

Hi,

I am writing to you in response to your call for comments on copyright and generative AI. The key takeaway is that using copyrighted material to train generative AI is absolutely infringing, and under no circumstance should outputs of the generative AIs should be eligible for copyright. At the end of this comment is a link to an article in which all of the posted questions are answered in detail. Copyright must always be respected, even if the copyright holder has not officially registered for it. Artists and other creative types are being deeply harmed by generative AI right now, and they should be in absolute control of whether or not their works is used to train generative AI. They must be asked for permission first to use their works to train generative AI.

What various AI corporations have done is scraping the web for a massive amount of copyrighted works and put them in datasets. The corporations have then used said works to train their own generative AI models without the consent of the copyright holders. (<https://theconversation.com/chatgpt-is-a-data-privacy-nightmare-if-youve-ever-posted-online-you-ought-to-be-concerned-199283>) They then release the models to the public forcing these copyright holders to compete against their own stolen work. (<https://www.artstation.com/artwork/o23lkw>) There is often little to no transparency about what the dataset used for training even consists of, making it extremely difficult for creative types to protect themselves from these exploitative practices. The corporations doing this include but are not limited to OpenAI, Stability AI, Midjourney, ElevenLabs, Microsoft, Adobe, Google, X (also known as Twitter), character.ai, and Meta.

Using copyrighted works to train generative AI is absolutely not fair use in any sense of the term. Passionate defenders of generative AI love to claim that AI somehow “learns” just like a human does, so using copyrighted works should fall under fair use, but there is ample evidence against this claim. There is a wonderful program named Glaze that artists can use to protect themselves from a generative AI mimicking their style. (<https://glaze.cs.uchicago.edu/>) The program takes an image and makes various changes to it that are barely noticeable to a human. However, when a generative AI is trained on images protected by Glaze, it is unable to mimic the artist’s art style. If generative AI truly was just like a human, why is it unable to ignore these changes that a human barely sees or can just ignore? Not to mention, illustrators typically develop their own art style when learning how to draw rather than imitate another artist exactly, so generative AI absolutely does not learn like a human or create like a human. At the end of the day, generative AI is just software, and a deeply exploitative type of software at that. All AI outputs are derivative of the training set, so there is no fair use when it comes to generative AI. There have been many cases of generative AI reproducing an image in its training set exactly or near exactly, and ChatGPT specifically has reproduced sentences verbatim from copyrighted books. (<https://www.businessinsider.com/openais-latest-chatgpt-version-hides-training-on-copyrighted-material-2023-8?international=true&r=US&IR=T>) Generative AI relies on a massive amount of copyrighted works in order to function, I do not believe pirated copies of every single Harry Potter book would fall under fair use in a training set. Without the training set, the generative AI cannot make anything, so with all of this said, using copyrighted works to train AI models is not fair use.

AI outputs should never be eligible for copyright. As of now, outputs are not eligible for copyright protection, and creative types are still experiencing substantial harm from these outputs. Already, artists are forced to compete with their own stolen labor on online social media platforms. If these outputs were eligible for copyright protection, the harms would be so much worse than they are now. Entire career paths like journalism, creative writing, illustration, and others are already at risk with exploitative generative AI existing. If these outputs had protection, these careers would be completely annihilated. Also, there is little to no human input when an AI generates something. Writing a prompt for an AI takes little to no skill, there is little to no practice or improvement involved. Even if one wants to argue that writing a prompt somehow has any artistic merit, the generative AI is still doing a substantial amount of the work. Even if a prompt for an image generator is bad or makes no sense, the AI is still designed to output an image that looks pretty, at least on a surface level. There are also many cases when the generator ignores parts or all of the prompt. So no, people who use generative AI should never receive copyright protection for the outputs.

When it comes to liability for copyright infringement where an AI output infringes upon copyright, I believe both the end user and the developers should be held liable, as well as whoever is hosting or sharing the infringing AI model. After all, anyone could get that exact same result using the same prompt and seed. I feel like how the responsibility would be distributed would probably be on a case-by-case basis. Perhaps the developers already trained the model on a massive amount of copyrighted material, or perhaps the user baited the AI into making something that looks too similar to a copyrighted character using image2image, even if the model was not trained on images of that character. This is why we desperately need new regulation for this technology.

AI that can imitate voices and people in general is not just a threat to creative types, it is literally a threat to society as a whole. Already, people are taking voices of people and using AI mimicry to scam loved ones of the person that voice belongs to. Already, people are using AI to make fake videos of people saying things they would not say. AI mimicry absolutely has to be made illegal and the people hosting the AI models that make this possible must be accountable, this technology is insanely dangerous and has caused serious harm, and will continue to cause serious harm if left unchecked.

Regarding the specific questions posted, I will leave a link to an article in which every question is answered. This person has done a very good job answering them, and I strongly agree with the points they make.

<https://luddite.pro/public-submission-for-usco-generative-ai-call-for-comment/>