
From: Matt Aimonetti <mattaimonetti@gmail.com>
Sent: Tuesday, November 12, 2019 4:18 PM
To: aipartnership
Subject: Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation

Hi, my name is Matt Aimonetti, I'm an engineer, technologist, entrepreneur and someone who's been spending time reflecting on the impact of ML/AI on the music industry as well as other creative fields. Thank you for giving me an opportunity to share my thoughts.

1. Should a work produced by an AI algorithm or process, without the involvement of a natural person contributing expression to the resulting work, qualify as a work of authorship protectable under U.S. copyright law? Why or why not?

In essence, I believe that there should be some protection for generated work but there are so many ways to abuse such a copyright law that I'm concerned about the consequences. Let's imagine for a moment that an algorithm generates and files for copyright all possible combinations of musical notes and chords or even hundred of thousands of lyrics for songs. Such an approach is clearly doable and will only get easier better and cheaper over time. It would be the equivalent of a Denial of Service when a perpetrator seeks to make a network resource unavailable to its intended users by temporarily or indefinitely disrupting services. It also means that whoever has the most technical resources has the most chance to be in a position that would prevent others from creating since the chance of authorship being contested would be high.

It would also mean that the work of the USPTO would be extremely complicated by a huge influx of generated copyrightable content to verify.

On the other end, there could be limitations put in place to prevent abuse and differentiate copyrights of automatically generated content and human created content (shorter duration, necessity to have the content publish in a meaningful way etc...)

2. Assuming involvement by a natural person is or should be required, what kind of involvement would or should be sufficient so that the work qualifies for copyright protection? For example, should it be sufficient if a person designed the AI algorithm or process that created the work;

contributed to the design of the algorithm or process;

chose data used by the algorithm for training or otherwise;

caused the AI algorithm or process to be used to yield the work;

or (v) engaged in some specific combination of the foregoing activities?

Are there other contributions a person could make in a potentially copyrightable AI-generated work in order to be considered an "author"?

I think that for someone's AI related work to be qualified for copyright protection they would need to have curated the data used by the algorithm, design the processing stages to yield the work and curated the outcome.

The last part being the key element missing in the question's suggestions. The curation process is very creative on its own which must limit the amount of content filed for copyright as well as bring a specific human selection.

3. To the extent an AI algorithm or process learns its function(s) by ingesting large volumes of copyrighted material, does the existing statutory language (e.g., the fair use doctrine) and related case law adequately address the legality of making such use? Should authors be recognized for this type of use of their works? If so, how?

As you all know, this is a very complex question due to the implicit nature of how ML algorithms vs humans work. If a human spends years listening to specific artists and create something influenced by their work, the creator isn't infringing any copyrights. As a matter of fact, that's this very process of influence that helps us move forward as a creative group. From a purely technical perspective, that's also extremely complicated since the training data is removed and not available in the output (weights are the results of training, they can't be reverse engineered to find the training sources). My current personal perspective is that training data shouldn't and can't efficiently protected by copyright laws but the output of the work however needs to be analyzed to see how closely it comes from copyrighted material. For instance, I could train a voice generator on Michael Jackson's voice and create a totally new song that sounds exactly like him. It's technically new work but it's clearly derivative work which should be permitted by his estate and licensed. AI generated work should not be considered remotely close to existing copyrighted material to be awarded protection. The work needs to be unique and innovative, not a remix/derivative work of clearly identifiable existing work.

5. Should an entity or entities other than a natural person, or company to which a natural person assigns a copyrighted work, be able to own the copyright on the AI work? For example: Should a company who trains the artificial intelligence process that creates the work be able to be an owner?

I think it's critical that we don't let companies training AI models become owners or we take the risk of stifling creativity. Companies have for primary goal to generate profits. They/we will find ways to adapt technology faster than you will be able to adapt copyright laws. In the meantime, creators will suffer competing on a race to the bottom against fast growing compute clusters. Such an approach would also increase litigation, once again hurting creators with less financial resources.

6. Are there other copyright issues that need to be addressed to promote the goals of copyright law in connection with the use of AI?

I think it's worth considering not allowing the copyrighting/patenting of technical approaches leading to model training and AI creation. While we are currently taking this approach with most algorithms being publicly and freely available, I'm afraid this might change quickly. AI can't be only serving the first movers and well financially backed companies. This is a delicate position that would also require owners to publish their technical approaches if they wanted their content to be copyrighted (without disclosing the training data nor going to deep into technical details).

7. Would the use of AI in trademark searching impact the registrability of trademarks? If so, how?

The biggest challenge here would be bias in the search algorithms. The second challenge being knowing if the search can be trusted to verify if a matching trademark does exist. If the search isn't reliable and doesn't match the verification algorithm, then the search is pointless. I would also strongly suggest to make the search and other AI implementations for trademark verification and implementation totally open for all to verify.

9. How, if at all, does AI impact the need to protect databases and data sets? Are existing laws adequate to protect such data?

As far as I'm aware there aren't any existing laws protecting data used for training/educational reasons. But even if there were, the enforcement of such laws would require whistleblowers to denunciate their companies. Companies around the world are training neural networks using copyrighted content coming from lots of different sources, from streaming service to pirated content or stolen information. This is not a secret to anyone and it seems unrealistic to think we can make that stop.

11. Do any laws, policies, or practices need to change in order to ensure an appropriate balance between maintaining trade secrets on the one hand and obtaining patents, copyrights, or other forms of intellectual property protection related to AI on the other?

I believe that to apply for a copyright relying on AI, the owner should be disclosing the technical approach taken. Patents on AI model design should not be granted and we should have a well documented, constantly updated automated process to check for copyright infringement of submitted content. Also AI copyrighted content should have a shorter protection life than other content and the amount of submitted content should be limited.

12. Are there any other AI-related issues pertinent to intellectual property rights (other than those related to patent rights) that the USPTO should examine?

I think there are two key AI-related issues pertinent to your work:

1. AI bias due to the training data being used. This is very important as you will more than likely create your own AI solution and might not realize the inevitable bias you will inject in your models.
2. What parts of the AI design and usage truly unique and therefore needing trade secret protection?