
THE DISINTEGRATION OF THE IDEA-EXPRESSION DICHOTOMY IN THE AGE OF GENERATIVE AI: AN INDIAN PERSPECTIVE

Manvi Khanduri, Delhi Metropolitan Education, affiliated to GGSIPU

Kerryn Pahuja, Delhi Metropolitan Education, affiliated to GGSIPU

ABSTRACT

Traditional copyright regimes are facing significant problems as creativity is being transformed by the emergence of generative artificial intelligence (AI). This disruption is centred on the idea-expression dichotomy, which separates abstract concepts that cannot be protected from expressions that can. AI's capacity to identify patterns in existing works and produce new outputs is putting growing pressure on this principle. It raises a critical question: whether the works generated through algorithmic processes can be sufficiently regulated by copyright law, which is historically designed to reward human creativity. This study looks at the Indian legal system, examining how authorship, ownership, and infringement are affected by AI-generated material, which also raises concerns about fairness, consent, and the economic rights of original creators. It focuses on important cases like *ANI v. OpenAI*.

The study, by contrasting the EU's legal framework, which includes transparency and opt-out methods for AI training, with the U.S. emphasis on human authorship and transformative application, situates India within the global debate. The report proposes specific changes to the Copyright Act of 1957, such as flexible fair use clauses, text and data mining exclusions, clarification of AI-assisted authorship, and sui generis protection for works created by AI. To address the breakdown of the idea-expression dichotomy and promote responsible AI-driven innovation, the study suggests that India has to implement a nuanced and adaptive legal framework.

Keywords: Copyright, Idea-Expression dichotomy, AI, framework, innovation, AI-driven innovation.

I. INTRODUCTION:

The essential premise that copyright protection only covers the particular manifestation of an idea, not the underlying idea itself, is established by the Idea Expression Dichotomy, which is one of the cornerstones of copyright law across the globe. Article 9.2 of the Agreement on *Trade Related Aspects of Intellectual Property Rights (TRIPS)* and other international instruments have codified this doctrine, which states that “copyright protection shall extend to expressions and not to ideas, procedures, methods of operation or mathematical concepts as such.”¹ This idea has been continuously maintained in India by judicial interpretation under the Copyright Act, 1957, most notably in the seminal ruling in *R.G. Anand v. M/s Deluxe Films*².

However, this well-established dichotomy is now facing formerly unheard of difficulties due to the development of generative artificial intelligence (AI). By training on enormous datasets of copyrighted content that has been scraped from the internet, generative AI systems, like *OpenAI's ChatGPT*, *Google's Gemini*, and *Anthropic's Claude*, are able to generate writing, graphics, code, and other outputs that mimic human ingenuity. These technological advancements raise serious concerns regarding the idea-expression distinction's ongoing relevance at a time when computers are capable of producing content on their own that can violate copyrights or produce works that are unclearly human authored.

These topics have progressed from theoretical discussion to actual litigation in India. India's first significant test case on the unlawful use of copyrighted information to train AI models was *Asian News International's (ANI)*³ November 2024 lawsuit against OpenAI in the Delhi High Court. In response to the lawsuit, the Indian government established an eight-member expert team in May 2025 to assess whether the Copyright Act of 1957 sufficiently handles the particular difficulties posed by generative artificial intelligence. This change indicates a growing understanding that the laws from the pre-digital period would need to be significantly modified in order to regulate a creative environment driven by artificial intelligence. This study uses doctrinal legal research based on analysis of statutes, case law, and international frameworks. It also uses comparative analysis between India, the United States, and the European Union.

¹ World Trade Organization, *Trade-Related Aspects of Intellectual Property Rights (TRIPS)*, Art. 9.2, available at https://www.wto.org/english/docs_e/legal_e/27-trips_04_e.htm

² *R.G. Anand v. M/s Deluxe Films & Ors.*, AIR 1978 SC 1613, 1978 (4) SCC 118

³ *ANI Media Pvt. Ltd. v. OpenAI Inc. & Anr.*, CS(COMM) 1028/2024 (Delhi H.C. 2024)

➤ LITERATURE REVIEW

Many academics believe that as AI challenges established copyright laws, particularly the idea-expression dichotomy, it is redefining creativity. Finding regulations that safeguard human creativity while promoting technological advancement is a difficult task. Achieving this equilibrium will determine whether AI strengthens or weakens the creative ecosystem. AI has caused major problems and presented difficulties in the field of copyright law. This essay explores the significance of AI in the creation of artistic, musical, and poetic works. The problems of authorship and deepfakes in AI-generated content will also be covered in the article.

This study explores the nuances and complexities that arise when copyright law comes into contact with AI-generated content, such as text, images, music, and code created by AI algorithms. As technology advances, conventional notions of authorship and ownership are in jeopardy because AI's role in creation makes these concepts more difficult to determine. It is unclear whether AI-generated works should be protected by copyright since conventional copyright law is predicated on human creativity. Even so, the copyright office refused to grant artificial intelligence authors copyright protection for their works.

Finding the owner of the intellectual property in these works is just as important as defining the "author" or figuring out who qualifies. The creativity and uniqueness of content generated by artificial intelligence are also being thoroughly examined. In order to produce results that mimic or replicate particular features of these datasets, AI systems frequently use large collections of previous publications. This calls into question the core concepts of originality and creativity that copyright laws aim to safeguard. The blurring of the lines between originality and plagiarism creates legal ambiguity and sparks discussions about the originality and calibre of content produced by AI.

II. THE IDEA-EXPRESSION DICHOTOMY: CONCEPTUAL FRAMEWORK AND INDIAN JURISPRUDENCE:

➤ Historical Origins and International Foundation:

The idea-expression distinction was first articulated by the U.S. Supreme Court in *Baker v.*

*Selden (1879)*⁴, where the Court determined that while Selden's copyright protected the particular depiction of his bookkeeping method, it did not protect the underlying system because doing so would go beyond the constitutional bounds of copyright. Later, this idea was incorporated into international law; TRIPS Article 9.2, which binds all WTO members, including India, explicitly states that copyright protects expressions, not ideas, procedures, methods, or mathematical concepts, and the Berne Convention implicitly acknowledges this distinction. The distinction serves important policy goals, including preventing monopolisation of ideas, guaranteeing public access to fundamental concepts for future creativity, striking a balance between exclusive rights and societal interest in knowledge sharing, and upholding copyright's constitutional goal of fostering innovation without restricting fundamental ideas.

➤ **Indian Legal Framework: Statutory Provisions and Judicial Interpretation:**

The Copyright Act of 1957 incorporates the idea-expression dichotomy through its provisions, despite not defining “idea” or offering a statutory test. Only unique expressions of ideas, not the ideas themselves, are protected by Section 13(1), which restricts copyright to original literary, dramatic, musical, and artistic works. In *R.G. Anand v. Deluxe Films (1978)*⁵, which concerned the play *Hum Hindustani* and the movie *New Delhi*, the Supreme Court explained this. Ideas, themes, stories, and historical details are not protected by copyright, according to Justice Fazal Ali; only the form, arrangement, and expression are. For there to be infringement, there must be more than just a thematic resemblance in terms of expression, treatment, circumstances, dialogue, and event sequence. Using this, the Court determined that there was no infringement because the movie had extra themes that weren't present in the play. According to *R.G. Anand*, the ultimate test is the overall impression of the works, and copyright only protects distinctive expression. Thematic similarity is insufficient.

Originality and the Modicum of Creativity Standard:

The *Eastern Book Company v. D.B. Modak (2008)*⁶ case developed the idea of originality, which forms the basis of the idea-expression dichotomy. The Supreme Court Cases volumes published by EBC contained edited judgements that were well-formatted, organised, and included headnotes. The respondents claimed the rulings were public domain and lacked

⁴ *Baker v. Selden*, 101 U.S. 99 (1879)

⁵ *R.G. Anand v. M/S. Delux Films & Ors.* (AIR 1978 SC 1613)

⁶ *Eastern Book Co. v. D.B. Modak*, 2008 (36) PTC 1 (SC)

enough originality, therefore they copied these versions without authorisation.

The Supreme Court ruled that although the rulings were in the public domain, EBC's versions were unique and protected by copyright because of the creative work that went into organising and editing them. The “modicum of creativity” test from *Feist Publications v. Rural Telephone Service Co. (1991)*⁷ was accepted by the Court. The “sweat of the brow” theory was rejected by the U.S. Supreme Court in the Feist case, which held that originality necessitates independent production and a minimum level of inventiveness; effort or a collection of information alone is inadequate.

III. GENERATIVE AI AND COPYRIGHT CHALLENGES: THE INDIAN CONTEXT

Key ways generative AI is dissolving this dichotomy include:

1. Inversion of Creativity: From Execution to Prompting

Conventional copyright rewards the effort put into carrying out a concept. This is changed by GenAI, which emphasises “asking the right questions” (prompting) over producing the finished product.

- The issue is that the AI performs the “expression” (execution) while the user supplies the “idea” (prompt). The output may not have enough human authorship to be protected if the prompt is judged to be too abstract, which would reverse the current copyright incentives and leave the final expression unprotected.
- **Merger Doctrine Strain:** The “merger doctrine”, which holds that an idea cannot be copyrighted if it can only be expressed in one way, is often activated. The idea and expression effectively merge because AI frequently finds the most statistically likely single way to express an idea.

2. The Blurring of “Substantial Similarity”

GenAI makes it challenging to apply the “substantial similarity” test, which is used to assess

⁷ *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340 (1991)

copyright infringement.

- **Input vs. Output:** AI models train on vast amounts of copyrighted expression to learn patterns, but not to copy them directly.
- **Output Similarity:** When AI produces an output that resembles a protected work, it is often hard to prove whether it copied the specific expression or just the underlying “idea” or “style” of the author, which is not protected.
- **Proposed Solutions:** Legal scholars suggest shifting the focus from the similarity of the output to whether the AI was trained on the specific, protected expression, or whether the user prompt specifically intended to reproduce that expression.

3. The Posthumanist View of Creation

Research suggests AI challenges the idea of authorship as a purely human endeavour.

- **Shared Process:** Creativity is increasingly seen as a collaborative act between human prompts and machine execution.
- **Uncertainty of Protection:** The European Parliament has highlighted the legal uncertainty regarding AI-generated content, noting that it creates a structural risk for the future of creativity because existing laws struggle to distinguish between AI-assisted and AI-generated works.

4. Input-Output Dichotomy Challenges

While input (training data) is argued by AI companies to be the use of facts/ideas, the output is undeniably a form of expression.

- **Mass Production:** GenAI enables the mass production of derivative works that can closely resemble copyrighted works, making it harder to determine where inspiration (unprotected idea) ends and infringement (protected expression) begins.

➤ The Nature and Operation of Generative AI Systems:

Programs that can produce writing, images, music, and other creative outputs when given

instructions are referred to as “generative artificial intelligence”. Models like OpenAI’s GPT-4, Google’s Gemini, and Anthropic’s Claude are trained on massive datasets that often contain copyrighted content that has been unlawfully taken from the internet.

The three primary steps in the training process are obtaining raw data from public sources, tokenising and vectorizing it into numerical representations, and then teaching the model to recognise statistical patterns that enable it to generate new outputs. These algorithms sometimes produce results that resemble protected works, which raises concerns about infringement even though they don’t save or replicate information exactly.

Unlike human writers who understand and reinterpret concepts, artificial intelligence (AI) analyses data mechanically and lacks creativity. This discrepancy challenges accepted copyright ideas and begs the question of whether AI’s data use should be treated differently from human learning. The lack of specific provisions for AI-generated works in the Copyright Act, 1957, creates uncertainty regarding authorship and ownership, which poses serious challenges for AI and copyright in India. “Deep fakes”, AI training on copyrighted data without authorisation, and whether human intervention is adequate to assert copyright over AI-generated outputs are important concerns.

➤ **Authorship and Ownership: The Human Creativity Requirement:**

Indian law restricts authorship to human or legal entities. This begs the question of who should be given credit for autonomous AI-generated works: the user, the AI, the author, or the creator? Since Section 2(d) restricts authorship to humans, AI is not permitted to be an author. Generative AI (GenAI) is quickly upending the traditional idea-expression dichotomy, a cornerstone of copyright law that protects an idea’s particular expression while leaving the underlying idea unrestricted. According to research papers and legal analyses, AI generates expressive outputs from conceptual prompts, turning ideas directly into finished works, making it harder to distinguish between unprotected ideas and protected expression.

In *Thaler v. Perlmutter* (2025)⁸, the U.S. Court of Appeals (D.C. Circuit) maintained the requirement for human authorship. Dr. Stephen Thaler asked for copyright for artwork created by his “Creativity Machine”, but the court decided that since copyright is meant to promote

⁸ *Thaler v. Perlmutter* (2025), No. 22-CV-384-1564-BAH

human creativity, it cannot apply to works created solely by machines.

There is no legally binding precedent in India, despite the Copyright Office's brief acknowledgement and subsequent withdrawal of AI co-authorship. The government created an expert panel to determine whether the Copyright Act adequately addresses AI authorship and ownership in response to this ambiguity.

➤ **Training Data and Infringement: The ANI v. OpenAI Litigation:**

In *ANI Media Pvt. Ltd. v. OpenAI Inc.* (Delhi High Court, 2024)⁹, ANI alleged that OpenAI trained ChatGPT using its news content without permission, producing outputs closely mirroring its reports and causing reputational harm, claiming infringement under Sections 14 and 51 of the Copyright Act, 1957. OpenAI challenged jurisdiction, argued that public availability did not require consent, and relied on Section 52's fair dealing exemption, asserting that AI training is a transformative, non-expressive use that processes patterns rather than replicating expressions. The court appointed amici curiae, with Professor Arul Scaria supporting non-expressive classification akin to human learning, while Advocate Adarsh Ramanujan distinguished data collection, tokenisation, and model training, contending that only tokenisation is non-expressive and the other stages constitute infringement, including vectorisation as adaptation under Section 14(a)(vi). As of October 2025, the case remains pending, prompting the Ministry of Commerce & Industry to establish an expert panel to examine AI-related gaps in the Copyright Act and propose reforms.

IV. THE DISINTEGRATION OF THE IDEA-EXPRESSION DICHOTOMY

The conventional idea-expression divide that forms the basis of copyright law is fundamentally upended by generative AI. The dichotomy presumes a distinct boundary between concepts that are not protected and expressions that are, yet artificial intelligence functions in a middle ground, drawing linguistic and structural patterns from pre-existing works to produce new outputs. AI that has been educated on copyrighted works not only replicates them but also recognises statistical correlations, which begs the question of whether this type of training amounts to copying expression or is just learning from concepts. Although temporary replication or storage of copyrighted works during training may still violate Section 14(a)(i), the procedure should not be considered an infringement if it merely includes the abstraction of

⁹ *ANI Media Pvt. Ltd. v. OpenAI Inc.*, CS(COMM) 1028/2024

patterns.

This conflict was at the heart of *ANI v. OpenAI*, where Advocate Ramanujan stressed that reproduction is inevitable at some points while Professor Scaria contended that AI training amounts to “non-expressive use”. When evaluating AI-generated outputs using the R.G. Anand “substantial similarity” test, the intricacy increases. Similarity may result from the AI's absorption of commonly accepted creative patterns rather than outright copying, making it more difficult to distinguish between concept and representation. One of the most significant challenges posed by generative AI is the ability of such systems to imitate the distinctive artistic styles of particular creators without directly copying any specific copyrighted work. Contemporary AI tools are capable of generating images and artworks that closely resemble the visual style associated with well-known studios such as Studio Ghibli, characterised by particular aesthetic elements, colour palettes, and narrative settings. While individual artworks produced by such studios are clearly protected under copyright law, the broader stylistic features that make these works recognisable are generally considered ideas or methods of artistic expression rather than protectable expressions themselves. As a result, AI-generated works that emulate stylistic elements may fall outside traditional copyright protection even though they evoke the creative identity of the original artist. This phenomenon illustrates how generative AI further destabilises the idea-expression dichotomy by producing outputs that replicate stylistic patterns without necessarily reproducing the protected expression of any single work. Furthermore, the merger doctrine which holds that concept and expression fuse due to a lack of expressive options may be generally applicable to AI-generated works, particularly in cases where certain structures are necessary for coherence.

This conundrum is further clarified by the idea of transformative usage, which is essential to American fair use law as demonstrated in *Authors Guild v. Google (2015)*¹⁰: Although AI training produces something novel and functionally unique, its results might rival those of original creations. However, India's Copyright Act lacks the adaptable transformative-use structure that may support AI innovation, restricting exceptions to certain “fair dealing” categories under Section 52. Although the 2011 *Super Cassettes v. MySpace*¹¹ decision acknowledged transformative applications in theory, India's existing legal system is still unprepared to handle AI's reconfiguration of authorship, creativity, and originality, which calls

¹⁰ *Authors Guild v. Google, Inc.*, 804 F.3d 202 (2d Cir. 2015)

¹¹ *Super Cassettes Indus. Ltd. v. MySpace Inc.*, FAO(OS) 540/2011, (2011) 47 PTC 1 (Del)

for a sophisticated, flexible revision of copyright doctrine.

V. INTERNATIONAL COMPARATIVE PERSPECTIVES

USA- In the United States, human authorship is a prerequisite for copyright protection, as affirmed by the Copyright Office and courts. In *Thaler v. Perlmutter*, the court held that works created entirely by AI without human input are not copyrightable, while AI-assisted works may qualify if they involve significant human creative contribution, such as arrangement or modification, beyond mere prompting, according to the Copyright Office's 2023 and 2025 guidance. Ongoing cases like *Andersen v. Stability AI*¹² and *The New York Times v. OpenAI & Microsoft*¹³ centred on whether AI training on copyrighted content constitutes fair use. Plaintiffs argue that AI outputs replicate original works and harm markets, whereas defendants contend that training is transformative, extracting patterns rather than copying expression. U.S. courts are developing a nuanced balance, supporting innovation while requiring demonstrable similarity to establish infringement.

European Union- In contrast, the European Union has taken a legislative stance via the AI Act (Regulation (EU) 2024/1689), which requires general-purpose AI models to adhere to copyright laws. To ensure accountability and transparency, Article 53 mandates that developers create copyright policies and make training data summaries publicly available. In addition, unless rights holders opt out through machine-readable notifications, the Digital Single Market Directive's (2019/790) copyright permits AI training on legally accessed content through an exemption for text and data mining. Although this framework respects artists' rights and provides legal clarity, it may also make compliance more difficult and limit the datasets that are available, which might hurt the EU's ability to compete.

India- These global advances have important implications for India, which is developing into a significant AI market and creative centre. The *ANI v. OpenAI* case from India emphasises the necessity of striking a balance between innovation and security. Transformative technologies could be better accommodated by implementing a more flexible fair use approach, similar to that in the United States. In order to protect artists, India may simultaneously adopt the EU's opt-out and transparency policies. Sui generis protection for AI-generated works and

¹² *Andersen et al. v. Stability AI Ltd.*, No. 3:23-cv-00201 (N.D. Cal. filed Jan. 13, 2023)

¹³ *The New York Times Co. v. Microsoft Corp.*, No. 1:23-cv-11195 (S.D.N.Y. filed Dec. 27, 2023)

clarification of the human authorship criteria might further guarantee that India promotes AI innovation while upholding fair copyright standards.

VI. TOWARD A REFORMED LEGAL FRAMEWORK: AUTHOR'S ANALYSIS

To balance innovation with creator protection, India's copyright laws must adapt to generative AI by clearly defining authorship, ownership, and permissible uses. The Copyright Act, 1957, should specify that works created solely by AI are not copyrightable, while AI-assisted works qualify if substantial human creativity is involved, with the claimant bearing the burden of proof. Ownership provisions should recognize contributions where AI serves as a tool, following the UK model for computer-generated works. A text and data mining (TDM) exception, inspired by the EU, could allow lawful AI training for research while enabling rights holders to opt out via machine-readable notices. India's strict fair dealing rules should evolve into a flexible "fair use" framework based on the U.S. four-factor test to accommodate transformative, non-exploitative AI uses. Sui generis protection for AI-generated works, granting limited economic rights without moral rights and shorter terms, could encourage innovation without creating monopolies. Institutional reforms, including mandatory AI disclosure in registrations, specialized IP tribunals, and collaboration with tech authorities, alongside international engagement with WIPO and regional partners, would help India establish a balanced, forward-looking framework that fosters responsible AI development while safeguarding creative rights.

VII. CONCLUSION

To sum up, generative AI fundamentally challenges the traditional idea-expression dichotomy, which has long served as a cornerstone of copyright law by clearly distinguishing between unprotected ideas and protectable expressions. Unlike human creators who consciously transform ideas into original expressions, AI systems operate by extracting statistical patterns, tokenizing data, and generating outputs that often occupy a grey area between ideas and expressions. This blurring complicates the application of Indian copyright law, which is historically grounded in concepts such as human authorship, originality, and the substantial similarity test established in landmark cases like *R.G. Anand v. Deluxe Films*.

The rise of AI-generated works, along with the diverse methods of training these systems and the transformative nature of AI outputs, exposes gaps in the existing legal framework, leaving

creators, AI developers, and courts uncertain about how infringement, ownership, and authorship should be determined. Comparative experiences from the United States and the European Union highlight the importance of clear guidelines regarding authorship, transparent mechanisms for rights management, and flexible exceptions such as fair use or fair dealing adapted to the AI context.

For India to adequately address these challenges, it must undertake a multi-faceted approach: amending the Copyright Act to define authorship and ownership for AI-assisted and AI-generated works, introducing sui generis protection to incentivize innovation without granting monopolistic control, and strengthening institutional capacity through specialized intellectual property tribunals, modernized registration systems, and collaboration with technology regulators. By implementing these reforms, India can establish a balanced legal and regulatory framework that simultaneously fosters technological innovation in AI and safeguards the rights and creative contributions of human authors, thereby ensuring that the benefits of AI coexist with the principles of copyright protection. India should create a hybrid model that combines flexible fair use like the U.S. with transparency rules like the EU.