31 West 34th Street · 8th floor
New York, NY 10001

(212) 791-3400
(212) 791-0333 (fax)

graphicartistsguild.org
info@graphicartistsguild.org

October 30, 2023

LIBRARY OF CONGRESS
Copyright Office

Docket No. 2023–6
(Federal Register / Vol. 88, No. 16)

Notice of Inquiry:
Artificial Intelligence and Copyright

Comments Submitted by the Graphic Artists Guild, Inc.

The Graphic Artists Guild welcomes the opportunity to submit comments in response to the Copyright Office's Notice of Inquiry on Artificial Intelligence and Copyright.

The Graphic Artists Guild is a trade association representing the interests of illustrators, designers, web developers, animators, and other visual artists. Our members are predominantly freelancers or microbusinesses who work in a tight market and rely on licensing. In fact, average income for illustrators has been stagnant for the past two decades. To demonstrate their skills and market their work, they must display their work online and publish it to social media. Their work is vulnerable to copyright infringement, which has a direct and significant impact on their income streams. In

The emergence of AI image generators has been met by both interest and consternation. Some visual artists have adopted generative AI as a tool in their creative process; others are deeply troubled by the technology. The major concerns are the use of visual artists' copyrighted images, without consent or compensation, in the training datasets; competition from the vast quantities of visual content created by AI image generators; and the generation of visual works which resemble their original works in the outputs; and the ability of users to ape a visual artist's unique style.

While backers of AI image generators state that the technology permits the democratization of artistic creation, we question the validity of this argument. Lowering the barrier to entering a creative profession is irrelevant if it is difficult, if not impossible, to achieve a livable wage in that profession. Comparisons are made to previous technological advancements, such as the development of photography, as a way of minimizing concern with generative AI technology. However, no previous technological advancement has enabled such volume of often high-quality images to be easily and quickly generated.

To protect the livelihoods of professional graphic artists, it is critical that copyright remain true to its purpose: to incentivize creators by assigning to them the exclusive rights to their works. Policies around generative AI must consider first and foremost the interests of the human creators, without whom this technology would not exist.

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

We appreciate the diligent effort on the part of the Copyright Office to hear the concerns of all stakeholders regarding generative AI. We request that the Office issue a study based on the results of this inquiry. We recognize that this poses a challenge because of the length of the inquiry and the high volume of responses the Office has received. However, a report, or reports on individual sections of this notice, would be very informative.

1. **As described above, generative AI systems have the ability to produce material that would be copyrightable if it were created by a human author. What are your views on the potential benefits and risks of this technology? How is the use of this technology currently affecting or likely to affect creators, copyright owners, technology developers, researchers, and the public?**

Illustrators and designers are early adopters of new technologies. Many were intrigued by the rollout of generative AI; some had already dabbled in early iterations of image generators.[1] However, many visual artists were alarmed by the emergence of a technology the permitted the generation of millions of images with no safeguards to protect the livelihoods of visual artists. The realization that this technology was developed from datasets of billions of images, many under copyright and scraped without the knowledge of consent of the visual artists, fueled outrage. The result is rift within the community between visual artists who have adopted generative AI, and visual artists who believe that AI image generators are based on an unethical and damaging business model.

Within the Guild, we have heard from both camps. Most graphic artists we've communicated with report that, if they use AI image generators, it is primarily for ideation. They are concerned about the ethical questions around the development of the technology. They are also reluctant to use AI image generators even in part to execute client work. They worry that the questionable copyrightability even AI derivative works would have negative consequences for their clients who seek exclusive licenses to those works.

Some graphic artists we spoke with do incorporate AI image generators into their workflow as a tool. In these instances, they're using AI image generators to produce minor elements of a larger work, for example, to create a background for an illustration. This permits them to license these larger works, since most of the illustration is original and copyrightable. These artists report that they can take on larger projects that would otherwise have been out of scope for their small studio, or that they were able to cut back on the number of freelancers they hired to help with large projects.

We've only heard directly from one artist whose career is as an AI illustrator. This artist creates derivative works using AI image generators, by inputting their own images into

[1] In his keynote address at the AI/ML Media Summit, illustrator Steven Zapata describes his interest in digital art and his early investigations int generative AI https://www.youtube.com/watch?v=h4ax8twEwgw

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

a training dataset to develop their own fine-tuning models, prompting, and curating successive rounds of outputs, developing the outputs off-platform in image editing software, or writing code for finetuning. They have told us they have a cognitive disability that prevents them from being able to finish artwork to as high a degree as AI image generators enable them to do.

From these reported experiences, the benefits AI image generators provide are easy and wide-ranging ideation, the ability to take on larger projects with fewer staff, the ability to generate work quickly, and mitigating physical and cognitive limitations. AI image generators also present visual artists with fresh the licensing opportunities. As more AI-generated images flood the public space, and as fine-tuning permits smaller datasets to be used to generate distinctive imagery, we envision that a licensing market for curated imagery will become increasingly important. However, such licensing opportunities would be curtailed by text and data mining exceptions to copyright that permit AI image generators to bypass licensing.

On the other hand, the negative impacts of AI image generators are significant. The issues raised by generative AI are the consequence of how the technology was developed and rolled out as a business model. The foremost concern is that the images used in the original dataset – the 5 billion+ LAION dataset which provided the fodder for the diffusion model many of the AI image generators are built on – were scraped with no thought for copyright status, privacy concerns, or origin. This is an outright copyright infringement; data (including creative works) can be and are licensed for use in machine learning. Visual artists were deprived of their right to license, or to refuse to license, their imagery for use in the original training dataset. This sets a terrible precedent for future licensing opportunities.

A related issue with AI image generators is the issue of the copying of visual artists' styles. The pretrained model contained millions of images by visual artists, labeled with their names. This permits users to enter the name of a visual artist into text prompts to generate images "in the style of" that visual artist. While some AI image generators removed the ability to use visual artists' names as text prompts, the problem persists. AI image generator users can develop fine-tuning models by uploading as few as 20 works by an illustrator. Fine-tuned models developed from the work of lesser-known visual artists are widely shared and even sold on AI generative marketplaces[2]. These models plagiarize visual artists' unique styles and are more likely to produce outputs which replicate the original images.

[2] In this example, user Laura_De_Martin generated images on a model fine-tuned on the artist Conrad Roset; his name is included in the set of text prompts provided. https://civitai.com/images/3053701?modelVersionId=191091&prioritizedUserIds=2303844&period=AllTime&sort=Most+Reactions&limit=20 This model has been downloaded over 2,000 times from the Civitai platform, a platform that purports to "…offer an environment where users can upload, share, and discover custom models, each trained on distinct datasets. With the aid of AI media software, these models can be leveraged as innovative tools for crafting your own unique creations." https://civitai.com/content/guides/what-is-civitai

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

This is a serious concern: visual artists spend years developing their unique styles. When thousands of images in a visual artist's style are released into the public space, the visual artist no longer has a unique presence; an entire oeuvre of a visual artist's work is drowned out by AI copycats[3]. We have even heard reports of anti-AI artists being retaliated against by having their works used for fine-tuning[4]. Once a model has been fine-tuned on a visual artist's unique style, that model can not only be used endlessly to generate images in that visual artist's style. It can also be used to generate harmful imagery – imagery with racists, violent, homophobic, or sexist messaging – that harms the reputation of the original visual artist[5].

AI image generators have implemented few and inadequate safeguards to protect visual artists from infringing activity and practices (such as style copying) which harm their reputations and ability to earn a living. While some AI image generators have limited the ability of users to prompt the names of visual artists, fine-tuned models as described above can circumvent that guardrail. There are also no measures taken that can effectively prevent users from uploading the AI image generators artwork they do not have permission to use. Once an image has been used for training, the information gleaned from the image cannot be easily disgorged. While the replication of a visual artist's work in an AI output occurs infrequently, it does happen.

We are concerned that, short of legislation and legal precedent set by any of the many lawsuits brought against generative AI companies, AI image generators will have little incentive to implement meaningful safeguards for visual artists, be transparent about the sources of the images used in datasets, or to abide by copyright law. Generative AI platforms dodge accountability for copyright infringement by utilizing technology developed from datasets scraped by academic research institutions. These academic institutions operate in countries or regions with broad text and datamining exceptions to copyright or use fair use arguments based in part on their non-profit status to excuse the scraping of copyrighted works. Additionally, all entities are not forthcoming on what images are used in image datasets, or the provenance of those images. As a result, AI image generators can distance themselves from accountability for copyright infringement.

---

[3] Greg Rutkowski summed up the issue in September 2022: "It's been just a month. What about in a year? I probably won't be able to find my work out there because [the internet] will be flooded with AI art,". This artist is dominating AI-generated art. And he's not happy about it, MIT Technology Review, September 16, 2022. https://www.technologyreview.com/2022/09/16/1059598/this-artist-is-dominating-ai-generated-art-and-hes-not-happy-about-it/

[4] Dr. Ben Zhao described how Japanese artists were being retaliated against by having their images uploaded for fine-tuning in a Graphic Artists Guild webinar, Learn How to Protect Your Artistic Style from Mimicry. Air date September 20, 2022 (statement begins at 53 minutes in) https://graphicartistsguild.org/product/protect-your-style-from-mimicry/

[5] "The Alt-Right Manipulated My Art. Then A.I. Claimed It." Sarah Anderson, *The New York Times*, December 31, 2022. https://www.nytimes.com/2022/12/31/opinion/sarah-andersen-how-algorithim-took-my-work.html

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

There are reports of job loss resulting from AI image generators, predominantly in freelancer and entry- or low-level jobs.[6] The job loss currently appears to be occurring primarily in the entertainment and concept art sectors, where it can be significant. One concept artist confided to us that, in six months, he lost $20,000 in income. However, it is difficult to gather evidence other than anecdotal. Many visual artists working in the most affected sectors are afraid to go public with their experiences, fearing a backlash from studios and agencies. Others sign non-disclosure agreements which prohibit them from sharing information about which projects they've been hired to work on and their compensation.

2. **Does the increasing use or distribution of AI-generated material raise any unique issues for your sector or industry as compared to other copyright stakeholders?**

Because of the ease with which visual content is generated by AI, and because skilled text-prompters can generate high-quality images, visual artists have been hit early and hard by the emergence of generative AI. Already in 2022, companies and publishers were investigating ways to integrate AI into their production cycles.[7] Today, the sheer volume of visual content generated by AI – estimated at 35 million per day[8] – is drowning out the original visual works created by visual artists. This reduces the online presence of visual artists – critical for gaining new clients and commissions – and competes with them for work opportunities.

Additionally, as we described above, damage is done when this AI generated content mimics the unique styles of visual artists. This is an issue that is particularly a concern for visual artists, for both the loss of presence and opportunity, and for the reputational damage that is done when poor quality images, or images promoting reprehensible beliefs, are associated with the visual artist.

The visual content output by generative AI will degrade as more and more AI-generated imagery saturate the public space[9]. This creates a need for licensing of high-quality images for curated datasets. However, AI image generator platforms have indicated that licensing from individual visual artists is difficult if not impossible, considering the high volume of images they require. This puts individual visual artists at a disadvantage

[6] "Workers are worried about AI taking their jobs. Artists say it's already happening." Tom Carter, Business Insider, October 1, 2023. https://www.businessinsider.com/ai-taking-jobs-fears-artists-say-already-happening-2023-10
[7] "The World's Smartest Artificial Intelligence Just Made Its First Magazine Cover" Gloria Liu, Cosmopolitan, June 22, 2021. https://www.cosmopolitan.com/lifestyle/a40314356/dall-e-2-artificial-intelligence-cover/
[8] According to Everypixel, AI image generators have output 15 billion images in the past years, and 34 million images are being generated daily. "AI Has Already Created As Many Images As Photographers Have Taken in 150 Years. Statistics for 2023" Everypixel Insights, August 15, 2023. https://journal.everypixel.com/ai-image-statistics
[9] "Combining Generative Artificial Intelligence (AI) and the Internet: Heading towards Evolution or Degradation?" Gonzalo Martínez, Ruiz de Arcaute, Lauren Watson, arXiv:2303.01255v1, February 17, 2023 https://www.semanticscholar.org/paper/Combining-Generative-Artificial-Intelligence-(AI)-Arcaute-Watson/a82d3924f488d74af91ba31ff1cf80a6c07ee206

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

in competing for licensing agreements against entities with large libraries of licensed imagery, such as publishing houses, media companies, and stock image agencies.

Collective licensing would provide a means by which individual visual artists would be remunerated for the use of their images in AI datasets by evening the playing field. However, we are only recommending collective licensing program for visual artists. There can be no one-fits-all collective licensing approach for all creative industries. Each creative industry has unique licensing considerations, work conditions, and issues within the generative AI environment, all of which would need to be addressed individually by a collective licensing program.

(Skipping Question 3)

4. **Are there any statutory or regulatory approaches that have been adopted or are under consideration in other countries that relate to copyright and AI that should be considered or avoided in the United States? How important a factor is international consistency in this area across borders?**

The broad text and datamining exceptions to copyright which have been implemented in some countries and in the EU[10], and are being considered elsewhere, are particularly concerning. Such exceptions undermine visual artists by creating such broad exceptions to copyright as to make it virtually impossible for visual artists to assert the rights to their work. The United States should not consider adopting such TDM exceptions.

The EU TDM extends broad TDM exceptions for dataset developed by academic and research institutes. Research for generative AI is hand-in-glove with the AI generative companies which implement the technology academic institutions develop. For-profit companies often fund the institutions developing AI/ML technology[11]. A TDM exception for research is not required to ensure the continued development of generative AI technology since there is already significant private sector support.

The EU TDM exceptions include provisions permitting the use of creative works in datasets for commercial purposes, provided an opt-out process is made available for copyright holders who do not wish to have their works used. This Is wholly inadequate to protect visual artists. First of all, an opt-out process would require visual artists to patrol for the emergence of datasets, and diligently opt-out. Secondly, by the time a visual artist becomes aware of a dataset they should opt out of, their works may very well have already been scraped and used in the training process. Such TDM

[10] "Text and Data Mining Exceptions Around the Globe", June 2023, Reed Smith LLP. https://www.reedsmith.com/en/perspectives/ai-in-entertainment-and-media/2023/06/text-and-data-mining-around-the-globe

[11] "This startup is setting a DALL-E 2-like AI free, consequences be damned", Kyle Wiggers, August 12, 2022 TechCrunch. https://techcrunch.com/2022/08/12/a-startup-wants-to-democratize-the-tech-behind-dall-e-2-consequences-be-damned/

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

exceptions harm the market for visual artists to license their images for use in datasets. We also believe that such TDM exceptions may very well be a violation of the Berne Convention, and joined other creators' organizations in raising these concerns.[12]

5. **Is new legislation warranted to address copyright or related issues with generative AI? If so, what should it entail? Specific proposals and legislative text are not necessary, but the Office welcomes any proposals or text for review.**

New legislation is required to address copyright and other issues with generative AI. Visual artists are left largely vulnerable to the continued infringement of their images in datasets and for fine-tuning. While there are a number of court cases addressing the copyright infringement of visual art in generative AI, these cases will take time to conclude and may not result in any clear precedent that will benefit visual artists. In the meantime, generative AI continues to develop apace. Legislation is required to set up guard rails to protect visual artists.

- Amendment of §1202 of the copyright statute: To prevent the stripping of metadata and other copyright management information from images, a change to §1202 should be made so that a copyright owner would only need to prove that their CMI was removed or altered knowingly or recklessly. Currently the statute requires that the CMI must have been removed or altered with the knowledge that doing so would "induce, enable, facilitate, or conceal infringement". This has created a loophole which has permitted companies and entities to remove the metadata containing CMI with little justification. It has become crucial to maintain metadata, since metadata can be used to determine whether a work has been ingested by an AI system, indicate the provenance of wholly AI or AI derivative works, and indicate whether an image may be ingested into a training dataset.

- Legislation to prevent the use of a visual artist's name, images, or other identifying materials utilized in training dataset: To prevent the creation of models which enable the production of images which closely mimic the unique styles of visual artists, it's necessary to prevent visual artists' works, name, and other identifying materials from being used without permission in generative AI systems. Such legislation should not include broad exceptions that render the protection meaningless. While a sui generis solution for protecting visual artists' styles has been proposed, we're concerned that such an IP protection may have unintended consequences. Adobe has proposed such legislation, the FAIR Act.[13]

- Legislation which addresses collective licensing: There are two paths to implementing collective licensing which we address further along in our comments.

[12] "Appeal for action on violations of the Berne Convention by the application to copying of creative works for AI development of the TDM exception in Articles 3 and 4 of the 2019 EU Directive on Copyright," July 2023. https://nwu.org/wp-content/uploads/2023/07/creators-coalition-AI-exceptions.pdf

[13] "The FAIR Act: A new right to protect artists in the age of AI" Dana Rao, Adobe Blog, September 12, 2023. https://blog.adobe.com/en/publish/2023/09/12/fair-act-to-protect-artists-in-age-of-ai

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

For collective licensing to be implemented, some legislation would be required depending on the collective licensing scheme:

- Antitrust legislation: express antitrust legislation would be required to permit visual artists to collectively bargain in a voluntary collective licensing scheme
- Extended collective licensing: in this scenario, a collective management organization (CMO) would be authorized by legislation to negotiate for licensing fees on behalf of an entire class of creators (in this case, visual visual artists).

- Opt-in for use of images in datasets: AI image generators should not be able to simple use images in datasets without consent or compensation. This use is a copyright infringement. While some companies such Open AI[14] have implemented opt-out measures, a wide adoption of opt-out would place an undue burden on visual artists. Legislation that requires AI image generators to abide by an opt-in system – one in which they could only use images which are proactively labeled for use in generative AI – would remove that burden from visual artists.

- Legislation requiring transparency on the part of AI image generators, including record-keeping of what works are ingested for training purposes. Barring any privacy concerns with images which have been properly licensed, the record should be made publicly available.

- Legislation requiring the labeling of purely AI generated works: Bad actors can edit the metadata on AI generated works in order to pass them off as copyrighted works. Labeling of purely AI generated works should be considered.

TRAINING

6. What kinds of copyright-protected training materials are used to train AI models, and how are those materials collected and curated?

The LIAON Database which was used to develop the diffusion model powering many of the AI image generators indiscriminately scraped over 5 billion images from online sources, including visual artists' portfolios, piracy websites, stock image agencies, and visual artists' forums.

Some AI image generators limit themselves to public domain and licensed works. These image generators include Bria, Adobe Firefly, and Shutterstock.

[14] OpenAI's "Artist and Creative Content Owner Opt Out" form. https://share.hsforms.com/1_OuT5tfFSpic89PqN6r1CQ4sk30

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

Images are copied into datasets by individual users developing fine-tuning models based on individual visual artist's works, or on curated collections. Thousands of such users share their fine-tuning models on platforms such as Civitai.

Meta is currently developing a large language model to challenge OpenAI's ChatGTP[15]. There are no reports that Meta is developing an AI image generator. However, Meta has implemented data deletion process for users to have their third party uploaded images removed from their social media platforms. That, and the fact that Meta claims rights to images uploaded to its platforms[16] raises questions about whether Meta also intends to develop an AI image generator platform.

(Skipping questions 6.1-7.4)

8. **Under what circumstances would the unauthorized use of copyrighted works to train AI models constitute fair use? Please discuss any case law you believe relevant to this question.**

It is difficult to square the scraping of images into datasets to train AI models with fair use. If the ingestion of images ended at providing data about the images that could be used for machine learning, with no further development beyond the academic research, then possibly that would be a fair use. The research into artificial intelligence and machine learning led directly to the development of generative AI platforms. These platforms generate visual content that mimics the images in the training dataset. The AI output competes with, if not replaces, the original images. Additionally, the market for licensing works for training datasets is circumvented.

Regardless of whether the AI image generator produces content that replicates an image in the dataset, an infringement has occurred when an image is copied into a dataset. The generative AI platforms argue that copying of images does not occur, and that only "non-expressive" elements are extracted from the images in the dataset. Without full transparency on how the images are utilized, it's difficult to gauge if this is an accurate characterization. However, it is difficult to believe that copying does not occur. Visual works must be accurately tagged before machine learning occurs for it to be effective. Copying of the image must be done for this step in the machine learning

[15] "Meta is developing a new, more powerful AI system, Wall Street Journal reports" Reuters, September 11, 2023; https://www.reuters.com/technology/meta-is-developing-new-more-powerful-ai-system-wsj-2023-09-10/

[16] From Facebooks's terms of use: "Specifically, when you share, post, or upload content that is covered by intellectual property rights on or in connection with our Products, you grant us a non-exclusive, transferable, sub-licensable, royalty-free, and worldwide license to host, use, distribute, modify, run, copy, publicly perform or display, translate, and create derivative works of your content (consistent with your privacy and application settings." https://m.facebook.com/legal/terms#

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

process to occur. Lastly, in comments made to the USPTO, OpenAI CEO Sam Altman stated that "copying must occur" [17] in generative AI.

Considering the use of images in AI training datasets along the four fair use factors falls short of establishing that the use is fair.

- Under the first factor, the purpose and character of the use, the argument is made that the "non-expressive" copying of the data which defines an image is in and of itself a "transformative" use because the use results in the development of an algorithm. This doesn't make sense since every line, every shape, the choice of color palette, the arrangement of shapes, etc. in the copied image is creative expression on the part of the visual artist. The visual content output by AI image generators mimic the creative expression of the original images in the dataset. More importantly, the visual content output by AI image generators serves the same purpose of the images used in the training dataset. This argues against considering the use of the original images to be transformative.

- Under the second factor, the nature of the copyrighted work, fair use is favored when the original work is factual rather than creative, and if that work was published. AI image generators require highly expressive, not factual, images in the training dataset in order to maintain the quality of the visual content output. Additionally, unpublished images displayed in public galleries and visual artists' portfolios were indiscriminately copied into the training dataset.

- Under the third factor, the portion and substantiality of the image used is considered, with fair use being favored when less of an image, or less critical areas of an image are used. For AI image generators, entire images are ingested into training datasets. In fact, the training wouldn't be accomplished effectively with partial images.

- Under the fourth factor, the effect of the use to the potential market of copyright value is considered. The market for images is affected in two ways when they are ingested into training datasets. First, the market for licensing the images for use in training datasets has been bypassed, and the visual artist has been denied the possibility of licensing income. Secondly, the images output by AI image generators serve the same purpose as the original ingested images and compete with those in the market.

[17] "Comment Regarding Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation" Comments submitted by OpenAI before the USPTO and Dept. of Commerce, Docket No. PTO–C–2019–0038 https://www.uspto.gov/sites/default/files/documents/OpenAI_RFC-84-FR-58141.pdf

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

8.1 In light of the Supreme Court's recent decisions in Google v. Oracle America and Andy Warhol Foundation v. Goldsmith, how should the ''purpose and character'' of the use of copyrighted works to train an AI model be evaluated? What is the relevant use to be analyzed? Do different stages of training, such as pre-training and fine-tuning, raise different considerations under the first fair use factor?

The Warhol decision confirmed that transformative use of a copyrighted work does not outweigh the other factors which must be considered when weighing whether a use is fair or not. Notwithstanding, as stated earlier, we have difficulty understanding how the use of copyrighted images in training datasets, at any stage of training, is "transformative."

8.2 How should the analysis apply to entities that collect and distribute copyrighted material for training but may not themselves engage in the training?

If copyrighted images are being collected and distributed for any reason, then whoever is engaging in that activity may be liable for copyright infringement. The determination on whether an infringement has occurred hinges on the facts: were copies of the images made, and do fair use defenses apply, and were the exclusive rights of the copyright holders infringed. Collecting images for use in AI training datasets can be an infringing activity. The fact that an individual or entity does not directly engage in the training does not absolve them from the legal consequences of any actions they take along the generative AI process.

8.3 The use of copyrighted materials in a training dataset or to train generative AI models may be done for noncommercial or research purposes. How should the fair use analysis apply if AI models or datasets are later adapted for use of a commercial nature? Does it make a difference if funding for these noncommercial or research uses is provided by for-profit developers of AI systems?

The fact that there are "non-commercial or research" purposes of collecting images for training datasets should not determine out of hand that this is a fair use. There is a strong relationship between the non-commercial, research-oriented development of generative AI technologies and the emergence of lucrative generative AI image market, valued at $302 million in 2022 and projected to grow at a rate of 17.5%[18]. When research is funded by industry, the "non-commercial" development of the technology is merely a waystation in the inevitable emergence of highly valued companies.[19]

---

[18] "AI Image Generator Market Size, Share & Trends Analysis Report By Component (Software, Services), By End-user (Media & Entertainment, Healthcare), By Region, And Segment Forecasts, 2023 - 2030", Grand View Research. https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-image-generator-market-report

[19] "How Midjourney Became a Billion-Dollar AI Art Startup in Less Than a Year", Ugo Peter, Medium, August 2, 2023. https://medium.com/@oluigbopeter/how-midjourney-became-a-billion-dollar-ai-art-startup-in-less-than-a-year-416663997f77

31 West 34th Street · 8th floor
New York, NY 10001

(212) 791-3400
(212) 791-0333 (fax)

graphicartistsguild.org
info@graphicartistsguild.org

8.4 What quantity of training materials do developers of generative AI models use for training? Does the volume of material used to train an AI model affect the fair use analysis? If so, how?

The LAION image database which powered much of the development of AI image generators consisted of over 5 billion images. New techniques permit AI machine learning to occur with smaller datasets. Datasets as low as 20 images can be used for fine-tuning. Databases of curated works – high quality images created by skilled visual artists – are becoming critical to prevent the eventual degradation of AI generated images occurring from the ingestion of mostly AI-generated images. This is a market where visual artists can license their works.

The volume of the images ingested into a training dataset should not have any bearing on a fair use analysis. Needing a high volume of images does not excuse generative AI companies from respecting the copyrights of the images they require.

(Skipping question 8.5)

9. Should copyright owners have to affirmatively consent (opt in) to the use of their works for training materials, or should they be provided with the means to object (opt out)?

Yes, under copyright law a copyright owner must affirmatively consent to the use of their work. An exception would be an extended collective licensing scheme in which images would be licensed for use in generative AI. However, an extended collective licensing scheme would require a simple, well-publicized opt-out for visual artists who do not want to have their images used in generative AI training databases.

Outside extended collective licensing, there should be no legislation which proposes an opt-out solution to permit generative AI platforms to use copyrighted works without prior consent. Such a regime would put an undue burden on visual artists, requiring them to stay abreast of generative AI platforms as they emerge to opt-out. Visual artists should not be forced to police for the potential infringement of their works by generative Ai platforms. A similar burden was created by the DMCA process.

Opt-outs implemented by companies may in fact be useless, to the point where visual artists question whether they are implemented in good faith or are mere window-dressing. For example, Meta has made available a data deletion form for users of their platforms who do not want their images uploaded to Meta's social media sites used for generative AI. The data deletion form requires visual artists to submit proof that Meta's AI systems were trained on the visual artists' work[20] – a requirement which is next to impossible to meet.

[20] "Artists Allege Meta's AI Data Deletion Request Process Is a 'Fake PR Stunt'", Kate Knibbs, October 26, 2023, Wired. https://www.wired.com/story/meta-artificial-intelligence-data-deletion

31 West 34th Street · 8th floor
New York, NY 10001

(212) 791-3400
(212) 791-0333 (fax)

graphicartistsguild.org
info@graphicartistsguild.org

9.1 Should consent of the copyright owner be required for all uses of copyrighted works to train AI models or only commercial uses?

Yes, outside a fair use defense, consent should be required for any use of copyrighted materials for training AI models.

9.2 If an ''opt out'' approach were adopted, how would that process work for a copyright owner who objected to the use of their works for training? Are there technical tools that might facilitate this process, such as a technical flag or metadata indicating that an automated service should not collect and store a work for AI training uses?

We do not support a blanket opt-out approach that would permit generative AI platforms to use images for AI training without prior consent. An opt-out approach ignores that fact that once an image has been scraped into a training dataset, any training cannot be "forgotten." The damage has been done. It's unreasonable to burden visual artists with the responsibility of proactively opt-out of every generative AI platform which emerges.

The only exception to this would be an extended collective licensing program. However, in this scenario, the visual artists would only be required to opt out once – with the CMO tasked with executing the extended collective license – and the opt out process should be simple to execute. As in the example we gave before of Meta's data deletion form, we are concerned that if opt outs were entrusted to generative AI platforms, they may not be always be executed in a fashion which is simple to use and effective.

An opt-in system should be simple for generative AI platforms to execute. For example, platforms should be able to implement means to read metadata indicating whether permission is given for an image to be used in a training dataset. However, for metadata to be leveraged adequately to indicate opt-in status, it needs to be preserved. This requires changes to §1202 prohibiting the removal of CMI.

9.3 What legal, technical, or practical obstacles are there to establishing or using such a process? Given the volume of works used in training, is it feasible to get consent in advance from copyright owners?

The argument that it is difficult to get permission in advance for the use of images in training data because of the huge volume of images required for machine learning is neither here nor there. Under copyright law, barring a fair use exception, consent is required before a copyrighted work can be used for any purpose.

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

The difficulty the AI platforms have in collecting and curating the extensive databases of images they require is not the problem of visual artists. However, the consequences of the use of unauthorized use of their images – loss of economic opportunities, loss of licensing opportunities to AI, loss of prospective clients and projects from the sheer volume of AI generated images, loss of a unique presences because of the aping visual artists' unique style – are the problems of visual artists.

9.4. If an objection is not honored, what remedies should be available? Are existing remedies for infringement appropriate or should there be a separate cause of action?

Ignoring a visual artist's objection to using their work for AI/ML training is equivalent to using a copyrighted image without consent. The existing remedies for copyright infringement should apply.

9.5. In cases where the human creator does not own the copyright—for example, because they have assigned it or because the work was made for hire— should they have a right to object to an AI model being trained on their work? If so, how would such a system work?

Under US Copyright Law, only the copyright holder has the legal means to address the unauthorized use of their work. We're not sure what legal remedies, if any, visual artists who create work-made-for-hire or assign their rights would have under our current system. Contracts should include terms addressing generative AI, for example, assigning all rights except for generative AI. We question whether contracts for work made for hire could include language that excludes use for generative AI.

10 If copyright owners' consent is required to train generative AI models, how can or should licenses be obtained?

Licenses for the use of work to train generative AI should be obtained the way licensing for any use of imagery is obtained: by negotiating with the visual artist, the visual artist's agent, or an entity empowered to negotiate on the visual artist's behalf.

10.1 Is direct voluntary licensing feasible in some or all creative sectors?

Generative AI platforms find it onerous to contact individual visual artists for direct licensing because of the sheer volume of images they required. However, there is a market for highly curated image datasets with limited images. It is entirely possible for AI image generators to engage in direct licensing with a visual artist or group of visual artists for this purpose.

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

10.2 Is a voluntary collective licensing scheme a feasible or desirable approach? Are there existing collective management organizations that are well-suited to provide those licenses, and are there legal or other impediments that would prevent those organizations from performing this role? Should Congress consider statutory or other changes, such as an antitrust exception, to facilitate negotiation of collective licenses?

A voluntary collective licensing scheme can work for visual artists. Considering the challenges facing visual artists – the difficulty bringing infringement lawsuits for images ingested into datasets, the uncertainty in how fair use defenses with be considered, the lack of negotiating power individual visual artists have, and the reluctance of AI image platforms to seek direct licenses – collective licensing may be the best option graphic artists have to be remunerated for the use of their work. Collective licensing should not deny visual artists the possibility of direct licensing; it must be voluntary.

In the United States, the American Society of Collective Rights Licensing (ASCRL) is a CMO which disburses international reprographic royalties to visual artists.

As we discussed previously, statutory changes should be considered to permit creators to collectively bargain as proposed by the Authors Guild and a coalition of creative organizations, including the Graphic Artists Guild.[21] These include amending the National Labor Relations Act to cover freelance professional creators, and creating an express antitrust exemption for professional creators.

10.3 Should Congress consider establishing a compulsory licensing regime? If so, what should such a regime look like? What activities should the license cover, what works would be subject to the license, and would copyright owners have the ability to opt out? How should royalty rates and terms be set, allocated, reported and distributed?

We do not support the establishment of a compulsory licensing scheme.

10.4 Is an extended collective licensing scheme a feasible or desirable approach?

An extended collective licensing scheme is a feasible approach for visual artists. Collective licensing in general – whether extended or not – is a desirable approach to address a market failure.

In an extended collective licensing scheme, a CMO would negotiate licensing with AI image generators on behalf of visual artists. The amount of the licensing fee a visual artists would receive would be determined by a number of factors, such the number of works published. These ensures that visual artists who are most vulnerable to harm by

[21] https://authorsguild.org/app/uploads/2023/04/Creators-Together-Collective-Action-Rights-Letter-4.27.23.pdf

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

generative AI are compensated accordingly. The extended licensing scheme would also cover visual artists who create works-made-for-hire.

Visual artists who choose to not license their works for AI training would need to opt out of the extended collective scheme. For this reason it is imperative that the opt-out process be simple and easy to access.

10.5 Should licensing regimes vary based on the type of work at issue?
Yes. Different creative disciplines have different considerations in how they works are licensed. Our answers only refer to licensing works of visual art.

11 What legal, technical or practical issues might there be with respect to obtaining appropriate licenses for training? Who, if anyone, should be responsible for securing them (for example when the curator of a training dataset, the developer who trains an AI model, and the company employing that model in an AI system are different entities and may have different commercial or noncommercial roles)?

The critical issue in obtaining licenses for generative AI for images is the sheer volume of licenses that are required for training. However, licenses should be obtained as licenses have always been obtained – by negotiating with the visual artist, the visual artist's agent, or an entity empowered to negotiate on the visual artist's behalf.

Anyone who uses copyrighted work in a way that implicates the copyright owner's exclusive rights is responsible for securing the license to that work. Acquiring a license for one purpose in the generative AI scheme – for example, to develop a curated database of images – does not cover further downstream uses of the image.

12. Is it possible or feasible to identify the degree to which a particular work contributes to a particular output from a generative AI system? Please explain.

One way to begin to assess the "degree" to which an image in the training dataset contributes to an output is by looking at outputs which resemble images in the dataset. This occurs rarely as a result of overfitting. However, when a model is fine-tuned on a limited dataset, the chance of an output replicating an image in the training dataset is increased.

Evaluating to what degree the image contributes to a particular output is to an extent beside the point. The fact that outputs can and do resemble images in the dataset is a clear indication that images have been copied into the dataset at some point in the training process. When this occurs, there could be a valid claim of copyright infringement based on the similarity of the output to the trained image. Regardless of whether an output resembles an image in the dataset, the use of the image for training without consent is in and of itself a copyright infringement.

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

13. What would be the economic impacts of a licensing requirement on the development and adoption of generative AI systems?

Regardless of the impact a licensing market would have on the adoption of any technology, copyrights must be respected. If AI image generators have hefty licensing fees to pay, that is a cost of doing business. No business should be able to sidestep copyrights and licensing responsibilities because the cost is high. In their first year, AI image generators have achieved remarkable growth. and many of companies have achieved a valuation around $1 billion[22]. It's difficult to imagine that licensing will pose an unsurmountable financial burden on these companies.

The question which should also be asked is what is the economic impact of *not* requiring licensing for creative works in AI training? What will be the cost to creators if AI generative platforms are permitted to continue to develop and do business without creators being remunerated for the use of their works? What will be the cost to the development of our culture and the creative works which contribute so much to our economy?

For generative AI to advance, the training dataset needs to ingest high quality creative works. If we no longer have a class of highly trained, highly educated, and highly skilled professional creators because individuals can no longer pursue viable careers with livable wages, less high quality imagery will be made, and generative AI will suffer.

(skipping question 14)

TRANSPARENCY & RECORDKEEPING

15. In order to allow copyright owners to determine whether their works have been used, should developers of AI models be required to collect, retain, and disclose records regarding the materials used to train their models? Should creators of training datasets have a similar obligation?

For visual artists to know whether their works have been used in training datasets, that information must be retained by any entity creating training datasets. Those records should be easily accessible and searchable.

15.1 What level of specificity should be required?

The information visual artists require to understand how their images were used is:

- The title and a thumbnail of the ingested image, or a link to the image source

[22] Within the first year of release, Stability is valued $1 billion (https://growjo.com/company/Stability_AI); MidJourney is valued at over $1 billion (https://medium.com/@oluigbopeter/how-midjourney-became-a-billion-dollar-ai-art-startup-in-less-than-a-year-416663997f77), and OpenAI, the owner of DALL-E, is projected to have revenue of over $1 billion (https://aibusiness.com/nlp/openai-on-track-to-top-1-billion-in-revenue).

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

- The date of the ingestion
- How the image was acquired (ie, by license, scraping, upload by a user)
- How the image was used
- Whether a copy of the image was kept and for how long
- Whether the image was passed on to additional parties

AI image generators should also make public the security measures they've put into place to protect images in training datasets.

15.2 To whom should disclosures be made?

This information should be publicly available and searchable. Visual artists should not have to request access and wait for a response in order to discover whether and how their images have been used in training datasets.

16. What obligations, if any, should there be to notify copyright owners that their works have been used to train an AI model?

AI image generators are obligated to obtain a license for copyrighted images they wish to ingest for AI training. The license serves to notify the copyright owner that their images have been used for AI training. For works which have already been ingested, a searchable database should be made available that will provide information about images or links to images which were ingested for training.

(Skipping question 17)

18. Under copyright law, are there circumstances when a human using a generative AI system should be considered the ''author'' of material produced by the system? If so, what factors are relevant to that determination? For example, is selecting what material an AI model is trained on and/or providing an iterative series of text commands or prompts sufficient to claim authorship of the resulting output?

We agree with the Copyright Office position that only human created works can be copyrighted. Images which are generated purely on AI image generator platforms should not be considered to be authored by a "human".

Current practices by users of AI image generators highlight the issue with considering humans to be the authors of works generated on an AI platform, even when using iterative series of text prompts. Such "prompt engineers" often generate AI artwork solely via sometimes complex, lengthy series of text prompts. They often share their prompts, via social media or AI art forums or user hubs. The question arises, when another individual takes those prompts and generates an image, who should be considered the author of that image? Is it the original author of the text prompts? Or is it the user who entered in the supplied text prompts to generating an image?

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

The same question arises when users share or license the fine-tuning models they create. Is the individual who created the fine-tuning model the author of any images generated from that model?

Those original elements which are created by humans and contribute to the generation of an AI image should be considered copyrightable. The copyrightable elements of AI derivative works are original works uploaded for ingestion, instructions written for fine-tuning, fine-tuning models created from non-infringing elements, and any creative work done off platform such as changes to composition, overpainting, addition of elements, changes in color palette, etc. Fine-tuning models developed from datasets of infringed artwork should not be copyrightable.

Extensive text prompts and iterative curation may also be considered copyrightable, just as extensive direction from an art director may also be considered copyrightable. However, a very high bar is set before an art director considers their contributions to a work to be copyrightable, rendering them a joint author of an image. Such a high bar should also exist for Ai images generated solely via text prompts and iterative curation.

Any reconsideration of the copyrightability of AI generated works must take into account foremost the impact of such a decision will have on human authors.

19. **Are any revisions to the Copyright Act necessary to clarify the human authorship requirement or to provide additional standards to determine when content including AI-generated material is subject to copyright protection?**

No revisions to the Copyright Act are required to clarify human authorship.

20. **Is legal protection for AI-generated material desirable as a policy matter? Is legal protection for AI-generated material necessary to encourage development of generative AI technologies and systems? Does existing copyright protection for computer code that operates a generative AI system provide sufficient incentives?**

Under the current Copyright Office guidance, original elements created by the human author which may be incorporated into an AI derivative image are copyrightable. We don't believe at this point in time additional legal protection for AI-generated images is necessary.

We also don't believe that additional legal protection for AI generated images is necessary to incentivize further development of the technology. Within one year, Stability AI's revenue has reached $16.8 million and the company is valued $1 billion[23];

[23] Stability AI Revenue and Competitors, growjo.com, https://growjo.com/company/Stability_AI

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

MidJourney has generated revenue of $250 million and is valued at over $1 billion[24], and OpenAI is projected to have revenue of over $1 billion[25]. While companies such as Stability AI have yet to generate a profit and are seeking funding to cover their significant development costs, it is dominates in user numbers. The global forecast for AI image generators is positive[26] without the intervention of increased legal protection for their outputs.

The goal of copyright is to protect creators, not companies which utilize copyrighted works but do not themselves create copyrighted works. An additional consideration is that AI image generators developed at the cost of the visual artists whose works were used without consent.

(Skipping question 20.1)

21. Does the Copyright Clause in the U.S. Constitution permit copyright protection for AI-generated material? Would such protection ''promote the progress of science and useful arts''? If so, how?

The Copyright Clause in the U.S. Constitution secures for a limited time exclusive rights for "Authors and Inventors", not non-human authors.

22. Can AI-generated outputs implicate the exclusive rights of preexisting copyrighted works, such as the right of reproduction or the derivative work right? If so, in what circumstances?

Where the output from an AI image generator closely resembles the image on which the model was trained, the output is infringing and implicates both the right of reproduction and the derivative work right. When close fitting occurs, the model has been overtrained on an image and Is more likely to replicate that image in the generated output. This can be the result of an image being overrepresented in the dataset. Foundational models can also be fine-tuned from limited datasets of as low as 20 images, and that resulting output can closely replicate the images in the training dataset.

Fine-tuning models developed from small training datasets frequently use images without consent, and can generate images which closely replicate images in the

[24] "How Midjourney Became a Billion-Dollar AI Art Startup in Less Than a Year," Ugo Peter, August 2, 2023, Medium https://medium.com/@oluigbopeter/how-midjourney-became-a-billion-dollar-ai-art-startup-in-less-than-a-year-416663997f77

[25] "OpenAI on Track to Top $1 Billion in Revenue," Deborah Yao, September 7, 2023. https://aibusiness.com/nlp/openai-on-track-to-top-1-billion-in-revenue

[26] "The global AI image generator market is expected to grow at a compound annual growth rate of 17.5% from 2023 to 2030 to reach USD 1,081.2 million by 2030." https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-image-generator-market-report

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

dataset. Sometimes that is the goal of a fine-tuned model: so that users can generate images which closely match the training images of an admired artist, or closely replicate their style. These models are widely shared or sold on AI image marketplaces, compounding the infringement.

23. Is the substantial similarity test adequate to address claims of infringement based on outputs from a generative AI system, or is some other standard appropriate or necessary?

"Substantial similarity" is the basic test by which courts determine whether a copyright infringement has occurred and can be used to determine infringement when the outputs of AI image generators closely replicate images in the training dataset. However, the unauthorized copying of entire images into training datasets is in and of itself an infringement of the copied artwork, whether or not the output resembles the training images.

Sui generis protection has been proposed to protect the unique styles of visual artists. We are not prepared to support such a protection at this time. A sui generis protection for style would need to be carefully crafted to not harm visual artists. Visual artists do copy the styles of other visual artists. This is often done as they develop their craft or explore new directions. However, significant pressures deter visual artists from widely marketing artwork which largely copies the style of another visual artist. Those pressures include professional and ethical considerations, the natural evolution of a personal style, and peer pressure. Additionally, visual artists may by chance develop a style similar to another visual artist's without having ever seen that visual artist's work.

As stated earlier, we would prefer a legislative approach prohibiting the use of existing visual artists' names, artwork, title or art, or other unique identifiers from being used in training datasets or prompting without permission. Such an approach would target the users of AI image generation platforms, who may not be working visual artists and may not share the professional and ethical considerations which deter visual artists from plagiarizing styles.

24. How can copyright owners prove the element of copying (such as by demonstrating access to a copyrighted work) if the developer of the AI model does not maintain or make available records of what training material it used? Are existing civil discovery rules sufficient to address this situation?

Short of having an AI image generator output which closely resembles their artwork, visual artists have no means to prove the element of copying. If AI image developers are not required to maintain and make available records of training materials, it is difficult to envision how existing civil discovery rules will be adequate for visual artists to prove that their work was used in a training dataset. Bringing an infringement lawsuit is out of reach for most visual artists, particularly if discovery in the case becomes a fraught process. The Copyright Claims Board, with its limited discovery process and

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

limited awards, is insufficient to address many claims, and respondents with deep pockets or legal counsel on retainer would most likely opt out of the CCB process.

25. If AI-generated material is found to infringe a copyrighted work, who should be directly or secondarily liable—the developer of a generative AI model, the developer of the system incorporating that model, end users of the system, or other parties?

Since copyright infringement lawsuits are fact specific, it stands to reason that any entity participating in the infringement should be held liable. Whether or not an entity – a platform developer, a model developer, a user, etc. – is primarily or secondarily liable should be determined by the facts of the case. As in any other copyright infringement, visual artists should be able to seek remedies from any party engaging in the infringing activity.

25.1 Do ''open-source'' AI models raise unique considerations with respect to infringement based on their outputs?

No, how visual artists are impacted by the infringement of their work does not change if the generative AI image model is open-source or not.

26. If a generative AI system is trained on copyrighted works containing copyright management information, how does 17 U.S.C. 1202(b) apply to the treatment of that information in outputs of the system?

§1202 of the Copyright Act has become even more critical as metadata is being used to identify, track, indicate licensing terms, and embed the copyright management information (CMI) of images. Requiring AI image generators to be transparent and publicly disclose images used in training datasets is pointless if means to identify images – image metadata – is stripped out. Image metadata and the contained CMI is routinely stripped by generative AI platforms, social media platforms, and others. This data is critical to identify and protect images and convey information.

Additionally, some AI image generator users alter metadata from AI generated works, for example to hide the image's AI origins to license those works to stock image platforms that prohibit AI works. Metadata is also removed or altered to mask the true copyright holder of an image. Removing and altering metadata can be done through image editing software such as Photoshop, or by using online tools developed specifically for that purpose.[27] This alteration of metadata can be a violation of §1202 of the Copyright Act.

[27] "ReMeta is a powerful tool for removing and reading metadata from PNG images generated by the Stable Diffusion Web UI." Released April 2023 on Github, https://github.com/geocine/remeta

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

As stated above, a revision to §1202 is required to prohibit the removal of CMI recklessly or intentionally, regardless of whether the removal is done with the knowledge that doing so would "induce, enable, facilitate, or conceal infringement".

(Skipping question 27)

LABELING OR IDENTIFICATION

28. Should the law require AI-generated material to be labeled or otherwise publicly identified as being generated by AI? If so, in what context should the requirement apply and how should it work?

A strong argument can be made to require the labeling of purely AI generated works. Clients commission works which they then license; it's critical for them to know whether the image they are commissioning is copyrightable or an AI-generated work that falls under public domain. As stated above, dishonest actors do hide the origins of purely AI generated work by altering metadata to pass it off as original work created by them, or to hide that they've infringed the works of visual artists in their training models. While labeling images created by AI is helpful, having a record of the images that are generated on an AI platform will also help in countering bad actors.

We're not prepared at this time to state that the labeling of derivative works containing AI generated outputs should be required in all cases. Questions as to what extent generative AI is used would need to be considered. For example, if a work is largely created off AI platform, but an AI content fill is used to fill in a portion of an illustration, does that mean that the resulting image is a derivative work that requires labeling? Does the size of the area filled using the AI tool need to be considered? What if the content fill is generated from original images the visual artist has created?

A best practice that is both responsible and ethical is for visual artists to voluntarily label their AI derivative works and keep a record of their provenance. This best practice looks after the interests of clients, and provides a record which will be useful when registering the copyrights of AI derivative works.

28.1 Who should be responsible for identifying a work as AI-generated?

Should the labeling of AI generated work be required, then the entity responsible for generating the AI image should be responsible for identifying the work as AI generated. That could be either the user generating the work, or the company employing an individual to generate the image. AI image generators should be required to embed metadata identifying works generated on platform, and the removal or alteration of that metadata should be prohibited.

28.2 Are there technical or practical barriers to labeling or identification requirements?

We do not have the knowledge to adequately address this question from the perspective of AI image generators. However, from a visual artists' perspective, applying metadata in order to identify and label AI generated images can be easily done. Metadata standards are being

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

developed by a number of organizations to record the provenance of AI generated images, including derivative images combining AI and non-AI works. For example, the Content Authenticity Initiative (CAI) has developed tools which make it easy for visual artists to embed the origins of derivative AI works, and have that information persist even when an image is grabbed via a screenshare.[28]

Utilizing such tools for labeling and identification does not pose a practical barrier for visual artists. Outside tools such as CAI, metadata can easily be applied off-platform by using image editing software. However, the simplest solution for visual artists and users would be for AI image generator platforms to autogenerate metadata for images created on-platform.

(Skipping questions 28.3-31)

32. Are there or should there be protections against an AI system generating outputs that imitate the artistic style of a human creator (such as an AI system producing visual works ''in the style of'' a specific visual artist)? Who should be eligible for such protection? What form should it take?

While not a copyright infringement, the copying of visual artists' style is a serious concern.
Visual artists invest years into developing their unique styles, and style also evolves naturally as visual artists practice their craft. Flooding the public space with images in their unique style of can essentially obscure the entire oeuvre of work of a visual artist. While some AI art generators have blocked users from using visual artists names into text prompts, on other platforms users can prompt for a visual artist's name, the title of a visual artist's work, or the names of characters developed by a visual artist.

Fine-tuning developed from visual artists' images as described previously produces works which plagiarize these unique styles. Such fine tuning has even been weaponized against visual artists who protest generative AI, with the understanding that the resulting outputs will resemble the visual artists' work or style and thus do market harm to them.[29]
Additionally, damage is done to the visual artist's reputation when AI models fine-tuned on their unique, identifiable styles is used to generate racist, homophobic, or otherwise offensive content.[30] Fine-tuning models developed from infringed images are widely shared or even sold on generative AI marketplaces.

---

[28] YouTube video: "How to bring more transparency to AI generated content | Adobe Firefly beta and Content Credentials", Content Authenticity Initiative, March 21, 2023. https://www.youtube.com/watch?v=UwBaPWYhHRg

[29] Artist Jon Lam on Twitter (now X): "Fights between Artists and Ai prompters are a given. However, when prompters start feeding an Artist's art into Ai to "Win" the argument, you've gone into behaviour inline with making revenge pr0n. Way to violate an artist publicly." The post includes a screenshot of an AI artists showing artwork he had trained on the works of an artist he had a dispute with.
https://x.com/JonLamArt/status/1612494765203009536?s=20

[30] "The Alt-Right Manipulated My Comic. Then A.I. Claimed It." Sarah Anderson, The New York Times, December 31, 2022. https://www.nytimes.com/2022/12/31/opinion/sarah-andersen-how-algorithim-took-my-work.html?ref=content-technologist.com

31 West 34th Street · 8th floor
New York, NY 10001
(212) 791-3400
(212) 791-0333 (fax)
graphicartistsguild.org
info@graphicartistsguild.org

An example of how visual artists' styles are being thoughtlessly appropriated and used to develop fine-tuning models is the "Style Capture and Fusion Model" contest currently running on civitai.com. The contest challenges model trainers to "go out and capture the unique styles of their favorite artists!"[31] The contest webpage includes screenshots of AI art generated in the styles of contemporary visual artists Banksy, Greg Rutkowski, and Adam Hughes.

Companies developing and operating generative AI can implements safeguards to prevent outputs from resembling inputs, but such safeguards are not widely nor consistently implemented. They also may not protect against users deploying own fine-tuned models. For example, although Stability Diffusion removed the ability of users to prompt for visual artist's names, fine-tuned models built on infringed artwork can still run on that platform and generate images that plagiarize the style if not the images of the infringed visual artists.[32]

Considering the difficulty in getting the generative AI platforms to implement workable solutions to prevent visual artists' works from being ingested without permission, there is concern that the industry will not voluntarily implement effective safeguards to prevent infringing outputs or protect visual artists from having their styles replicated. Licensing, including collective licensing, could provide a means whereby the AI image generators would be required to implement safeguards as part of the licensing agreement.

We previously suggested that legislation which prohibits the use of a visual artist's images, name, or titles of works. Such legislation would target the users of AI image generators. Additionally, recognizing that the use of images in training datasets is a licensable use not protected by fair use will deter users from developing fine-tuning models based on infringing works.

(Skipping question 33)

34. Please identify any issues not mentioned above that the Copyright Office should consider in conducting this study.

The emergence and wide adoption of AI image generators have spotlighted deficiencies in the copyright ecosystem which have long plagued visual artists. It is increasingly critical that visual artists be able to register their copyrights easily and affordably. The extreme limitation imposed by the GRUW registration option – a group of 10 unpublished works – requires graphic artists to pick and choose which works they can afford to register. We require a larger registration option for graphic artists

[31] "Style Capture & Style Fusion Contest(s)!" on civitai, https://civitai.com/articles/2755

[32] From the civitai "Style Capture & Style Fusion Contest(s)!" webpage: "One of the most amazing things about Stable Diffusion is the ability to collect and deploy styles with the same attention a painter may use color or texture." https://civitai.com/articles/2755

31 West 34th Street · 8th floor
New York, NY 10001

(212) 791-3400
(212) 791-0333 (fax)

graphicartistsguild.org
info@graphicartistsguild.org

which reflects how we work and is affordable. Limiting group registrations to unpublished works creates additional problems since the guidance of what connotes "publication" does not reflect the real-world complexity of the online environment.

Permitting visual artists to register works under a provisional or deferred registration option, as we have advocated for in the past, would permit visual artists to initiate the registration of more works without imposing an immediate burden on the Copyright Office. We request that the Office reconsider its dismissal of deferred registration proposals.

The Copyright Office's guidance on registering AI generated provided some needed instruction. However, we're concerned that bad actors will abuse the system by registering AI generated works as original works, or by neglecting to disclaim AI elements properly.

We're also concerned that visual artists with the best intentions will be confused by the registration process for AI derivative works, and may either avoid registering works, or disclaim elements improperly. In particular, there is some confusion about what constitutes generative AI elements which must be disclaimed. For example, AI tools can add content aware fills, adjust lighting, provide alternate layouts, and change color palettes or line weights. Do these tools constitute generative AI, and if so, should the use of these tools be disclaimed? Is there a difference if such a tool is applied to the entire image or a portion of the image?

Ideally the registration process would include technical means by which metadata which records the provenance of AI and AI derivative works could be read. This would deter bad actors from falsifying registration records and relieve creators from confusion when registering their derivative AI works. We recognize that a significant commitment of funds would be required to develop and implement such forward-thinking technical measures in the registration process.

Thank you for the opportunity to weigh in on these important matters.

Rebecca Blake
Advocacy Liaison
Graphic Artists Guild