VOLUME 6 | ISSUE 4

INTERNATIONAL JOURNAL OF ADVANCED LEGAL RESEARCH

AI AND COPYRIGHT IN INDIA: RETHINKING THE BASICS

- Prachi Garg[1]

INTRODUCTION

The Rampant contemporary growth of generative artificial intelligence has raised several dilemmas surrounding copyright violations and interpretation of copyright laws in light of this development. The most prevalent argument that copyright owners contend in litigation across jurisdictions, be it the United States or India, is that AI's use of copyrighted works to train its generative software amounts to infringement of their legal rights and therefore must fall within the 'fair use' or 'fair dealing' exceptions envisaged within the law in order to be valid. This very contention is the assumption that is shaping the current jurisprudence around generative AI and copyright violation.

This blog insists that this argument is fundamentally disconcerted, especially in the Indian context. We need not begin with the presumption that use of copyrighted works by generative AI models for training purposes is infringement per se, rather the provisions of the Indian legislation are very much determinative of the scope of the protection it grants. Section 14, of the copyright Act 1957, does not confer a general right over all uses of a work. It specifically grants certain exclusionary rights in terms of reproduction, adaptation, economic exploitation of the work etc. to the copyright owner. Hence, if a use neither reproduces the expression, nor seeks to adapt the work substantially, it falls outside the very domain of copyright protection.

Therefore, use of copyrighted works by generative AI platforms for training their models by extracting statistical patterns rather than copying the 'expression' of a copyrighted work, falls outside the scope of s.14. By treating this act as infringement, current debates fail to recognise the explicit distinction between 'expression' and 'information'. This poses a significant risk by expanding the domain of copyright beyond its intended limits. A more cohesive approach would be to adjust this debate in light of idea-expression dichotomy and

[1] Student, LL.M (IP Law & Management), National Law University, Delhi
For general queries or to submit your research for publication, kindly email us at ijalr.editorial@gmail.com

https://www.ijalr.in/

the statutory limits of s.14, rather than relying on external defences to call this act infringement.

COPYING V. COPYRIGHT INFRINGEMENT

The very starting point of this contemporary debate is that just because AI systems make copy of copyrighted works, have access to these works for the purpose of training, they must be infringing copyright per se. Presuming that copying amounts to infringement shuts the scope for asking the relevant question as to whether it even engages with copyrighted expression at all or not. Indian copyright law lays down a clear legal premise that does not prohibit all types of copying, rather, s.14 protects specific uses of a work, especially those that involve reproduction of the protected expression or its communication to the public. [2]Indian legal jurisprudence has often laid down that it is this act of communication of the work to the public by adding more recipients to the expression that amounts to infringement if not done via authorised or statutory means. Where this human recipient angle is absent, it may amount to copying but it does amount to copyright infringement. This doctrinal distinction becomes more relevant in context of AI training. Generative AI platforms scrape the web for large volumes of data, the expression of some of which may indeed by copyrighted, then use it to extract statistical patterns to train the software. However, nowhere does this process involve communication of the copyrighted work's expression to public. There is no reproduction of such expression. It is merely extraction of certain patterns from already available information rather than violation of copyright. [3]Unlike conventional acts of infringement, AI training does not make copyrighted expression available for consumption, nor does it increase public access to it, it rather generates 'metainformation' from larger datasets to extract patterns and train Gen AI from it.

SECTION 14 AND THE SCOPE OF COPYRIGHT PROTECTION

A deeper perusal and interpretation of s.14 of the Copyright Act 1957 reflects how the AI-Copyright debate is often misconcepted in India. Infringement of copyright in India arises only when it is in violation of the rights explicitly granted by s.14. The Indian courts have repeatedly distinguished between ideas and expression contained in a work, hence, when a

2 Tips Indus. Ltd. v. Wynk Music Ltd., MANU/MH/0862/2019 (Bom. H.C. Apr. 23, 2019)
3 Oren Bracha, Generative AI's Two Information Goods, YALE L.J. (forthcoming)

use does not lead to reproduction of the expression of copyrighted work, it does not fall within the scope of this right. Indian courts have also narrowed down the interpretation of adaptation rights. [4]In Barbara Taylor Bradford v. Sahara Media Entertainment Ltd., the Court rejected the contention that any form of alteration would amount to adaptation. It held that "alteration" must be construed narrowly, only covering minor or recognizable alterations. It does not include alterations which are of such a nature that the resulting product is totally different. [4]In E.M. Forster v. A.N. Parasuraman, as well, the court held that themes, plots and ideas are not copyrightable, rather their expression is. The guidebook in this case was found not to be reproducing the copyrighted expression and was an independent creation, hence infringement was not made out. These judicial interpretations have time and again reiterated that such transformative works do not constitute copyright infringement as the work created is substantially different in nature, purpose and character from the original work, hence falling beyond the scope of s.14. Likewise, extraction of patterns from copyrighted work and use of it for training generative AI could also be seen as a transformative work in the Indian context, thereby not amounting to infringement. However, this definitely does not except liability at the output stage where if the output generated reflects substantial copying of the copyrighted expression, it might amount to infringement. But this non-expressive training process remains outside the scope of infringement.

The confusion around AI and copyright largely comes from mixing up technology with legal rules. Just because AI systems use large amounts of data does not automatically mean there is copyright infringement. Under Indian law, copyright does not give complete control over all uses of a work. It only protects specific rights, mainly related to the use of the *expression* of a work. If a use does not involve copying or sharing that expression, it may fall outside copyright altogether.

In the case of AI training, systems analyse data to identify patterns, rather than reproduce or communicate the original content. Treating such use as infringement would wrongly extend copyright protection to cover information itself, which has never been protected. This could

[4] Barbara Taylor Bradford v. Sahara Media Ent.Ltd., MANU/WB/0106/2003 (Cal. H.C. July 16, 2003) [4] E.M. Forster v. A.N. Parasuram, AIR 1964 Mad 331 (Mad. H.C.)

also create serious problems for innovation, as many new technologies depend on access to large datasets to function and improve.

At the same time, it is important to recognise that not all concerns around AI can or should be addressed through copyright law. Issues like job displacement, ethical use of data, and market impact require separate legal and policy solutions. Expanding copyright beyond its intended limits is not the right way to deal with these challenges.

Therefore, instead of assuming that AI training is infringement and then trying to justify it through exceptions, the correct approach under Indian law is to first ask a simpler question: does this use even fall within the rights granted under Section 14? If it does not involve protected expression, then copyright law may not apply at all.

CONCLUSION

A brief comparison with the United States law highlights the clarity in Indian position. In the US, courts treat transformation as part of copyright, but then excuse it through fair use. This creates confusion, as something can be both infringing and permissible at the same time. Recent cases such as [5]Andrea Bartz v. Anthropic and [6]Richard Kadrey v. Meta show how courts are still struggling with this issue. Indian law avoids this problem because it deals with transformation at the level of grant of rights themselves. This simplifies the debate in India, because here the real question is not whether use of copyrighted work for AI training could be justified via an exception but the fact of it not constituting infringement could be ruled out at a statutory level itself. As discussed above, training does not reproduce protected expression or create a recognisable version of any work. Instead, it works by identifying patterns resulting into creation of an altogether transformative work. The problem with contemporary debate is treatment of all copying as infringement and failure to distinguish between expression of a copyrighted work and information already in public domain. This often tends to expand copyright law beyond its intended scope. Concerns of market competition, labour displacement etc. are indeed real and relevant, but they cannot be addressed by expanding the scope of copyright protection. A more balanced approach would be to keep copyright limited

[5] Andrea Bartz and ors.v. Anthropic PBC, No. C 24-05417 WHA (N.D. Cal. June 23, 2025)
[6] Richard Kadrey v. Meta Platforms, Inc., No. 23-cv-03417-VC (N.D. Cal. Mar, 2025)

to its core function i.e. protecting expression and incentivising creation. But, the process of AI training and use of large datasets for it, falls outside the scope of copyright protection. This approach is not only consistent with the Indian statutory interpretations but is also in consonance with contemporary technological advancements.