GENERATIVE AI AND THE LOOPHOLES IN FAIR USE: A LEGAL ANALYSIS

Saloni Kumari, Campus Law Centre, University of Delhi

ABSTRACT

The rapid integration of generative AI into creative industries has exposed significant legislative gaps in India's copyright framework, which was not designed to address non-human creative agents or AI-driven data use. The Copyright Act, 1957 contains no explicit provisions governing AI-generated works or the use of copyrighted content as training data, leaving courts and stakeholders without clear guidance. Lastly, the essay will propose a set of possible solutions that aim to protect the rights and livelihoods of human artists while still encouraging the responsible and ethical development of AI tools in creative industries.

Hayao Miyazaki founded Studio Ghibli to create meaningful, handcrafted films. In these movies, he explored complex human emotions and themes like nature, pacifism, and childhood innocence. With his captivating storytelling, visual aesthetic, and animation style, he intrigued everyone.

But last year, a trend took over social media, with people using AI tools like ChatGPT to instantly generate images in the cherished Ghibli style. This sparked discontent among viewers who understood the art and had developed a deep connection with the films; and instigated a much-needed discussion about copyright infringement by Generative AI.

While the rapid advancement of AI has democratized creativity, making it accessible to people worldwide, it has also raised complex issues that are difficult to navigate and regulate. On one hand, this accessibility has led to innovative expressions and broader participation in creative fields. On the other hand, this accessibility has sparked legitimate concerns among artists, creators, and audiences who feel that the essence of human-made art is being diluted or even replaced. This development has challenged our understanding of the

copyright laws that protect the work of artists and their livelihoods as AI threatens to replace them.

As AI continues to evolve and become increasingly capable of mimicking human creativity, it puts pressure on existing copyright frameworks that were not designed to deal with non-human creators. This disruption has exposed the inadequacies of current copyright laws, particularly in their ability to define ownership, consent, and infringement in the context of AI-generated works.

Understanding Copyright and Fair Use

Both the creative industries and AI sector are trying to navigate the copyright laws to protect their work and prevent infringements, with unclear guidelines, lack of precedents, and no explicit provisions.

In order to understand the complexity of the current legal uncertainties, it becomes important to revisit what copyright law entails and the objectives it was designed to serve.

Copyright laws are intended to protect original artistic works, including literary, dramatic, musical, etc., by giving the owner exclusive rights to reproduce, adapt, distribute, and publicly perform or display a work. Copyright does not protect ideas but the expression of ideas. The work to be protected should be original and should have a minimal level of creativity.

These rights of copyright owners are protected by different statutes provided within the specific jurisdiction under which the work was created and the infringement took place. Due to the internet, this issue has become global, with anyone being able to access content worldwide; international treaties become important. Berne Convention, 1883, is the most important one, with more than 180 countries as its signatories. It provides any content produced in one of these countries' protection from infringement in all other countries as well. Its basic principles are national treatment, automatic protection, independence of protection, and minimum standards of protection.

The use of copyrighted work can be done if it falls under the definitions of "fair use" in the USA or "fair dealing" in the UK or India. Whereas the protection in the USA is broader and

more subjective, the fair dealing in India is very objective and leaves very limited space for judicial discretion.

Factors to be considered while determining fair use under US law[1] are—

1. The purpose and character of use, i.e., for nonprofit educational purposes, research, criticism, commentary, reporting news, etc.

2. The nature of copyrighted work: the higher the creativity involved, the stronger the protection.

3. The amount and substantiality of the portion used in relation to the copyrighted work as a whole

4. The effect of use upon the potential market or value of the copyrighted work, if the work is directly competing and adversely affecting the market of the original work, it most likely won't be fair use.

All the factors are weighed together while determining for use, but transformative nature and its impact on the potential market are more significant. Even if the work is transformative, the court might not consider it fair use if it hurts the potential market of the original work.

In the UK, fair dealing laws are governed by Section 29[2] and 30[3] of the Copyright Designs and Patents Act, 1998. -

1. When used for research or private study

2. When used for the purpose of criticism or quotation

3. When used to report current events (with the exception of photographs)

Article 9(2) of the Berne Convention[4] introduces the Three-Step Test for *any* exception to

[1] U.S. Copyright Act, 17 U.S.C. § 107 (1976)
[2] Copyright, Designs and Patents Act 1988, s. 29 (UK)
[3] Copyright, Designs and Patents Act 1988, s. 30 (UK)
[4] *Berne Convention for the Protection of Literary and Artistic Works*, art. 9

copyright, including fair use or fair dealing:

1. Certain special cases only (exceptions like fair use or fair dealing must not be too broad or vague)

2. No conflict with a normal exploitation of the work.

3. No unreasonable prejudice to the legitimate interests of the author.

India takes a similar stand as UK regarding 'fair dealing' in Section 52 of the Indian Copyright Act, 1957[5].

The work produced by AI must comply with established copyright laws to ensure that, its use of the content on which it has been trained, falls within the boundaries of fair use. Since AI systems rely on vast datasets, often sourced from existing copyrighted works, it is crucial to determine whether the AI's output infringes upon the intellectual property rights.

How AI and Generative Models Function

Artificial intelligence refers to computer systems that can do complex tasks that used to require a human level of intelligence; these can learn and perform programs without being explicitly programmed. They learn from experience; by perceiving the environment and the responses received, this makes them very valuable for any task.

AI models are trained by using diversified datasets to build machine learning models. This model then makes predictions using the input data and then is provided feedback to adjust its predictions accordingly, the training and predictions are stored by the model. This cycle of data input and feedback is continued until the desired level of optimization.

Artificial intelligence models have been used for years for many tasks, like optimizing the algorithms on apps like YouTube and Instagram for personalized content and targeted ads, in voice assistants like Siri, in e-commerce apps for product recommendations, Google Maps for real-time updates, in educational apps for adaptive learning, and many more.

[5] Indian Copyright Act, 1957, sect 52

Among the various types of artificial intelligence categorized based on capabilities, functionality, and application, one in particular—generative AI—has attracted many users and legal scrutiny. This is due to its capability of creating new content like texts, images, music videos, etc.; most popular examples of generative AI are ChatGPT, Midjourney, Deepfake tools, Amper, and Runway ML.

Generative AI is different from traditional AI, as traditional AI analyzes, predicts, classifies, recommends, and optimizes based on input, whereas generative AI constructs something novel. This ability gives creative access to the masses but also creates problems. As due to its purpose to develop something new based off the dataset provided to it, during training, it ends up giving results that are similar to that dataset or are adaptations of it based on the input given.

This leads to a legal evaluation of the methods by which data is sourced for training the models. The process of forming a training dataset includes various techniques, like web scraping (the automated extraction of data from websites), utilization of public datasets, proprietary data generated by a company from its operations, data collection by crowdsourcing, and creation of synthetic datasets.

Of these methods, web scraping is contentious from the legal perspective, as it involves the extraction of copyrighted content without obtaining the necessary permissions. This raises important questions about whether it falls under the fair use provision of copyright laws.

Should companies be prohibited from training AI using such data? Considering that growth in such data over past years and its global availability are major reasons for the rapid development seen in AI, it is argued that restricting the use of this data will be perilous for the AI sector.

Is AI work really Fair Use

When determining fair use, there are two positions to consider: whether the training of an AI model using artistic works is fair use and whether the content generated by AI is fair use.

The training data used for AI models has doubled every nine months since 2010; we must move forward keeping in mind such large amounts of human-generated data being used for training. It is argued that the training of AI comes under ‘fair use’ or ‘fair dealing’ as it is

for the purpose of research and development of a technology. While this is true, it needs to be understood that the core issue here is that, here R&D is done with commercial intentions. The developers of AI are gaining profits without reasonably compensating the original data creators. A landmark decision of Thomson Reuters[6] in the US recognized this stance by declaring the use of the company's content to create a competing AI-based legal platform impermissible.

The question of whether the outputs given by AI are fair use or fair dealing has increasingly become relevant due to the growth and popularity of generative AI. The data used to train generative AI generally manifests in the work created by AI, sometimes holding a close resemblance. The consequence of which people start to visit the chatbots for that content instead of their websites. For instance, in New York Times vs OpenAI[7], plaintiffs argue that the results of AI are competing in the same market, and as they were based off of their content, it is copyright violation.

Transformative use in US

Under US jurisdiction, transformative use of copyright is considered fair use. Hence, it is also argued by AI companies that the content created is highly transformative, making both the training of AI models and its outcomes fair use.

Transformative use involves giving the work a new meaning or purpose. As decided in the DABUS case (U.S)[8], AI can't be an inventor, due to the reasoning that AI is not a human and it lacks consciousness, intention, and understanding, the qualities also required for creativity. If the courts are not imparting a mental process on AI, it can also not be capable of adding value to any preexisting creation on its own and hence can't make transformative use.

The court also considers the market impact and the nature of the original work while determining the legitimacy of the defence of transformative use. Here, not only does the work of AI not minimize the financial harm to the original creator, but it also acts as a

[6] Thomson Reuters Enterprise Centre GMBH and West Publishing Corp. v. Ross Intelligence Inc., No. 1:20- cv-613-SB
[7] *The New York Times Co. v. Microsoft Corp. et al.*, (U.S. District Court for the Southern District of New York, Ongoing
[8] *Thaler v. Vidal*, 43 F.4th 1207 (Fed. Cir. 2022)

substitute for the original work. If the new content can act as a substitute for the original one, there isn't much innovation there.

This brings into focus the need to examine international standards, particularly those under the Berne Convention. As the statutes governing copyright should maintain reasonability within Article 9 of the Berne Convention[9]. Article 9 allows exceptions only in special cases, which AI no longer is, given the great scale on which the training is done. Further, by not compensating the original artists, it is violating the clause of 'No conflict with a normal exploitation of the work.'

Reimagining Copyright for the Age of AI

The fast-paced evolution of AI technologies has outpaced the development of legal safeguards. There needs to be a balanced approach in order to protect the rights of creators without hampering the development of AI. Without clear guidelines and policy interventions, the gap between innovation and regulation will only widen.

Interpretation of laws by courts needs to widen existing laws in order to include works done by AI under the purview of it. This can be done by requiring a minimum threshold of transformation in such works. Given that many AI-generated works currently fail the transformative use test due to their substitutive nature and minimal originality, courts may consider establishing a stricter standard of transformation tied to user-directed input. This approach would help distinguish between passive reproduction and active creation, allowing only sufficiently modified outputs to be considered for fair use.

Policy reforms are essential for ensuring better compliance and accountability. These policy changes need to be made at an international level, as AI infringements are not a territorial issue, and there is a need for a uniform mechanism relating to AI-generated works. Hence, the Berne Convention, which governs international copyright protection, should be revised to include provisions regarding AI-generated works.

Rethinking the limitations within copyright laws is also needed regarding the protection that can be provided to styles or ideas and its distinction from expression, albeit in a restricted form. Artistic styles not being protected under copyright laws is the reason for allowing the

[9] *Berne Convention for the Protection of Literary and Artistic Works,* art. 9

recreation of art by AI that took an artist enormous times to learn and perfect. The question arises: if this is not an infringement by humans, then why AI? The answer lies in the fundamental nature of recreation. Where on one hand, AI is a tool that statistically replicates patterns from its training data, humans recreate the idea or style by applying their minds to it and make it their own expression with a different context. While tangential to the main argument, it is relevant to mention that an expression of a similar idea by a human is tolerable on an economic landscape, but AI, with its capability of faster creation, disrupts the market for the original artist, and the more artistic value of the work, the stronger enforcement of this is required.

Developers can also receive directives to establish licensing agreements with copyright holders, guaranteeing fair compensation. This has recently been done by many companies, like Reddit striking a deal with Google to allow Google to use its content for training AI models, or between OpenAI and Hearst, a major media company, showing this is a way to resolve major ongoing lawsuits. This approach enables development while ensuring compensation for entertainment companies and artists. Additionally, it will inform users if the result given by AI includes a significant portion of someone's work. Such an approach will allow users to make informed decisions before using that content commercially. The process will also help clear out the liability in a future infringement suit.

Technology is a useful servant but a dangerous master. – Christian Lous Lange. The aim of artificial intelligence should be to empower human creativity, not to undermine it. Hence, we must tread the way for the future, keeping in mind that human creations are to be protected. The future of creativity depends not on what AI can do, but on what we allow AI to do and at what cost.