

**THE CONUNDRUM OF CHATGPT AND ITS INTRINSIC COPYRIGHT MALADY: A  
HUMAN RIGHTS**

**Vignesh Bhise** Fourth Year B.A.LLB Student, Thakur Ramnarayan College of Law, Mumbai  
**Seona D'mello** Fourth Year B.A.LLB Student, Thakur Ramnarayan College of Law, Mumbai

**Abstract:**

*The usage of artificial intelligence systems such as ChatGPT cannot be justified under Indian copyright law. This article discusses the legal and ethical difficulties that come from the creation and distribution of derivative works utilising AI platforms. This raises questions about the legal tenets of fair use, the development of derivative works, and the legality of data collection and use. The use of information provided for the aim of training and improving AI models raises serious concerns about potential copyright breaches. Thus, this paper articulates proposals for protecting the best interests of owners of copyright and rivals, while also addressing legal difficulties and accelerating AI growth. Furthermore, this study examines the ChatGPT platform as a case study to investigate the necessary changes that copyright regulations must undergo in order to adequately address the complexities of creators and ownership in the realm of AI-generated artistic works, thereby upholding an author's basic rights.*

**Keywords:**

ChatGPT, Artificial Intelligence, Copyright, Freedom of Expression, Human Right Approach.

**Introduction:**

ChatGPT, or generative pre-trained transformer (GPT), is a key player in the new internet era and getting work done quickly and more effectively, as it is necessary in the modern technologically advanced world. From translating to efficiently completing a tonne of other paperwork. The story doesn't end here. If you can clearly express the problem, this AI-powered software, which is a math wizard, can solve even the trickiest algebraic puzzles. It will indicate where you missed the comma in the coding if you are attempting to troubleshoot your incorrect language code. With a few instructions, it can compose an essay on any topic, regardless of how difficult it may be, and if you have already written one, it will grade it.

This tool has shown that ChatGPT is a true benefit to humanity. However, despite all the benefits that the ChatGPT wizard has brought, there have been numerous instances of first-hand authors' rights being violated, as their work is being used without their consent and because they stand to gain so little from it. As a result, the integrity of the original work is being compromised.

People all around the world, have discovered that ChatGPT is capable of responding to user questions and completing specific tasks, from providing solutions to various problems to producing complex text or writing computer programs. These tasks range from providing smoothly written and high-quality translations.

ChatGPT and the technology have the potential to be widely employed in business as tools to carry out tasks that, in the past, could only be carried out by humans. The improved GPT-4 model, which is currently being used in the next iteration, has been announced. It was trained utilizing a vastly increased amount of data and numerous additional starting parameters.

**The Strong Language model of the AI tool:**

By producing automated text-based responses, the ChatGPT software program was created to help and support interactive communication. The process entails combining data from numerous sources, including books, journals, websites, and articles, leading to the creation of unique and interesting discourse. The application of computational techniques to simplify the comprehension and analysis of massive volumes of data in a way that mimics human language is known as **natural language processing (NLP)**. By utilizing NLP functionalities, computer systems are capable of obtaining significant amounts of data or carrying out tasks. It's important to remember that ChatGPT's operating

capabilities are restricted to syntax comprehension. The method used includes a thorough examination of the sentence's syntactic structure and the determination of the ChatGPT software program was created to help and promote interactive conversation by producing text-based responses automatically. Information from various interdependencies among its component parts is integrated into the process. ChatGPT continues to rely on statistical methods for analysis and comprehension even after extensive training in handling large datasets.

The statement above implies that the system has the capacity to spot trends in the usage of particular terminology and idiomatic expressions and makes use of these patterns to provide educated guesses about the intended meaning of a given sentence or phrase. It is vital to recognize that this supplied answer could seem clear and insightful, yet it might not fully understand the philosophical tenets and theories that underlie the question. A sophisticated chatbot called ChatGPT uses natural language processing (NLP) to analyse enormous amounts of data and generate responses that closely mimic human speech. At the moment, its abilities are restricted to syntactic understanding, which means that it can analyse sentence structure and understand how words and phrases relate to one another. Therefore, for the purpose of avoiding future issues, it is crucial that individuals order to verify that the designs adhere to predetermined requirements for accuracy and efficiency for particular usage scenarios, they should be reviewed and modified.

### **Ethics of ChatGPT from Intellectual Property Point of View**

To err is human, and since humans have given birth to AI, we cannot expect it to be error-free either. This actuality exposes us to the big bad world of legal and ethical dilemma concerning its use. The utmost view of ChatGPT on intellectual property is a highly analytical topic as well as highly debatable consultation and numerous other beginning conditions. Machine learning is a sensitive topic that has already prompted the adoption of certain legal measures, including litigation against the developers of AI systems. Large-scale systematic analysis and processing of data, particularly data that contains works protected by copyright laws in different jurisdictions, is necessary for training the concerned system. The fundamental point of contention is whether or not users of systems that "learn" are permitted to use databases that are available to the public. The main criticism directed at providers of these kinds of systems is web scraping, which refers to the illegal automatic collection and processing of enormous amounts of publicly available data for the purpose of content creation. These are people who assert the original writers' or other right holders' rights. It's significant that owners of databases used by AI systems to "learn" can make such claims. If ChatGPT is deemed to be capable of creating original literary work, it can be regarded as an author, and if it is regarded so, there arises the issue of copyright infringement liability as well as human rights violations as well. However, the terms of OpenAI seem to indicate that infringement liability appears to shift onto the subsequent owner of the output generated, namely, the user (remember that OpenAI assigns ownership of the generated output to the user who entered the respective input).

### **ChatGPT and Copyright issue:**

Content published by the AI chatbot ChatGPT may appear to have been written by a person. Although the technology has many potential applications, some crucial issues regarding content ownership are raised by its outstanding capabilities. The fundamental tenet of copyright law is that only works produced by people are eligible for protection. The algorithms that power ChatGPT were created at OpenAI, therefore it would seem that the business has copyright protection over those. But perhaps not for chatbot responses subject to any agreement to the contrary, the employer is the initial owner of any copyright in any literary, dramatic, musical, or artistic work created by an employee while they are employed.

The possible legal protection of a "creative" output created by AI technologies under copyright or comparable intellectual property (hence referred to as "the output") has been a major focus of the

current legal discussion around generative AI and copyright<sup>1</sup>. It's crucial to understand, though, that using copyrighted data to train and create AI systems raises serious copyright concerns as well.

More innovations of this nature are certain to occur, and early adopter businesses have the opportunity to profit from the current scenario by employing AI to boost the productivity of their operations. The "first-mover advantage" refers to the potential advantage that businesses may experience when they are the first to launch a good or service to a market.

One could foresee a scenario where an AI is classified as the author and the creators of that AI as the owners of the output if copyright law in the future evolves away from its focus on human action. This might lead to a scenario in which a small number of extremely influential AI firms have enormous power. They can wind up with hundreds of thousands of pieces of copyrighted content, including songs, books, images, and other digital assets. One could argue that this could eventually result in a dystopian scenario where the majority of newly created works are produced by AI and owned by corporations.

Books, articles, webpages, and social media posts are just a few of the materials used as input by ChatGPT as a language model. The copyrighted works that are utilized to train ChatGPT's language processing algorithms may be found in these sources. It is possible that ChatGPT will encounter similar difficulties given the legal considerations around copyright and the usage of training data for machine learning. The conundrum is whether and under what circumstances using copyrighted information to train ChatGPT's language processing algorithms is legal, similar to other AI systems. This raises a serious legal challenge because ChatGPT and other generative AI systems rely largely on vast volumes of training data, which may contain copyrighted works. ChatGPT uses "training" as a mechanism to generate the input data it uses. The model receives a large corpus of textual data during the training phase, which is used to guide the speech-processing algorithms.

The corpus under review demonstrates the ability to include a wide variety of text-based sources, including but not limited to books, articles, websites, social media posts, and comparable items. The specific job or use case for which the model has been trained determines the type of data used to provide instructions to ChatGPT. Customer conversation transcripts or online reviews may be used as the training data corpus if ChatGPT is designed to handle customer service enquiries in a certain language. Any information used for the dissemination of must be obtained to ensure compliance with copyright laws. This can entail getting authorization to use content that is protected by copyright or using information that is open to the public. Fair use or other legal exemptions could be pertinent in certain situations. It is important to remember that this is a dynamic and varied area of law, therefore each individual case must be carefully examined.

A smart writer of creative content must cite their sources in order to support the accuracy of their work and prevent plagiarism. Why are the language models used by ChatGPT exempt from this demand? The solution is simple: language models lack strong opinions and the ability to come up with original thoughts. They are able to produce texts thanks to comprehensive training on a variety of data sources. However, it is imperative that all writers, whether they be humans or machines, properly credit and reference their sources. This procedure not only ensures the trustworthiness and accuracy of the material but also helps to curtail some instances of plagiarism. The methods used for editing and citing sources in human writing and language models (such as ChatGPT) differ significantly. Human authors frequently express their culpability and see the source as a moral duty. Citing sources is one way to substantiate your arguments, and you must take responsibility for the accuracy of your claims.

### **Legal Limitations with Output Allocation from the Lens of Copyright Laws in India**

Allocating output to the user seems challenging technically and practically, and it would be failed on many levels. Let's evaluate the technology issues from the standpoint of India. First off, assuming that ChatGPT immediately transfers ownership of whatever work it produces to OpenAI is not legally tenable. India is required to have human authors according to Section 17, as shown by: Sec. 16 of the Act states that "no person" is entitled to copyright except as provided by law, while Form XIV for the

---

<sup>1</sup> European journal of Risk Regulation 2023

Application for Registration of Copyright (Form-XIV) requests disclosure of name, nationality, and address, and Sec. 2(d)(vi) of the Act permits authorship to be granted to the person "who causes the work to be created" in the case of computer-generated works.

Such artificial intelligence models, like ChatGPT, produce material that draws from a variety of literary and other works as well as the writings of other writers. If a previously published work served as inspiration for the content, there was no copyright violation. It is not immediately clear if the material was produced by the input source, the OPEN AI, or both. The user is likely to retain ownership of any copyright in any submissions, even if there could be certain restrictions depending on the specific terms and conditions. Users should thus carefully inspect them before sending them to an AI writing tool.

The user is responsible for the input and output, "including for ensuring that it complies with all applicable laws," as stated in the ToU (Terms of Use). What transpires when a third party's copyright is unintentionally breached even when the Output is not protected by copyright? When pressed, ChatGPT acknowledges that a sizable quantity of text data was used in its training. However, no credit is given to ChatGPT for the original creation of the Output. It is unclear if consent has been obtained for such data. Because the usage could not even be regarded as fair dealing, this presents a difficult issue for the user who claims copyright over the Output. In cases involving infringement before the Calcutta High Court determined in *Saregama India Ltd. Ors vs Alkesh Gupta & Ors on 1 October, 2013 GA No. 3009 of 2013 in CS No.347of 2013* and *Delhi High Court (Tips Industries Ltd v Wynk Music Ltd and Anr., N.M(L) 197/2018 in C.S. I.P(L) 114/2018, 23 April 2019*<sup>2</sup>, that streaming sound recordings (protected under copyright law) for a "fee or revenue may be from its sponsors or from third parties" was commercial exploitation, and the Delhi High Court reached a similar conclusion. This was the case where the respective Defendant websites/OTT platform allowed streaming of sound recordings protected under copyright law and claimed exception under Section 52(1)(a)(i) of the Act, 1957 (i.e. private or personal use).

### **General Issue Related to ChatGPT and Copyright Law from Human Rights Approach:**

The possibility for the text produced by ChatGPT to be regarded as a derivative work is one concern with ChatGPT and copyright legislation. A work that is based on another work may be considered a derivative work and may enjoy the same copyright protections as the original work. In the case of ChatGPT, the text produced by the model could be based on sources that are protected by copyright, which might make the content subject to legal safeguards and limitations.

Another issues is with the ability for the model to be trained on copyrighted items without the owner's consent is another problem with ChatGPT and copyright law. If this happens, it can be considered copyright infringement, which is when anything is used without permission. The fair use doctrine, which permits the restricted use of copyrighted content for purposes like criticism, comment, news reporting, teaching, scholarship, or research, is an exemption to the copyright law.

Concerns about the usage and dissemination of the content created by ChatGPT are present in addition to these problems. There may be extra legal problems connected to copyright law, for instance, if ChatGPT is used to create text for commercial reasons, such as marketing or advertising. Similar to this, there can be issues with regard to ownership and control of content created by ChatGPT if it is shared online or through other channels.

### **Conclusion and Way Forward!!**

In this digital age, new technologies that allow for process automation are emerging quickly. Without a doubt, ChatGPT and similar tools will make technical work simpler, more effective, and faster, but we must be aware that these tools cannot take the place of human innovation.

The arrival of AI-generated writing has made it necessary to take copyright laws into consideration. As a result of ChatGPT's progress, it is now able to create original material. Therefore, it will be necessary for the stakeholders to consider how new technologies may impact copyright infringement.

---

<sup>2</sup> <https://indiankanoon.org/doc/121369009/>

In India, AI-created text and copyright laws provide a number of difficulties, including how to implement these regulations and safeguard artists against copying. There have been claims of copyright infringement as a result of ChatGPT's usage of AI. According to Open AI, it would make an effort to protect its consumers from illegal content and compensate original creators. The uncertainty around the legal standing of language generated by AI, however, raises questions concerning the reliability of this site and how India's copyright laws are applied.

The law must now be changed to reflect contemporary shifts in technology and culture, given the rising use of AI and its derivatives, which pose a danger of violation. Additionally, the user should be aware of Open AI's disclaimer, which notes that it is not intended to provide advice and that it may occasionally create offensive or biased content in addition to false or misleading information. The user must exercise independent discernment prior to consuming the content. Additionally, although they are still in their infancy, there are already a lot of ethical considerations to be made when using them. Technology developments are fantastic as long as it's used responsibly, therefore ChatGPT has to be controlled as well, like other AI, but it could take some time for governments to implement effective legislation.

**Reference:-**

- <https://spicyip.com/2023/03/ai-and-copyright-law-analysing-the-impact-of-chatgpt.html>
- [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4483390](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4483390)
- [https://www.researchgate.net/publication/373496346\\_ChatGPT\\_A\\_Case\\_Study\\_on\\_Copyright\\_Challenges\\_for\\_Generative\\_Artificial\\_Intelligence\\_Systems](https://www.researchgate.net/publication/373496346_ChatGPT_A_Case_Study_on_Copyright_Challenges_for_Generative_Artificial_Intelligence_Systems)
- <https://www.barrons.com/articles/chatgpt-will-unleash-copyright-chaos-artificial-intelligence-51674780407>
- <https://www.theglobeandmail.com/amp/opinion/article-ai-programs-like-chatgpt-are-built-on-mass-copyright-infringement/>
- <https://www.ihrb.org/focus-areas/information-communication-technology/we-asked-chatgpt-about-its-impact-on-human-rights-on-business-heres-what-it-told-us>