

Comment: US Copyright Office Artificial Intelligence Study

Background:

This comment is submitted to the US Copyright Office anonymously and in good faith. The writers of this document do not represent the holders of any of the intellectual property within. This Intellectual Property is used in demonstration and should not be assumed to be owned by the writers. This document was authored by a group of machine learning experts, engineers, artists, and social media experts who have been observing the current state of the AI landscape.

Summary:

This submission describes ways in which *Generative AI* should be considered infringing for various reasons. Additionally, this submission contains evidence that many users of Generative AI blatantly and knowingly create potentially infringing Generative AI programs and that there exists a market for these programs funded by well known entities. Additionally, we conclude that this implies an awareness by aforementioned entities (such as a16z) that regulation of this technology on the basis of Copyright is inevitable.

Key Conclusions:

This submission concludes that the training of Datasets without Copyright Holders permission is Copyright Infringement as it integrates Intellectual Property into part of a computer program.

This submission concludes that there is a well funded group of Tech investors who are paying and encouraging infringing Training.

This submission concludes that Generative AI Outputs can be infringing as they clearly and plainly usurp the market for the original work and/or creator

Contents

1. ***Training AI is Not Fair Use***
 - 1.a ***What is Training?***
 - 1.b ***How can Training Infringe Copyright?***
 - 1.c ***Generated Outputs can Demonstrate Infringement in Training.***
 - 1.d ***OpenAI is Losing Staff Over Copyright Concerns.***
2. ***Outputs of Generative AI Inherently Impact the Market Value of Works on Which it was Trained.***
3. ***There exists a well funded network of Generative AI Users knowingly engaged in and encouraging Training on Copyrighted works without permission***

[1]

Training AI is not Fair Use

1.a What is Training?

Training turns raw data, that is to say: any *digital work, file, or digital information* into a machine readable *Dataset* or *Artificially Intelligent Model*. Almost all machine learning algorithms do this by *associatively compressing some data by its similarity to other data*. [2][3] Datasets are required for Generative Models to produce an output. It is tempting to say the machine appears to “learn” associations on its own, but this is actually due to the effort of humans manually labeling data.[4] Without human driven data labeling and human driven data creation, Generative AI algorithms would not be possible. When a Dataset or Generative Model is trained, the computer builds *mathematical functions (like the kind in Algebra Class) in relation to “Concepts”* regarding the labels. Concepts are boolean functions (true/false statements) related to labeled data. We conclude that nothing about “Artificial Intelligence” is artificial, that, while the process of training is automated, all data and labeling is human in origin; and therefore subject to Copyright.

1.b How can Training Infringe Copyright?

Examples of infringing training can be observed in popular visual Generative AI like *Dalle* and *Stable Diffusion*. Dalle for example, uses the CLIP Dataset, a Dataset created by ‘*400 million (image, text) pairs collected from the internet.*’ (We urge regulators to note the phrase ‘collected from the internet.’) OpenAI, the creators of the CLIP, have been less than transparent about the origin of all 400 million images contained within the set. We also urge regulators to note that *Stability AI is presently being sued by GettyImages for using watermarked Stock Images to train Stability AI’s Dataset*. In both cases, the dataset was developed by crawling the internet for images along with their Alt-Text’ **without regard to origin or Copyright Owners.**

Another dataset known as ImageNet (a pre-trained Dataset which CLIP is based on) tells researchers not to use the dataset in commercial application. It is also important to note even the ImageNet dataset appears to skirt the law by linking to images rather than providing a live copy of the dataset. Staggeringly however, LAION-5B, another dataset on which Stable Diffusion has been based, is trained on a staggering **five billion different image-text pairs**. This dataset, like ImageNet, also requires that the end user download the images to be used in training. Like OpenAI, LAION offers a paid model, which contains even more image-text pairs. Because this Dataset has obvious commercial applications and is often used in such applications.

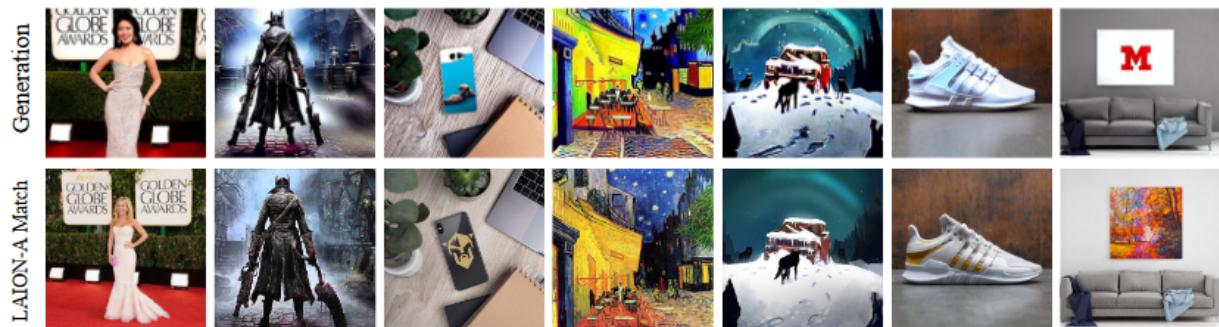
Does ImageNet own the images? Can I download the images?

No, ImageNet does not own the copyright of the images. ImageNet only compiles an accurate list of web images for each synset of WordNet. For researchers and educators who wish to use the images for non-commercial research and/or educational purposes, we can provide access through our site under certain conditions and terms. For details click [here](#).

[Fig 1-A] ImageNet’s note about image ownership, possibly demonstrating an awareness of the legally dubious nature of training

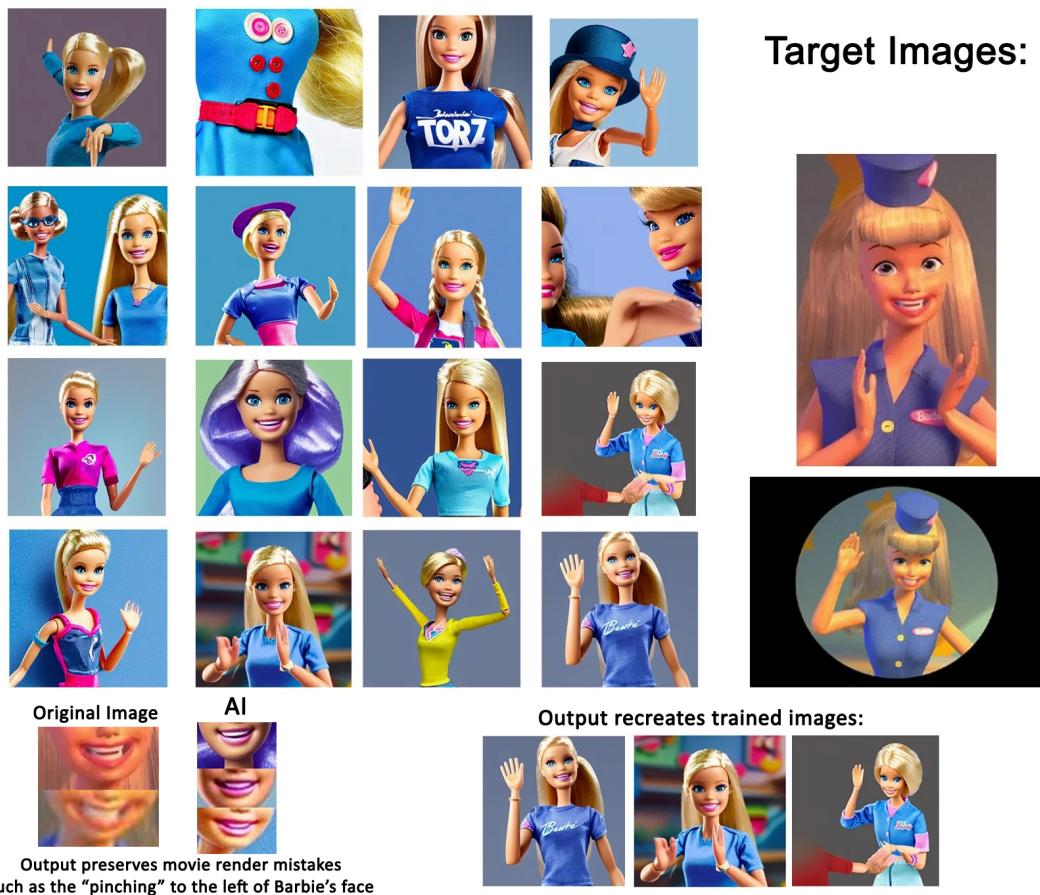
1.c Generated Outputs can Demonstrate Infringement in Training.

Despite the claims by entities like OpenAI, the process of training a Dataset saves the original images in a presently unknown way. [1] This has been detailed extensively in our sources 1, 3, 5, and 6. Below is a figure from this source [1] indicating how Stable Diffusion can be used to expose this training data.



[Fig 1-B]

Figure taken from "Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models" [1]

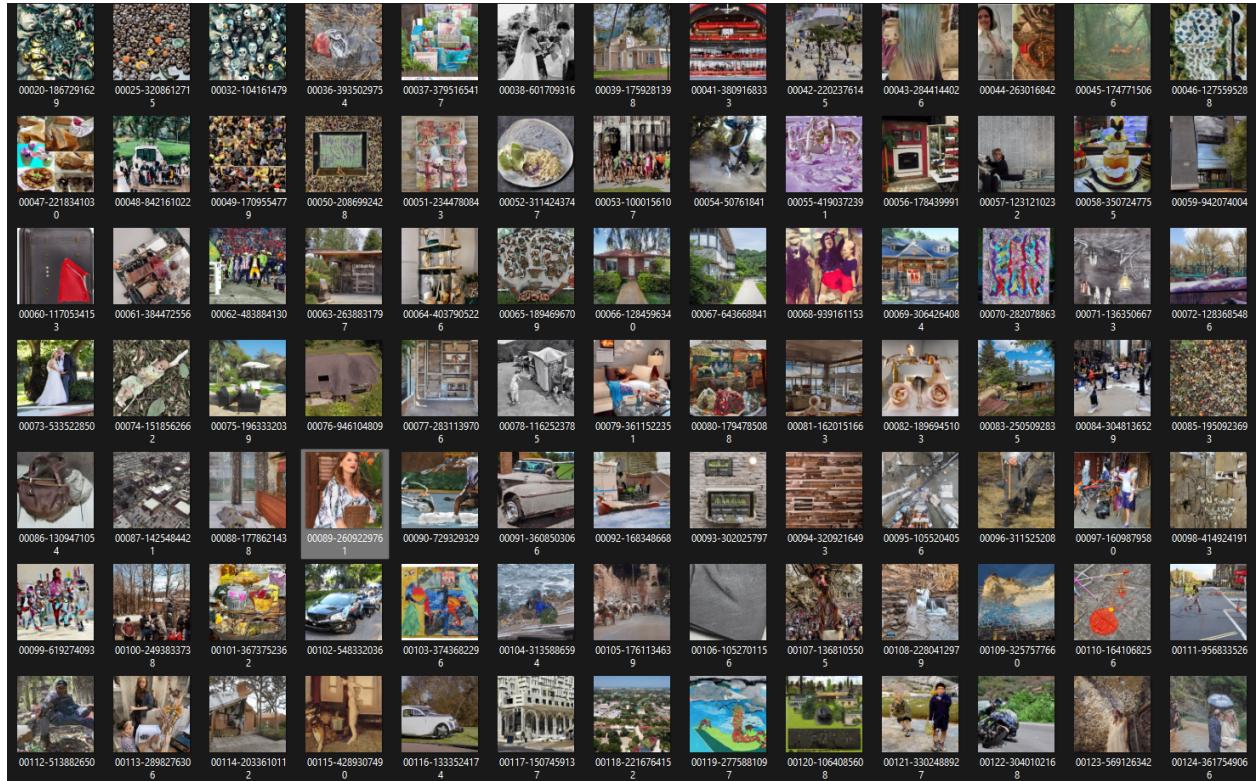


[Fig 1-C]

Our experiment confirms that Copyrighted material exists within the CLIP Dataset and is used by AI in a significant way. Notice that unique and specific visual details are preserved.

1.d Generative AI Stores and Uses Data on Which It Has Been Trained.

Figure 1B demonstrates the ability of Stable Diffusion to faithfully recreate images it's Dataset. In this case, the experimenters used the LAION-A dataset, which is similar to LAION-5B Dataset mentioned in section 1, but it has been specially trimmed and selected for "Aesthetics." (We once again ask regulators to consider what application other than commercial outputs can possibly come from an 'Aesthetically Tuned Image Library.') Likewise, in our own experimentation we were able to verify that Copyrighted Imagery can be extracted from the standardized CLIP Dataset (Fig 1C.) Likewise, we found that when not entering a prompt Stable Diffusion is likely to generate an image that resembles stock photography, or Google Street View images (Fig 1D.) We believe these experiments once again verify that trained data is not "learned" but compressed in a way we don't yet understand. [1][4]



[Fig 1-D]

Without a prompt, AI Image Generation produces images from "latent space" that often resemble internet memes, celebrities, stock photos, and Google Street View.

It is important to note that Generative AI does not learn like a human as OpenAI and Stability have continued to suggest. It is true that Neural Networks resemble a brain (hence the name,) this is only functionally true in the most very basic of ways. While Neural Networks are a model

of brain-like structures, they are still computers, performing computation tasks assigned to them. Likewise, the model of the brain on which Neural Networks are based, has long since been seen as basic and limited. It is indeed tempting to assign human qualities to these programs, they are simply that, a computer program. Generative AI seems so human because it was designed to mimic humanity. Convincing a mimic some forms of Generative AI may be, it is nothing but the result of a machine, blindly averaging data. It is a mathematical algorithm which is to be engaged and arrive at an expected result. To suggest Generative AI is anything but a computer program is disingenuous and intentionally misleading. Therefore, we ask that the Office should conclude that training a Dataset or any AI Model on works without permission is no different in legality than a pirated library of stock photographs. (At this time, we also would like to remind the Office that OpenAI, StabilityAI, and MidJourney have outstanding cases against them for training AI on pirated materials.)

Once again, despite the claims of many Generative AI advocates, outside the scope of its training. This is particularly true for Image Generating AI like Stable Diffusion and Dalle. For example, if we task a human being with drawing a made up creature with a made up name we give it, the human can readily do so. However, Generative AI will produce a random image, or, an image of something real with similar spelling to the aforementioned made up name. For the AI to produce a made up creature, a human would have to first draw or describe it for the machine to process into an image. This indicates that Generative AI can only generate outputs based on the data it was trained on. Therefore, all outputs from AI are inherently derivative works.

[2]

Outputs can be Infringing.

2.a Generative AI Outputs Directly Interfere With the Market and Reputation of Copyrighted Material and Copyright Holders It was Trained On.

The unfathomable scope of works trained without permission between CLIP and LION has gone on under the dubious guise of 'Research.' However, proliferation and use of Generative AI by Individuals, and Commercial entities would indicate something else entirely. As Figure 2D indicates, these datasets and models are indeed being used in Commercial Settings. We urge regulators to ask themselves; 'What other possible application exists for these Datasets and models, trained especially for aesthetics, other than for a commercial application?' At this point in time, no reasonable person can conclude that more recent models and datasets have been developed 'purely for research.' These models and datasets are made explicitly to be sold or produce something to be sold; in direct competition with the works it was trained on, without permission.

Outputs produced by AI can also create confusion about and do damage to a Copyright Holder's reputation and their Copyrighted Works. As indicated by a recent, viral trend of generating risque Disney-PIXAR movie posters for hypothetical films. Disney is presently taking action over these generations. This ability to damage reputation is not limited to Intellectual Property, but personal reputation too, indicated by the growing number of cases of deepfakes and generated revenge porn. Many people now also find their voices have been cloned and made to say unspeakable things, or their voice has been used to scam their own friends and family. Given Generative AI's undeniable ability to mimic a person, art, literature, and even voice; it is impossible to deny Generative AI's ability to damage the reputation and market for Copyrighted Works it has been trained on.

2b. There exists a culture of Generative AI Users knowingly engaged in and encouraging training on Copyrighted works without permission:

The community CivitAI as of the creation of this document is hosting a "Style Capture" contest. This contest is encouraging members of the community to "Capture Styles" of artists. In numerous cases this will be carried out without permission of the creators of the art in question. The resulting LoRA and models will be proliferated en masse. [5][6]

Additionally, the capture of this data is strongly funded by venture capital firms like a16z, who as of writing this document, has funded the aforementioned "style capture contest" and lists job postings for CivitAI on their website. [5]

Likewise, many Generative AI users are now actively trying to prevent the deployment of *Nightshade*. There are existing guides and recommendations on how to defeat Glaze. *Nightshade* and *Glaze* are technologies meant to protect visual artworks from being trained on. This is a clear and obvious circumvention of *Technological Copyright Protection Measures (Title 17 - Chapter 12)*

The screenshot shows a dark-themed website for CIVITAI. On the left, there's a banner for a contest. It features a cartoon robot icon, a Steam gift card icon, and two blue buttons labeled "2x Gift Card" and "6x Civitai Sub". Below the banner is the CIVITAI logo. On the right, there's a sidebar with a "Table of Contents" section. The sidebar includes links to "Style Capture & Fusion Contest", "Style Capture - LoRAs", "Part 2: Style Fusion - Images", "The Fine Print", "Part 1: Style Capture (Oct 27th - Nov 3rd)" (which is highlighted in blue), and "Part 2: Style Fusion (Nov 3rd - Nov 10th)".

Style Capture & Fusion Contest

Greeting Civitans! We've decided to do something a bit different than the classic Halloween contest. We are excited to host a two-part contest that promises to be both challenging and fun!

Part 1: Style Capture (Oct 27th - Nov 3rd)

In the first part of the contest, we are looking for some awesome art style LoRAs! Ever felt like there has been a void of style LoRAs for artists that you love? From the primal beauty of cave paintings to the modern times today, we're calling for our best trainers to go out and capture the unique styles of their favorite artists! 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. More styles can only mean more fun!

[Fig 2-b]
CivitAI encouraging style capture for compensation.

2.c Copyright Law as it Applies to Humans Cannot be Applied to Generative AI

As we have established so far, we can safely say that Generative AI is made up of the works it has ingested. As such, aspects of intellectual product that are not traditionally Copyrightable now must be considered part of a protected work. In the case of illustration, an individual or Group may wish to protect an artstyle that is part of their or their products identity. As we established, training Generative AI to extract an artstyle provides direct market competition to the works it has been trained on. Below, Figure 2F demonstrates that highly unique artstyles are captured and mimicked by using the creators name. Therefore, "artistic

"style" in all of its forms (musical, lyrical, visual, literary) should be a protected form of intellectual property in regards to Generative AI. Moreover this also means one's own appearance, voice and all other aspect of oneself that can be trained into a machine must also be protected from unauthorized use in Training.

Closing Comment:

Generative AI is a computer program and must be considered as such. Generative AI and its outputs can only be created and improved by consuming the work of Copyright holders and the data of millions of everyday people. The changes this brings to the concept of Intellectual Property are unprecedented. Every single American is potentially a victim of infringement of anything they have ever posted online thanks to the past or future training of AI. The scale of possible abuses from the ability to clone a voice, a face, place, musical style, artistic style, and literary style are vast in scope. These abuses not only range from economic, to defamatory; they can span sexual, and potentially endanger national security. As users need only type a few words to bring the aforementioned abuses to life, the onus of legal responsibility falls on the creators and users of these tools. We urge strong regulation of past and future trained models, and outputs.

References:

1. **Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models**
(Gowthami Somepalli, Vasu Singla, Micah Goldblum, Jonas Geiping, Tom Goldstein) - 2022
https://openaccess.thecvf.com/content/CVPR2023/papers/Somepalli_Diffusion_Art_or_Digital_Forgery_Investigating_Data_Replication_in_Diffusion_CVPR_2023_paper.pdf
2. **White-Box Transformers via Sparse Rate Reduction: Compression Is All There Is?**
(Yaodong Yu, Sam Buchanan, Druv Pai, Tianzhe Chu, Ziyang Wu, Shengbang Tong, Hao Bai, Yuexiang Zhai, Benjamin D. Haeffele, Yi Ma) - 2023
<https://arxiv.org/abs/2311.13110>
3. **The Illustrated Stable Diffusion**
(Jay Alammar) - 2022
<http://jalammar.github.io/illustrated-stable-diffusion/>
4. **Learning Transferable Visual Models From Natural Language Supervision**
(Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, Ilya Sutskever) - 2021
<https://arxiv.org/abs/2103.00020>
5. **We've Raised \$5.1M in Seed Funding Led By Andreessen**
<https://civitai.com/articles/2976>

*Note: Authors of this submission do not represent all of the rights holders in this document, and their works are **not** submitted on their behalf. Certain outputs have been selected from well known rights holders as examples of **potentially infringing** outputs, they should not be viewed as specific claims of rights infringement.*

