International Journal of Contemporary Legal Research

Authorship, Ownership and Liability in AI-Generated Works: A Critical Analysis of Human Authorship and Training Data Regimes in Generative AI

Rizwan Ahmad [1]* **and Abu Faizan** [1]

Abstract

The advanced growth of generative artificial intelligence has fundamentally challenged traditional dogmas of intellectual property (IP) law, especially the concepts of authorship, ownership and liability. The legal issues raised by AI-generated works that are critically analyzed in this paper are the changing definition of human authorship and the controversial concept of training data regimes. The period of 2023 to 2025 such as the case of *Thaler v. Perlmutter* (2023-2025) and the U.S. Copyright Office AI guidelines, and the litigations of *Getty Images v. Stability AI* and The *New York Times v. OpenAI* indicate that the judiciary continues to insist on human creativity as a precondition to receiving copyright protection while disclosing the unresolved tensions in the use of copyrighted data in training AI. The research claims that the current IP structures, based on anthropocentric conceptualization of creativity, become less effective to deal with practices of algorithmic production of content and massive ingestion of data. It also questions the fair use, consent, and liability doctrines available today whether they are doing enough to balance the innovations with the rights of producers, particularly in an area such as India where the regulation remains unclear. The paper focuses on the emergent global response to the issue by taking a comparative and doctrinal look at the EU AI Act and policy discussions regarding data governance. The paper has concluded that the IP law needs a recalibration, one that redefines the authorship criteria, defines the ownership rights, sets clear and fair standards of training data to make sure not only the technological progress but also legal responsibility is present in the era of generative AI.

Keywords: Artificial Intelligence; Copyright Law; Human Authorship; Training Data; Liability

1. Introduction: The Disruption of Traditional Copyright Paradigms

The rapid evolution of generative artificial intelligence (AI) has thrown copyright law's traditional foundations - premised on human creativity, originality and authorship into disarray. ChatGPT, DALL·E, and Midjourney can create text, images, music and even computer code that mimic human creations.[1] This development is challenging the orthodox view of copyright protection in the products of human creativity.[2] As a result, legal frameworks face complex issues of authorship, ownership and liability in the automation of creativity.

[1] MANUU Law School, Maulana Azad National Urdu University, Hyderabad, India

Correspondence: mohdrizwan1890@gmail.com

[1] *OpenAI, ChatGPT: Optimizing Language Models for Dialogue* (2022); *OpenAI, DALL·E: Creating Images from Text* (2021); *About Midjourney*, *MIDJOURNEY*, https://www.midjourney.com (last visited Apr. 28, 2026).

[2] *Feist Publications, Inc. v. Rural Telephone Service Co.*, 499 U.S. 340, 345 (1991).

This upheaval is underpinned by an incongruence. Copyright law aims to reward human ingenuity and safeguard individual expression, while generative AI is based on massive data processing, pattern matching and probabilistic predictions - even without human authorship.[3] This signals a need to re-examine whether current approaches are sufficient or need to be overhauled.

1.1. Rise of Generative AI and Creative Autonomy

Generative AI represents a shift from previous AI technologies in that they are able to autonomously create outputs.[4] These systems learn from large data sets of copyrighted and non-copyrighted works, and reproduce patterns in these works.[5] They are not simply tools, but have a degree of "creative autonomy" in generating outputs through algorithms, rather than human will.[6]

1.2. Tension Between Innovation and Legal Orthodoxy

Generative AI heightens the clash between technological innovation and legal traditions. It provides efficiency, ease of use and democratisation of creativity, but could infringe on human creative rights, particularly through the inclusion of copyrighted materials in datasets.[7]

This is evident in the courts. *Thaler v. Perlmutter* reaffirms human creativity, ruling out works solely created by AI.[8] And, in *New York Times v. OpenAI*, issues of training data and market disruption are at stake.[9] Both show an evolution in the adaptation of legal doctrines to technological change.

1.3. Research Methodology

The research has a doctrinal and comparative approach. It examines legislative, judicial and policy developments in various countries, such as India, the US and the EU. The article also involves a critical analysis of new case laws and existing litigations to gain insight into the evolution of the legal approach towards generative AI. It draws upon secondary sources like academic papers, policy documents and institutional guidelines to inform the debate and highlight the deficiencies of the existing law.

2. Human Authorship in the Age of Artificial Creativity

2.1. Classical Doctrine of Authorship: "Originality" and Human Intellectual Effort

Traditionally, copyright protection has been extended to works that are "original" and originality has been tied to human intellectual labour.[10] In some countries, like the United States, originality is equated with a low threshold of creativity and independent authorship, as set out in *Feist Publications v. Rural Telephone Service*.[11] Similarly, in India, the case law, especially *Eastern Book Company v. D.B. Modak*, follows the "modicum of creativity" doctrine, rather than the previous "sweat of the brow" doctrine.[12]

---

3 *World Intell. Prop. Org., WIPO Conversation on Intellectual Property and Artificial Intelligence* (2019).

4 Stuart Russell & Peter Norvig, *Artificial Intelligence: A Modern Approach* 1020–25 (4th ed. 2021).

5 *Eur. Parl., Intellectual Property Rights and Artificial Intelligence (2020).*

6 Margaret A. Boden, *AI: Its Nature and Future* 145–50 (2016).

7 *Org. for Econ. Co-operation & Dev., Artificial Intelligence, Machine Learning and Big Data in the Digital Economy* (2019).

8 *Thaler v. Perlmutter*, 687 F. Supp. 3d 140 (D.D.C. 2023).

9 *Complaint*, *The New York Times Co. v. OpenAI, Inc.*, No. 1:23-cv-11195 (S.D.N.Y. Dec. 27, 2023).

10 *Feist*, supra note 2, at 345.

11 *Id.*

12 *Eastern Book Co. v. D.B. Modak*, (2008) 1 S.C.C. 1.

The human intellectual labour doctrine has both justificatory and practical implications: it justifies the conferral of exclusive rights and offers a reliable means of identification of the author. But, in the context of works that are produced through AI systems which work on the basis of data-driven pattern recognition, rather than human creative effort, the relevance of these norms is questionable.[13] Lack of human intention in the creative process also undermines the human-centred notion of originality.[14]

Table: 1 Academic Writing Before vs After Generative AI (Observed Trends 2018–2025)

No.	Parameter	Pre-AI (Approx. 2018–2020)	Post-AI (Approx. 2023–2025)	Empirical Indicator / Data Trend	Interpretation
1	Average Writing Time (Research Paper)	3–6 months	2–8 weeks	Studies show 30-60% reduction in drafting time	AI accelerates production
2	Literature Review Time	4–8 weeks	1–3 weeks	AI summarization tools reduce reading load by 50%	Efficiency ↑, depth ↓
3	Draft Generation	100% human written	40–70% AI-assisted drafts	Surveys: majority students use AI for drafting	Hybrid writing model
4	Language Quality Errors	High variation (5-15% error rate)	Reduced to 2–5%	AI grammar tools improve linguistic accuracy	Standardization ↑
5	Plagiarism Detection Rate	15-25% flagged cases	5–15% flagged (but hidden AI paraphrasing)	Reports show drop in direct plagiarism rise in AI-paraphrased text	Detection complexity ↑
6	Publication Output (Papers/year)	Moderate growth	Significant surge (20-40% increase)	Large datasets (millions of papers) show output rise	Quantity ↑ vs quality debate
7	Authorship Pattern	Single / limited co-authors	Increased multi-author + AI-assisted writing	Studies show change in writing style consistency	Authorship blurred
8	Time to Complete Thesis	3–5 years (PhD avg.)	Reduced by 10–20%	Observed faster drafting & revisions	Completion speed ↑

[13] WIPO, supra note 3.

[14] Boden, *supra* note 6, at 145–50.

No.	Parameter	Pre-AI (Approx. 2018–2020)	Post-AI (Approx. 2023–2025)	Empirical Indicator / Data Trend	Interpretation
9	Student AI Usage	<10% (negligible)	60–90% (survey-based adoption)	Multiple global surveys confirm high adoption	AI normalization
10	Supervisor Dependency	High	Moderate reduction (20–30%)	Students rely more on AI tools	Mentorship shift
11	Revision Cycles	4–6 drafts	2–4 drafts	Faster editing using AI	Iteration speed ↑
12	Conceptual Understanding	Deep engagement	Mixed (surface + assisted)	Studies report **over-reliance risk**	Critical thinking concern
13	Detection Tools Use	Plagiarism only	Plagiarism + AI detection tools	Turnitin AI & similar tools adopted globally	Evaluation shift
14	Ethical Violations Type	Copy-paste plagiarism	AI ghostwriting / automation	Shift in misconduct type observed	Ethics complexity ↑
15	Research Skill Profile	Writing + analysis	Prompting + editing + validation	Skill transition noted in academia	New competency model

The table above shows empirical trends based on recent surveys, institutional reports and large publication data sets.[15] The numbers are approximate ranges rather than institutional numbers, as there is no centralised reporting on the use of AI in academic writing.

2.2. Judicial Position on Non-Human Authorship

The courts' responses to AI authorship have highlighted the importance of human authorship. The US courts definitively concluded that copyright law protects only human-made works, thus dismissing the argument that an AI could be considered the author of an original work of authorship in *Thaler v. Perlmutter*.[16] In the years since then (through 2026), including the refusal to grant a rehearing, this has been confirmed, establishing the idea that non-humans cannot be authors under copyright law.[17]

This position is part of the courts' aversion to anthropocentric legal norms. Judges have emphasised copyright law's role in rewarding human creativity and incentivising human authors, not computers.[18] This means that AI works without sufficient human input are typically not protected, and are considered to be in the public domain.[19]

[15] *Org. for Econ. Co-operation & Dev.*, *OECD Digital Economy Outlook* (2023).
[16] *Thaler*, supra note 9, at 146–48.
[17] *Thaler*, supra note 9, at 146–48.
[18] *Id.*
[19] *U.S. Copyright Off.*, *Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence* (2023).

2.3. The "Prompter as Author" Debate

A recent point of contention is whether the "prompter" who inputs prompts into an AI system may be regarded as the author of the work created. In favour of the "prompter as author" approach, it is argued that prompting is an activity that demands skill, creativity and judgement, especially in sophisticated generative systems in which the prompts determine the nature of the result.[20]

But there are doctrinal issues with this argument. Unlike other forms of authorship, in which an author has direct control over the expressive components of a work, prompting can involve only indirect control over an independent system.[21] The connection between the prompt and the outcome may be too tenuous to fulfil the authorship test. Additionally, there may be different outcomes from the same prompt, diminishing the sense of control.

3. Ownership Complexities in AI-Generated Outputs

3.1. Competing Ownership Models: User, Developer, or Public Domain

There are three main models of ownership for AI generated content that have been put forward in the literature and policies:[22]

Firstly, User as author and owner: This view suggests the human who provides the prompts or inputs should be considered the author and owner of the AI-generated work. This model is based on the notion that users have some level of creative agency, especially when it comes to iterative prompting and refinement of AI-generated content.[23] This model is in keeping with the incentive theory of copyright by incentivising human interaction with AI.[24] But this approach is seen as exaggerating the role of the user, particularly in cases where the product is mainly driven by algorithmic processes rather than user expression.

Secondly, the developer-as-owner model focuses on the AI system's designers, suggesting that ownership should be with the designers, trainers and maintainers of the system.[25] It is argued that developers invest significant efforts and resources, especially in establishing training data and model structures.[26] However, this approach has implications for excessive rights concentration, potentially lead to technology companies gaining undue control over a large chunk of content.[27]

Finally, the public domain model is the third option suggesting AI-generated works without human authorship should be placed in the public domain. This view aligns with the traditional human creativity requirement, and has been implicitly supported in countries that do not protect non-human works.[28] This approach is conducive to access and prevents monopolies, but could also undermine investment in AI creativity.[29]

3.2. Copyright Law in India: Gaps under the Copyright Act, 1957

The Indian law does not explicitly deal with ownership of AI works, resulting in interpretative problems. The Copyright Act, 1957,[30] typically assigns authorship to humans, with little acknowledgment of non-human authorship. For example, Section 2(d)(vi), which

---

20 Andres Guadamuz, *Artificial Intelligence and Copyright*, 7 W.I.P.O. J. 1 (2017).
21 *Id.*
22 *World Intell. Prop. Org.*, *supra* note 3.
23 Guadamuz, *supra* note 20, at 1.
24 William M. Landes & Richard A. Posner, *The Economic Structure of Intellectual Property Law* 37–42 (2003).
25 *Org. for Econ. Co-operation & Dev.*, *supra* note 15.
26 *Id.*
27 *Eur. Parl.*, *supra* note 5.
28 *U.S. Copyright Off.*, *supra* note 19.
29 *World Intell. Prop. Org.*, *supra* note 3.
30 Copyright Act, 1957 (India).

mentions "computer-generated works", defines the author as the person who causes the work to be created.[31]

But the provision was drafted in an environment before the advent of generative AI, when "computer-generated" referred to deterministic processes, rather than autonomous data-driven creativity.[32] When applied to today's AI systems, this provision presents interpretive challenges, as it is not clear who "causes" the work to come into existence - the user (prompt provider), the designer (system architect) or the platform provider (host).

However, the focus on human ingenuity in cases like *Eastern Book Company v. D.B. Modak* indicate that courts may be wary of providing protection where the work is not substantially human created.[33] This leaves a gap in the law, highlighting the need for clarity in the law or judicial creativity to provide solutions to the challenges of generative AI.[34]

3.3. International Context: US, EU and Emerging Jurisdictions

Comparative insights show a variety of perspectives on ownership of AI-generated works, but with a consistent focus on human authorship.

US, the courts and the Copyright Office have refused to protect works without human authorship, as recently reiterated in *Thaler v. Perlmutter*.[35] The general trend is that AI-generated works (without substantial human involvement) are not protected by copyright.

The European Union follows suit, by stressing the need for "author's own intellectual creation.[36] However, the EU has adopted a more regulatory stance with tools like the EU Artificial Intelligence Act, which, although not specifically dealing with copyright ownership, introduces transparency and accountability requirements that indirectly impact the regulation of AI-generated content.[37]

United Kingdom, The UK Copyright, Designs and Patents Act acknowledges "computer-generated works" with the author being the person who has made the arrangements for its creation.[38] While this seems to be more inclusive, it is unclear if this provision will apply to advanced generative AI.[39]

3.4. Threshold of Human Creativity and Control

Underpinning the debate about ownership is the question of the degree of human involvement that is necessary for copyright protection. This has given rise to a "threshold of human creativity and control" as a possible doctrine for the law.[40]

This would mean ownership hinges on the degree of human involvement, exercise of creative judgment, in determining what is ultimately expressed.[41] Relevant considerations might include the detail of the instruction, the level of iteration, the curation and placement of the generated elements and the human's level of control. If the human's role is significant and impacts the expressive features of the work, there may be a copyright claim.[42]

On the other hand, if the AI system has considerable autonomy and the human intervention is confined to the activation of the system, then the threshold may not be

---

31 *Id.* § 2(d)(vi).
32 P. Bernt Hugenholtz, *Copyright and Artificial Creation* (2019).
33 *Eastern Book Co.*, supra note 12.
34 *NITI Aayog, Responsible AI for All* (2021).
35 *Thaler*, supra note 9, at 146–48.
36 *Infopaq Int'l A/S v. Danske Dagblades Forening*, Case C-5/08, 2009 E.C.R. I-6569.
37 Artificial Intelligence Act, Regulation (EU) 2024/1689, 2024 O.J. (L 1689).
38 Copyright, Designs and Patents Act 1988, § 9(3) (U.K.).
39 *U.K. Intell. Prop. Off., AI and IP Consultation Response* (2022).
40 Boden, *supra* note 6, at 145–50.
41 Jane C. Ginsburg, *The Concept of Authorship in Comparative Copyright Law*, 52 DePaul L. Rev. 1063 (2003).
42 *Id.*

reached.[43] This approach is consistent with copyright law's aim of only protecting those things deserving of protection.

4. Training Data Regimes and Copyright Infringement Concerns

4.1. Nature and Sources of Training Data in Generative AI

Generative AI models are trained on massive data sets that may contain a combination of publicly accessible, licensed and scraped data.[44] Such training data can include anything from books to news articles, artworks, social media postings, and other forms of content.[45] The volume and variety of data is vital to the nature of AI systems to learn and produce meaningful outcomes.[46]

But there are legal and ethical issues associated with the training data. Often, copyrighted material is included in training data without permission of the copyright owner.[47] Some rely on data that is publicly available, but this does not necessarily mean that it is publicly domain.

4.2. Fair Use Doctrine vs Unauthorized Data Scraping

A key legal argument in the context of training data is whether AI training data use can be defended under the doctrines of fair use (in the United States) or fair dealing (in countries like India and the United Kingdom).[48]

Arguments in favour of fair use suggest that the training of AI is transformative, in the sense that it does not copy the works in their original form, but incorporates statistical and mathematical relationships. They contend the purpose of the training - to allow machines to produce new works - is distinct from the purpose of the original works.[49] This view is supported by previous court decisions allowing for intermediate copying for technological reasons.[50]

But others maintain that the massive unauthorised data mining for training purposes violates copyright, especially if it involves copying whole works.[51] They note that generative AI models can generate works that are alternatives to the copyrighted works, thus displacing them and harming the authors. The volume, automation, and for-profit use of AI in training are viewed as having a greater impact on the market than traditional fair use scenarios, with concerns of unjust enrichment.[52]

4.3. Ongoing Global Litigations

The issue of training data has also been at the vanguard of legal debate through recent cases, and judges are increasingly having to decide upon the use of data.[53]

In the case of *New York Times v. OpenAI*, plaintiffs claim that the use of their news articles in the training data sets used to produce AI has allowed the AI to produce works

---

43 Hugenholtz, *supra* note 32.
44 *Org. for Econ. Co-operation & Dev.*, *supra* note 15.
45 *Id.*
46 U.S. Copyright Off., *Artificial Intelligence Policy Guidance*, https://www.copyright.gov/ai (last visited Apr. 29, 2026).
47 U.S. Copyright Off., *Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence* (2023), https://www.copyright.gov (last visited Apr. 29, 2026).
48 Copyright Act, 1957 (India); Copyright, Designs and Patents Act 1988 (U.K.).
49 *Authors Guild v. Google, Inc.*, 804 F.3d 202 (2d Cir. 2015).
50 *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417 (1984).
51 *Eur. Parl.*, *supra* note 6.
52 Mark A. Lemley, *Fair Use and Artificial Intelligence* (2023).
53 World Intellectual Property Organization, *supra* note 3.

that are substantially similar to the news articles, thus impinging on their subscription model.[54]

Likewise, *Andersen v. Stability AI* raises allegations by visual artists that their artworks have been used to train image generators to produce works that replicate distinctive artistic styles.[55] This case presents challenges for protection of artistic style and the extent to which style can be protected by copyright.[56]

In *Kadrey v. Meta*, writers are suing over the use of copyrighted books to train large language models, claiming that it constitutes mass copyright infringement.[57] The litigation highlights the problem of large-scale digitization without permission, and echoes previous cases involving digital libraries.

Together, these cases suggest a trend towards courts focusing on the inputs of AI systems, in addition to their outputs.[58] They also suggest that there is a potential for a split of opinions based on interpretations of fair use, market harm, and transformative use.

4.4. Economic Rights of Authors vs Technological Advancement

The training data debate is ultimately about the balance between authors' economic rights and technological progress.[59] The aim of copyright protection is to reward authors for their creative efforts through the provision of exclusive rights, such as reproduction and distribution. But generative AI undermines this approach by requiring broad access to existing works to spur innovation.

From the copyright owners' viewpoint, the uncompensated inclusion of copyrighted works in training data can be seen as free-riding. It devalues the works and has potential chilling effects on creativity.[60] Several authors and creators have recently demanded licensing and compensation models to compensate them for the use of their works in the training of AI systems.[61]

A compromise solution may be to create hybrid licensing frameworks, such as mandatory licensing, collective rights management or data trusts, that attempt to strike a balance between authors' rights and innovation.[62] These models will seek to balance compensation of authors while enabling the innovative use of generative AI.

5. Liability in AI-Generated Works: Attribution and Accountability

5.1. Absence of Clear Liability Frameworks

Current legal frameworks are not well-equipped to handle liability issues related to AI-generated works.[63] Copyright law is mainly concerned with authorship and ownership, and tort and intermediary liability concepts are not entirely suited to address the multi-party, decentralised nature of AI systems. This leaves a regulatory blind spot where various actors - developers, users and platforms - may play a role in the generation and distribution of harmful content, but may not be held clearly liable according to current norms.[64]

---

54 *Complaint*, *The New York Times Co. v. OpenAI, Inc.*, No. 1:23-cv-11195 (S.D.N.Y. Dec. 27, 2023).
55 *Complaint*, *Andersen v. Stability AI Ltd.*, No. 3:23-cv-00201 (N.D. Cal. 2023).
56 *Id.*
57 *Complaint*, *Kadrey v. Meta Platforms, Inc.*, No. 3:23-cv-03417 (N.D. Cal. 2023).
58 Pamela Samuelson, *Implications of AI for Copyright Law* (2024).
59 Landes & Posner, *supra* note 24.
60 *Authors Guild*, *Statement on AI and Copyright* (2023).
61 Authors Guild, *supra* note 60.
62 *Eur. Comm'n*, *AI Policy Framework Discussions* (2024).
63 *World Intell. Prop. Org.*, *supra* note 3.
64 *Id.*

Many legal systems require liability to be based on intent or negligence, which may be hard to prove in the case of autonomous AI systems.[65] The nature of generative outputs also adds to the complexity, as developers and users may not be in full control of the outputs. As a result, courts are increasingly faced with instances where there is harm, but uncertainty as to who is responsible.

5.2. Developer Liability vs User Responsibility

In this case the developer should be liable because responsibility for the risks of using AI should lie with the designers and trainers.[66] Developers have control over key aspects of the system, such as the data, model and safety measures. When damage occurs due to biased training data, insufficient protections or oversight, or negligent design, developer liability is consistent with product liability law.[67]

But making developers liable in a wide range of scenarios could stifle innovation, especially in the rapidly changing field of technology. It also challenges the notion of whether developers can foresee and guard against all possible uses or abuse of their systems.

Conversely, the case for user liability stresses the importance of the users of AI systems in particular contexts. Users have some level of control over the system via prompts, instructions and choices about using and sharing its results. If harm occurs as a result of purposeful misuse such as creating defamatory material or copyright infringements, it seems reasonable to hold users responsible.

5.3. Platform Accountability and Intermediary Role

In India, the Information Technology Act, 2000 and the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 provide for liability of intermediaries, with a focus on due diligence and timely action to remove illegal content.[68] But the application of these laws to generative AI is unclear because the content produced is not "third-party" but is produced by the platform.

Internationally, there are regulatory trends towards increased platform responsibility. The EU Artificial Intelligence Act provides for risk-based requirements for transparency, record-keeping and risk mitigation measures to avoid harmful outcomes.[69] While not explicitly liability rules, these requirements mark a trend towards platform accountability for the risks associated with AI technologies.

5.4. Emerging Judicial Trends on AI Misuse and Accountability

Courts are showing a trend towards ensuring accountability in the deployment of AI systems, despite the lack of specific legislation. Judges have started to highlight the need for human oversight - especially professionals - in validating the results of AI systems.

For example, in the case of *Mata v. Avianca, a U.S.* court reprimanded lawyers for filing AI-generated content with made-up citations, highlighting the need for diligence in using AI technologies.[70] This decision is part of a trend towards holding individuals responsible for their actions despite the use of AI.

---

[65] *Id.*

[66] Ginsburg, *supra* note 41.

[67] *Id.*

[68] Information Technology Act, 2000 (India); Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 (India).

[69] Artificial Intelligence Act, Regulation (EU) 2024/1689 (EU).

[70] *Mata v. Avianca, Inc.*, No. 22-cv-1461 (S.D.N.Y. 2023).

Likewise, Current cases like *New York Times v. OpenAI* - while primarily centred on copyright infringement - also highlight issues with the liability of AI developers and platforms for generating content that is claimed to copy protected material.[71]

Table 1: ***Authorship, Ownership and Liability in AI-Generated Works: Legal Evolution & Emerging Framework***

Year	Case / Development	Legal Issue	Key Holding / Position	Significance
2008	**Eastern Book Company v. D.B. Modak**	Originality Standard	Modicum of creativity required for copyright	Established human creativity threshold in India; foundation for rejecting purely AI works
2013	**Feist Publications v. Rural Telephone Service**	Human Authorship	Copyright requires minimal creativity + human authorship	Global doctrinal base-AI lacks "human intellect"
2023	**Thaler v. Perlmutter**	AI as Author	AI cannot be recognized as author	Landmark ruling-firm rejection of non-human authorship
2024	**Andersen v. Stability AI**	Training Data Use	Use of artists works without consent challenged	Raised issue of style imitation + consent in AI training
2024	**Kadrey v. Meta**	Data Scraping	Alleged unauthorized use of books for AI training	Focus shifted from output - training data legality
2024–25	**New York Times v. Open AI**	Market Harm	AI outputs may substitute original works	Introduced "market displacement" argument in AI copyright
2025	**EU Artificial Intelligence Act**	Regulatory Framework	Imposes transparency & accountability obligations	First comprehensive AI regulation; indirect impact on copyright
2025	**Mata v. Avianca**	AI Liability	Lawyers held accountable for AI-generated false citations	Established principle: AI use and escape from human liability

[71] Complaint, *The New York Times Co. v. OpenAI, Inc.*, supra note 54.

Year	Case / Development	Legal Issue	Key Holding / Position	Significance
2025–26	**ANI v. Open AI**	Indian Position	Challenge to use of copyrighted content in AI training	First major Indian case; may shape future jurisprudence
2026	**Global Policy Shift**	Ownership & Governance	Focus shifting from "ownership" to "accountability and regulation"	Marks transition toward hybrid AI governance model

6. Indian Legal Developments and Emerging Jurisprudence

6.1. Absence of Direct Statutory Recognition of AI Authorship

The Indian copyright law (Copyright Act, 1957) does not specifically acknowledge artificial intelligence as an author.[72] The Act is built on the idea of human authorship and the author is generally a natural person, or in some cases, a juristic person (such as employers).

Section 2(d)(vi), which deals with "computer-generated works" states that the author shall be "the person who causes the work to be created.[73] This provision seems flexible but was passed in a technological paradigm that is different from current generative AI.[74] At the time, "computer-generated" works were thought to be rule-based and deterministic, rather than independently capable of creating original and creative content.

The interpretation and application of this provision in the context of generative AI is not clear.[75] It is not clear whether the "person who causes" the creation is:

- The user of the prompts,
- The programmer of the system, or
- The user of the system.

In addition, the Indian copyright jurisprudence has consistently focused on human creativity - as in *Eastern Book Company v. D.B. Modak*, in which the Supreme Court applied a "modicum of creativity" test.[76] This may mean that Indian courts may hesitate to confer authorship on works devoid of any significant human creative effort, thus barring AI-only creations from copyright protection.

6.2. Pending Litigation and Policy Debate

The first cases relating to copyright and AI in India are now emerging, particularly in *ANI v. OpenAI*.[77] This is a landmark case, given it raises issues concerning the use of copyrighted works in the training of AI and possible infringement of economic rights.

The case is representative of issues faced by the media and creative sectors in India such as:

- Infringements arising out of the use of copyrighted material for training AI,
- Possibility of AI outputs competing with existing markets, and
- No remuneration for the rights-holders.

---

72 Copyright Act, 1957 (India).
73 *Id.* § 2(d)(vi).
74 Hugenholtz, *supra* note 32.
75 *Id.*
76 *Eastern Book Co. v. D.B. Modak*, supra note 12.
77 *ANI Media Pvt. Ltd. v. OpenAI, Inc.*, No. CS (Comm) ___/2024 (Del. H.C. 2024).

While courtrooms are not the only place where things are happening. Regulatory bodies and think tanks are discussing the need for special regulations for AI, including transparency and ethical use, and intellectual property protection.[78] But these are in their early stages and no major reforms have been introduced yet.

7. Rethinking Copyright: Policy, Ethics and the Future of AI Creativity

7.1. Data Colonialism and Exploitation Concerns

A key ethical concern for generative AI is what is referred to as "data colonialism".[79] This involves the mass collection and use of data often without permission and without a payday from creators, especially in the Global South or marginalised groups.

Generative AI uses large datasets of cultural and intellectual works and, in a sense, steals the public good's creativity to produce valuable products.[80] This brings into question issues of exploitation, power imbalances and inequitable gain, particularly when tech firms profit from AI systems developed using publicly available but privately-owned materials.

If not carefully managed, generative AI has the potential to exacerbate economic inequality by centralising profits among a small number of corporate players, at the expense of individual artists.

7.2. Regulatory Responses (EU AI Act, Transparency Norms)

To address these issues, regulatory proposals are starting to appear, notably in the European Union. The EU Artificial Intelligence Act proposes a risk-based approach to regulation, with obligations on providers (and deployers) of AI systems, especially high-risk systems.[81]

While the Act doesn't directly address copyright ownership issues, it includes transparency measures that have significant implications for the training data challenge.[82] This includes the need to provide information on the use of copyrighted materials and to comply with intellectual property law. This helps improve transparency and allows rights owners to more effectively enforce their rights.

Outside the EU, there is a trend towards transparency standards, such as data set documentation, auditability of AI systems and disclosure of AI content.[83] These measures indicate a move away from ex post regulation (litigation) to ex ante to avoid harms.[84]

But the lack of regulatory harmony across countries poses challenges, highlighting the need for international harmonisation of AI regulation.

7.3. Towards a Hybrid Model: Human-Centric Authorship with AI Accountability

In this model, copyright protection would still require "human contribution", thus preserving the anthropocentric nature of the legal system.[85] However, AI systems and their creators would be held to higher standards of accountability, particularly with respect to the data used to train the AI system, transparency in its operation and mitigation of risks.[86]

Elements of this blended approach may include:

---

78 *NITI Aayog*, *Responsible AI for All* (2021).
79 Nick Couldry & Ulises A. Mejias, *The Costs of Connection* (2019).
80 *Id.*
81 Artificial Intelligence Act, Regulation (EU) 2024/1689 (EU).
82 Artificial Intelligence Act, Regulation (EU) 2024/1689 (EU).
83 *Org. for Econ. Co-operation & Dev.*, *supra* note 15.
84 World Intellectual Property Organization, *supra* note 3.
85 *Feist*, supra note 2, at 345.
86 *World Intell. Prop. Org.*, *supra* note 3.

1. Threshold-based creativity standards for authorship that take account of human creativity and control;
2. Mandatory licensing/remuneration for the use of copyrighted material in training data;
3. Labelling of AI-generated works to ensure transparency and awareness of usage;
4. Joint liability schemes, distributing responsibility between developers, consumers and platforms, as appropriate.

This shift is emblematic of a shift in thinking - from the allocation of exclusive rights to the allocation of rights as part of a governance system to manage creativity in the AI era.[87]

8. Suggestions and Recommendations

8.1. Clarifying the Legal Status of AI-Generated Works

Parliaments should make clear the position of AI-generated works under copyright. In India, the Copyright Act, 1957 needs to be amended or clarified to:

1. Specify AI-assisted and AI-generated works;
2. Clarify copyright exists only if there is a human intellectual contribution;
3. Offer interpretational guidance in relation to Section 2(d)(vi) ("person who causes the work to be created") in relation to generative AI.[88]

This will eliminate uncertainty and promote judicial certainty.

8.2. Establishing a Regulated Training Data Framework

Since training data plays a pivotal role in most cases, a regulatory framework is needed. Policymakers should consider:

1. Compulsory licenses for training with copyrighted content;
2. Collective management of rights to compensate authors;
3. Compulsory dataset disclosure and transparency standards, in line with international developments, such as the EU Artificial Intelligence Act.

8.3. Strengthening Judicial and Institutional Capacity

As AI-related disputes will be highly technical, it is important to increase institutional capacity by:

1. Setting up AI and technology law benches or expert panels;
2. Encouraging judicial education on AI technologies, data practices and digital forensics;
3. Supporting the creation of technical standards and guidelines for expert testimony. [89]

8.4. Promoting Ethical AI and Protecting Against Data Exploitation

To tackle issues of data colonialism, exploitation and abuse, regulations should include ethical protections such as:

1. Acknowledgment and respect of community rights over cultural and creative data;
2. Rectification of inequalities through fair distribution with creators whose art or data is used to train the datasets;
3. Promotion of ethical development of AI (fairness, accountability, transparency - FAT principles).[90]

---

87 Ginsburg, *supra* note 41.
88 *Id.* § 2(d)(vi).
89 *NITI Aayog*, *supra* note 78.
90 Nick Couldry & Ulises A. Mejias, *supra* note 79.

9. Conclusion

The explosion of generative artificial intelligence has led to deep legal uncertainties as to authorship, ownership and liability under copyright law. As this paper has shown, copyright law's traditional foundation - centred on human creativity, originality and human authorship - cannot effectively address works created by autopilot, data-driven technologies.[91]

The ownership analysis demonstrates a diverse range of ownership models - user, developer, and public domain that reflect different normative values, but do not offer a single best answer. Likewise, the conundrum posed by training data regimes, as reflected in the ongoing cases, such as *New York Times v. OpenAI*, *Andersen v. Stability AI*, and *Kadrey v. Meta*, provides another perspective, which reveals that the true battleground is not just the outputs of AI but also the inputs.[92] These cases exemplify the tension between authors' economic rights and the possibilities afforded by artificial intelligence.

Liability-wise, unclear attribution mechanisms also add to the legal complexities. The decentralised nature of AI (including the role of the developer, user and platform) complicates the application of traditional concepts of intention and control. Recent court developments indicate that although AI might change the modes of creation, it does not remove the role of humans, thus requiring contextual and flexible approaches to liability.

In India, the absence of specific provisions for AI generated works under the Copyright Act, 1957, and the emergence of new litigation like *ANI v. OpenAI* underlines the need for reform.[93] The current recourse to courts, while helpful, is not adequate to tackle the systemic and policy-oriented issues with generative AI.

This paper has suggested that the issues surrounding AI go beyond doctrinal issues to normative issues such as data exploitation, power imbalances and undermining of traditional justifications for copyright.[94] In this regard, an exclusive focus on ownership is not enough. Rather, it is necessary to reframe copyright as an element of a governance system that strikes a balance between innovation, accountability and benefit sharing.

A hybrid legal framework is needed, with a continued emphasis on human authorship, but also with the inclusion of necessary safeguards for AI systems. This model would include threshold tests for authorship, fair and transparent data practices for training AI systems, and appropriate liability measures.[95] It would also demand legislators to act, and for nations to work together.

91 *World Intell. Prop. Org.*, *supra* note 3.
92 Complaint, *The New York Times Co. v. OpenAI, Inc.*, supra note 54.
93 *ANI Media Pvt. Ltd. v. OpenAI, Inc.*, supra note 77.
94 Nick Couldry & Ulises A. Mejias, *supra* note 79.
95 *Feist*, supra note 2, at 345.