

# BA840 DATA ETHICS

Project Report Group A2-G7

Sricharan Mahavadi

Riris Grace Korolina

Gaurangi Agrawal

Adarsh Prajapat

Ananya Anand

Ya Chu Shu



May 24, 2024

## Exploring Copyright Laws wrt AI Generated images

### Introduction:

While it may seem like these new AI tools can conjure new material from the ether, that's not quite the case. Generative AI platforms are trained on data lakes and question snippets — billions of parameters that are constructed by software processing huge archives of images and text. The AI platforms recover patterns and relationships, which they then use to create rules, and then make judgments and predictions, when responding to a prompt.

*(GenAI Intellectual Property Problem)*

Innovations in artificial intelligence (AI) are raising new questions about how copyright law principles such as authorship, infringement, and fair use will apply to content created or used by AI. So-called “generative AI” computer programs—such as Open AI’s DALL-E and ChatGPT programs, Stability AI’s Stable Diffusion program, and Midjourney’s self-titled program—are able to generate new images, texts, and other content (or “outputs”) in response to a user’s textual prompts (or “inputs”). These generative AI programs are trained to generate such outputs partly by exposing them to large quantities of existing works such as writings, photos, paintings, and other artworks. This Legal Sidebar explores questions that courts and the U.S. Copyright Office have begun to confront regarding whether generative AI outputs may be copyrighted and how generative AI might infringe copyrights in other works.

*(GenAI Copyright Law)*

### Filed Lawsuits:

Plaintiffs have filed multiple lawsuits claiming the training process for AI programs infringed their copyrights in written and visual works. These include lawsuits by the Authors Guild and authors Paul Tremblay, Michael Chabon, Sarah Silverman, and others against OpenAI; separate lawsuits by Michael Chabon, Sarah Silverman, and others against Meta Platforms; proposed class action lawsuits against Alphabet Inc. and Stability AI and Midjourney; and a lawsuit by Getty Images against Stability AI. The Getty Images lawsuit, for instance, alleges that “Stability AI has copied at least 12 million copyrighted images from Getty Images’ websites . . . in order to train its Stable Diffusion model.” This lawsuit appears to dispute any characterization of fair use, arguing that Stable Diffusion is a commercial product, weighing against fair use under the first statutory factor, and that the program undermines the market for the original works, weighing against fair use under the fourth factor.

*(GenAI Copyright Law)*

### Who owns the copyright work and how does US government deal with it:

US: No ownership

The US Copyright Office denied protection for images created using the generative AI art platform Midjourney, but it allowed registration for the text, and the selection and arrangement

of images and text, where the applicant, Kris Kashtanova, attested to sole responsibility for those elements. In [the same guidance statement](#), the US Copyright Office also reaffirmed that protection remains for underlying original works, even where they're enhanced by technological tools.

*(Copyright Ownership)*

### **Countries having copyright laws for AI generated images:**

1. EU
2. UK
3. China

### **Consequences:**

1. Some image-hosting platforms have banned AI-generated content for fear of legal blowback. And several legal experts have cautioned generative AI tools could put companies at risk if they were to unwittingly incorporate copyrighted content generated by the tools into any of the products they sell.

*(Legal Cases)*

2. Businesses that were solely dependent on AI suffered from it. Some had to shut down.

### **Is it Ethical ?**

- **According to a Consequentialist:** A consequentialist, who judges the morality of an action based on its outcomes, might argue that if AI-generated images promote creativity, increase access to art, and do not harm anyone, then they are ethical. They would focus on the positive impacts, such as democratizing art creation and providing new tools for expression. However, if the use of AI-generated images leads to negative consequences like copyright infringement or economic harm to artists, a consequentialist might view it as unethical.
- **According to Deontologist:** A deontologist, who bases ethics on adherence to rules or duties, might argue that the ethicality of AI-generated images depends on whether they respect the moral rights of creators and adhere to copyright laws. If AI-generated images are created without violating any creators' rights or legal statutes, a deontologist might consider them ethical. Conversely, if they are produced by disregarding the rights of human creators or by using copyrighted material without permission, a deontologist would likely view them as unethical.

### **Projection for copyright law**

The projection of the impact of law implementation to protect the copyright of images used by AI for the stakeholders and artists.

#### **1. Stakeholders**

The implementation of the copyright law may impact for some of the stakeholders, they are:

**a. AI company**

- License cost: AI companies would need to pay the license to use the copyrighted images. This could affect the pricing of AI products and services, thereby impacting the profitability of the companies.
- Data access: If copyright laws are implemented, it may reduce access to some sources of images. This may affect the barrier to AI image-generated improvement.
- Legal risk: AI companies need to assess the legal risk and ensure compliance with copyright laws. This requires legal teams to monitor and manage the use of copyrighted images, thereby increasing the overhead cost.
- Innovation slowdown: As access to some sources would be limited, the innovation of AI-generated images would be slow down.

**b. Business sectors**

- Some businesses, such as marketing, advertising, and content creator, which rely on AI to generate their advertisement, may impacted by the law. They might face legal risks if they use the copyrighted image from an AI that does not have the proper license, and they may have to pay more because AI companies are likely to increase costs for users. Consequently, this may lead to an increase in production costs for these companies.

**c. Legal and compliance sectors**

- Demand of services: Legal and compliance sectors, such as law firms, intellectual property attorneys, and legal consultants, will experience increased demand from both the artist and AI companies regarding compliance with copyright laws.
- Dispute resolution: There would be the increasing in disputes cases and litigation related to copyright infringement.

**d. Governments**

- As the policymakers, the Government needs to balance the law for various stakeholders, such as AI companies, consumers, and artists, to create fair use and effective copyright protections.

**2. Artist**

**a. Legal protection**

Artists will receive legal protection for their work from unauthorized users, reducing the infringement of their art creation.

**b. Compensation**

Artists would benefit from the licensing fees paid by AI companies for the use of their work, thereby providing financial incentives for the artists.

**c. Market changes**

There will be an increase in human-created art due to the copyright protection law, as artists feel more valued and can increase their income from these licenses.

### **3. Individual Consumers/Users**

#### **1. Cost of product**

Consumers might face the increase in AI services fees due to the increased cost of licensing images. This may affect the number of AI users, potentially leading to a decrease in the number of users.

#### **2. Access to content**

Consumers may have limited access to content due to the availability of content access by AI becomes restricted or regulated. If the users want to access the images, they may have to pay more for the services.

#### **3. Creativity**

Consumer creativity in art might also decrease because they have limited references from AI to generate their art.

### **Sources of Data for AI Image Generators**

AI image generators use data from various sources to create their images. Some common sources include:

- Publicly available image datasets: These datasets contain large collections of images from diverse sources and can be used for training AI image generators.
- Online image repositories: AI image generators can scrape images from websites, social media platforms, and other online sources for training purposes.
- Custom datasets: Some AI image generators allow users to upload their own datasets or use specific images to train the generator, creating unique and personalized images.

*(Sources)*

### **Solutions:**

As the world is moving towards Artificial Generated Intelligence (AGI) anything that is uploaded or digitalized online is not protected. At the same time, OpenAI is pursuing [content licensing deals](#) with publishers and platforms and has [said](#) that it would "be impossible to train today's leading AI models without using copyrighted materials." According to statistics:

- More than **15 billion images** created using text-to-image algorithms since last year. To put this in perspective, it took photographers 150 years, from the first photograph taken in 1826 until 1975, [to reach the 15 billion mark](#).
- Since the launch of DALL-E-2, people are creating an average of **34 million images per day**.

- The fastest-growing product is Adobe Firefly, the suite of AI algorithms built into Adobe Photoshop. It reached [1 billion images](#) created in just three months since its launch.
- Midjourney has [15 million users](#), the largest user base of any image generation platform for which we have publicly available statistics. Compare: Adobe Creative Cloud, which consists of Adobe Photoshop and other graphic design and video editing software, including the generative AI tool Adobe Firefly, has 30 million users, according to [Prodesigntools](#).
- **Approximately 80% of the images (i.e. 12.590 billion)** were created using models, services, platforms, and applications based on Stable Diffusion, which is open source.

*(The Street Article on Open AI & ImagesCreated using AI)*

To resolve this issue we suggest a solution for a problem. To amend the copyright laws. Current copyright laws are not in regards to the AI Generated images as it does not have any human intervention. So there is no such law to protect the work of authors. If we compare with other countries Japan has allowed to use the copyright work without the permission of the owner if it is used to train the models for the purpose of AI not leisure. According to the copyright laws there has to have a human intervention incase of issuing a copyright.

In fiscal 2023, the Office issued more than 441,500 registrations and collected more than \$37 million in registration application fees.

*(Copyright Official Site)*

As we consider where we are now and where we want to go, we have identified four overarching goals: copyright for all, continuous development, impartial expertise, and enhanced use of data. The Office has long been engaged in furthering these goals in some form, in alignment with those of the Library of Congress, but we have broadened their scope and sharpened their focus as we look to the future.

What if each artist is given a royalty on the copyright image being used by the company itself.

**Advance against royalties** = an upfront, lump sum payment made to the copyright owner, which is typically credited against future royalty payments

Copyright owners can sometimes bargain for a guaranteed minimum royalty payment.

*(Copyright Agreement)*

One more potential solution is to implement a system of digital footprinting for AI-generated content. This would involve assigning a unique digital identifier or "watermark" to each piece of AI-created content.

The digital footprint could contain key information such as:

- The source datasets and models used

- Licensing and attribution details for any copyrighted material
- Identification of the human prompts that led to the output
- Timestamp and metadata about the creation process

By embedding this digital footprint, it would become easier to trace the provenance and legitimate use of AI-generated content. This could help differentiate lawful use of copyrighted training data from outright infringement.

The footprinting could be implemented through technical standards or regulations overseen by a governing body. AI companies could be required to add the footprint as part of their content generation process.

Digital footprinting has the potential to balance the needs of AI innovators, content creators, and end-users. It could enable new licensing frameworks where artists are automatically compensated when their work is used to train AI. Overall, it offers a path to protect intellectual property rights in the age of generative AI.

## **Conclusion**

The rise of generative AI tools like DALL-E, Stable Diffusion, and Midjourney has raised significant copyright challenges. Lawsuits have been filed alleging that the training process for these AI models infringes on copyrights in the underlying images and text. Existing copyright laws do not clearly address the ownership of AI-generated content.

While some countries like Japan have updated their copyright laws to allow use of copyrighted material for AI training, the US has not done so yet. This leaves a legal gray area that could hamper innovation in AI while also failing to adequately protect the rights of artists and creators.

To balance the competing interests, potential solutions include updating copyright laws to provide a royalty system where artists are compensated when their work is used to train AI, as well as clearer guidelines around fair use exceptions for AI training. Proactive engagement between AI companies, artists, and policymakers will be crucial in developing a sustainable and equitable framework for AI-generated content and intellectual property rights.

## References:

## Copyright Bill Passed:

<https://natlawreview.com/article/generative-ai-copyright-disclosure-act-2024-balancing-innovation-and-ip-rights>

GenAI Copyright Law: <https://crsreports.congress.gov/product/pdf/LSB/LSB10922>

## GenAI Intellectual Property Problem:

<https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem>

### Legal Cases:

<https://techcrunch.com/2023/01/27/the-current-legal-cases-against-generative-ai-are-just-the-beginning/>

United States Copyright Office: <https://www.copyright.gov/docs/zarya-of-the-dawn.pdf>

## The Street Article on Open AI:

<https://www.thestreet.com/technology/copyright-lawsuits-against-openai-microsoft-chatgpt>

Copyright Ownership:

<https://www.cooley.com/news/insight/2024/2024-01-29-copyright-ownership-of-generative-ai-outputs-varies-around-the-world>

### Copyright Laws in various Countries:

- <https://www.cooley.com/news/insight/2024/2024-01-29-copyright-ownership-of-generative-ai-outputs-varies-around-the-world>
- <https://medium.com/@inyoungcheong/ai-art-the-us-uk-and-china-take-different-approaches-to-copyright-protection-9678e93a795d>

Copyright Act of Japan:

[https://www.wipo.int/about-ip/en/artificial\\_intelligence/strategy-search.jsp?territory\\_id=87&policy\\_id=24](https://www.wipo.int/about-ip/en/artificial_intelligence/strategy-search.jsp?territory_id=87&policy_id=24)

Sources:

<https://autogpt.net/how-are-ai-images-even-generated/#:~:text=AI%20image%20generators%20use%20data,for%20training%20AI%20image%20generators.>

## Japan's GenAI Policy:

[https://insights.manageengine.com/artificial-intelligence/the-us-should-look-at-japans-unique-approach-to-generative-ai-copyright-law/#:~:text=Unlike%20the%20U.S.%20and%20E.U.,AI%20\(gen%20AI\)%20models.](https://insights.manageengine.com/artificial-intelligence/the-us-should-look-at-japans-unique-approach-to-generative-ai-copyright-law/#:~:text=Unlike%20the%20U.S.%20and%20E.U.,AI%20(gen%20AI)%20models.)

### Japan's Amendment:

[https://www.wipo.int/about-ip/en/artificial\\_intelligence/strategy-search.jsp?territory\\_id=87&policy\\_id=24](https://www.wipo.int/about-ip/en/artificial_intelligence/strategy-search.jsp?territory_id=87&policy_id=24)

Copyright Question in different countries:

<https://news.bloomberglaw.com/ip-law/ais-thorny-copyright-questions-create-international-patchwork>

Copyright Official Site: <https://www.copyright.gov/about/>

Copyright Agreement:

<https://www.justia.com/intellectual-property/copyright/copyright-licensing/international-license-agreements/>

ImagesCreated using AI: <https://journal.everypixel.com/ai-image-statistics>

<https://crsreports.congress.gov/product/pdf/LSB/LSB10922#:~:text=URL%3A%20https%3A%2F%2Fcrsreports.congress.gov%2Fproduct%2Fpdf%2F LSB%2F LSB10922%0AVisible%3A%200%25%20>