TRAINING AI ON COPYRIGHT WORKS: IS INDIA MOVING TOWARDS A COMPULSORY LICENSING REGIME?

Ananya Karan, Maharaja Agrasen Institute of Management Studies, GGSIPU

ABSTRACT

The rapid development of generative artificial intelligence (AI) has led to the worldwide debate of the legality of training AI models on copyrighted content. The US goes mostly by the principle of fair use, the European Union has a kind of opt out system, and India has taken a different path altogether. In a move that kind of surprised everyone, India backed a royalty - based compulsory blanket licensing system called One Nation, One Licence, One Payment Model (ONLP) by which AI developers must pay copyright owners every time they use their works for AI training purposes.

This piece delves into the regulatory landscape that India is trying to create around generative AI (GAI), looking at the legal basis, economic consequences, and policy aims of the government. Locating India's position relative to the other global frameworks and the court rulings as they are progressing, the article points out that compulsory licensing would theoretically give licensing rights and thus creators the upper hand and also clear up regulatory issues, but at the same time such a measure would be associated with higher innovation costs, the problem of administration, and the trouble of getting approval at the level of constitutional law.

Keywords: AI models, Copyright, Generative AI, Global Framework, Legality, AI Training.

1. Introduction

Artificial Intelligence (AI), especially generative models like large language models (LLMs) and diffusion based image generators, is basically dependent on the availability of sufficiently large training datasets. The majority of these datasets are composed of copyrighted literary, artistic, musical and audiovisual works which have been indiscriminately gathered from the web. This has led to a worldwide legal turmoil situation, wherein lawmakers have to deal with the problematic issue of copyright laws and algorithmic training.

India, with one of the largest digital economies and creator ecosystems in the world, is now facing the regulatory battle head on. At the end of 2025, the Department for Promotion of Industry and Internal Trade (DPIIT) of India, proposed a royalty, based licensing regime for AI training datasets, which would make it compulsory for companies like OpenAI and Google to pay Indian artists for the use of copyrighted content in training their AI systems[1] . This step is significantly different from the lenient "fair use" strategy employed in the US, and the opt out solution granted by the EU.

This paper examines whether India is going to adopt a viable compulsory licensing system for AI training and if such a government intervention would be an appropriate mix of innovation facilitation and intellectual property protection.

2. The Copyright Challenge in AI Training

Generative AI models [2]extensively explore datasets to discover linguistic patterns, visual relationships, and semantic associations. This practice of learning involves the copying, storing, and processing of enormous quantities of copyrighted works. Conceptually, such large scale ingestion from a copyright angle is capable of raising unauthorized reproduction, adaptation, and storage concerns.

Content owners worldwide have been filing a series of lawsuits against AI companies accusing them of infringement. In India, this conflict gained a judicial hearing through the

[1] Department for Promotion of Industry and Internal Trade (DPIIT), Generative AI and Copyright: Working Paper, Government of India, December 2025.
https://www.dpiit.gov.in

[2] Times of India, Government Extends Deadline for Public Feedback on Generative AI and Copyright Working Paper, 8 January 2026.
https://timesofindia.indiatimes.com/technology/tech-news/government-extends-deadline-for-public-feedback-on-generative-ai-and-copyright-working-paper/articleshow/126374366.cms

groundbreaking ANI v OpenAI case before the Delhi High Court, where the news agency ANI alleged that OpenAI had used copyrighted journalistic content for AI training without permission[3]. The Court's inclination to accept jurisdiction over offshore AI entities demonstrated a harsher regulatory posture . As AI models are gradually generating outputs that have great commercial value, the economic significance of data appropriation can no longer be ignored.

The major core of the dispute is that there is a central legal question whether training AI on copyrighted works is a violation or if it can be considered as a permitted exception like fair dealing or text, and, data mining?

3. The Indian Legal Position: Statutory Ambiguities and Doctrinal Gaps

3.1 The Copyright Act, 1957

India's Copyright Act, 1957, grants the copyright holder exclusive rights to reproduce, store, adapt, and communicate the work to the public. Section 52[4] provides for certain limited exceptions under the fair dealing doctrine, which mainly pertain to research, private study, criticism, review, and reporting.

Indian copyright laws are currently devoid of any precise statutory instructions on the legality of AI training, hence the courts and regulators have to deal with doctrinal uncertainties.

When AI training is defined as reproduction at scale, a remuneration right ensues as a matter of principle even in the absence of satisfying the traditional substantial similarity test in Section 51 of downstream outputs.

Nevertheless, AI training quite hardly reconciles with these traditional grounds for exemption. The commercial use of AI generated contents serves to negate the assertions of non profit research or educational usage. Besides, the large-scale scraping of data for algorithmic training is distinctly different from personalized human learning.

[3] *The Future of Indian Copyright Legislation in the Wake of ANI Media v. OpenAI*, IAM MEDIA (Jan. 15, 2026), https://www.iam-media.com/guide/india-managing-the-ip-lifecycle/2026/article/the-future-of-indian-copyright-legislation-in-the-wake-of-ani-media-v-openai

[4] Exceptions to Infringement Under Copyright Act, 1957https://copyright.gov.in/Exceptions.aspx

3.2 Emerging Judicial Signals

The ANI case depicts a judicial leaning towards the imposition of Indian copyright rules on AI developers. The Delhi High Court, by dismissing the jurisdictional plea, highlighted that the digital operations that impact the Indian market cannot escape domestic regulatory scrutiny.

Such a trend in court decisions means that, in the absence of legislative detail, the Indian courts might gradually view AI training more as a matter of rights holders.

4. India's Proposed Compulsory Licensing Regime

In December 2025, the Department for Promotion of Industry and Internal Trade (DPIIT) released a Working Paper on Generative AI and Copyright which proposed a fundamentally new regulatory framework[5] for the use of copyrighted works to train artificial intelligence systems. The main element of the proposal is the introduction of a mandatory blanket licensing system[6], whereby AI developers would be automatically granted legal access to copyrighted datasets, in return for the payment of royalties to a centralized collecting body representing the rights holders after commercialisation.This system represents a departure from the conventional mode of enforcement through litigation after the event to authorization by statute before the event, thus decreasing the risk of litigation and at the same time guaranteeing a regular flow of compensation to the creators. The main points of the proposal are:

- Automatic training access to all lawfully available copyrighted works
- Statutory royalty obligations arising from commercial use of AI models
- Centralized collection and distribution mechanisms, analogous to those of collective rights management organizations under copyright law.

The proposal is a clear policy move away from negotiating licenses with individual copyright owners, which is often unworkable due to the massive amounts of data needed to train large language models. Instead, it supports collective management of copyright, following the

[5] TechCrunch, India Proposes Charging OpenAI, Google for Training AI on Copyrighted Content, 9 December 2025.https://techcrunch.com/2025/12/09/india-proposes-charging-openai-google-for-training-ai-on-copyrighted-content/

[6] IIPRD Blog, Innovation under Licence: A Critical Analysis of DPIIT's Generative AI Copyright Proposal, January 2026.https://www.iiprd.com/innovation-under-licence-a-critical-analysis-of-dpiits-generative-ai-copyright-proposal/

example of the music, broadcasting and reprography sectors where such mechanisms have been in place for a long time.

In fact, DPIIT kept the public consultation on the following framework open till 6 February 2026[7], thus showing the government's readiness for stakeholder feedback and possible regulation adjustment. Such a participatory measure demonstrates the understanding that there is a complicated relationship between the incentives to innovate, the rights of creators, the constitutional safeguards, and the competition in the market.

4.1 Regulatory Motivation

The idea was born as a result of increasing lawsuits and regulatory uncertainty worldwide. The New York Times v. OpenAI[8] and Sarah Andersen v. Stability AI Ltd[9]. lawsuits in the United States have challenged the notion that AI training is non-infringing fair use. These cases dispute whether mass ingestion of copyrighted material without permission or remuneration can really be called "transformative."

These incidents have led to government regulation, pointing to the insufficiency of current copyright laws to deal with the scale and secrecy of AI training pipelines. More fundamentally, the suggested scheme strives to balance the two sides of Indian constitutional law namely Article 19(1)(a) (freedom of expression and dissemination of knowledge) and Article 300 A (right to property) by providing extensive access to information while at the same time securing compensation for the exploitation of property[10].

In addition, the compulsory licensing policy may be concerning under Article 19(1)(g) when the creators use the rights of their works for moral or reputational reasons and raise the question of the freedom of occupation and expression[11].

[7] Economic Times, DPIIT Extends Generative AI-Copyright Consultation Deadline to February 6, January 2026.https://economictimes.indiatimes.com/news/company/corporate-trends/dpiit-extends-generative-aicopyright-consultation-deadline-to-february-6/articleshow/126356281.cms
[8] The New York Times Company v. OpenAI, Inc., No. 1:23-cv-11195 (S.D.N.Y. 2023).https://nytco-assets.nytimes.com/2023/12/NYT_Complaint_Dec2023.pdf
[9] Sarah Andersen v. Stability AI Ltd., No. 3:23-cv-00201 (N.D. Cal. 2023).https://jipel.law.nyu.edu/andersen-v-stability-ai-the-landmark-case-unpacki
[10] INDIA CONST. Article 19 in Constitution of India, Indian Kanoon, Article 19 in Constitution of India
[11] INDIA CONST. art. 19; *see also* Article 19 in Constitution of India, Indian Kanoon, Article 19 in Constitution of India

5. Commercial and Economic Implications

India is a major growing market for generative AI platforms, in fact, OpenAI's CEO Sam Altman has publicly referred to it as the company's second- largest market globally. Setting up a royalty regime would essentially change the economics of businesses, raising the costs of compliance while at the same time, validating AI training practices.

From the viewpoint of the broader economy, the scheme has the potential to open up data monetization to a wider base, enabling creators to garner a fair participation in the AI value chain. However, in the absence of tiered licensing mechanisms, the negative impact may be felt most by domestic startups who would then need to bear the costs, thus the dominance of capital, rich multinational corporations would be further strengthened.

Furthermore, given that the rates for licensing remain commercially viable, regulatory certainty could be a driver for attracting the right kind of investments. If royalties are too high, it may discourage innovators whereas, if the price is too low, it will not be able to ensure the adequate protection of creators. Therefore, finding this balance is of paramount importance for a sustainable development of the market.

6. Conclusion

India's initiative to enforce compulsory licensing for AI training data may be one of the most ambitious regulatory measures in the world. Through this, India wants to put creators first and give them a fair share of the revenue, at the same time providing a clear regulatory framework, thus rebalancing the economic power between technology platforms and content producers.

But the success of the regulation will largely depend on the institutional framework, its economic impact and the possibilities given by technology. A stiff licensing system might be counterproductive to India's goal of making it a global AI hub. On the other hand, there is a real risk of the creative industries being jeopardized if data is taken without any regulation.

In principle, India's One Nation, One License, One Payment concept, if carefully carried out, could serve as a world reference point for ethical AI governance reconciling the need for innovation with intellectual property rights.