GENERATIVE AI AND COPYRIGHT INFRINGEMENT

Taligattu Subhash, Damodaram Sanjivayya National Law University

ABSTRACT

The emergence and development of Generative Artificial Intelligence (GenAI) has changed the creation and distribution of literary, artistic and informational content completely. Training GenAI on the basis of vast amounts of data, including copyrighted works, has caused serious issues relating to the unlawful reproduction and use of derivative works and infringement of the exclusive rights of the authors. The current legal regulation based on the Indian Copyright Act, 1957 which was developed during the pre-digital age does not provide proper solutions concerning the problems related to AI output and machine learning through the use of copyrighted materials. This paper explores the connection between generative AI and copyright infringement paying attention to the issues of authorship, originality and ownership under Indian copyright law. In this context, through a comparative analysis of the current international practices of GenAI regulation in the USA, the UK and the EU the study will try to assess the possibility of balancing technological development and copyright in India**.**

Keywords: Artificial Intelligence, Generative Artificial Intelligence, Copyright, Infringement, Copyright law.

I. Introduction:

The advent of Artificial Intelligence (AI) has heralded an epoch of unprecedented technological advancements, radically altering the way information is generated, processed, and distributed. Among the diverse applications of AI, Generative Artificial Intelligence (GenAI) stands out as one of the most disruptive technological developments of the twenty-first century. Unlike conventional AI programs that only analyse and manipulate data, GenAI can generate original-looking content.

The operation of generative AI relies on training based on huge datasets, which include books, journal articles, pictures, artworks, and other creative outputs, many of which are copyright-protected. This issue has prompted heated debates about whether the use of copyrighted materials to train AI systems infringes on copyright laws. creators claim that the use of their work without permission compromises their rights. However, tech giants assert that using copyrighted materials to train AI is a transformative use of the existing works and vital to progress and innovation.

The problem adds importance within the Indian legal environment. In India, the Copyright Act, 1957, which predates the advent of AI technology, sought to govern conventional methods of creation and authorship. There is no clear provision within the Act that explicitly provides any guidance on the issue of the use of copyrighted materials for training AI technology, nor does the act deal specifically with the ownership of AI created materials. It thus stands to reason that the traditional copyright principles of authorship, originality, and fair dealings among others, have never been more challenged than now.

The need to balance innovation with creators rights necessitates a reconsideration of India's copyright framework

Against this background, this paper will seek to discuss the issue of Generative AI technology in the context of copyright infringement and to explore whether the copyright system in India is ready for it.

II. Legal framework and background

The swift growth of generative artificial intelligence (AI) has posed new challenges to copyright law around the globe. Generative AI refers to AI systems which can generate texts,

songs, artworks, computer programming languages, and other types of content that are similar to the creations made by humans. Such systems are trained using massive data sets, which include copyrighted content drawn from books, research papers, photographs, web pages, and other digital libraries. As such, issues relating to ownership, authorship, violation, and liabilities are becoming central to discussions on copyright law in modern times.

Copyright law today can be dated back to the Statute of Anne in 1710 and has since then been developed through various international treaties like the Berne Convention and the TRIPS agreement, which call on member countries to offer minimum levels of protection of copyright.[1]

The Copyright law in India is mainly regulated by the Copyright Act, 1957, which is amended from time to time, specifically by the Copyright (Amendment) Act, 2012. Copyright protection is provided to the following types of works under the Act: original literary, dramatic, musical, and artistic works, cinematograph films and sound recordings. Section 13 of the Act provides the categories of works eligible for copyright protection whereas Section 14 deals with the exclusive rights of the copyright owners.[2]

One of the essential aspects of copyright protection in India is that of originality. In Eastern Book Company v. D.B. Modak (2008),[3] the Supreme Court embraced the idea of the "modicum of creativity," stating that in order for a work to receive copyright protection, it has to reflect some level of creativity and intellectual efforts of its creator.

With the advent of AI technology, the notion of originality has become even more complicated. In accordance with Section 2(d)(vi) of the Copyright Act,[4] in the case of computer-generated works, the author of such a work shall be the person who causes the work to be created. This provision was adopted in the time when there were no independent generative AI systems, and thus the main purpose of this provision was to regulate computer-assisted works. Therefore, it becomes unclear whether it is the user, the developer, or even the AI system itself that is the

[1] Statute of Anne, 1710, 8 Ann., c. 19 (Eng.); Berne Convention for the Protection of Literary and Artistic Works, Sept. 9, 1886, as revised at Paris, July 24, 1971, 1161 U.N.T.S. 3; Agreement on Trade-Related Aspects of Intellectual Property Rights art. 9, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1C, 1869 U.N.T.S. 299.
[2] The Copyright Act, No. 14 of 1957, §§ 13–14, India Code (1957), as amended by the Copyright (Amendment) Act, No. 27 of 2012.
[3] Eastern Book Co. v. D.B. Modak, (2008) 1 SCC 1 (India).
[4] The Copyright Act, No. 14 of 1957, § 2(d)(vi), India Code (1957).

author of an AI-generated work.

A third critical area is the use of copyrighted works during the training of AI systems. The models used by generative AI are often trained using huge amounts of data, including copyrighted works, without the explicit consent of the copyright holder. This gives rise to problems related to the unlicensed reproduction of copyrighted works and the violation of the exclusive rights granted by Section 14 of the Copyright Act.[5] On the other hand, the developers of such technologies claim that the use of copyrighted works in this context would amount to transformative use and may be considered an exception to the principle of fair dealing.

In Indian copyright law, there are very limited exceptions available under Section 52 of the Copyright Act, which is known as fair dealing exceptions.[6] The problem here is that these exceptions have not been framed with AI technology in view and there is nothing much about data mining and machine learning in those exceptions. This leaves judges and policy makers with a tough job of figuring out how to apply the copyright laws to the reality of AI.

The lack of a legal framework for AI training and creation indicates that we have to think again about traditional notions of authorship, originality, ownership, and infringement.

III. Main Analysis and Arguments - Balancing Innovation and Creative Rights:

The emergence of generative AI technology may be seen as one of the most disruptive technological developments in intellectual property law after the introduction of the internet. While other innovations served only to facilitate the work of human beings, the new technology allows producing essays, artworks, music compositions, computer programs, and even legal research papers automatically. It has shaken the very basic premises upon which copyright law was founded. First, traditional copyright law was based on two major premises: the fact that creative expression arises out of intellectual labor and, thus, its author should have exclusive rights in order to encourage further creativity. However, generative AI challenges both assumptions.

The above dispute brings into question the possibility of a technology gaining tremendous commercial success on the basis of copyrighted works without the creators' permission.

[5] The Copyright Act, No. 14 of 1957, § 14, India Code (1957).
[6] The Copyright Act, No. 14 of 1957, § 52, India Code (1957).

Traditionally, copyright protection has guaranteed that authors are shielded from the reproduction of their work since their economic and moral rights are seen as vital for the development of culture and knowledge. The use of copyrighted works for AI training without the permission of the author might be detrimental to the economic interests of the author and may decrease the incentive to create content. The authors, artists, and musicians become afraid that the AI system trained using their creations becomes a competitor to them.

On the other hand, overly protective policies might inhibit the development of technology. Artificial intelligence may change the face of such fields as health care, science, and creative industry. Overly stringent conditions required for obtaining a license to develop the AI system might make it too costly to develop, which would allow only a few companies to afford such development.

Another problem that arises relates to the problem of authorship. Anthropocentric copyright laws prevail all over the globe; the law considers that authors are natural persons able to conduct intellectual activities. As per the Copyright Act, 1957, especially Section 2(d)(vi), the author of any computer-generated work shall be the person by whom the work is caused to be created. But this law was passed during a period when computers were just tools for humans rather than independent systems with the ability to produce creative works.

The issue of who the author of an AI-generated work is still unclear. Awarding authorship to the developers overlooks the role of the users, whereas giving authorship to the users undervalues their contribution. As AI does not have legal standing, there is need for legislation concerning the ownership of AI-generated works.

The originality criterion recognized in the case of Eastern Book Company v. D.B. Modak is hard to establish for AI-generated content as the creative effort involved might have been little human intellect at all.[7]

However, AI-generated outputs have high potential for reproducing copyrighted expressions, thus bringing up tough issues concerning infringement and liability. Since copyright infringement assumes some form of human activity, it is not quite clear who can be held responsible.

[7] Eastern Book Co. v. D.B. Modak, (2008) 1 SCC 1, supra note 3.

The lack of clear guidance in legislation in India is especially worrying. After all, the Copyright Act, 1957, was passed at a time when artificial intelligence and machine learning technologies were unthinkable. Even though it has been amended many times since then, it fails to mention any problems connected with text and data mining, training processes in AI, ownership of the materials generated by AI, or machine-generated infringement. Only judicial interpretation may fail to deal with all the new problems.

Some useful information can be found from the experience of other countries. For example, the EU has created some exceptions regarding text and data mining, but on the other hand, it allows rightsholders to opt-out of having their works used for AI training.[8] In the USA, much litigation has taken place about whether using copyrighted material for AI training is a fair use.[9] All of this means that countries around the world are facing the same problems.

India must not go to either of these extremes. Excessive freedom in the regulation of AI can harm the creators' rights and well-being. At the same time, an overly strict system can hold back technological development and hinder India's ambitions to become one of the world leaders in AI technology. What needs to be done is to find a balanced approach to the problem.

First, it is necessary to provide in the legislation the rules regarding the use of copyrighted material in the process of AI training and create a system of transparent licensing in cases when it is needed. Secondly, it is necessary to clearly define in the Copyright Act what will be considered the work of AI and require at least some participation of humans in order to protect the works from infringement. Finally, special provisions regarding the responsibility for copyright infringement committed by AI need to be introduced.

In essence, the question that is raised by generative artificial intelligence is not just a technical one, but a juristic one, which forces us to rethink the very purpose of copyright law altogether. Copyright is not only about rewarding the creator for his creative work, but also ensuring that people learn from the creations and spread the knowledge gained. In the era of artificial intelligence, this purpose remains the same, but its means need to change. The law can no longer stay rigid when the technological revolution changes the very nature of creation and innovation. Indian copyright law needs to be reconstructed in order to safeguard innovation

[8] Directive 2019/790, of the European Parliament and of the Council of 17 April 2019 on Copyright and Related Rights in the Digital Single Market, art. 4, 2019 O.J. (L 130) 92, 103.
[9] Complaint, Authors Guild v. OpenAI Inc., No. 1:23-cv-08292 (S.D.N.Y. filed Sept. 19, 2023); Complaint, N.Y. Times Co. v. Microsoft Corp., No. 1:23-cv-11195 (S.D.N.Y. filed Dec. 27, 2023).

and human creativity in general.

IV. Recommendations :Towards an AI-Responsive Copyright Framework in India

With the exponential progress in the realm of generative artificial intelligence, it becomes imperative to make a complete evaluation of the copyright law of India. The Copyright Act of 1957 was framed during a period when the very notion of machine-generated art was not possible and hence does not have adequate provisions to resolve the legal problems posed by modern-day AI technology.

First, the legislature can formulate provisions regarding the use of copyrighted material for AI learning purposes. A statutory licensing system could be established under which AI developers will be able to access copyrighted material by making appropriate compensation to the owner.

Second, the Copyright Act must define the legal status of the works generated by AI. Only if the creation contains sufficient human effort can it qualify for copyright protection. Otherwise, such works must be considered a part of the public domain.

Third, there needs to be transparent data governance in which AI developers would have to make known the origin of the training data and also keep record of any copyrighted material used for training purposes. This will go a long way in promoting accountability, allowing copyright holders to find out any unauthorized use of their works.

Fourth, specific clauses regarding liability of any infringement by artificial intelligence need to be included in the law. The law has to make it clear the conditions in which developers, platform operators, and users may be considered liable for any infringements committed by artificial intelligence systems.

Lastly, India should take an approach that promotes innovation while at the same time protecting the interests of the creators. As has been mentioned already, the goal of copyright law is the encouragement of creativity, knowledge, and technological advancements. Therefore, the rethinking of Indian copyright law in the age of artificial intelligence becomes essential.

V. Conclusion

Generative artificial intelligence is a revolutionary event that has occurred in the realm of intellectual property law. Technologies able to create literary, artistic and musical works challenge some of the basic premises upon which copyright legislation has been based. Such notions as authorship, originality, ownership and infringement have been formed within an environment where creativity is perceived as being an activity unique to humans only. Generative artificial intelligence technology poses challenges and raises questions related to copyright laws.

The main problem copyright legislation faces is connected to the need to reconcile two equally important tasks. Copyright aims at protecting the economic and moral rights of authors who are granted exclusive rights to their creations. However, the law also needs to promote innovation, technological progress and dissemination of knowledge for the public good. Too restrictive a system can inhibit the development of artificial intelligence and deprive society of the tremendous possibilities it presents. At the same time, too permissive an approach might jeopardize the rights of authors because unauthorized usage of copyrighted material to train AI and competition between creators and machines created on the basis of their works might occur.

However, the Indian Copyright Act, 1957, in its current version does not appear to be adequate in dealing with the problems mentioned above. First of all, the Act lacks provisions regulating the use of copyrighted content in order to train AI systems. Moreover, the issue of the authorship and ownership of AI generated works also remains unaddressed. Hence, legislative changes have become a matter of necessity in this case.

It is crucial to note that as the technology keeps developing, the legislation should change accordingly. Copyright law throughout its history has managed to accommodate new forms of human creativity and methods of expression that have emerged thanks to various inventions and discoveries. This means that the era of generative artificial intelligence requires the same process of legal development in order to ensure fair treatment of technological advancement in this context.

Thus, reimagining Indian copyright law in light of the rise of the generative AI should not only be a question of legal obligation, but also a normative requirement in terms of ensuring further progress of the law.

Bibliography

Cases

Eastern Book Co. v. D.B. Modak, (2008) 1 SCC 1 (India).

Statutes

The Copyright Act, No. 14 of 1957, India Code (1957), as amended by the Copyright (Amendment) Act, No. 27 of 2012.

Statute of Anne, 1710, 8 Ann., c. 19 (Eng.).

Treaties and International Instruments

Agreement on Trade-Related Aspects of Intellectual Property Rights, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1C, 1869 U.N.T.S. 299.

Berne Convention for the Protection of Literary and Artistic Works, Sept. 9, 1886, as revised at Paris, July 24, 1971, 1161 U.N.T.S. 3.

Directive 2019/790 of the European Parliament and of the Council of 17 April 2019 on Copyright and Related Rights in the Digital Single Market, 2019 O.J. (L 130) 92.

Litigation Materials

Complaint, Authors Guild v. OpenAI Inc., No. 1:23-cv-08292 (S.D.N.Y. filed Sept. 19, 2023).

Complaint, N.Y. Times Co. v. Microsoft Corp., No. 1:23-cv-11195 (S.D.N.Y. filed Dec. 27, 2023).