

Outsourcing AI Creation: The Business Risks of Generative AI Under Indian and U.S. Copyright Law

Ayan Roy Chowdhury¹

¹Doctoral Candidate, NMIMS (Deemed to be University), Mumbai, India

Abstract

Enterprises in the United States increasingly buy creative and technical work (software, marketing copy, training manuals, design assets, research summaries) from Indian service providers whose delivery teams now produce much of that work with generative AI. The contracts governing these relationships were written for a world in which a human employee of the vendor authored the deliverable and a signed assignment moved title to the client. That world is disappearing, and the copyright rules on both sides of the arrangement have not kept pace. This Article maps the resulting exposure. It argues that outsourced AI creation generates four distinct legal risks: liability for the inputs used to build and adapt the model, liability for memorized or retrieved material in the output, an authorship gap that may leave the deliverable without any copyright to transfer, and output infringement that the client cannot shift upstream. The two legal systems handle these risks very differently. The United States insists on human authorship, a position fixed by the Supreme Court's refusal in March 2026 to review *Thaler v. Perlmutter*, yet its courts remain divided on whether training is fair use. India, by contrast, has since 1994 named "the person who causes the work to be created" as the author of computer-generated works, and in July 2026 the Delhi High Court provisionally treated LLM training as fair dealing in *ANI Media v. OpenAI*. These asymmetries, together with India's strict assignment formalities, the narrowing of secondary liability in *Cox Communications v. Sony Music*, and the conditional nature of vendor indemnities, mean that the risk lands on the parties least able to price it. The Article proposes a framework of Layered Provenance Allocation that assigns each risk to the party holding the cheapest information about it, replaces title warranties with process warranties, and treats documented human contribution as a balance-sheet asset.

Keywords: Generative AI, copyright, outsourcing, authorship, fair use, fair dealing, India, work made for hire, indemnity, AI governance

Introduction

Consider a composite but ordinary transaction. A Minneapolis medical-device company engages a Bengaluru technology vendor to build a patient-education portal: several thousand lines of code, two hundred pages of plain-language explanations, a library of illustrations, and a set of short explainer videos. The master services agreement, negotiated in 2019 and renewed each year since, provides that all deliverables are "works made for hire" and, to the extent they are not, are hereby assigned to the client. It includes a warranty of non-infringement and an uncapped intellectual property indemnity. Nothing in the agreement mentions artificial intelligence.

The vendor, pressed by the client to cut costs by a quarter, reorganizes delivery around generative AI. Engineers use a coding assistant licensed from a U.S. provider. Writers prompt a large language model and edit its drafts. The illustrations come from an image generator, lightly retouched. A retrieval pipeline pulls text from medical websites to ground the model's answers. The portal launches on time and under budget.

Eighteen months later three things happen. A competitor copies the illustrations wholesale, and the client's lawyers discover that they may not be protected by copyright at all. A medical publisher sends a demand letter asserting that several explanatory passages reproduce its articles nearly word for word. And the vendor, when asked to honor its indemnity, points to its own upstream AI license, which covers only outputs generated with the provider's content filters switched on, filters that the vendor's engineers had disabled to improve speed.

None of this is hypothetical in its parts. Indian IT firms are rewriting contracts around AI-driven productivity; Reuters reported in August 2026 that roughly eighty percent of Tata Consultancy Services' business-services contracts now tie fees to outcomes, and that clients are demanding spending cuts of twenty-five to thirty percent.[1] The authorship question has reached the end of the line in the United States: the Supreme Court denied certiorari in *Thaler v. Perlmutter* on March 2, 2026, leaving intact the D.C. Circuit's holding that the Copyright Act requires a human author.[2] Three weeks later, the Court narrowed contributory copyright liability in *Cox Communications, Inc. v. Sony Music Entertainment*, a ruling whose most important consequences for AI may be felt not by platforms but by their customers.[3] And in India, the Delhi High Court on July 24, 2026 refused interim relief to the news agency ANI against OpenAI, holding on a prima facie basis that storing news articles to train a model was reproduction but that it fell within fair dealing for private use and research.[4] An appeal is now pending before a Division Bench.[5]

The scholarship on generative AI and copyright is already large, but most of it asks one of two questions: whether training is lawful, or whether machine output deserves protection.[6] Both are framed around the model developer and the end user. Very little work examines the intermediary: the service provider that sits between a foundation model and a commissioning enterprise, and that now performs a large share of the world's software and content production. That gap matters more for India than for any other jurisdiction, because India's IT and business-process sector is built on precisely this intermediary role, and it matters for American businesses because they are its largest customers.

Minnesota's corporate base illustrates the stakes. Several of the state's largest employers, including Target, UnitedHealth Group through Optum, and 3M, maintain substantial technology and content operations in India, and many mid-sized medical-device and agricultural-technology firms rely on Indian development partners.[7] When these enterprises commission software, patient-education materials, marketing assets, or training content from Indian vendors that now rely on generative AI, the four risks mapped in this Article cease to be abstractions and become diligence and balance-sheet questions for Minnesota boards, general counsel, and the venture investors who fund the region's software and health-technology companies. The framework proposed in Section 8 is therefore not only a comparative exercise but a governance tool for that community.

This Article offers three contributions. The first is analytical. It breaks outsourced AI creation into a four-layer supply chain (source works, model developer, outsourcing vendor, enterprise client) and identifies four risk interfaces where copyright doctrine creates exposure. Drawing on the supply-chain approach developed by Lee, Cooper, and Grimmelmann,[8] it adds the layer their account largely leaves implicit, the contracting intermediary, and shows that the intermediary is where risk tends to pool.

The second contribution is comparative. The Article sets the U.S. and Indian regimes side by side at each interface. The comparison yields a result that should unsettle practitioners: the two systems are not simply different but mismatched. A deliverable can be owned under Indian law and unprotected under U.S. law; an act of training can be provisionally lawful in Delhi and actionable in Delaware; an

assignment valid in New York can lapse in Mumbai for want of exercise. The Article traces how these mismatches interact in a single cross-border deal.

The third contribution is normative and practical. Section 8 develops a framework, which I call Layered Provenance Allocation, for distributing these risks by contract and, where contract cannot reach, by policy. Its central claim is that because copyright title in AI-heavy deliverables may not exist, parties should stop warranting title and start warranting process: what tools were used, with what settings, on what data, and with what documented human contribution.

The Article proceeds in seven Sections. Section 2 describes the economics and architecture of outsourced AI creation. Section 3 analyzes the authorship gap. Section 4 examines how ownership fractures in the outsourcing contract, including a conflict-of-laws problem that has received almost no attention. Section 5 addresses input liability for training, fine-tuning, and retrieval. Section 6 turns to output liability and to the secondary-liability shift after *Cox*. Section 7 evaluates the contractual tools the market currently relies on and explains why AI indemnities are often less than they appear. Section 8 sets out the proposed framework and recommendations for both legislatures.

2. The Anatomy of Outsourced AI Creation

2.1 From Labor Arbitrage to Model Arbitrage

For three decades, the Indian outsourcing model rested on labor arbitrage: skilled engineers and writers in Bengaluru, Pune, or Hyderabad produced work at a fraction of U.S. cost. The legal architecture that supported the model was simple. The vendor's employees created the work in the course of employment, the vendor became first owner by operation of law, and a contractual assignment carried title to the client. Both countries' statutes accommodated this sequence without strain.[9]

Generative AI replaces labor arbitrage with what might be called model arbitrage. The vendor's margin now depends less on the hourly cost of its people than on how much of each deliverable a licensed model can generate, and on how little human time is needed to review it. The Reuters reporting cited above captures the commercial side of the shift: clients expect the productivity gains to be passed through, and outcome-based pricing makes the vendor's revenue depend on maximizing automation.[10] The legal side of the shift has received less attention. Each increment of automation removes a human contribution on which the traditional ownership chain depended, and each new tool inserts another licensor, with its own terms, into the chain of title.

The incentive structure deserves emphasis. Under a time-and-materials contract, the vendor gains nothing by automating; the client bears the cost of human hours and receives human authorship. Under an outcome-based contract, the vendor gains from every hour of human work it eliminates, while the client bears the legal consequences of that elimination (weaker copyright, higher infringement exposure) without seeing them. The contract thus separates the party that decides how much AI to use from the party that absorbs the copyright cost of that decision. A lawyer trained in agency theory will recognize the pattern: a principal who cannot observe the agent's effort, and an agent whose compensation rewards cutting the very effort that protects the principal.

2.2 A Four-Layer Supply Chain

Figure 1 models the arrangement. At Layer 0 sit the source works: the books, news articles, code repositories, images, and song lyrics that make up training corpora, some licensed, many scraped, and in several litigated cases pirated. Layer 1 is the model developer, typically a U.S. company that trains a foundation model and licenses it through terms of use or enterprise agreements. Layer 2 is the

outsourcing vendor, which adapts the model through fine-tuning or retrieval-augmented generation, engineers prompts, reviews outputs, and delivers finished work. Layer 3 is the enterprise client that commissions and commercializes the deliverable. Layer 4 is the market, a term I use broadly to include customers, regulators, rights holders, and competitors.

Figure 1. The outsourced generative-AI supply chain.

Contractual promises (licenses, warranties, indemnities) flow downstream but are conditioned and capped; liability claims flow upstream but are blunted by limits on secondary liability, concentrating direct exposure at the vendor and client layers. Source: author's analysis.

Four risk interfaces arise between these layers. Risk 1 (R1) is input liability: the reproduction of protected works to build or adapt a model. Risk 2 (R2) concerns memorization and retrieval, where the model or its retrieval layer reproduces protected expression in its output. Risk 3 (R3) is the authorship gap and the related fracture in the chain of title: the possibility that the deliverable has no author recognized by law, or that ownership cannot be transferred as the parties assume. Risk 4 (R4) is output infringement and the loss of exclusivity, where the deliverable either infringes a third party's work or cannot be protected against copyists.

The model is a simplification. Real deployments often involve several models, open-weight components hosted on third-party clouds, and subcontractors within the vendor's own organization. But the four-layer structure captures the feature that matters most for legal analysis: the party that makes the key technical choices (Layer 2) is neither the party that built the model nor the party that owns the business risk of the deliverable.

2.3 Why Copyright Is the Load-Bearing Risk

Outsourced AI raises data-protection, confidentiality, product-liability, and export-control questions as well, and a complete governance program must address them. This Article concentrates on copyright for a structural reason. Most outsourcing value is delivered in the form of works that copyright law protects: source code is a literary work in both countries,[11] and documentation, marketing copy, illustrations, and audiovisual content fall squarely within its subject matter. When the client pays for a deliverable, it is typically paying for the exclusive right to exploit it. If copyright in the deliverable is weak, uncertain, or absent, the client has paid for something less than it believes. And if the deliverable infringes, the client, as the publisher, is usually the most visible defendant. Copyright is therefore the body of law that determines whether the core asset of the transaction exists and whether it is safe to use.

3. Authorship Without an Author: The Output Problem

The authorship gap (R3) is the risk least visible to business teams and most consequential for the value of the deliverable. It is also where U.S. and Indian law diverge most sharply. This Section examines each system and then asks what the divergence means in theory and in practice.

3.1 The United States: Human Authorship as a Settled Threshold

3.1.1 Doctrinal Foundations

The Copyright Clause empowers Congress to secure to "Authors" exclusive rights in their "Writings." [12] The Supreme Court has never squarely held that an author must be human, but its cases consistently assume so. In *Burrow-Giles Lithographic Co. v. Sarony*, the Court upheld copyright in a photograph because it reflected the "original intellectual conceptions of the author," describing the author as the person "to whom anything owes its origin." [13] The case is instructive for the AI debate because the camera was the disruptive technology of its day, and the Court located authorship in the

photographer's choices of pose, costume, lighting, and arrangement, not in the machine. *Feist Publications, Inc. v. Rural Telephone Service Co.* then fixed originality, meaning independent creation plus "at least some minimal degree of creativity," as the constitutional floor.[14] In *Naruto v. Slater*, the Ninth Circuit held that a crested macaque lacked statutory standing to sue over the "monkey selfie," reasoning that the Act's references to an author's children, widow, and heirs presuppose human beings.[15]

3.1.2 Thaler and the End of the Autonomous-Author Argument

Stephen Thaler sought registration of *A Recent Entrance to Paradise*, an image he said was created autonomously by his "Creativity Machine," listing the machine as author and himself as owner under the work-made-for-hire doctrine. The Copyright Office refused, the district court affirmed, and in March 2025 the D.C. Circuit affirmed again.[16] The court's reasoning was statutory rather than constitutional. It read multiple provisions of the 1976 Act (on duration measured by the author's life, on inheritance, on signatures, on nationality and domicile) as coherent only if the author is a human being. It also rejected the work-made-for-hire theory: a machine cannot be an "employee," and in any event the doctrine determines who is treated as author of a work that already has one; it does not create an author where none exists.[17] The Supreme Court's denial of certiorari in March 2026 ended the case.[18]

Commentators rightly note that *Thaler* decided less than it seemed to. Thaler had insisted that no human contributed to the work, so the court did not have to decide how much human input suffices.[19] That is the question that matters for outsourcing, because no commercial vendor claims its deliverables emerged without human involvement. The live issue is where along the spectrum of human participation protection begins.

3.1.3 The Copyright Office's Framework

The Copyright Office has filled the space *Thaler* left open. Its March 2023 registration guidance requires applicants to disclose AI-generated material that is more than de minimis and to exclude it from the claim.[20] In *Zarya of the Dawn*, the Office cancelled and reissued a registration for a graphic novel, protecting the human-written text and the selection and arrangement of images but not the Midjourney-generated images themselves.[21] The Review Board applied the same logic to refuse Jason Allen's *Théâtre D'opéra Spatial*, which was produced through more than six hundred prompts and later edited in Photoshop, because Allen had not disclaimed the AI-generated portions.[22] Allen's challenge remains pending in federal court in Colorado; cross-motions for summary judgment on whether iterative prompting can amount to authorship were briefed in early 2026, and no decision had been reported as of this writing.[23]

The Office's January 2025 report on copyrightability consolidates this approach.[24] It concludes that prompts alone generally do not give the user sufficient control over the expressive elements of the output, because the same prompt can yield many different results and the model, not the user, determines how the instruction is executed. It recognizes three routes to protection: human expression that is perceptible in the output (for example, a user's own drawing that the model modifies), creative selection and arrangement of AI-generated material, and creative modification of that material. It declines to recommend legislation.

3.1.4 Assessment

The Office's position is doctrinally defensible, but it rests on a theory of control that sits uneasily with the history of other tools. A photographer does not control each photon, and a director does not control

every gesture of an actor, yet both are authors. The better explanation of the Office's line is not control as such but predictability of expression: the photographer's choices map onto the image in a way the photographer can anticipate and revise, whereas a prompt's relation to the output is stochastic. Ginsburg and Budiardjo made a similar point before the current generation of models, arguing that authorship requires both conception and execution by a human, and that a person who supplies only a general idea while a machine supplies the expressive execution has not authored anything.[25]

For the outsourcing client, the practical implication is stark. The more efficiently a vendor uses AI, the less of the deliverable is protected. A codebase written largely by an assistant and accepted with light review may contain protectable human contributions in its architecture, selection, and arrangement, but its individual functions may be free for anyone to copy. The client cannot register what it does not own, and an unregistered U.S. work cannot support statutory damages or attorney's fees for infringement that begins before registration.[26]

3.2 India: The Uncertain Promise of Section 2(d)(vi)

3.2.1 Text and Origin

India's Copyright Act appears to answer the question that U.S. law leaves open. Section 2(d)(vi), inserted by the Copyright (Amendment) Act, 1994, provides that the "author," "in relation to any literary, dramatic, musical or artistic work which is computer-generated," means "the person who causes the work to be created." [27] The drafting tracks section 9(3) of the United Kingdom's Copyright, Designs and Patents Act 1988, which vests authorship of computer-generated works in "the person by whom the arrangements necessary for the creation of the work are undertaken." [28] Neither provision was written with generative models in mind; both were responses to the computer-assisted works of the 1980s, such as weather maps and database outputs, where it seemed natural to credit whoever set the machine running. Pamela Samuelson's early analysis of computer-generated works anticipated the difficulty of choosing among programmer, user, and machine, and warned that allocating rights to the user was the least bad option only because the programmer had already been rewarded through copyright in the program.[29]

3.2.2 Three Problems of Interpretation

Section 2(d)(vi) raises at least three unresolved questions, each of which matters directly to an outsourcing deal.

The first is whether the provision assumes that a computer-generated work is protectable or merely identifies the author if it is. Section 13 extends copyright only to "original" works.[30] In *Eastern Book Co. v. D.B. Modak*, the Supreme Court of India rejected both the "sweat of the brow" approach and the high novelty standard, adopting instead a "skill and judgment" test that requires a minimal degree of creativity, drawing on the Canadian Supreme Court's decision in *CCH Canadian*. [31] If originality requires the author's own skill and judgment, and the author of a computer-generated work is by definition someone who did not exercise creative judgment over its expression, then section 2(d)(vi) may name an author for a work that section 13 does not protect. The Delhi High Court's interim ruling in *ANI* restated the *Modak* standard in the course of rejecting protection for bare facts, which suggests that the courts will not abandon it lightly.[32]

The second question is who "causes" the work to be created. In a generative-AI workflow, candidates include the model developer who built the system, the vendor that configured it and wrote the system prompt, the vendor's engineer who typed the user prompt, and the client whose specification set the task. A causal reading that looks for the proximate human trigger would favor the engineer; an

organizational reading that looks for the party that made "arrangements" would favor the vendor. The word "causes" is broader than the U.K. formulation, and Indian courts have not interpreted it.

The third question is how section 2(d)(vi) interacts with first ownership under section 17. If the vendor's employee is the author because she caused the work to be created, section 17(c) makes the vendor first owner where the work was made "in the course of the author's employment under a contract of service." [33] If instead the vendor is itself the "person" who causes creation, it is both author and first owner. Either way, the client acquires nothing without an assignment, a point developed in Section 4.

3.2.3 Administrative Practice and Policy Signals

The Indian Copyright Office's handling of *SURYAST* shows how unsettled the position is. In 2020 the Office registered the artwork with two authors, the applicant Ankit Sahni and the "RAGHAV Artificial Intelligence Painting App." In November 2021 it issued a notice seeking withdrawal of the registration and asking the applicant to explain the legal status of the AI tool; the applicant responded that the Act gave the Registrar no power to review a registration once granted. [34] No conclusive public resolution has followed. The episode demonstrates that registration in India offers little assurance: the Office itself doubted its own grant, and registration there is in any event not a precondition to ownership or suit.

At the policy level, the Department-related Parliamentary Standing Committee on Commerce recommended in 2021 that the government revisit the Copyright Act and the Patents Act to address AI-generated works. [35] The Department for Promotion of Industry and Internal Trade (DPIIT) published the first part of a working paper on generative AI and copyright in December 2025, dealing with training. [36] A second part on outputs, authorship, and moral rights was expected within months but, as of this writing, has not been released. [37]

3.3 The Divergence Summarized

Table 1 compares the two regimes along the dimensions most relevant to a commercial buyer. Figure 2 then plots where each system's protection ends along a spectrum of declining human control.

Table 1: Authorship and Ownership of AI-Involved Works: United States and India Compared

Issue	United States	India	Business consequence
Statutory anchor	Copyright Clause; 17 U.S.C. §§ 101, 102(a), 201	Copyright Act, 1957, §§ 2(d)(vi), 13, 17	Different starting points for the same deliverable
Authorship of AI output	Human author required; <i>Thaler</i> (cert. denied 2026)	"Person who causes the work to be created" is author of computer-generated work	A work may have an owner in India and none in the U.S.
Originality standard	Independent creation plus minimal creativity (<i>Feist</i>)	Skill and judgment plus minimal creativity (<i>Modak</i>)	Indian protection for pure AI output remains doubtful despite § 2(d)(vi)
Effect of prompting	Prompts alone generally insufficient (Part 2 Report)	Untested; "causes" may reach the prompter or the operator	Prompt logs may matter more in India than in the U.S.

Issue	United States	India	Business consequence
Registration	Disclosure and disclaimer of AI material required; prerequisite to suit for U.S. works	Voluntary; <i>SURYAST</i> withdrawal notice shows instability	U.S. registration strategy must be planned at delivery
Employer or commissioner ownership	Work made for hire: employees, or nine commissioned categories with signed writing	§ 17(c) contract of service; § 17(b) commissioned photographs, paintings, portraits, engravings, films	Contractor-made code and text need express assignment in both systems
Policy trajectory	No legislative change recommended; case-by-case	DPIIT Part 2 on outputs pending	Contract drafting must anticipate change in India

Figure 2. The authorship spectrum.

Green indicates protection is likely; amber indicates genuine uncertainty; red indicates protection is likely unavailable. Source: author's analysis of the authorities discussed in Section 3.

3.4 Theory: Causal Authorship and Creative Authorship

The divergence reflects two competing conceptions of authorship. The first, which I will call creative authorship, locates authorship in the mental conception and execution of expression. It underlies *Burrow-Giles*, *Feist*, and *Thaler*, and it fits naturally with both natural-rights theories, in which the work is an extension of the author's personality, and with Lockean labor theory, in which the author mixes her intellectual labor with the commons. The second, causal authorship, locates authorship in the person responsible for bringing the work into existence, whether or not that person shaped its expression. It underlies section 2(d)(vi) and its U.K. counterpart and fits an incentive theory of copyright, in which the question is simply whom the law must reward to ensure the work gets made.

Neither conception is adequate on its own terms for generative AI. Creative authorship risks leaving economically valuable output in the public domain, which may reduce incentives to invest in AI-enabled production and encourages parties to overstate human contribution. Causal authorship, meanwhile, rewards investment rather than creativity and has no principled way to choose among the many humans in the causal chain. It also sits awkwardly with the constitutional-style originality floor that Indian courts adopted in *Modak*. Gervais has argued that the incentive rationale does not support protecting machine output at all, because the machine needs no incentive and the human investment is already rewarded through markets for the model and its services.[38] Lemley has gone further, suggesting that generative AI inverts copyright's traditional premise: when expression becomes cheap, the valuable human contribution shifts to the idea and the prompt, which copyright has never protected.[39]

For present purposes, the theoretical dispute matters less for which side is right than for what it predicts. Because India's statute embodies causal authorship while its case law embodies creative authorship, Indian outcomes are likely to be unpredictable until the Supreme Court or Parliament resolves the tension. Because U.S. law has committed to creative authorship at every level, its outcomes are predictable in principle and uncertain only at the margin, in the zone of selection, arrangement, and

modification. A risk manager should therefore treat U.S. protection of AI-heavy deliverables as low but calculable, and Indian protection as possibly higher but unreliable.

4. Ownership in the Outsourcing Contract: Where Title Fractures

Even where a deliverable is protected, the outsourcing client does not own it unless the law or a valid transfer puts title in the client's hands. Traditional outsourcing agreements rely on two devices: the U.S. work-made-for-hire doctrine and a back-up assignment clause. This Section shows that both devices perform poorly for AI-involved work, and that Indian assignment formalities add traps that U.S. drafters routinely overlook.

4.1 The U.S. Work-Made-for-Hire Doctrine and Its Limits

Under section 101 of the Copyright Act, a work made for hire is either a work prepared by an employee within the scope of employment or a specially ordered or commissioned work that falls within one of nine enumerated categories (a contribution to a collective work, part of a motion picture or other audiovisual work, a translation, a supplementary work, a compilation, an instructional text, a test, answer material for a test, or an atlas) and that the parties expressly agree in a signed writing to treat as made for hire.[40] In *Community for Creative Non-Violence v. Reid*, the Supreme Court held that "employee" carries its common-law agency meaning, to be determined by a multi-factor test that emphasizes the hiring party's right to control the manner and means of production.[41]

An Indian vendor's engineers are not the U.S. client's employees under *Reid*. Their deliverables therefore qualify as works made for hire only if they fit one of the nine categories. Stand-alone software, marketing copy, and most illustrations do not. Some documentation might qualify as an "instructional text" or a "supplementary work," and some video content as part of an "audiovisual work," but the fit is uncertain and the categories were not designed for these products. The work-made-for-hire clause in a typical master services agreement is therefore largely ineffective for the core deliverables, and ownership depends on the assignment clause.

Generative AI introduces a second difficulty. The work-made-for-hire doctrine, as *Thaler* confirmed, allocates authorship of a work that already has a human author; it cannot manufacture one.[42] If a portion of the deliverable was generated without sufficient human authorship, no contractual label can make the client its author. The same is true of assignment. A transfer of copyright under section 204(a) must be in a writing signed by the owner,[43] but a signed writing that purports to assign copyright in unprotectable material assigns nothing. The client receives the physical code and files, and whatever contractual rights the agreement grants against the vendor, but not an exclusive right enforceable against the world.

This is the first half of the title fracture: the chain of title is not merely weak but, for some components, empty. The consequences show up in diligence. An acquirer reviewing a target's software assets, a lender taking a security interest in intellectual property, or an investor relying on a representation that the company "owns all right, title, and interest" in its products may discover that a portion of those assets is not property at all. Representations and warranties drafted before 2023 routinely assume otherwise.

A less obvious U.S. risk arises where the parties fail to execute an effective assignment for the human-authored components. The Ninth Circuit's decision in *Effects Associates, Inc. v. Cohen* holds that a commissioning party that pays for a work without a written transfer receives only a nonexclusive implied license, while the creator keeps the copyright.[44] In multi-tier outsourcing, where the vendor

uses subcontractors or freelance prompt engineers, a missing sub-assignment at any link leaves the client with a license rather than ownership.

4.2 India: Contract of Service, Commissioned Works, and Section 19

4.2.1 First Ownership Under Section 17

Section 17 makes the author the first owner, subject to provisos. Proviso (c) gives the employer first ownership of works made "in the course of the author's employment under a contract of service or apprenticeship."^[45] The distinction between a contract *of* service (employment) and a contract *for* services (independent contracting) thus determines whether the vendor owns its personnel's output in the first place. For salaried engineers the answer is usually clear. For the freelancers and gig workers increasingly used in content operations, it is not.

Proviso (b) is more interesting for AI deliverables and has received comparatively little attention in the commentary. It provides that in the case of "a photograph taken, or a painting or portrait drawn, or an engraving or a cinematograph film made, for valuable consideration at the instance of any person," that person is first owner "in the absence of any agreement to the contrary."^[46] The Supreme Court in *Indian Performing Right Society Ltd. v. Eastern India Motion Pictures Ass'n* read section 17's provisos as vesting first ownership in the commissioning producer where their conditions were met.^[47] If AI-generated images or videos are "paintings," "photographs," or "cinematograph films" under the Act's broad definitions, and if they are protected at all, a commissioning client may be first owner by operation of law. The categories do not, however, cover code, text, or music, which are the bulk of outsourced deliverables. The upshot is an odd patchwork: a U.S. client commissioning AI illustrations from an Indian studio may have a stronger statutory claim to ownership in India than in the United States, while its claim to the accompanying code is weaker in both.

4.2.2 Assignment Formalities as Hidden Traps

Where ownership must pass by assignment, Indian law imposes requirements that are stricter than section 204(a) and that U.S.-drafted agreements frequently fail to meet. Section 19 requires that an assignment be in writing, signed by the assignor, and that it "identify the work, and specify the rights assigned and the duration and territorial extent of such assignment."^[48] Three default rules follow, each of which can defeat the parties' expectations:

First, under section 19(4), if the assignee does not exercise the assigned rights within one year of the assignment, the assignment "shall be deemed to have lapsed" unless the instrument provides otherwise.^[49] A client that commissions a large library of content and publishes only part of it in the first year may find that rights in the unused portion have reverted to the vendor.

Second, under section 19(5), if the duration is not stated, the assignment is deemed to be for five years.^[50] Agreements that say "hereby assigns all right, title, and interest" without stating a term, a common formulation in U.S. forms, may therefore transfer only a five-year interest under Indian law.

Third, under section 19(6), if the territorial extent is not specified, it is presumed to extend only to India.^[51] For a U.S. client, whose principal market is outside India, this default is particularly damaging.

A further trap lies in the third proviso to section 18(1), added in 2012, which provides that an assignment does not extend to "any medium or mode of exploitation of the work which did not exist or was not in commercial use at the time when the assignment was made, unless the assignment specifically referred to such medium or mode."^[52] A legacy assignment that predates the commercial use of generative AI

may not carry the right to use the assigned work as training data or as grounding material in a retrieval system, because those modes were not in commercial use when the assignment was made. This matters where a client wants to use years of vendor-produced documentation to fine-tune its own model.

Finally, section 57 confers on authors moral rights of attribution and integrity that survive assignment.[53] The Delhi High Court's decision in *Amar Nath Sehgal v. Union of India* shows that Indian courts take these rights seriously, awarding damages to a sculptor whose mural the government removed and damaged.[54] Whether AI-assisted deliverables carry moral rights for the vendor's human contributors, and whether contractual waivers of those rights are enforceable, remain open questions that DPIIT's pending second paper is expected to address.[55]

4.3 The Conflict-of-Laws Problem

The comparative analysis so far has treated each legal system separately. In a cross-border outsourcing deal, however, the two systems apply simultaneously, and their interaction produces a problem that has received very little attention. Suppose an Indian vendor's engineer in Pune produces an AI-generated illustration with minimal creative input. Under section 2(d)(vi), read literally, the engineer or the vendor may be its author, and the work may be protected in India. Under U.S. law, it is not protected. Who owns what, and where?

U.S. courts do not apply U.S. ownership rules mechanically to foreign works. In *Itar-Tass Russian News Agency v. Russian Kurier, Inc.*, the Second Circuit held that questions of ownership are governed by the law of the country with the most significant relationship to the work and the parties, usually the country of origin, while questions of infringement are governed by the law of the place where the infringement occurs.[56] The Berne Convention, to which both countries adhere, provides that the extent of protection is governed exclusively by the law of the country where protection is claimed.[57]

Combining these rules yields a split result. Whether copyright *subsists* in the United States is a question of U.S. law, and *Thaler* forecloses protection for material lacking human authorship. Whether a *person* owns whatever rights subsist may be a question of Indian law. For the purely AI-generated components, the answer in the United States is therefore that nothing subsists to be owned. For the human-authored components, however, Indian law may determine initial ownership, which means that section 17 and section 19 can govern title to rights the client will seek to enforce in U.S. courts. A U.S. client that relies solely on a New York-law assignment clause may find that a U.S. court looks to Indian law to decide whether the vendor ever had the rights it purported to transfer, and whether the transfer satisfied Indian formalities.

The practical lesson is that choice-of-law clauses in outsourcing agreements do not settle ownership. Parties can choose the law that governs their contract, but they cannot contract out of the rule that a court may apply the law of the country of origin to determine initial ownership. The only reliable solution is to draft an assignment that satisfies both systems simultaneously: signed, work-specific, rights-specific, perpetual (or for the full term of copyright), worldwide, expressly covering future and not-yet-existing media including machine-learning uses, and expressly excluding the section 19(4) lapse rule.

5. Input Liability: Training, Fine-Tuning, and Retrieval

Input liability (R1) is usually discussed as a problem for model developers, and in its largest form it is. But outsourcing vendors increasingly perform their own input operations: they fine-tune models on client or third-party corpora, and they build retrieval-augmented generation (RAG) systems that copy documents into vector databases and feed them to the model at query time.[58] These operations create

input liability at Layer 2 that does not depend on the lawfulness of the foundation model. This Section examines the U.S. and Indian positions and then isolates the risks peculiar to the vendor.

5.1 The United States: Fair Use in Flux

5.1.1 The District Court Split

Three district court decisions in 2025 frame the U.S. position, and they do not point in one direction.

In *Thomson Reuters Enterprise Centre GmbH v. Ross Intelligence Inc.*, Judge Bibas, sitting by designation, granted partial summary judgment against a legal-research start-up that had used Westlaw headnotes, obtained through a third party, to train a non-generative AI search tool. He held that the use was not fair because it was commercial, not transformative, and aimed at building a competing product.[59] The Third Circuit accepted an interlocutory appeal on originality and fair use and heard argument on June 11, 2026; a decision is expected later this year.[60]

In *Bartz v. Anthropic PBC*, Judge Alsup reached a different result for generative training. He described the use of lawfully acquired books to train a large language model as "exceedingly transformative" and fair, and he treated the scanning of purchased print books as a fair format change. He refused, however, to extend fair use to the company's retention of millions of books downloaded from pirate libraries to build a permanent central library, holding that such piracy was not justified by a later transformative use.[61] The parties then settled a certified class action for \$1.5 billion, roughly \$3,000 per work for more than 400,000 books; the court granted final approval in July 2026.[62]

In *Kadrey v. Meta Platforms, Inc.*, Judge Chhabria granted summary judgment to Meta, but in terms that offered developers little comfort. He concluded that the plaintiffs had failed to develop evidence of market harm, while stating that in many cases training on copyrighted works without permission would be unlawful because generative models can flood the market with competing works and "dilute" demand for the originals.[63]

The largest case is still pending. In the consolidated litigation in the Southern District of New York, including *The New York Times Co. v. Microsoft Corp.*, Judge Stein heard argument on summary judgment in September 2026, with the Department of Justice filing a brief supporting the defendants' fair use position.[64]

5.1.2 The Copyright Office's Middle Path

The Copyright Office's pre-publication report on training, released in May 2025, rejects both extreme positions. It treats training as a potentially transformative use but emphasizes that transformativeness depends on the purpose of the deployed model; training a model to generate outputs that compete with the originals in the same market is less likely to be fair. It gives weight to whether the copies were lawfully accessed and identifies market dilution and lost licensing markets as cognizable harms under the fourth factor.[65] The report's approach tracks the Supreme Court's instruction in *Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith* that the first factor asks whether the specific use at issue shares the purpose of the original, not whether the defendant's product is new in some general sense.[66]

5.1.3 What the Split Means for a Vendor

For a business buying outsourced AI work, the U.S. position has two clear lessons and one unclear one. First, provenance matters: *Bartz* and the Copyright Office agree that pirated or unlawfully accessed sources weaken or destroy a fair use defense, and the size of the *Bartz* settlement shows the price. Second, purpose matters: a fine-tuning run that trains a model to write in the style and substance of a

particular competitor's publications, for a product that substitutes for them, looks more like *Ross* than *Bartz*. What remains unclear is how appellate courts will treat market dilution, and until the Third Circuit and other courts of appeals rule, any risk assessment must include the possibility that one of them adopts a dilution theory broad enough to capture ordinary commercial fine-tuning.

5.2 India: Fair Dealing Without a Text-and-Data-Mining Exception

5.2.1 The Statutory Framework

India has no specific exception for text and data mining or machine learning. The general reproduction right in section 14(a)(i) expressly includes "the storing of it in any medium by electronic means."^[67] The exceptions in section 52 are enumerated rather than open-ended. Section 52(1)(a) permits fair dealing with a work (other than a computer programme) for private or personal use, including research; criticism or review; and reporting current events.^[68] Sections 52(1)(b) and (c) permit transient or incidental storage in the technical process of electronic transmission and for the purpose of providing links or access, subject to conditions.^[69] Indian courts have long insisted that the section 52 list is exhaustive, although the Delhi High Court's decision in the *DU Photocopy* case treated fair dealing as a user right to be read expansively in light of its purpose.^[70]

5.2.2 ANI Media v. OpenAI

ANI, an Indian news agency, sued OpenAI in November 2024, alleging that OpenAI had copied its articles to train ChatGPT and that the model reproduced its content in responses. Publishers, music labels, and industry bodies sought to intervene, and the court appointed amici curiae. On July 24, 2026, Justice Amit Bansal declined interim relief.^[71] The ruling contains four holdings of direct importance to businesses operating in India.

First, on jurisdiction, the court rejected OpenAI's argument that the absence of servers in India placed training outside the court's reach. Following its earlier decision in *Neetu Singh v. Telegram FZ LLC*, it treated storage on offshore servers as only the "terminal step" in a chain that began with access to the plaintiff's content from India.^[72]

Second, the court held on a prima facie basis that electronic storage of works in the course of training is reproduction within section 14(a)(i), whether temporary or permanent, and that deleting the data after training does not undo the reproduction.

Third, and most significantly, the court held that the storage was prima facie protected by section 52(1)(a)(i) as fair dealing for private or personal use, including research. It reasoned that commercial use does not by itself exclude this purpose, that the training corpus remained internal and inaccessible to third parties, that no market substitution had been demonstrated, and that the public benefits of the technology were substantial.^[73] The court confined its reasoning to freely available material; it did not decide the position for paywalled or pirated sources.

Fourth, the court distinguished memorization from retrieval. ANI's examples of infringing outputs involved articles published after the relevant training cut-offs, and the court concluded that such responses were likely produced through retrieval of live web content rather than recall of trained data. Responses "not based on trained data cannot be used against Open AI in respect of claims based on memorisation"; they must instead be tested as independent acts of reproduction.^[74]

ANI has appealed. A Division Bench of the Delhi High Court has issued notice, directed OpenAI to respond to the renewed application for interim restraint, and listed the matter for December 5, 2026; an industry body has also applied to intervene.^[75] Two contingencies therefore remain open as of this

writing: the Division Bench may narrow or reverse the fair dealing holding, and DPIIT's second working paper, which is expected to address outputs, authorship, and moral rights, has not yet been published.[76]

5.2.3 Critical Assessment

The interim ruling is a significant victory for AI developers, but its reasoning is vulnerable on appeal, and businesses should not treat it as settled law.

The weakest link is the classification of commercial model training as "private or personal use." The phrase was introduced to protect individuals, and its most natural reading refers to non-public use by a person for their own purposes. Treating the internal operations of a global company that serves hundreds of millions of users as "private" stretches the term well beyond that core. The court's answer, that the training corpus is not itself disclosed to third parties, would render almost any internal corporate copying "private," a result difficult to reconcile with the enumerated and purposive structure of section 52. Aparajita Lath has criticized a related move: the court's elevation of "public interest" into a freestanding consideration in the fairness analysis, which she argues lacks limiting principles and could favor any socially attractive use over the rights the statute protects.[77]

A second tension is institutional. DPIIT's December 2025 working paper proposed a "hybrid model" under which AI developers could train on lawfully accessed works without permission, but would pay statutory remuneration set as a percentage of global revenue and distributed through a new collective, the Copyright Royalties Collective for AI Training (CRCAT).[78] The proposal rests on the premise that training implicates the reproduction right and requires compensation. The *ANI* ruling, if affirmed, would make that premise largely irrelevant for freely available material, because a use that is fair dealing requires neither permission nor payment. Critics of the DPIIT proposal, including the Takshashila Institution, have warned that a global-revenue royalty would be difficult to administer and could disadvantage Indian developers.[79] Businesses therefore face two plausible Indian futures that could hardly be further apart: free training under an expansive fair dealing doctrine, or mandatory remunerated licensing under new legislation.

5.3 Fine-Tuning and Retrieval: The Vendor's Own Exposure

The foundation-model litigation tends to overshadow the input operations that vendors perform themselves, even though those operations carry their own risk.

Fine-tuning a model on a narrow corpus, such as a competitor's technical manuals, a particular publisher's articles, or a style guide, differs from training a general model on billions of documents in ways that matter to fair use and fair dealing. The corpus is smaller, so each work contributes more to the model's behavior; the purpose is narrower, so the resulting output is more likely to substitute for the source; and the risk of memorization rises as the dataset shrinks. Research on extraction attacks has shown that both image and language models can regurgitate training examples verbatim, and that the risk is concentrated in duplicated or distinctive data.[80] A Munich court has already held that memorization of song lyrics in model parameters is itself a reproduction under German law.[81] The English High Court reached a different conclusion in *Getty Images v. Stability AI*, holding that model weights that do not store the training images are not an "infringing copy," though Getty has been granted permission to appeal.[82] A vendor that fine-tunes for a multinational client can thus face different answers to the same technical question in each market the client serves.

Retrieval is riskier still. A RAG system copies source documents into an index and inserts passages from them into the model's context at query time, so that the output is often a close paraphrase or a

verbatim excerpt. None of the training-stage reasoning in *Bartz* or *ANI* applies straightforwardly to this process. *ANI* itself separated retrieval-based responses from memorization claims and left them to be judged as ordinary reproduction.[83] In India, the transient-storage exceptions in sections 52(1)(b) and (c) were drafted for network transmission and linking, not for building persistent indexes of full-text works, and they are unlikely to cover a vendor's vector database.[84] In the United States, a RAG output that reproduces substantial portions of a news article for a user who would otherwise read the article looks less like the intermediate copying in *Authors Guild v. Google, Inc.*, where the court emphasized that Google Books displayed only snippets that did not substitute for the books,[85] and more like a substitute. The European Union's AI Act adds a regulatory overlay for general-purpose model providers, requiring a copyright-compliance policy and a public summary of training content, obligations that can attach to a downstream party whose modifications are significant enough to make it a provider in its own right.[86]

For a client, the key point is that a vendor's RAG pipeline is a copying operation for which the vendor, not the model developer, is primarily responsible, and for which no foundation-model indemnity is likely to provide cover.

6. Output Liability and the Shift Downstream

6.1 Direct Infringement by Outputs

When an AI-generated deliverable is substantially similar to a protected work, the person who reproduces, distributes, or publicly displays it may be liable for direct infringement, and in both countries that liability does not depend on intent. In the United States, a plaintiff must show copying and substantial similarity; access can be inferred where the work was in the training data, and memorization research makes that inference more plausible for widely duplicated works.[87] In *Andersen v. Stability AI Ltd.*, the court allowed artists' claims to proceed on the theory that a diffusion model may itself embody protected expression and that its developer may be liable for inducing infringement, while dismissing other claims.[88] The Second Circuit's volitional-conduct requirement, under which a system operator is not a direct infringer when copies are made automatically at a user's command, tends to push direct liability toward the human who issues the command and uses the result.[89] In the outsourcing chain, that human is the vendor's employee, and the party that publishes the result is the client.

Indian law is similar in structure but harsher in remedy. Section 51 treats as infringement any unlicensed exercise of the exclusive rights, and section 51(a)(ii) extends liability to anyone who permits a place to be used for profit for infringing communication unless unaware of it. The Delhi High Court's Division Bench in *MySpace Inc. v. Super Cassettes Industries Ltd.* read "place" to include online spaces but required actual knowledge for platform liability.[90] More importantly for business planning, knowing infringement is a criminal offense under section 63, punishable by imprisonment of six months to three years and a fine, and section 69 extends liability to company officers in charge of the business unless they prove lack of knowledge or due diligence.[91] Criminal exposure for managers changes the calculus of an Indian vendor's leadership in a way that has no close U.S. analogue.

6.2 Cox v. Sony and the Downstream Migration of Risk

In March 2026 the Supreme Court decided *Cox Communications, Inc. v. Sony Music Entertainment*, reversing a holding that an internet service provider was contributorily liable for its subscribers' infringement, a holding that had supported a \$1 billion jury verdict. The Court held that contributory copyright liability requires either inducement of infringement or the provision of a service tailored to

infringement, and that "mere knowledge that a service will be used to infringe is insufficient to establish the required intent." [92] The decision builds on the Court's earlier framework in *Sony Corp. of America v. Universal City Studios, Inc.* and *Metro-Goldwyn-Mayer Studios Inc. v. Gorkster, Ltd.*, which protected technologies capable of substantial noninfringing uses while imposing liability on those who actively promote infringement. [93]

Cox was not an AI case, but its logic applies directly to AI providers. A general-purpose model with substantial legitimate uses, marketed without encouragement of infringement, will rarely be "tailored to infringement." Rights holders who cannot reach the model developer through secondary liability will look to direct infringers instead. In the outsourcing chain, those are the vendor that generates and delivers the infringing output and the client that publishes it. *Cox* thus has a distributive effect that its business commentary has largely overlooked: by narrowing liability for the upstream technology provider, it concentrates liability at the downstream layers where the human act of copying and distribution occurs. The upstream provider's contractual indemnity, examined in Section 7, becomes the main route by which that liability can be shifted back up the chain, which makes the conditions attached to that indemnity far more important.

6.3 Adjacent Risks: Personality, Voice, and Synthetic Media Rules

Outsourced content production also raises risks that sit next to copyright and are often litigated alongside it. Indian courts have developed a robust personality-rights jurisprudence in response to generative tools. The Delhi High Court restrained the unauthorized use of the actor Anil Kapoor's name, image, voice, and likeness, including through AI-generated deepfakes, [94] and the Bombay High Court granted similar relief to the singer Arijit Singh against platforms offering AI voice-cloning of his voice. [95] A vendor that uses a "sound-alike" voice for a client's advertising campaign may therefore face injunctive relief in India even if no copyright is infringed. This is not a marginal scenario in outsourced production. Much of the marketing, onboarding, and training video that Indian vendors now deliver is voiced or presented synthetically, and the clearance question is rarely addressed in the statement of work: a synthetic narrator that resembles a recognizable performer, or a demonstration video built around a celebrity-like likeness, exposes the client to personality-rights claims that its copyright warranties do not touch and its AI provider does not indemnify. The resulting risk interface belongs to the same family as the authorship and infringement problems analyzed here, but it is governed by a different body of law, and it is the subject of a growing Indian jurisprudence on synthetic performance and identity. [96] Outsourcing agreements should accordingly treat likeness, voice, and performance clearance as a separate warranty and indemnity item, with named-talent consents flowing through to the client.

In February 2026, the Ministry of Electronics and Information Technology amended the intermediary rules to regulate "synthetically generated information," requiring prominent labels on AI-generated audio-visual content that appears authentic and shortening takedown deadlines to three hours for unlawful content and two hours for non-consensual intimate imagery. [97] Comparable transparency duties now apply in the European Union, where providers and deployers of systems that generate or manipulate content must mark that content as artificially generated and, for deep fakes, disclose it. [98] These are platform and provider obligations, but they flow into vendor contracts, because clients that operate platforms will demand that deliverables arrive labelled and that vendors support rapid takedown.

7. The Indemnity Mirage: Contractual Risk Allocation

Faced with legal uncertainty, the market has turned to contract. Model developers offer copyright indemnities; vendors pass through warranties; clients demand uncapped intellectual property indemnities. This Section argues that these instruments provide considerably less protection than their headline terms suggest, and explains why.

7.1 Upstream Commitments and Their Conditions

In September 2023 Microsoft announced its Copilot Copyright Commitment, later extended as the Customer Copyright Commitment, under which it undertook to defend commercial customers against third-party copyright claims arising from the use of outputs of covered services.[99] OpenAI followed in November 2023 with "Copyright Shield" for ChatGPT Enterprise and its API, excluding the free and individual consumer tiers.[100]

The details matter. For Azure OpenAI services, Microsoft's commitment applies only if the customer implements specified mitigations: a system message instructing the model to avoid infringement, a documented testing and evaluation program, and, for text and code generation, protected-material filters and jailbreak protections switched on. Content flagged asynchronously is not covered, and customers must implement new mitigations within six months of their publication to preserve coverage.[101] Such conditions are reasonable from the provider's perspective; they reward responsible deployment. From the perspective of an outsourcing client, however, they create a compliance obligation that the vendor, not the client, controls. If the vendor's engineers switch off a filter to reduce latency, or fail to maintain an evaluation record, the indemnity may be lost without the client ever knowing.

Output-ownership terms follow a similar pattern. Major providers' terms typically assign to the customer whatever rights the provider may have in the output, while warning that outputs may not be unique and that other users may receive similar content.[102] An assignment of rights "if any" is an honest description of the authorship gap discussed in Section 3: the provider transfers what it has, which may be nothing. It is not a warranty that any copyright subsists, and it does not protect the client against a competitor who copies the deliverable.

7.2 The Indian Contract Law Overlay

When an Indian vendor gives the client a warranty or indemnity, Indian contract law may govern its scope, and it is narrower than U.S. drafters often assume. Section 124 of the Indian Contract Act, 1872 defines a contract of indemnity as a promise to save the other party from loss caused by the conduct of the promisor or of any other person.[103] Indian courts have read the statutory definition as non-exhaustive and have enforced broader contractual indemnities, but the careful drafter will specify that the indemnity covers third-party claims regardless of fault, defense costs, settlements, and losses arising before judgment. Limitation-of-liability clauses in Indian vendor agreements commonly cap total liability at fees paid over a period, and unless intellectual property indemnities are expressly carved out of the cap, the client's recovery for a major infringement claim may be a small fraction of its exposure.

Enforcement is a separate problem. The United States is not a "reciprocating territory" under section 44A of the Code of Civil Procedure, so a U.S. judgment against an Indian vendor cannot be executed directly in India; the client must bring a fresh suit on the judgment, subject to the defenses in section 13.[104] Arbitral awards made in a New York Convention country are, by contrast, enforceable under Part II of the Arbitration and Conciliation Act, 1996. For outsourcing indemnities, arbitration is therefore not merely a preference for confidentiality; it is often the only practical route to recovery.

7.3 A Risk Matrix

Table 2 draws the preceding analysis together. The severity and likelihood ratings are qualitative judgments for a typical enterprise deployment of the kind described in the Introduction; they will vary with the type of content, the tools used, and the markets served.

Table 2: Risk Matrix for Outsourced Generative-AI Deliverables (as of September 2026)

Risk	Primary locus	U.S. exposure	India exposure	Rating (likelihood / severity)	Priority mitigation
R1 Input liability (foundation model)	Layer 1	Fair use unsettled; <i>Bartz</i> settlement shows cost of pirated sources	<i>ANI</i> interim ruling favors training; appeal pending; DPIIT levy possible	Medium / High	Use providers with lawful-provenance commitments and indemnities; monitor appeals
R1 Input liability (vendor fine-tuning and RAG)	Layer 2	<i>Ross</i> and dilution theory apply with more force to narrow corpora	No TDM exception; § 52(1)(b)-(c) unlikely to cover persistent indexes	High / Medium	Licensed or client-owned corpora only; logged data lineage
R2 Memorization and retrieval	Layers 1-2	Direct liability for verbatim outputs; volitional conduct points to user	<i>ANI</i> treats RAG outputs as ordinary reproduction; criminal liability if knowing	Medium / High	Similarity screening; retrieval citation and quotation limits
R3 Authorship gap	Layers 2-3	No protection without human authorship; registration disclaimers required	§ 2(d)(vi) uncertain; <i>Modak</i> originality threshold	High / Medium	Documented human contribution; disclosure schedules
R3 Title fracture	Layers 2-3	Work-made-for-hire categories rarely fit; missing sub-assignments	§ 19(4)-(6) lapse, five-year, and India-only defaults; § 18 new-media proviso	High / High	Dual-compliant assignment; confirmatory deeds; subcontractor flow-downs
R4 Output infringement	Layers 2-3	<i>Cox</i> narrows upstream secondary liability; client as publisher	§§ 63, 69 criminal exposure; personality-rights injunctions	Medium / High	Conditioned upstream indemnity maintained by vendor; carve-out from caps
R4 Loss of exclusivity	Layer 3	Competitors may copy unprotected components	Protection possible but unreliable	High / Medium	Trade secret, contract, and trademark protection layered over copyright

7.4 Why Traditional Clauses Fail

Table 3 sets out the standard outsourcing clauses, explains why each performs poorly once generative AI enters the delivery process, and proposes replacement language in summary form.

Table 3: Re-drafting the Outsourcing Agreement for Generative AI

Clause	Traditional formulation	Failure mode with generative AI	Recommended approach
Ownership	"All deliverables are works made for hire and are hereby assigned to Client"	WMFH categories rarely fit; nothing to assign in unprotected output; Indian § 19 defaults	Dual-compliant assignment (worldwide, full term, all media including ML uses, § 19(4) disappplied) plus a license to any unprotectable material and a covenant not to reuse it for third parties
Title warranty	"Vendor owns and has the right to assign all IP in the deliverables"	Warranty is false or unverifiable for AI components; breach claim does not create title	Process warranty: disclosure of tools, versions, settings, data sources, and human contribution for each deliverable
Non-infringement warranty	"Deliverables will not infringe third-party rights"	Probabilistic outputs; vendor cannot guarantee; invites disputes over knowledge	Warranty of compliance with an agreed AI-use protocol (filters on, similarity screening, licensed corpora) plus indemnity
IP indemnity	Uncapped indemnity for third-party IP claims	Vendor's upstream indemnity is conditional and narrower; vendor may be judgment-proof or unenforceable in India	Indemnity carved out of the cap, backed by insurance and by an obligation to maintain upstream indemnity eligibility; arbitration seated in a New York Convention state
Subcontracting	Consent required for subcontractors	AI providers and freelance prompt engineers are not treated as subcontractors	Approved-tools list; flow-down of assignment and confidentiality to all human contributors; prohibition on consumer-tier tools
Confidentiality	Client data used only for the services	Inputs may be retained or used for model improvement by the AI provider	Enterprise tiers with no-training commitments; data residency and deletion terms

Clause	Traditional formulation	Failure mode with generative AI	Recommended approach
Audit	Financial and security audit rights	No visibility into prompts, logs, or model settings	Right to inspect generation logs, evaluation reports, and dataset lineage for a defined period

The common thread is a shift from promises about results to promises about conduct. The vendor cannot truthfully promise that an AI-generated deliverable is owned or non-infringing, but it can promise that it followed an agreed protocol, disclosed what it did, and preserved the records needed to defend the work.

8. Toward Layered Provenance Allocation

The preceding Sections show that the risks of outsourced AI creation are real, that they are distributed unevenly across the supply chain, and that neither legal system currently distributes them well. This Section proposes a framework for doing better. I call it Layered Provenance Allocation (LPA). It rests on four principles, each linked to a risk interface, and it is accompanied by recommendations for Indian and U.S. policymakers.

8.1 Principle One: Allocate Each Risk to the Cheapest Information Holder

The economic analysis of liability has long held that losses should fall on the party best placed to avoid them at least cost. Calabresi and Hirschoff framed the question as identifying the party best placed to weigh the costs of an accident against the costs of avoiding it, and to act on that judgment.[105] In the AI supply chain, the relevant cost is informational: who knows, or can most cheaply find out, where the risk comes from?

Applied to the four interfaces, the principle produces a clear allocation. The model developer holds the best information about training-data provenance (R1 at the foundation level) and should bear that risk through indemnities that are not conditioned on facts the customer cannot verify. The vendor holds the best information about its own fine-tuning corpora, retrieval sources, prompt histories, and tool settings (R1 at the adaptation level and R2), and should bear those risks. The vendor and client together hold the information about human contribution (R3), and the contract should require them to generate and preserve it. The client holds the best information about how and where the deliverable will be exploited (R4), including which markets matter and which rights holders are likely to object, and should bear the residual risk of exploitation choices it makes against advice.

This allocation departs from current practice in one important respect. Today, the client typically demands a broad indemnity from the vendor, and the vendor, unable to obtain an equivalent upstream promise, prices the gap into its fees or quietly accepts it. LPA would instead make the vendor's liability turn on the things the vendor controls, and would require the vendor to preserve its eligibility for upstream protection as a contractual obligation owed to the client.

8.2 Principle Two: Warrant Process, Not Title

If copyright may not subsist in AI-generated components, a warranty that the vendor owns and transfers "all right, title, and interest" is a promise about a legal status that neither party can verify. Breach of such a warranty gives the client a damages claim, not the missing right. LPA replaces title warranties

with process warranties: the vendor warrants that it used only approved tools on approved settings, used only licensed or client-owned data for adaptation, ran agreed similarity checks, and recorded the human contribution to each deliverable.

Process warranties have three advantages. They are verifiable, because compliance can be audited against logs. They are insurable, because underwriters can price adherence to a protocol more readily than an open-ended guarantee of non-infringement. And they align with the direction of regulation, which in both the European Union and India is moving toward documentation and disclosure obligations rather than guarantees of outcome.[106] Content-provenance standards such as those of the Coalition for Content Provenance and Authenticity offer a technical vocabulary for recording the process in a portable form.[107]

8.3 Principle Three: Treat Documented Human Contribution as an Asset

The authorship gap is partly a problem of evidence. Under the Copyright Office's approach, human selection, arrangement, and modification can support protection, but only if the applicant can identify and describe them. Under section 2(d)(vi), whether a particular person "caused" the work may depend on records of who configured, prompted, and curated it. In both systems, the difference between a protected and an unprotected deliverable may lie in whether anyone wrote down what the humans did.

LPA therefore treats human-contribution records as a balance-sheet asset, created at the time of production and delivered with the work. For software, this means version-control histories that distinguish AI-suggested code from human-written or human-modified code. For text and images, it means retaining prompts, drafts, and edit histories, and a short authorship statement describing the creative decisions made by named individuals. These records serve four purposes at once: they support registration in the United States, they support a claim of authorship or causation in India, they evidence independent creation in an infringement defense, and they satisfy diligence requests in financing and acquisition transactions.

A caution is necessary. Documentation requirements can encourage what might be called authorship theater: token human edits made only to create a record. The Copyright Office's approach guards against this by asking whether the human contribution is itself creative, and courts in both countries will likely do the same. The purpose of the records is to capture real contribution, not to manufacture it. Businesses should also recognize the cost. Logging and review consume some of the productivity that motivated the use of AI in the first place, and a rational client will require detailed records only for deliverables whose exclusivity it actually values.

8.4 Principle Four: Hedge Across Jurisdictions

The mismatch between U.S. and Indian law cannot be eliminated by a choice-of-law clause. LPA therefore calls for jurisdictional hedging. Assignments should satisfy both section 204(a) and section 19. Deliverables intended for the U.S. market should meet U.S. authorship standards even if Indian law would protect them more generously. Adaptation datasets should be assembled on the assumption that the more restrictive of the two regimes (currently, for narrow commercial corpora, likely U.S. law after *Ross*, and potentially Indian law if *ANI* is reversed on appeal) will apply. And dispute-resolution clauses should be chosen with enforcement in India in mind.

8.5 Application: Diligence, Financing, and the Startup Balance Sheet

The framework matters most where intellectual property is the principal asset. For venture-backed content and software companies, the authorship gap and the title fracture translate directly into

representation-and-warranty exposure at financing, and into lower valuations where exclusivity cannot be demonstrated. Model financing documents have already begun to respond: the National Venture Capital Association's forms ask a company to confirm that generative AI tools have not been used to develop material intellectual property in a way that could compromise the company's ownership, and the British Private Equity and Venture Capital Association's approach goes further, expecting documented processes for disclosing and auditing data provenance and compliance with the license terms of third-party data.[108] A founder who cannot answer those questions is not merely disclosing a risk; she is disclosing that the company may not own what investors are pricing.

Layered Provenance Allocation supplies the answers in a form diligence can consume. Human-contribution records substantiate the ownership representation for the components that are protectable and identify, honestly, the components that are not. Dual-compliant assignments satisfying both section 204(a) and section 19 close the chain of title through offshore vendors and freelance contributors, which is where diligence most often finds gaps. Process warranties from vendors, backed by audit rights, give the company something to stand on when it makes its own warranties to investors or acquirers. These should become standard closing conditions, and the underlying records should be maintained from the first commit rather than reconstructed under time pressure in a data room.

There is a second-order effect worth naming. Where diligence standards harden around provenance, the vendors that can evidence their process will command a premium, and those that cannot will compete only on price. That is a healthier market outcome than the present equilibrium, in which the cheapest vendor is often the one that documents least.

8.6 Recommendations for Policymakers

8.6.1 India

First, DPIIT's forthcoming paper on outputs should resolve the interpretive questions surrounding section 2(d)(vi). The most coherent solution is to confirm that the provision identifies the author of a computer-generated work only where that work independently satisfies the originality standard in *Modak*, and to define "the person who causes the work to be created" as the person who makes the creative arrangements, not merely the person who presses the button. A narrower alternative would be to adopt a sui generis right of limited duration for AI-generated works lacking human originality, which would protect investment without distorting the concept of authorship.

Second, Parliament should soften the section 19 default rules for commercial assignments between business entities. The one-year lapse rule, the five-year default term, and the India-only default territory protect individual authors against overreaching publishers, a legitimate concern. They serve no evident purpose in a negotiated contract between a multinational client and a large IT company, and they impose avoidable transaction costs on India's most important export sector.

Third, if the *ANI* ruling is affirmed, the government should decide deliberately whether it wishes to replace fair dealing with the hybrid remuneration model DPIIT proposed. The worst outcome for Indian businesses would be a prolonged period in which courts treat training as free fair dealing while policy documents treat it as a compensable use, leaving vendors unable to price either future.

8.6.2 The United States

First, the Copyright Office should supplement its registration guidance with a documentation safe harbor: a standard form of authorship statement that, if completed accurately, creates a presumption that the identified human contributions are sufficient for registration of the claimed elements. This

would reduce the cost and uncertainty of the case-by-case approach without altering the human authorship requirement. The same instrument could be adapted in India. Because the originality threshold in *Modak* and the identity of the person who "causes" a computer-generated work both turn on evidence of human input, a standard authorship statement issued as DPIIT guidance would do double duty, serving registration in the United States and proof of authorship or causation in India.[109]

Second, Congress and the courts should be cautious about reading *Cox* to shelter AI providers entirely from responsibility for outputs. The Supreme Court's standard of tailoring or inducement is appropriate for general communications infrastructure. It fits less comfortably with systems whose developers have detailed knowledge of memorization risks and the technical ability to reduce them. A market in which the parties with the best information bear the least liability is unlikely to invest efficiently in safety.

Third, federal procurement, which already shapes commercial software contracting, could adopt process-warranty and disclosure terms for AI-generated deliverables. Doing so would establish a market standard that private buyers, including Minnesota's large corporate purchasers of outsourced services, could adopt at low cost.

8.7 Objections

Three objections deserve a response. The first is that the framework overstates the risk because litigation has targeted model developers, not vendors or clients. That is true today, but *Cox* changes the incentives of rights holders, and the history of copyright enforcement suggests that plaintiffs go where liability is clearest and defendants are solvent. Large enterprise clients meet both criteria.

The second objection is that process warranties shift too much risk to clients, who lose the comfort of an outcome guarantee. The answer is that the comfort is often illusory. An outcome guarantee from a vendor that cannot satisfy it, backed by an upstream indemnity that the vendor may have voided, protects the client less than a verifiable process combined with a well-structured indemnity and adequate insurance.

The third objection is that the framework is too costly for small businesses. LPA is scalable. A small company commissioning low-value marketing content need not require generation logs; it should, at minimum, prohibit consumer-tier tools, require disclosure of AI use, and obtain a dual-compliant assignment. The full framework is warranted where the deliverable is core intellectual property.

9. Conclusion

Generative AI has changed who creates the work that outsourcing delivers, but the law and the contracts that govern outsourcing still assume a human author employed by the vendor. That assumption now fails in three ways. It fails in the United States because material lacking human authorship cannot be protected, a rule the Supreme Court declined to revisit in 2026. It fails in India because a statute that seems to answer the authorship question in fact raises new ones, and because assignment formalities designed to protect authors can defeat the expectations of commercial parties. And it fails across both systems because the risks of training, retrieval, and output infringement concentrate in the vendor and the client, the layers whose contracts were written for a different technology.

The Article's central claim is that these risks are manageable, but only if the parties stop relying on promises about legal status that neither can verify. Allocating each risk to the party with the cheapest information, warranting process rather than title, treating documented human contribution as an asset, and drafting for both legal systems at once will not remove uncertainty. Pending appeals in *ANI*, *Ross*, and the New York litigation, together with DPIIT's unfinished work, guarantee that the ground will

keep moving. What these measures offer is a structure that remains sound as the doctrine changes. For the businesses that commission, produce, and depend on outsourced creative and technical work, that is the most the law can currently offer, and more than most of their contracts now provide.

Appendix

Table 4: Status of Principal Proceedings Cited (as of September 17, 2026)

Proceeding	Forum	Core issue	Status	Relevance to outsourcing
<i>Thaler v. Perlmutter</i>	D.C. Cir.; U.S. Supreme Court	Autonomous AI authorship	Affirmed 2025; cert. denied Mar. 2, 2026	Confirms human authorship requirement (R3)
<i>Allen v. Perlmutter</i>	D. Colo.	Iterative prompting as authorship	Cross-motions pending	Will define protection for prompt-heavy deliverables (R3)
<i>Cox v. Sony Music</i>	U.S. Supreme Court	Contributory liability standard	Decided Mar. 25, 2026	Shifts exposure downstream (R4)
<i>Thomson Reuters v. Ross</i>	3d Cir.	Fair use of training data	Argued June 11, 2026	First appellate ruling on AI training (R1)
<i>Bartz v. Anthropic</i>	N.D. Cal.	Training fair use; pirated libraries	Settlement approved July 2026	Prices provenance failure (R1)
<i>Kadrey v. Meta</i>	N.D. Cal.	Market dilution	Summary judgment for Meta, June 2025	Dilution theory for narrow corpora (R1)
<i>In re OpenAI Copyright Litigation</i> (incl. <i>NYT v. Microsoft</i>)	S.D.N.Y.	Training and output infringement	Summary judgment argued Sept. 2026	Output and RAG exposure (R1, R2)
<i>ANI Media v. OpenAI</i>	Delhi High Court	Training, fair dealing, RAG	Interim relief refused July 24, 2026; appeal listed Dec. 5, 2026	Indian training and retrieval law (R1, R2)
<i>Getty Images v. Stability AI</i>	High Court of England and Wales	Model as infringing copy	Secondary infringement claim dismissed Nov. 2025; permission to appeal granted	Cross-border fine-tuning risk (R1)
<i>GEMA v. OpenAI</i>	Regional Court Munich I	Memorization as reproduction	Judgment for GEMA, Nov. 2025	Memorization liability in EU markets (R2)

References

- [1] Sai Ishwarbharath B, Abhirami G & Haripriya Suresh, *AI Reshapes India's IT Services Sector Contracts as Clients Demand More for Less*, Reuters (Aug. 20, 2026), <https://finance.yahoo.com/technology/ai/articles/ai-reshapes-indias-services-sector-230207347.html> (reporting outcome-linked pricing at TCS, gain-sharing structures at Cognizant and HCLTech, and a twenty percent decline in the Nifty IT index during 2026).
- [2] Thaler v. Perlmutter, 130 F.4th 1039 (D.C. Cir. 2025), *cert. denied*, No. 25-449 (U.S. Mar. 2, 2026); see *Thaler v. Perlmutter* (25-449), SCOTUSblog, <https://www.scotusblog.com/cases/thaler-v-perlmutter/>; Holland & Knight, *The Final Word? Supreme Court Refuses to Hear Case on AI Authorship and Inventorship* (Mar. 2026), <https://www.hklaw.com/en/insights/publications/2026/03/the-final-word-supreme-court-refuses-to-hear-case-on-ai-authorship>.
- [3] Cox Commc'ns, Inc. v. Sony Music Ent., No. 24-171, 607 U.S. ____ (Mar. 25, 2026), https://www.supremecourt.gov/opinions/25pdf/24-171_bq7d.pdf.
- [4] ANI Media Pvt. Ltd. v. OpenAI OpCo LLC, I.A. 45300/2024 in CS(COMM) 1028/2024, 2026:DHC:5900 (Del. HC July 24, 2026) (Bansal, J.), <https://indiankanoon.org/doc/93327052/>.
- [5] *Delhi HC Seeks OpenAI Response in ANI Case: Here's the Timeline*, MediaNama (Sept. 2026), <https://www.medianama.com/2026/09/223-delhi-hc-openai-ani-case/> (noting that the Division Bench issued notice on ANI's appeal and listed the matter for December 5, 2026).
- [6] For the training question, see Mark A. Lemley & Bryan Casey, *Fair Learning*, 99 Tex. L. Rev. 743 (2021), <https://texaslawreview.org/wp-content/uploads/2021/03/Lemley.Printer.pdf>; Matthew Sag, *Copyright Safety for Generative AI*, 61 Hous. L. Rev. 295 (2023), <https://houstonlawreview.org/article/92126-copyright-safety-for-generative-ai>. For the authorship question, see Jane C. Ginsburg & Luke Ali Budiardjo, *Authors and Machines*, 34 Berkeley Tech. L.J. 343 (2019), https://btlj.org/data/articles2019/34_2/01_Ginsburg_Web.pdf; Daniel J. Gervais, *The Machine as Author*, 105 Iowa L. Rev. 2053 (2020), <https://ilr.law.uiowa.edu/print/volume-105-issue-5/the-machine-as-author>.
- [7] See, e.g., *Target in India*, Target Corp., <https://india.target.com/>; *Optum India*, UnitedHealth Group, <https://www.optum.in/>; *3M India*, 3M Co., https://www.3mindia.in/3M/en_IN/company-in/.
- [8] Katherine Lee, A. Feder Cooper & James Grimmelmman, *Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain*, 72 J. Copyright Soc'y U.S.A. 251 (2025), <https://arxiv.org/abs/2309.08133>.
- [9] See 17 U.S.C. §§ 101, 201(b), 204(a), <https://www.law.cornell.edu/uscode/text/17/201>; The Copyright Act, 1957, §§ 17(c), 18, 19 (India), <https://www.indiacode.nic.in/handle/123456789/1367>.
- [10] Ishwarbharath B et al., *supra* note 1.
- [11] 17 U.S.C. § 101 (defining "computer program"), <https://www.law.cornell.edu/uscode/text/17/101>; The Copyright Act, 1957, § 2(o) (India) (including computer programmes within "literary work"), <https://www.indiacode.nic.in/handle/123456789/1367>.
- [12] U.S. Const. art. I, § 8, cl. 8, <https://constitution.congress.gov/browse/article-1/section-8/clause-8/>.
- [13] Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53, 58 (1884), <https://supreme.justia.com/cases/federal/us/111/53/>.
- [14] Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 345-46 (1991), <https://supreme.justia.com/cases/federal/us/499/340/>.
- [15] *Naruto v. Slater*, 888 F.3d 418, 425-26 (9th Cir. 2018), <https://cdn.ca9.uscourts.gov/datastore/opinions/2018/04/23/16-15469.pdf>.

- [16] Thaler v. Perlmutter, 130 F.4th 1039 (D.C. Cir. 2025), <https://media.cadc.uscourts.gov/opinions/docs/2025/03/23-5233.pdf>, *aff'g* 687 F. Supp. 3d 140 (D.D.C. 2023).
- [17] *Id.*
- [18] *See supra* note 2.
- [19] *See Thaler Is Dead. Now for the AI Copyright Questions That Actually Matter.*, Copyright Lately (Mar. 2026), <https://copyrightlately.com/thaler-is-dead-ai-copyright-questions/>.
- [20] Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence, 88 Fed. Reg. 16,190 (Mar. 16, 2023), <https://www.federalregister.gov/documents/2023/03/16/2023-05321/copyright-registration-guidance-works-containing-material-generated-by-artificial-intelligence>.
- [21] Letter from Robert J. Kasunic, Assoc. Register of Copyrights, U.S. Copyright Office, to Van Lindberg, Re: *Zarya of the Dawn* (Registration # VAu001480196) (Feb. 21, 2023), <https://www.copyright.gov/docs/zarya-of-the-dawn.pdf>.
- [22] U.S. Copyright Office Review Board, Decision Affirming Refusal of Registration of *Théâtre D'opéra Spatial* (Sept. 5, 2023), <https://www.copyright.gov/rulings-filings/review-board/docs/Theatre-Dopera-Spatial.pdf>.
- [23] Complaint, Allen v. Perlmutter, No. 1:24-cv-02665 (D. Colo. Sept. 26, 2024), <https://www.courtlistener.com/docket/69198079/allen-v-perlmutter/>; *see* Copyright Lately, *supra* note 19 (describing the Office's January 2026 cross-motion and Allen's February 2026 reply).
- [24] U.S. Copyright Office, Copyright and Artificial Intelligence, Part 2: Copyrightability (Jan. 2025), <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-2-Copyrightability-Report.pdf>.
- [25] Ginsburg & Budiardjo, *supra* note 6.
- [26] 17 U.S.C. §§ 411(a), 412, <https://www.law.cornell.edu/uscode/text/17/412>.
- [27] The Copyright Act, 1957, § 2(d)(vi) (India), *inserted by* The Copyright (Amendment) Act, 1994, No. 38 of 1994, <https://www.indiacode.nic.in/handle/123456789/1367>.
- [28] Copyright, Designs and Patents Act 1988, c. 48, §§ 9(3), 178 (U.K.), <https://www.legislation.gov.uk/ukpga/1988/48/section/9>.
- [29] Pamela Samuelson, *Allocating Ownership Rights in Computer-Generated Works*, 47 U. Pitt. L. Rev. 1185 (1986), <https://www.law.berkeley.edu/wp-content/uploads/2024/01/Pam-Samuelson-Allocating-Ownership-Rights-in-Computer-Generated-Works.pdf>.
- [30] The Copyright Act, 1957, § 13(1)(a) (India), <https://www.indiacode.nic.in/handle/123456789/1367>.
- [31] Eastern Book Co. v. D.B. Modak, (2008) 1 SCC 1 (India), <https://indiankanoon.org/doc/1062099/>.
- [32] *ANI Media*, *supra* note 4, ¶¶ 86-90.
- [33] The Copyright Act, 1957, § 17 & proviso (c) (India), <https://www.indiacode.nic.in/handle/123456789/1367>.
- [34] *Exclusive: Indian Copyright Office Issues Withdrawal Notice to AI Co-Author*, Managing IP (Dec. 13, 2021), <https://www.managingip.com/article/2a5bqtj8ume32iwlao5y/exclusive-indian-copyright-office-issues-withdrawal-notice-to-ai-co-author>; *The Divergence in Copyright Recognition for AI-Generated Works: An In-Depth Analysis of Ankit Sahni's Case in the US and India*, K&S Partners, <https://ksandk.com/intellectual-property/divergent-copyright-recognition-ai-generated-works-sahnis-case-us-vs-india/>.

- [35] Dep't-Related Parliamentary Standing Comm. on Commerce, 161st Report: Review of the Intellectual Property Rights Regime in India (July 2021), <https://prsindia.org/policy/report-summaries/review-of-the-intellectual-property-rights-regime>.
- [36] Dep't for Promotion of Indus. & Internal Trade, Working Paper on Generative AI and Copyright (Part 1) (Dec. 2025), <https://www.dpiit.gov.in/static/uploads/2025/12/ff266bbeed10c48e3479c941484f3525.pdf>.
- [37] *DPIIT's Second Paper on Copyrighted Content in AI Likely in Two Months*, Bus. Standard (Dec. 11, 2025), https://www.business-standard.com/technology/tech-news/dpiit-s-second-paper-on-copyrighted-content-in-ai-likely-in-two-months-125121100700_1.html; *AI Training & Copyright in India: DPIIT Hybrid Model Explained*, Intepat, <https://www.intepat.com/blog/ai-training-copyright-india-dpiit-hybrid-model> (reporting that Part 2 remains pending).
- [38] Gervais, *supra* note 6.
- [39] Mark A. Lemley, *How Generative AI Turns Copyright Upside Down*, 25 Colum. Sci. & Tech. L. Rev. 190 (2024), https://law.stanford.edu/wp-content/uploads/2024/09/2024-09-30_How-Generative-AI-Turns-Copyright-Upside-Down.pdf.
- [40] 17 U.S.C. § 101 (definition of "work made for hire"), <https://www.law.cornell.edu/uscode/text/17/101>.
- [41] Cmty. for Creative Non-Violence v. Reid, 490 U.S. 730, 751-52 (1989), <https://supreme.justia.com/cases/federal/us/490/730/>.
- [42] *Thaler*, 130 F.4th 1039; *see* discussion *supra* Part II.A.2.
- [43] 17 U.S.C. § 204(a), <https://www.law.cornell.edu/uscode/text/17/204>.
- [44] *Effects Assocs., Inc. v. Cohen*, 908 F.2d 555, 558-59 (9th Cir. 1990), <https://law.justia.com/cases/federal/appellate-courts/F2/908/555/168957/>.
- [45] The Copyright Act, 1957, § 17 proviso (c) (India), <https://www.indiacode.nic.in/handle/123456789/1367>.
- [46] *Id.* § 17 proviso (b).
- [47] *Indian Performing Right Soc'y Ltd. v. E. India Motion Pictures Ass'n*, (1977) 2 SCC 820 (India), <https://indiankanoon.org/doc/331232/>.
- [48] The Copyright Act, 1957, § 19(1)-(2) (India), <https://www.indiacode.nic.in/handle/123456789/1367>.
- [49] *Id.* § 19(4).
- [50] *Id.* § 19(5).
- [51] *Id.* § 19(6).
- [52] *Id.* § 18(1) third proviso.
- [53] *Id.* § 57.
- [54] *Amar Nath Sehgal v. Union of India*, 2005 (30) PTC 253 (Del.) (Del. HC Feb. 21, 2005) (India), <https://indiankanoon.org/doc/1990275/>.
- [55] *See supra* note 37.
- [56] *Itar-Tass Russian News Agency v. Russian Kurier, Inc.*, 153 F.3d 82, 90-91 (2d Cir. 1998), <https://www.courtlistener.com/opinion/757300/itar-tass-russian-news-agency-v-russian-kurier-inc/>.
- [57] Berne Convention for the Protection of Literary and Artistic Works art. 5(2), Sept. 9, 1886, as revised at Paris July 24, 1971, <https://www.wipo.int/wipolex/en/text/283698>.

- [58] Patrick Lewis et al., *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*, 33 *Advances in Neural Info. Processing Sys.* 9459 (2020), <https://arxiv.org/abs/2005.11401>.
- [59] Thomson Reuters Enter. Ctr. GmbH v. Ross Intel. Inc., No. 1:20-cv-613-SB, ECF No. 770 (D. Del. Feb. 11, 2025), https://www.ded.uscourts.gov/sites/ded/files/opinions/20-613_5.pdf.
- [60] Thomson Reuters Enter. Ctr. GmbH v. Ross Intel. Inc., No. 25-2153 (3d Cir. argued June 11, 2026); *see Baker Botts, Third Circuit Hears Oral Argument in Ross v. Reuters AI Training Copyright Case* (July 2026), <https://www.bakerbotts.com/thought-leadership/publications/2026/july/third-circuit-hears-oral-argument>.
- [61] Order on Fair Use, Bartz v. Anthropic PBC, No. 3:24-cv-05417-WHA, ECF No. 231 (N.D. Cal. June 23, 2025), <https://copyrightalliance.org/wp-content/uploads/2025/06/Bartz-v.-Anthropic-Order.pdf>.
- [62] *Judge Approves Record \$1.5 Billion AI Copyright Settlement Involving Anthropic*, *Jurist* (July 2026), <https://www.jurist.org/news/2026/07/judge-approves-record-1-5-billion-settlement-involving-anthropic/>; *see also Bartz v. Anthropic Settlement Receives Final Approval*, *Authors Alliance* (July 21, 2026), <https://www.authorsalliance.org/2026/07/21/bartz-v-anthropic-settlement-receives-final-approval/>.
- [63] Kadrey v. Meta Platforms, Inc., No. 3:23-cv-03417-VC, ECF No. 598 (N.D. Cal. June 25, 2025), <https://law.justia.com/cases/federal/district-courts/california/candce/3:2023cv03417/415175/598/>.
- [64] *Historic NYT v. OpenAI Copyright Battle Heats Up*, *Axios* (Sept. 8, 2026), <https://www.axios.com/2026/09/08/nyt-openai-microsoft-copyright-lawsuit>.
- [65] U.S. Copyright Office, Copyright and Artificial Intelligence, Part 3: Generative AI Training (pre-publication version May 2025), <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf>.
- [66] Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith, 598 U.S. 508 (2023), <https://supreme.justia.com/cases/federal/us/598/21-869/>.
- [67] The Copyright Act, 1957, § 14(a)(i) (India), <https://www.indiacode.nic.in/handle/123456789/1367>.
- [68] *Id.* § 52(1)(a).
- [69] *Id.* § 52(1)(b)-(c).
- [70] The Chancellor, Masters & Scholars of the Univ. of Oxford v. Rameshwari Photocopy Servs., RFA(OS) 81/2016 (Del. HC Sept. 16, 2016) (India) (Division Bench), <https://indiankanoon.org/doc/135895592/>.
- [71] *ANI Media, supra* note 4.
- [72] *Id.* ¶¶ 51-55 (citing Neetu Singh v. Telegram FZ LLC (Del. HC Aug. 30, 2022) (India), <https://indiankanoon.org/doc/164684887/>).
- [73] *See Vishno Sudheendra, ANI v. OpenAI: User Rights, Fair Dealing, and the Future of AI in Indian Copyright Law (Part I)*, *SpicyIP* (July 2026), <https://spicyip.com/2026/07/ani-v-openai-user-rights-fair-dealing-and-the-future-of-ai-in-indian-copyright-law-part-i.html> (discussing paragraphs 176-219 of the judgment).
- [74] *ANI Media, supra* note 4, ¶¶ 83-84.
- [75] *See MediaNama, supra* note 5 (reporting the September 14, 2026 hearing before Justices Avneesh Jhingan and Manmeet Pritam Singh Arora, the direction to OpenAI to reply, and the December 5, 2026 listing).
- [76] *See Bus. Standard, supra* note 37; *supra* note 36. Parties negotiating Indian AI clauses in the interim should assume that both the judicial and the legislative position may change within the life of a typical three-year outsourcing contract.

- [77] Aparajita Lath, *Public Interest, Copyright and Fair Dealing: ANI v OpenAI*, Sup. Ct. Observer (Aug. 19, 2026), <https://www.scobserver.in/journal/public-interest-copyright-and-fair-dealing-ani-v-openai/>.
- [78] DPIIT Working Paper, *supra* note 36; *see also Govt Panel Proposes Blanket Licence for AI Firms to Use Copyrighted Works, Seeks Public Views*, Deccan Herald (Dec. 2025), <https://www.deccanherald.com/business/govt-panel-proposes-blanket-licence-for-ai-firms-to-use-copyrighted-works-seeks-public-views-3824802>.
- [79] Takshashila Institution, *Policy Advisory: Working Paper on Generative AI and Copyright (Part 1)* (Feb. 6, 2026), <https://takshashila.org.in/content/publications/20260206-generative-ai-and-copyright-consultation.html>.
- [80] Nicholas Carlini et al., *Extracting Training Data from Diffusion Models*, Proc. 32d USENIX Sec. Symp. (2023), <https://arxiv.org/abs/2301.13188>; Milad Nasr et al., *Scalable Extraction of Training Data from (Production) Language Models* (2023), <https://arxiv.org/abs/2311.17035>.
- [81] GEMA v. OpenAI, Landgericht München I [Regional Court Munich I], 42 O 14139/24 (Nov. 11, 2025) (Ger.); *see Germany: Court Prohibits Memorization and Reproduction of Copyrighted Song Lyrics in AI Models*, Libr. of Cong. Glob. Legal Monitor (Jan. 13, 2026), <https://www.loc.gov/item/global-legal-monitor/2026-01-13/germany-court-prohibits-memorization-and-reproduction-of-copyrighted-song-lyrics-in-ai-models/>.
- [82] Getty Images (US) Inc. v. Stability AI Ltd. [2025] EWHC 2863 (Ch) (Eng.), <https://caselaw.nationalarchives.gov.uk/ewhc/ch/2025/2863>; *see Permission to Appeal Granted in Getty Images v Stability Copyright Claim*, IPKat (Jan. 2026), <https://ipkitten.blogspot.com/2026/01/permission-to-appeal-granted-in-getty.html>.
- [83] *See supra* note 74.
- [84] *See supra* note 69.
- [85] Authors Guild v. Google, Inc., 804 F.3d 202 (2d Cir. 2015), <https://law.justia.com/cases/federal/appellate-courts/ca2/13-4829/13-4829-2015-10-16.html>.
- [86] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence, art. 53(1)(c)-(d), 2024 O.J. (L 1689), <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>.
- [87] *See* Carlini et al., *supra* note 80.
- [88] Andersen v. Stability AI Ltd., No. 3:23-cv-00201-WHO (N.D. Cal. Aug. 12, 2024), https://bpb-us-e2.wpmucdn.com/sites.uci.edu/dist/d/2220/files/2024/09/Andersen-v-Stability-AI-Ltd_Redacted.pdf.
- [89] Cartoon Network LP v. CSC Holdings, Inc., 536 F.3d 121, 131 (2d Cir. 2008), https://law.justia.com/cases/federal/appellate-courts/ca2/07-1480/07-1480-cv_opn-2011-03-27.html.
- [90] MySpace Inc. v. Super Cassettes Indus. Ltd., FAO(OS) 540/2011 (Del. HC Dec. 23, 2016) (India) (Division Bench), <https://indiankanoon.org/doc/12972852/>.
- [91] The Copyright Act, 1957, §§ 63, 69 (India), <https://www.indiacode.nic.in/handle/123456789/1367>.
- [92] Cox, *supra* note 3; *see* Mayer Brown, *Supreme Court Significantly Narrows Liability for Contributory Copyright Infringement* (Mar. 2026), <https://www.mayerbrown.com/en/insights/publications/2026/03/supreme-court-significantly-narrows-liability-for-contributory-copyright-infringement>.

- [93] Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417 (1984), <https://supreme.justia.com/cases/federal/us/464/417/>; Metro-Goldwyn-Mayer Studios Inc. v. Gorkster, Ltd., 545 U.S. 913 (2005), <https://supreme.justia.com/cases/federal/us/545/913/>.
- [94] Anil Kapoor v. Simply Life India, 2023 SCC OnLine Del 6914, CS(COMM) 652/2023 (Del. HC Sept. 20, 2023) (India), <https://indiankanoon.org/doc/113724486/>.
- [95] Arijit Singh v. Codible Ventures LLP, Interim Application (L) No. 23560 of 2024 (Bom. HC July 26, 2024) (India), <https://indiankanoon.org/doc/103091928/>.
- [96] On the emerging Indian framework for synthetic performances and the case for a purpose-limited licensing right in performer identity, see the companion treatment in Ayan Roy Chowdhury, *The Synthetic Star: Reconstructing Performer Rights for Generative AI in India's Entertainment Economy* (forthcoming). For the statutory background, see The Copyright Act, 1957, §§ 38, 38A, 38B (India) (performers' rights and performers' moral rights), <https://www.indiacode.nic.in/handle/123456789/1367>.
- [97] Khaitan & Co., *MeitY Notifies the IT Amendment Rules 2026* (Feb. 2026), <https://www.khaitanco.com/thought-leadership/MeitY-notifies-the-IT-Amendment-Rules-2026> (noting notification on February 16, 2026 and effect from February 20, 2026).
- [98] Regulation (EU) 2024/1689, art. 50, <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>.
- [99] Brad Smith & Hossein Nowbar, *Microsoft Announces New Copilot Copyright Commitment for Customers*, Microsoft On the Issues (Sept. 7, 2023), <https://blogs.microsoft.com/on-the-issues/2023/09/07/copilot-copyright-commitment-ai-legal-concerns/>.
- [100] *OpenAI's "Copyright Shield" Broadens User IP Indemnities for AI-Created Content*, Proskauer (Nov. 2023), <https://www.proskauer.com/blog/openais-copyright-shield-broadens-user-ip-indemnities-for-ai-created-content>.
- [101] *Customer Copyright Commitment Required Mitigations*, Microsoft Learn, <https://learn.microsoft.com/en-us/azure/foundry/responsible-ai/openai/customer-copyright-commitment>.
- [102] See, e.g., *Terms of Use*, OpenAI, <https://openai.com/policies/row-terms-of-use/> (providing that, as between the parties and to the extent permitted by law, the user owns the output and that OpenAI assigns its right, title, and interest, if any, in the output).
- [103] The Indian Contract Act, 1872, §§ 73, 124-125 (India), <https://www.indiacode.nic.in/handle/123456789/2187>.
- [104] Nishith Desai Assocs., *Enforcing U.S. Arbitral Awards and Judgments in India* (2020), <https://www.nishithdesai.com/Content/document/pdf/ResearchPapers/Enforcing-U.S.-Arbitral-Awards-and-Judgments-in-India-2020.pdf>.
- [105] Guido Calabresi & Jon T. Hirschoff, *Toward a Test for Strict Liability in Torts*, 81 Yale L.J. 1055 (1972), https://openyls.law.yale.edu/bitstream/20.500.13051/1261/2/Toward_a_Test_for_Strict_Liability_in_Torts.pdf.
- [106] See Regulation (EU) 2024/1689, *supra* note 86, art. 53; Khaitan & Co., *supra* note 97.
- [107] Coal. for Content Provenance & Authenticity, *C2PA Specifications*, <https://c2pa.org/specifications/>.
- [108] See Orrick, *AI, Risk, and Reps: A Practical Guide for Start-ups Raising International Venture Capital* (Apr. 2026), <https://www.orrick.com/en/Insights/2026/04/A-Practical-Guide-for-German-Startups-Raising-International-Venture-Capital-Part-1> (describing NVCA and BVCA model provisions on generative AI, training-data provenance, and output ownership).

- [109] *See* U.S. Copyright Office, *supra* note 24; Eastern Book Co., *supra* note 31. Interoperability matters here: a single form that satisfies both systems reduces compliance cost for exactly the cross-border deliverables this Article examines.