

**USPTO Request for Comments on Intellectual Property Protection
for Artificial Intelligence Innovation - Public Comment by the
Electronic Frontier Foundation**

January 10, 2020

Kit Walsh
Electronic Frontier Foundation
815 Eddy Street
San Francisco, CA 94109
(415) 436-9333
kit@eff.org

The Electronic Frontier Foundation (EFF) is grateful for the opportunity to comment on the impact of artificial intelligence (AI) on intellectual property rights. In this comment, we highlight the importance of fair use in promoting the public interest and detail how existing law applies to AI.

EFF is the leading nonprofit organization defending civil liberties in the digital world, with over 34,000 active donors. For nearly thirty years, EFF has worked to ensure that technology and technology law enhance liberty rather than restraining it. A critical element of this work is advocating for intellectual property laws to serve the public interest rather than maximizing the control and rents for rightsholders to the detriment of both innovation and the general public.

Technical Background: Relevant Elements of an AI System

Artificial Intelligence is a broad and imprecise term. These comments use “AI” to refer to machine learning systems and to different components of a machine learning system as appropriate.

In general terms, the relevant components are:

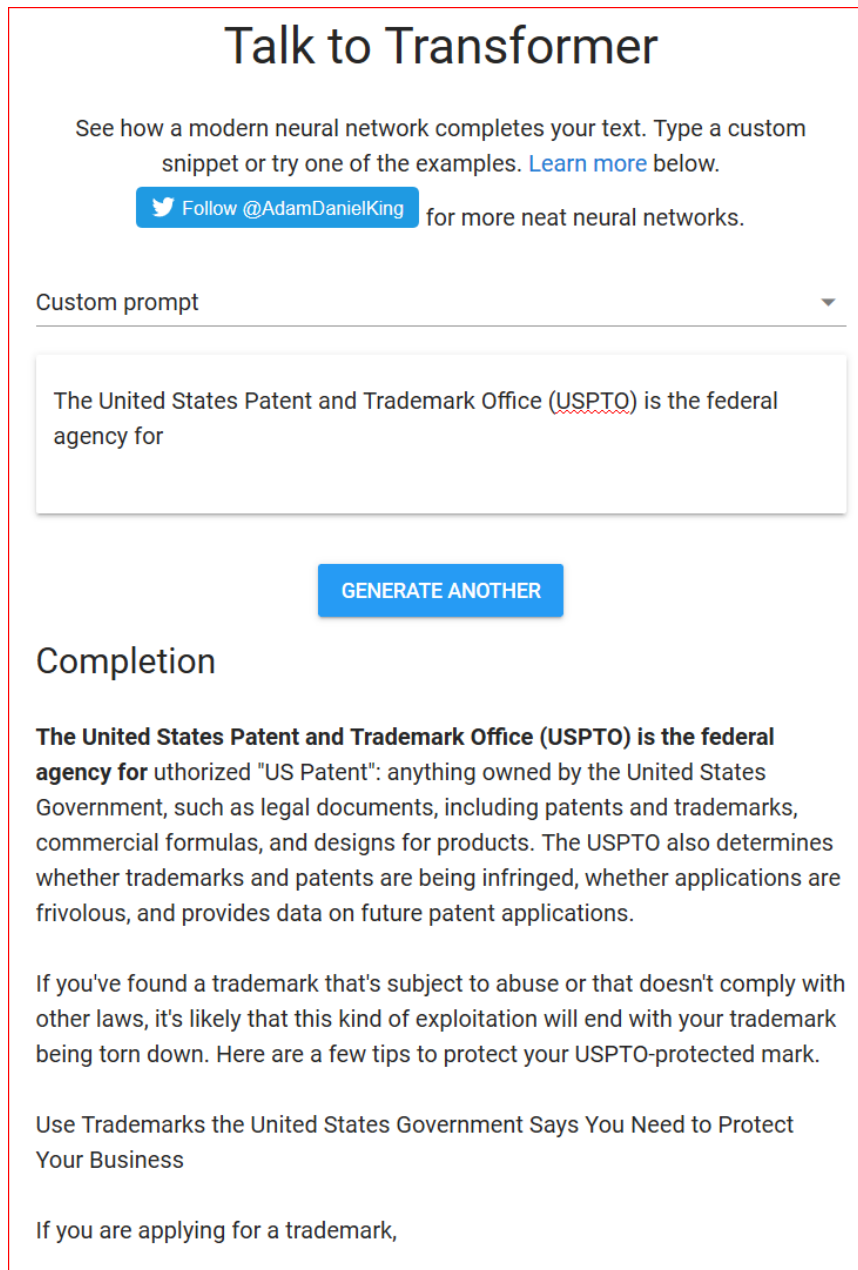
1. *Training data* - any kind of digital information that can be processed by a computer. Sometimes this data includes tags applied through human judgment (for instance, when you specify whether or not an image contains a car or traffic light).
2. *Machine learning algorithms or training algorithms and software implementing those algorithms* - software instructs the computer in how to analyze the training data. In some cases, humans conduct testing and tweaking of parameters of the software implementing the training algorithm to optimize the models that result.
3. *Trained model* - a description of statistical relationships that results from operating the machine learning software on the training data (i.e. ‘training’ the system).

4. *Outputs* - the end result when the model is executed based on new inputs or a random seed. This may be as simple as a prediction that it will rain, or as detailed as an original image that the model would categorize as a cat.¹

¹ See Tangerman, "A New AI Draws Cats, and They're Utterly Grotesque," available at <https://futurism.com/cat-doesnt-exist-ai>.

Technical Background: An Example

To illustrate these elements, consider talktotransformer.com: a website that invites regular Internet users to type in a prompt, which will be operated on by a trained model in order to generate a new paragraph of text.



The screenshot shows the 'Talk to Transformer' website interface. At the top, the title 'Talk to Transformer' is displayed. Below it, a description reads: 'See how a modern neural network completes your text. Type a custom snippet or try one of the examples. [Learn more](#) below.' A blue button with a Twitter icon says 'Follow @AdamDanielKing for more neat neural networks.' Below this is a 'Custom prompt' section with a dropdown arrow. The prompt text is: 'The United States Patent and Trademark Office (USPTO) is the federal agency for'. A blue button labeled 'GENERATE ANOTHER' is positioned below the prompt. Underneath, the 'Completion' section shows the generated text: 'The United States Patent and Trademark Office (USPTO) is the federal agency for uthorized "US Patent": anything owned by the United States Government, such as legal documents, including patents and trademarks, commercial formulas, and designs for products. The USPTO also determines whether trademarks and patents are being infringed, whether applications are frivolous, and provides data on future patent applications. If you've found a trademark that's subject to abuse or that doesn't comply with other laws, it's likely that this kind of exploitation will end with your trademark being torn down. Here are a few tips to protect your USPTO-protected mark. Use Trademarks the United States Government Says You Need to Protect Your Business If you are applying for a trademark,'

Figure 1: A neural network trained on millions of webpages, operating on a sentence fragment to generate several sentences of predicted text. A screen-reader-compatible version is provided in the addendum.

In this case, the “training data” consisted of millions of webpages, which were screened for quality by only selecting pages linked from reddit.com with a certain level of quality (‘karma,’ generated by the voting of reddit users).²

Researchers at OpenAI generated and tested a series of models using a set of new training algorithms, altering the parameters of the training to optimize for different features. This set of training algorithms, and the resulting trained models are called “GPT-2.”

A company called Hugging Face released a software implementation of GPT-2’s trained model, which was modified by Adam King, the creator of talktotransformer.com.

When a person visits the site, they select or create input text which is then transformed using the site’s implementation of the GPT-2 model. The software generates a paragraph of text based on the input text by reference to the model’s distillation of observed structure and probability from the millions of webpages that made up the training data.

The text that is generated typically follows some topical themes from the input text, and typically is text that has never been seen before (i.e. it is not a quotation from any element of the training data).³

The output’s blend of sense and nonsense has an aesthetic appeal that has led to the creation of new works created using talktotransformer.com. One example is *Neuralwheel* by Janie Jaffe W, a deck of storytelling game prompts with evocative and sometimes surreal verbal imagery created using Talk to Transformer.⁴

2 OpenAI, “Better Language Models,” at fn1, available at <https://openai.com/blog/better-language-models/#fn1>.

3 Radford et al., “Language Models are Unsupervised Multitask Learners,” Section 4, available at https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf.

4 Available at <https://janiejw.itch.io/neuralwheel>.

Responses to the USPTO's Specific Questions 1-4

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?*

It is unclear what real-world use case would give rise to this hypothetical, but as a theoretical matter if no natural person makes any copyrightable contribution to a work then it falls within the public domain.

In reality, any work generated using an AI process is the result of human interventions. Given a proper understanding of AI, no new copyright law is necessary to analyze authorship.⁵ In the hypothetical case where no human creativity is involved, not even in the creative selection of rules to be followed by a machine, then there is no copyrightable work of authorship. *See Bridgeman Art Library, Ltd. v. Corel Corp.*, 25 F. Supp. 2d 421, 426-27 (2d Cir. 1998) (applying US precedent regarding originality).

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 (i) designed the AI algorithm or process that created the work; (ii) contributed to the design of the algorithm or process; (iii) chose data used by the algorithm for training or otherwise; (iv) 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"?*

The software used to train a machine learning system may be copyrightable to the same extent as any other work of software.

The statistical relationships observed by a machine learning system, making up the trained 'model,' are a purely factual analysis of the features observed in the input data. They are, therefore, unlikely to include copyrightable, creative elements. A particular software implementation of a model may be copyrightable to the same extent as any software implementing uncopyrightable statistical formulas.

When a model is used to transform a particular set of input data into a new output, authorship will depend on the facts of each case. There are at least three categories of potential authors: (1) authors of works used

5 Grimmelmann, James, *There's No Such Thing as a Computer-Authored Work - And It's a Good Thing, Too* (December 6, 2015). 39 *Columbia Journal of Law & the Arts* 403 (2016); U of Maryland Legal Studies Research Paper No. 2016-06. Available at SSRN: <https://ssrn.com/abstract=2699862>.

as training data, (2) authors creatively selecting training parameters, and (3) authors making creative choices of input data and parameters for the trained model (if any). (If a person in category (3) selects someone else's copyrightable work as input data, a fourth potential author is involved).

At one extreme, imagine that the set of training data is extremely small and the training is simplistic, such that the system winds up essentially reproducing one of the input works. Absent fair use, this is likely to be an infringing use, and the author of the original work is the only 'author' as far as copyright is concerned.

If, however, the system makes sufficient changes that the final product is a derivative work, then the traditional principles of authorship apply. The original author is the author of their work, and the subsequent contributors are the authors of their contributions.

Another possibility, using this simplistic machine learning system, is that the contributors work together sufficiently to make them all joint authors.

More realistic machine learning programs are far less likely to create outputs that are closely-matched to any single copyrightable work in the training data. In this case, if there is no subsequent creative step in the selection of inputs and parameters for the trained model, then either the developers of the system made creative contributions, in which case they are the sole authors of the resulting work, or they did not, in which case there is no copyright in the resulting work.

Finally, the person authoring or selecting inputs for a trained model, such as a user of "Talk to Transformer," is another potential author of the ultimate output. To the extent that the selection of input is overwhelmingly the determining factor in the ultimate output, then that user is likely to be a sole author of that output. If, however, the *copyrightable*, creative choices of the trainers of the model are heavily reflected in the output, they may be joint authors. This is, again, a fact-specific inquiry, but it is not so novel that existing law is unprepared to adjudicate it.

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?

The law of fair use recognizes that it is transformative to use copyrighted works for analysis of their contents or functional characteristics. *E.g.* *Authors Guild v. Google Inc.*, 804 F.3d 202 (2d Cir. 2015); *Sega Enters.*

v. Accolade, Inc., 977 F.2d 1510, 1524 (9th Cir. 1992). This policy promotes the constitutionally-mandated purpose of copyright law: to promote the progress of science and the useful arts. Machine learning systems ranging from text prediction to search engines could not exist if they needed to license each element of training data.

When a machine learning system ingests large numbers of works, it is both unlikely that any individual work has a great impact on the trained model and impractical to identify all authors.

4. Are current laws for assigning liability for copyright infringement adequate to address a situation in which an AI process creates a work that infringes a copyrighted work?

Existing law, correctly applied to the technology at issue, can determine whether someone has infringed copyright using an AI system.

First of all, AI training software and trained models typically are capable of substantial noninfringing uses and publishers of these works therefore are not contributory infringers when a third party employs those works in the course of infringement. *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417, 442 (1984).

The possibility of direct infringement requires the application of other well-established doctrines.

First of all, since actual copying is required for infringement, the use of AI trained on a limited universe of works creates new opportunities for precise determination of whether copying in fact occurred. Presumptions based on access or publication are less significant to the extent that the complete universe of inputs can be cataloged and potentially said to exclude a given work.

Second, since machine learning is a process of identifying statistical patterns, these tools may provide new ways of identifying patterns that constitute *scènes-à-faire*. This is particularly so when the tools are trained on large bodies of work with multiple contributing authors, such as an entire genre, or publications from a given time period. To the extent that such systems' *outputs* reproduce these common patterns, the existing law of *scènes-à-faire* and merger will help inform courts as to whether a work that winds up being substantially similar to one of the input works is, in fact, similar with respect to copyrightable elements, or if the similarities are a result of the factual observations made by the machine learning system.

Third, to the extent that a work is produced with a machine learning tool that was trained on a large number of copyrighted works, the degree of copying with respect to any given work is likely to be, at most, *de minimis*.

And, of course, fair use protects the creation of many works generated using AI tools that might otherwise be infringing derivative works. The analysis will necessarily be fact-specific, but it should be informed by the considerations above.

Addendum: Screen-Reader-Compatible Description of Figure 1 and Its Content

The figure depicts the interface of “Talk to Transformer,” where the input prompt reads “The United States Patent and Trademark Office (USPTO) is the federal agency for.”

The system generated the following continuation of that text:
The United States Patent and Trademark Office (USPTO) is the federal agency for uthorized ‘US Patent’: anything owned by the United States Government, such as legal documents, including patents and trademarks, commercial formulas, and designs for products. The USPTO also determines whether trademarks and patents are being infringed, whether applications are frivolous, and provides data on future patent applications.

If you've found a trademark that's subject to abuse or that doesn't comply with other laws, it's likely that this kind of exploitation will end with your trademark being torn down. Here are a few tips to protect your USPTO-protected mark.

Use Trademarks the United States Government Says You Need to Protect Your Business

If you are applying for a trademark,

(The provided text ends with that incomplete sentence).