

ARTIFICIAL INTELLIGENCE AND COPYRIGHTS: A QUAGMIRE WAITING TO UNRAVEL

Abstract

This PoV explores the complex landscape of intellectual property rights in the context of AI-generated content. There are some key challenges and implications surrounding copyright ownership, licensing, and fair use as AI language models continue to evolve, urging for a balanced approach to foster innovation while protecting creative works.



With the advent of generative AI, artificial intelligence to our doorsteps. AI is not a new concept; however, generative AI's biggest differentiator is generative pre-trained transformer which generates realistic outputs based on its training data set. This has given AI the human-like characteristic of reasoning and logic. Generative AI is a category of AI techniques and models that create outputs learnt from terabytes of information fed into it.

Intricate AI models utilize neural networks that recognize underlying connections in data sets through a process quite similar to how neurons relay information in our brains.

OpenAI has been leading the charge with the development of ChatGPT, Dall.E 2, Whisper, and Codex. ChatGPT is a language model that was trained on predefined algorithms and a vast amount of data. Large language models have been a game changer in the industry; they have been able to improve their performance using

reinforced learning with human feedback as the tool is publicly available. As more people use these language models, they will improve their performance through their feedback.

However, these refinements and humanlike characteristics have already led to several unique legal, moral, and ethical issues. Copyright, data ownership, and consent have been the areas of greatest debate and concern with respect to the sources used and the works produced by these tools.

This paper seeks to highlight some of the copyright issues and concerns being raised by both intellectual property (IP) owners and AI creators. Before we delve into them, here are some commonly used terms:

1. Artificial intelligence refers to the enablement of a computer system to perform cognitive functions like a human being; these include perceiving visual data, speech recognition, and decision making, through the process of machine learning.

2. Corpus refers to a collection of written or spoken texts that have been compiled and organized for the purpose of linguistic analysis, natural language processing, or machine learning. The datasets contained in the corpus are key to an AI's knowledge base.

3. Datasets contained in the corpus can either be open-source information available for free on the internet; copyrighted material from books, articles, blogs; or any material normally available subject to licensing terms.

4. AI-assisted output is the output generated with human input and intervention. The output involves intellectual exercise by a person that has contributed to the conception of work.

5. AI-generated output refers to the output generated entirely by AI, by using algorithms and machine learning techniques, without any human intervention.



Copyright issues involving artificial intelligence

Issue 1: Ownership in AI-generated output

The grant of protection to the creator of an intellectual property coupled with recognition and financial incentives has pushed the boundaries of human innovation. However, with the output generated solely by AI, the issue of ownership and copyright protection is much debated throughout the world. In the absence of any specific statute, law, or guideline, this has posed challenges not only to law makers and law enforcement agencies but also to big corporates who have and continue to invest in these AI engines.

The juxtaposition of human ability to create using intellectual exercise and artificial intelligence's ability to perform tasks using machine learning is ambivalent. Humans have come a long way from the age of hunter-gatherers to the threshold of becoming an interplanetary species. We

are most certainly on our way to create machines that can replicate our minds.

Some civil law countries hold that an author's/creator's personality should be imprinted on their work, thereby denying legal protection to AI-generated work. Granting a juristic person status to an AI would consequently mean making AI competent to enter into contracts, and more importantly to sue and be sued under law. The Federal Court in Germany found a way around the limitation posed by legislation; it ruled that the named inventor in a patent application must be a natural person, but that the AI responsible for the invention can be additionally named¹. However, this approach fails to give AI tools their due credit, and makes humans liable for the output.

The United States Copyright Office has clarified that works generated by a machine without input or intervention from a person may be copyrighted;

however, it must still have to satisfy the legal requirements of originality and creativity. In *Naruto v. Slater*², popularly known as the "Monkey Selfie" case, a US court held that photographs clicked by the monkey cannot be granted copyright protection, which can only be assigned to a human and not to animals. Although there is no clear legal framework for granting copyright to works created by machines, this case will have a lot of bearing on the question of importance of human intervention.

Section 9(3) of the UK Copyright, Designs and Patents Act has a section for content generated by computers. The Act grants the status of author or first owner rights either to the programmer who wrote the code that generated the work or to the person who commissioned the work. However, the Act requires that the work be original and created by a human author to qualify for copyright protection.

¹Out-Law News article on Pinsent Masons; [German court considers AI generated inventions \(pinsentmasons.com\)](https://www.pinsentmasons.com/news/german-court-considers-ai-generated-inventions); last referred on February 13, 2023.

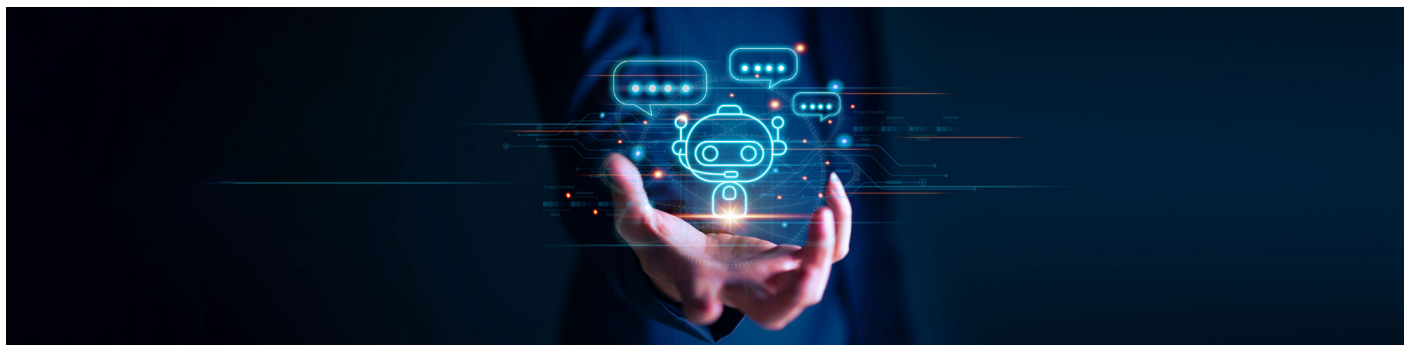
²*Naruto v. Slater*; Citaton: 888 F.3d 418 (9th Cir. 2018).

Northern District of California, San Francisco Division. It has been alleged in the lawsuit that the training data includes data in many publicly accessible GitHub repositories, which are limited by licenses. It is also alleged that Copilot, which is owned by Microsoft, plainly reproduces code that can be traced back

to open-source repositories or licensees, thus violating those licenses. The code reproduced by Copilot never includes any attribution to the underlying authors⁵.

Creators would consider this as a major infringement of their rights; they want to avoid the elimination of their professions by an AI system trained by their own work,

according to an official complaint filed at the court. Although the law is unclear on the rights of IP owners and subsequent duties of AI creators, the scale is, in our opinion, tilted towards the IP owners by the usage of due credit and licenses.



Issue 3: Copyrighted/original works in output.

AI creators accept that generative AI can create output that infringes copyrighted work. AI creators have been advocating for identical rights to humans while entertaining infringing claims against them for the outputs. Creators argue that output generated is transformative. Work based on an old work is transformative, if it uses the source in completely new or unexpected ways; this theory was first introduced by the US Supreme Court in *Campbell v. Acuff-Rose Music*. Courts in America have allowed defense of transformative uses, even when statutory factors would weigh against fair use. Since then, gaming, music⁶ and other creative industries have taken the benefit granted by judicial precedence. In *Authors Guild v. Google*⁷, Google digitized millions of books without

taking permission from authors for inclusion in their searchable database. Users, while searching the database using some keywords or phrases, small original “snippets” of text from those books would show up. The Second Circuit Court upheld this as fair usage since this is transformative in nature and didn't supplant authors' expression in their original works. Transformative use communicates something new or different from the original work and extends the utility of such work. The court also gave crutches to the argument by holding that transformative fair use served a different function from that of the original works.

Moreover, companies that have rolled out their AI generative tools are not willing to accept their liability for the output being generated. They have put the onus on the user to be careful of the copyright material

that may be generated subsequent to their inputs. Further, authors of originals works have been permitted to raise their concerns if they come across usage of their copyrighted or trademarked work.

Australian artist Kim Leutwyler accused *Lensa*, the app that uses artificial intelligence to generate self-portraits, of stealing her content⁸. She claims that in the portraits generated by *Lensa*, some of the artists have their exact styles replicated in brush strokes. Leutwyler searched for her own work among the 5.8 billion images used to train *Lensa* and found many of her portraits in the database. Authors and artists are skeptical about the work that is being generated using their copyrighted material; moreover, they are aware that copyright laws have lost pace with the speed technology is moving at.

⁵Emma Roth: Microsoft, GitHub, and OpenAI ask court to throw out AI copyright lawsuit; <https://www.theverge.com/2023/1/28/23575919/microsoft-openai-github-dismiss-copilot-ai-copyright-lawsuit>.

⁶*Campbell v. Acuff-Rose Music*, 510 US 569 (1994).

⁷804 F.3d 202 (2d Cir. 2015).

⁸Cait Kelly article in *The Guardian* on December, 11; “Australian artists accuse popular AI imaging app of stealing content, call for stricter copyright law”. [Australian artists accuse popular AI imaging app of stealing content, call for stricter copyright laws | Artificial intelligence \(AI\) | The Guardian](#).



Issue 4: Moral rights.

There are two moral rights that are generally accorded to a creator: the right of paternity and the right of integrity. The right of paternity is the right of a creator to be associated with and named for their work. Right of integrity enables a creator to protect their work from any distortion that is prejudicial to their reputation.

The Berne Convention for the Protection of Literary and Artistic Works has identified moral rights for the artistic works. Moral rights provided for the right to claim authorship of the work and the right to object to any mutilation, deformation, or other modification of or other derogatory action in relation to, the work that would be prejudicial to the author's honor or reputation⁹. United States ratified the Convention in 1989. The Virtual Artists Rights Act of 1990 (VARA) was passed by U.S Congress granting moral rights in relation to work only of visual art.

France was one of the first nations to adopt the Convention. French law has

developed complexities to protect moral rights. This was illustrated in the case *Turner Entertainment Co. v. Huston*. The legal heirs of the aggrieved brought action against Turner Company to stop the exhibition of a colorized version of one of his films, originally shot in black and white. The highest court in France held that the colorization of a black and white film created by the complainant was a violation of his moral rights as protected under French law, even though the rights of performance to the film were acquired legally. The court made the distinction between moral rights of an artist, which cannot be transferred, and the economic rights of an artist, which can be transferred. The court held that colorization was not simply an adaptation. Instead, the choice of black and white related to the "aesthetic" aspect of the work; the "interplay of black and white" created a certain "atmosphere." The court drew corollary between director's choosing black and white film and sculptor's choice to carve marble or clay as opposed to bronze to create a work

of art¹⁰.

There is a possibility that AI tools could be considered as autonomous systems depending upon the level of human intervention involved while generating the output. Policymakers are faced with the difficult question whether granting protection to the AI tool is equivalent to having a personality, like a human being. We will need to differentiate between IP rights and moral rights. Moral rights are not driven by any economic interests but are rather based on ideological values of an author, rooted in the ethical and moral principles. Since, we are not sure of the technological development that may take place within AI space, AI systems cannot be put on the same pedestal as people where they can claim moral rights on AI-generated output. Issues will also be raised with respect to the moral rights of artistes whose original work is used in the data set and undergoes transformation and deformation.

⁹[Summary of the Berne Convention for the Protection of Literary and Artistic Works \(1886\) \(wipo.int\)](#).

¹⁰*Turner Entertainment Co. v. Huston*, CA Versailles, Civ. Ch., No. 68, Roll 615/92 (1994), No. 16 10 ENT. L. REP. 3 (1995)

Conclusion

Allowing generative AI to use copyrighted material to train under the guise of fair use should not be detrimental to the copyrighted work of the creator who has undergone immense intellectual exercise to create the work. Courts can further develop jurisprudence while expanding arguments in favor of transformative material with limitation. Authors also fear that if AI uses so much of their original work, it can create a “competing substitute” available to public. For instance, picture generative tools can create work that resemble the great artist who are only

served by their great legacies. Industry has allayed fear of dilution of their work.

Our society has benefited from the protection extended by intellectual property rights which has stimulated progression and diffusion of innovative solutions based on designs, patents, engineering solutions, creative or other intangible assets. This is applicable to artificial intelligence as it is to human beings. It’s urgent for lawmakers to bridge the gap to ensure creators are incentivized, and their works protected, while recognizing the need to regulate AI

tools in terms of transparency, fairness, and accountability. Unequivocal distinction for ownership will drive innovation.

To keep up with the progress of AI, a more practical approach is required by lawmakers and experts. A separate branch of intellectual property protection needs to be enforced to cater to the needs of AI-generated output. A “sui generis” system should dispense with the issue of similar or identical outputs and not subjugate human creativity.

Authors



Bahul Kalra
Team Lead - Legal Process



Kodaneth Sunil Kumar Nair
Solution Design Head – Legal Process

* For organizations on the digital transformation journey, agility is key in responding to a rapidly changing technology and business landscape. Now more than ever, it is crucial to deliver and exceed on organizational expectations with a robust digital mindset backed by innovation. Enabling businesses to sense, learn, respond, and evolve like a living organism, will be imperative for business excellence going forward. A comprehensive, yet modular suite of services is doing exactly that. Equipping **organizations with intuitive decision-making** automatically at scale, actionable insights based on real-time solutions, anytime/anywhere experience, and in-depth data visibility across functions leading to hyper-productivity, [Live Enterprise](#) is building connected organizations that are innovating collaboratively for the future.

For more information, contact infosysbpm@infosys.com

Infosys[®]
Navigate your next

© 2023 Infosys Limited, Bengaluru, India. All Rights Reserved. Infosys believes the information in this document is accurate as of its publication date; such information is subject to change without notice. Infosys acknowledges the proprietary rights of other companies to the trademarks, product names and such other intellectual property rights mentioned in this document. Except as expressly permitted, neither this documentation nor any part of it may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, printing, photocopying, recording or otherwise, without the prior permission of Infosys Limited and/ or any named intellectual property rights holders under this document.