention.
- IIC – International Review of Intellectual Property and Competition Law, 49(2), 131–135.
- Grimmelmann, J. (2016). Copyright for literate robots. Iowa Law Review, 101(2), 657–681.
- Hugenholtz, P. B., & Quintais, J. P. (2021). Copyright and artificial creation: Does EU copyright law protect AI-assisted output? IIC – International Review of Intellectual Property and Competition Law, 52(9), 1190–1216.
- India TV Independent News Service Pvt. Ltd. v. Yashraj Films Pvt. Ltd., 2012 SCC OnLine Del 4298.
- Kadrey v. Meta Platforms, Inc., No. 3:23-cv-03417 (N.D. Cal. 2025).
- Lemley, M. A., & Casey, B. (2021). Fair learning. Texas Law Review, 99(4), 743–785.
- Ministry of Electronics and Information Technology, Government of India. (2025, November). India AI governance guidelines.
- Narayanan, P. (2017). Intellectual property law (3rd ed.). Eastern Law House.
- Navigators Logistics Ltd. v. Kashif Qureshi, 2018 SCC OnLine Del 11321.
- Neetu Singh v. Telegram FZ LLC, 2022 SCC OnLine Del 2637.
- NITI Aayog. (2025, October). Roadmap on AI for inclusive societal development. Government of India.
- Quintais, J. P. (2025). Generative AI, copyright and the AI Act. Computer Law & Security Review, 56, 106107.
- R.G. Anand v. Delux Films, (1978) 4 S.C.C. 118.
- Reddy, T. P., & Chandrashekar, S. (2017). Create, copy, disrupt: India's intellectual property dilemmas. Oxford University Press.



# International Journal of Engineering, Science and Humanities

An international peer reviewed, refereed, open access journal

**Impact Factor: 8.3** [www.ijesh.com](http://www.ijesh.com) **ISSN: 2250 3552**

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence, art. 53.

Rupendra Kashyap v. Jiwan Publishing House Pvt. Ltd., 1996 PTC (16) 439 (Delhi High Court).

Sag, M. (2019). The new legal landscape for text mining and machine learning. *Journal of the Copyright Society of the U.S.A.*, 66(2), 291–367.

Sag, M. (2023). Copyright safety for generative AI. *Houston Law Review*, 61(2), 295–347.

Samuelson, P. (2023). Generative AI meets copyright. *Science*, 381(6654), 158–161.

Scaria, A. G. (2014). *Piracy in the Indian film industry: Copyright and cultural consonance*. Cambridge University Press.

Senftleben, M. (2023). Generative AI and author remuneration. *IIC – International Review of Intellectual Property and Competition Law*, 54(10), 1535–1560.

Sobel, B. L. W. (2017). Artificial intelligence’s fair use crisis. *Columbia Journal of Law & the Arts*, 41(1), 45–97.

Super Cassettes Industries Ltd. v. Hamar Television Network Pvt. Ltd., 2010 SCC OnLine Del 2086.

Thaler v. Perlmutter, No. 23-5233 (D.C. Cir. 2025), cert. denied (U.S., 2 March 2026).

The Chancellor, Masters and Scholars of the University of Oxford v. Rameshwari Photocopy Services, 2016 SCC OnLine Del 6229 (Division Bench).

Thomson Reuters Enterprise Centre GmbH v. ROSS Intelligence Inc., No. 1:20-cv-00613-SB (D. Del. 2025), interlocutory appeal argued, No. 25-2153 (3d Cir., 11 June 2026).

Tips Industries Ltd. v. Wynk Music Ltd., 2019 SCC OnLine Bom 13037.

United States Copyright Office. (2025). Copyright and artificial intelligence, Part 2: Copyrightability; Part 3: Generative AI training (pre-publication version). Library of Congress.

Zydu Lifesciences Ltd. v. E.R. Squibb & Sons LLC, FAO(OS)(COMM) 120/2025 (Delhi High Court, 12 January 2026).