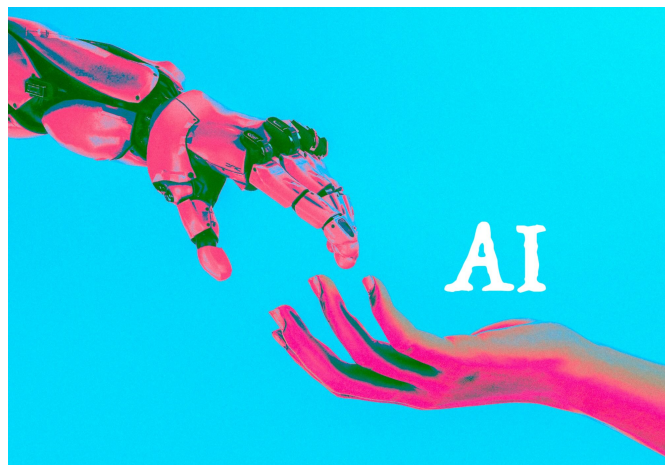


Gen AI & Copyright: India's Pragmatic Middle Path

The rapid evolution of large language models (LLMs) and generative AI systems has intensified longstanding tensions in copyright law. AI systems derive their capabilities from access to massive volumes of text, images, audio and audiovisual content, much of which is protected by copyright. While AI developers argue that publicly available content should be freely usable for training, creators and publishers increasingly contest uncompensated use, citing economic harm and erosion of creative incentives. Globally, this conflict has resulted in fragmented litigation and regulatory uncertainty.



Against this backdrop, the Department for Promotion of Industry and Internal Trade (DPIIT) constituted a committee on April 28, 2025, to examine AI-copyright concerns, assess the adequacy of existing legal frameworks, and recommend reforms. The first of its two working papers on Generative AI and Copyright was released on December 8, 2025, and marks a significant policy intervention, proposing a hybrid model that seeks to avoid extremes on either side of the debate.

Global Context: Why Existing Models Fall Short

The DPIIT Committee undertook a comparative review of global approaches, including those in the US, UK, EU, Japan and Singapore. Each jurisdiction reflects a different balance between innovation and copyright protection.

In the United States, AI developers rely heavily on the doctrine of fair use, arguing that AI training is “transformative”. However, recent litigation - such as *Thomson Reuters v. Ross Intelligence*, *Kadrey v. Meta* and *Bartz v. Anthropic* - has produced divergent judicial outcomes, with no settled appellate guidance. Japan and Singapore permit broad text and data mining (TDM) exceptions, subject to conditions such as lawful access and non-prejudicial use. The EU and UK have adopted TDM exceptions with opt-out mechanisms, though these have been widely criticised for being ineffective and difficult to enforce at scale.

The Committee concluded that none of these models, in isolation, adequately serve India’s objectives: promoting AI innovation, ensuring access to high-quality and diverse datasets (to reduce bias and hallucinations), and safeguarding the economic interests of copyright holders - particularly informal creators in unorganised sectors. Zero-priced licensing undermines creator incentives; opt-out TDM regimes are impractical; voluntary licensing is logistically unworkable at scale; and full rightsholder control risks hold-outs and data fragmentation.

The Hybrid Model: Mandatory Access with Statutory Remuneration

Rejecting an “all or nothing” approach, the Committee proposes a Hybrid Model centred on mandatory blanket licensing coupled with a statutory remuneration right.

Under this framework, AI developers would be entitled, as a matter of right, to train models on all lawfully accessed copyrighted works. In return, creators and rightsholders would receive statutory remuneration through a single-window, government-designated non-profit collective - the proposed Copyright Royalties

Collective for AI Training (CRCAT). Notably, rightholders would not have the option to withhold their works from AI training; compensation, rather than consent, becomes the primary safeguard.

The blanket licence would apply across classes of works, with CRCAT comprising copyright societies and collective management organisations (CMOs) representing different categories such as literary, musical and audiovisual works. Commercial AI developers would make annual payments to CRCAT.

Royalties would be triggered not at the training stage but upon commercialisation of the AI system. This design choice is intended to avoid upfront barriers to innovation, particularly for start-ups and research-driven developers. Rates would be fixed by a government-appointed committee comprising legal, economic, financial and technical experts, and would follow a flat, revenue-share model linked to the AI developer's global revenues from the commercialised system. Rates would be reviewed periodically and remain subject to judicial review.

Importantly, the Committee recommends that the royalty obligation apply retroactively to AI systems already trained on copyrighted content, once such systems achieve commercial success.

To support royalty allocation, AI developers would be required to submit a "Sufficiently Detailed Summary" of training datasets - disclosing broad categories, nature and lawful sources of content - through a light-touch disclosure format. The intent is to provide enough transparency to protect rightholders' interests without imposing burdensome compliance obligations that could deter AI development in India.

Royalties collected by CRCAT would be distributed across different classes of works proportionally, based on disclosed usage. Copyright owners seeking remuneration would need to register their works on a Works Database for AI Training Royalties maintained by the relevant copyright society or CMO, providing basic details such as author name, work title, nature of the work and publication details. Both organised and unorganised creators would be eligible to receive remuneration, and unclaimed royalties would flow into a welfare fund for under-represented sectors.

Why the Model Matters

The Committee's proposal reflects a pragmatic recognition that waiting for courts - domestic or foreign - to resolve AI-copyright disputes is neither efficient nor equitable. Litigation between AI firms and content producers is ongoing worldwide, but judicial outcomes remain fragmented. In the meantime, the absence of a clear framework disproportionately benefits large AI players capable of absorbing legal uncertainty.

By prioritising statutory remuneration over consent, the Hybrid Model attempts to preserve creative incentives while avoiding structural barriers to AI innovation. While concerns remain - particularly around rate-setting fairness and proportional distribution - the Committee correctly acknowledges that a flawed but functioning system is preferable to regulatory vacuum.

Stakeholder comments on the working paper are invited until the first week of January 2026. Whether the Hybrid Model ultimately becomes law or evolves through consultation, it signals India's intent to chart an independent, balanced course in the global AI-copyright debate- one that recognises both data as an engine of innovation and creativity as an economic right.