October 30, 2023

Re: Artificial Intelligence Study (88 FR 59942)

I welcome the opportunity to submit the following comments in response to the notice of inquiry published by the U.S. Copyright Office (the "Office") in the Federal Register on August 30, 2023, regarding artificial intelligence (AI) and copyright.

I approach this topic from the perspective of being a published author (reg. no. TX0006970593), a software developer with an emphasis on building artificial intelligence systems, and an intellectual property lawyer specialized in the copyright issues associated with computers. As a "content creator," machine learning developer, and attorney, I have experience with many facets of AI that may not be apparent to those without the same background.

One of the features of the US legal system is that the law is never analyzed in a vacuum. Legal opinions start by discussing the relevant facts of a case. Based on these facts, legal principles from previous cases are applied using logic and analogy, extending the law to new circumstances.

Applying copyright law to machine learning should follow the same process. A technology-first analysis of *each* different phase of AI model training and use is necessary to understand the relevant facts for the legal analysis. As highlighted in the U.S. Supreme Court's recent decision in *Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith et al.,* ("*Warhol*"),[1] some uses (like the licensing of Orange Prince for a magazine cover) may be infringing, whereas other uses (such as creating, displaying, or selling the Orange Prince work) may be considered fair use.[2]

Especially in complex areas of the law, the legal analysis can be heavily influenced by the metaphors chosen to simplify and highlight the relevant questions. But creating those metaphors and performing legal analysis without a correct understanding of the actual processes can lead to inconsistent and incorrect conclusions.

[1] Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith et al., ______ U.S. ______, ______ S. Ct._____ (2023), slip. op available at https://www.supremecourt.gov/opinions/22pdf/21-869_87ad.pdf.

[2] *Id*. at 21 (Only [AWF's commercial licensing of Orange Prince to Condé Nast] . . . is alleged to be infringing. We limit our analysis accordingly. In particular, the Court expresses no opinion as to the creation, display, or sale of any of the original Prince Series works"). *See also* fn. 10 ("Congress has directed courts to examine the purpose and character of the challenged 'use.' . . . Had AWF's use been solely for teaching purposes, that clearly would affect the analysis").

Answers to Questions

Question 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?*

Generative AI is the first *general-purpose* software tool we have developed. But unlike previous general purpose technologies[3] that replaced physical work, generative AI replaces mental work. Accordingly, we should expect disruption in fields that rely on mental capabilities, even as we see greater levels of creation and creativity.

This puts creators and copyright owners at severe risk. Not because AI systems are being trained on works already created, but because creators are heading into a period of disruption and competition **without having enforceable copyrights available in works that they create using AI tools**.[4] This is the greatest risk to the copyright system and the economic force it enables.

The history of letterpress printing provides a likely scenario. Once ubiquitous, letterpress printing fell out of favor in the 1970s because of the rise of computerized printing methods. Computers made printed material cheaper and more abundant.

From the perspective of the letterpress industry, computerized printing was a disaster. Thousands of print shops closed. But from the perspective of society, however, computerized printing was amazingly positive. The amount and quality of printed material increased. New people were enabled to create works that would not have been created without these new tools.

Letterpress printing did not go away. Today it has been revitalized as a specialty art form and high-end customized product. Humans can deliver something that couldn't be replicated in a computer printout.

Knowledge and creative workers are now facing some of the same pressures that letterpress printers did in the 1970s and 1980s. There will be a consolidation and a reduction of some types of jobs that currently exist. On the other hand, there will be an explosion of new works as AI tools enable more people to create works and make new types of works possible, such as

---

[3] Bresnahan, T.F. and Trajtenberg, M. General purpose technologies 'Engines of growth'? *J. Econom. 65*, 1 (Jan. 1995), 83–108, 83. Available at https://doi.org/10.1016/0304-4076(94)01598-T ("Whole eras of technical progress and growth appear to be driven by a few 'General Purpose Technologies' . . . such as the steam engine, the electric motor, and semiconductors.")

[4] This issue is discussed in detail in the answer to question 18.

image-QR code hybrids that would be extremely difficult for people to create unaided by tools. But without *prima facie* protection for AI-assisted works, we will be failing to secure for these authors the exclusive right to their works.[5]

Question 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?*

The software industry is arguably the place where the professional use of AI tools is the most widespread. Current estimates are that more than 90% of the companies building software have evaluated using AI-assisted coding tools (like GitHub Copilot, Amazon CodeWhisperer, etc.) and are planning on making some sort of AI assistance available as a regular tool. Even in organizations where AI coding assistants are not officially recognized, AI-generated code is coming "in through the bottom" based upon the independent choice of software developers.

This gives the software industry a unique view on how AI systems affect an industry. The initial responses are positive:

- **AI-assisted coding tools enhance, not replace, human expertise.** AI-assisted coding tools have not removed the need for skilled humans. Knowledge and domain expertise are needed to guide the tools and make sure that their outputs are not only correct, but actually solve the underlying problems that lead to the need for software.

- **More productive software development leads to induced demand**. It is possible that AI-assisted coding tools will reduce the number of software developers needed overall, but so far the experience has been the opposite: the effective cost for software development is going down, leading to constant or increasing need for skilled human developers. This is consistent with earlier experience in the software industry, when developer tools first developed limited code generation facilities.

- **AI-assisted coding tools remove the least-fulfilling aspects of software development**. Every job has some parts that are engrossing and rewarding, and some parts that are difficult and tedious. AI-assisted coding tools perform their best with difficult and tedious work. They do not do well with the more engrossing and creative problem-solving that humans do. One study found that between 60–75% of users reported they felt more fulfilled with their job, were

[5] Article I, Section 8, Clause 8 ("[The Congress shall have Power . . .] To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.").

> less frustrated when coding, and are able to focus on more satisfying work when using AI-assisted coding tools.[6]

While not every industry will be affected the same way, it is reasonable to think that some of these positive results will also be reflected in other creative industries.

There have been some offsetting negative effects. Companies are needing to perform “snippet scans” to avoid inadvertent copyright infringement, and AI tools are a new vector for confidential data loss.

One unnecessary negative effect is the need for human revision of even correct code. Because the current policy of the Office is to only recognize copyright in post-generation human activities, some companies are creating policies that require developers to revise all code that is AI-generated, even if it is already correct. **This is a frustrating task performed solely to ensure the copyrightability of the code in light of the Office’s current policy regarding AI-assisted works**.

Question 3. *Please identify any papers or studies that you believe are relevant to this Notice. These may address, for example, the economic effects of generative AI on the creative industries or how different licensing regimes do or could operate to remunerate copyright owners and/or creators for the use of their works in training AI models. The Office requests that commenters provide a hyperlink to the identified papers.*

Lindberg, Van, *Building and Using Generative Models Under US Copyright Law* (May 30, 2023). 18 Rutgers Bus. L.R. No. 2 (2023)., Available at SSRN: https://ssrn.com/abstract=4464001.

This paper is consistent with these comments but includes additional information and analysis not included here.

Question 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.*

U.S. copyright law is robust and has reacted appropriately to many changes in technology. For the most part, no changes are needed. The current law is the result of a careful balance between creators’ interests and the public interest. That balance must be maintained.

---

[6] GitHub, *Research: Quantifying GitHub Copilot’s impact on code quality*, GitHub Blog (Oct. 10, 2023), https://github.blog/2023-10-10-research-quantifying-github-copilots-impact-on-code-quality/

If any changes are contemplated, they must be equally evenhanded. For example, one change that might resolve the tension between copyright owners and AI system developers is a statutory declaration that training of AI systems for purposes of generative AI systems is a reserved right under copyright, paired with the introduction of a digital deposit requirement for registered works and the imposition of a periodic copyright maintenance fee for copyright registrations. This would preserve the copyright value of recent and commercially valuable works, while simultaneously expanding and strengthening the public domain, making most AI model training and research possible.

Separately, it may be helpful to consider the protection of AI model "weights" – the tables of statistical probabilities that result from the AI model training procedure. Under current U.S. intellectual property law, public use of models makes them likely unprotectable:

- Under copyright, these weights are not chosen by a human. They are facts–statistical probabilities–generated by the observations during training. Further, the arrangement and content of the weights are not the result of human decision-making or creativity. They are the result of a fixed mechanical process and are, at present, uninterpretable to humans.

- Under patent, these weights do not fall into any protected statutory class. They are unpatentable as printed matter.

- The weights could be a trade secret, but public use of an AI system is known to reveal the underlying system weights over time, making the trade secret discoverable by others even without direct disclosure.

Despite the possible erosion of the trade secret, most weights are currently kept confidential. To encourage the public dissemination of AI model weights, it may be helpful to statutorily create quasi-copyright protections for the weights, similar to the way the Vessel Hull Design Protection Act and Semiconductor Chip Protection Act enabled protection for boat hulls and mask works.

I encourage the Office to consider and request comment on whether a quasi-copyright right in AI model weights would be helpful and what it would include.

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

Any AI system that generates possibly-copyrightable material was trained on copyrighted material. AI systems learn about language by hearing and reading language, just like humans. AI systems learn about art by observing art, just like humans. AI systems learn about music by hearing and reading music, just like humans. There is no other way for the system to "learn."

What is unknown is the amount of registered material used in the training of various AI systems. Copyright vests in creative expressions immediately upon fixation–but only a very small amount of all copyrighted material is registered. And some material that was previously registered is now in the public domain.

Question 6.1. *How or where do developers of AI models acquire the materials or datasets that their models are trained on? To what extent is training material first collected by third-party entities (such as academic researchers or private companies)?*

Historically, most datasets used for AI model training have been put together by research teams attempting to learn about different aspects of AI systems. Papers with Code[7] a repository of scholarly papers with matching technical data, lists 8694 different datasets created by researchers, covering images, text, audio, video, medical, 3D, and graph data.

The greatest source of this training material is unregistered, permissively licensed, or public domain material from the Internet. For example, essentially every large language model includes Reddit posts (unregistered text), Wikipedia (licensed under the Creative Commons Attribution-ShareAlike 4.0 license), and texts from Project Gutenberg (public domain), While registered material is also used in training AI systems, it is frequently due to that material being posted on the Internet alongside unregistered material.

For example:

- ImageNet[8] and CIFAR-10[9] were developed to help scientists develop and test image recognition systems. They have each since been cited more than 25,000 times in other papers.

- More recently, image datasets like LAION-5B[10] have not included the images themselves. Instead, these datasets consist of URLs to publicly-posted images with explanatory text.

- The Common Crawl[11] is the predominant source for the text used in every large language model. It is a public dataset of 250 billion web

---

[7] https://paperswithcode.com/

[8] J. Deng et al., "*ImageNet: A large-scale hierarchical image database,*" 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.

[9] Krizhevsky, Alex. "*Learning Multiple Layers of Features from Tiny Images*" Canadian Institute for Advanced Research (2009).

[10] Schuhmann et al., "*LAION-5B: An open large-scale dataset for training next generation image-text models,*" 36th Conference on Neural Information Processing Systems (NeurIPS 2022).

[11] https://commoncrawl.org/

pages spanning 16 years curated by the CommonCraw Foundation, a 501(c)(3) nonprofit organized to enable researchers to understand and analyze language.

Many new datasets are being created by compiling, converting, and annotating previously available training material into new, larger datasets.

Question 6.2. *To what extent are copyrighted works licensed from copyright owners for use as training materials? To your knowledge, what licensing models are currently being offered and used?*

There have traditionally been a few providers of training materials for use in AI systems. Training materials licensed from data providers are typically of higher quality or more specialized than the broadly-sweeping data most used for machine learning. However, the cost associated with these specialized datasets has historically precluded their use in machine learning research.

With the recent interest in AI, many companies are examining how to provide copyrighted, previously-developed content as training materials for AI systems. Usually, however, access to the machine-readable versions suitable for AI model training frequently include contractual clauses prohibiting the distribution of weights resulting from using the copyrighted training material, making these services a “dead end” for many types of collaboration or later development by third parties.

Finally, some copyright owners are trying to categorically restrict the use of their content for machine learning. For example, the New York Times recently updated its terms of service to forbid the scraping of its content to train a machine learning or AI system. This includes copyrightable text and images, but also restricts information that is not protectable under current copyright law, including “look and feel,” and metadata, such as author attribution.

Question 6.3. *To what extent is non-copyrighted material (such as public domain works) used for AI model training? Alternatively, to what extent is training material created or commissioned by developers of AI models?*

Public domain works are used as much as possible, provided that they have been gathered together and are accessible. For example, the public domain Project Gutenberg materials are used in the training of every large language model. Music models almost always include a wide range of classical music. However, public domain works are not as comprehensively digitized as works that have been created more recently.

While training material is sometimes created by AI researchers, the volume of training material needed for modern AI systems usually makes direct creation impractical.

Question 6.4. *Are some or all training materials retained by developers of AI models after training is complete, and for what purpose(s)? Please describe any relevant storage and retention practices.*

Retention practices vary widely. Datasets are large and expensive to retain. If the dataset is easily available from a third party, they are more likely to delete the training materials. On the other hand, developers are more likely to retain materials if they anticipate further training or a technical need to evaluate the previous training.

To the extent that training materials are retained, it is most similar to a local backup of the training materials than any other use.

Question 7. *To the extent that it informs your views, please briefly describe your personal knowledge of the process by which AI models are trained.*

I have personally built and trained machine learning models for the past fifteen years, with a focus on patent documents as a training corpus. The majority of those models were classification models, but since 2017 I have been creating generative models, including interactive models similar to ChatGPT. Additionally, I read as much of the technical literature as I can and am reasonably familiar with the state of the art.

This section aims to provide an accurate and easily understandable description of the mechanics of machine learning at a level of detail appropriate for legal analysis.

The creation of a generative AI system involves two phases. The first phase is the training of a "model". The second phase is using the “model” to make new outputs, such as new sentences, pictures, or code. These phases need to be examined separately because training and generation happen at different times, usually by different parties, and they involve different outputs.

In addition, the model itself is an independent object, different in kind than both the training material and the generated output.

Two Analogies for AI Training

One persistent misunderstanding some people have is how AI applications can recreate familiar objects. These people think of a machine learning model as just a complicated type of storage that saves everything it sees and then brings forth bits and pieces of memorized material to mash together into a collage.[12] In contrast, the power of machine learning is that it helps the computer identify meaningful correlations that are too attenuated or esoteric to be expressed by software developers. In other words, the model isn’t

[12] *See, e.g.*, Anderson v. Stability., No. 3:23-cv-00201, at *3 (N.D. Cal. filed Jan. 13, 2023) (describing Stable Diffusion as “merely a complex collage tool.”).

memorizing the copyrightable expression in an input. Rather, it is evaluating and recording factual relationships between different elements of the expression.

Before diving into the mechanics of AI system training, there are two analogies that may be helpful in developing a mental model of how training works: the art inspector and the law student. Both of these analogies illustrate the mechanics of model training, but in slightly different ways.

The Art Inspector

Imagine a newly hired art inspector given the job of examining every painting in the Louvre. This inspector has no background or experience in art and so has no preconceived ideas about art (what's beautiful or repugnant) or what is significant about any particular painting (what makes a Picasso a Picasso).

Lacking any guidance, the inspector studies each painting by measuring everything about it, such as the number of brushstrokes, paint thickness, average space between brushstrokes, size of the painting, and the thickness of lines. He includes *every* piece of information he can–the age of the painter, the date the painting was made, and in which corner the artist signed their name. The inspector measures aspects of the paintings that seem bizarrely random or unimportant, such as the number of consonants in the artist's name and the relationship between colors that are six inches apart. He is meticulous in his approach. Nothing is left untouched in the exhaustive analysis. Everything is recorded in the inspector's database.

As the inspector studies each painting, he tries to make his job more interesting by turning each measurement into a guessing game. Before he makes each measurement, he tries to predict what the answer will be, using the information he has gathered already. "How many brushstrokes are in this painting?" he wonders. "Well, it's a Rembrandt from the middle third of his career. I'd guess... 84 brushstrokes per square inch." The inspector then checks the measurement and records how good his prediction was before moving on to the next measurement and the next prediction. When the inspector begins to play, his answers are usually wrong. But as he takes more and more measurements, his predictions are increasingly correct.

After studying thousands (or millions) of paintings, the inspector is the world's foremost authority on validating paintings. He is regularly asked his opinion as to whether various newly-discovered paintings are legitimate. His ability to pinpoint which artist created a painting and to predict other things about each painting is unparalleled. Where before, the inspector took the artist's name and information to predict the measurements of their paintings, now the inspector uses the measurements to predict the painter.

The Law Student

When a student begins law school, they are frequently told that their job isn't just to learn the law–their job is to learn how to "think like a lawyer." As a result, legal teaching is structured differently than many other types of professional training. Learning the rules isn't enough; they must learn how to apply the law to new situations.

One common way of teaching legal reasoning is the *case method*. The case method involves studying judicial decisions, or cases, in order to understand the legal principles and rules that govern a particular area of law. Rather than simply memorizing legal rules and statutes, students learn to analyze and apply the law through a close examination of real-life legal disputes.

In practice, law students are given a set of facts, usually from a court decision, and then asked to analyze the legal issues raised. They look at the relevant laws, the arguments made by the parties, and the reasoning behind the decision. By examining the evidence and arguments, students develop their own understanding of the legal principles at play.

Law professors usually pair the case method with the Socratic method. In the Socratic method, the professor asks questions instead of providing answers. As the law students struggle to imitate previous "correct" answers to similar questions, they begin to derive legal principles from the various scenarios. The professor provides feedback–validation of a correct analysis, or correction of a wrong answer–which the students then use to further refine their understanding.

When the time comes for the exam, a successful law student is able to take a hypothetical situation and generate a new analysis that nevertheless incorporates the correct principles, even though the student was never *explicitly* taught which principles to use. The student may not have a specific reason to weigh one factor over another, emphasize certain facts, or avoid certain arguments. She just knows, based on evaluation of example cases, how courts have weighted various facts and principles in the past.

In contrast, imagine a second student who attempts to master the material by memorizing all the facts and holdings from every case discussed in class. This second student does well when asked to describe the facts and holding of an important case, but fails to apply the principles of the law to new situations.

In short, the successful law student has a mental model of how the law is "supposed" to work based upon her analysis of the many cases studied during the class. Unlike the second student who just memorized facts, she can predict how courts would analyze new facts and new situations. She has learned to "think like a lawyer."

Examining the Two Analogies

Training an AI model is similar to the processes of the art inspector and the law student. In both cases, the basic steps are the same: *receive* an example; *predict* the relationship between the different elements of the example; *check* the result, and *adjust* to improve future predictions. Those commonalities apply to the mechanical process performed by a computer during model training. However, there are some individual points from each analogy worth emphasizing.

The art inspector analogy is better at showing how AI models start with a clean slate. Before models are trained they are literally filled with random data. All associations that come out of AI applications were learned by observation. The art inspector's measurement of small, random details is also closer to the fixed process that occurs during AI model training. However, although the art inspector recorded and saved of all his measurements (inputs), what is actually recorded during AI model training is instead the changing *probabilities* associated with different inputs.

The law student analogy is better at showing how unifying principles are the product of inference. Just like the law student is never explicitly taught the correct legal principles, AI model training processes are not instructed what any of their inputs "mean." Instead, the "meaning" that is observed in an AI application is actually a complex probability function with millions or billions of parameters.

The law student analogy also demonstrates how direct memorization of inputs is actually antithetical to the goals of model training. Avoiding direct memorization is so important that AI model training processes almost universally involve removing part of the training information to force the model to engage in the inference process. This is sometimes referred to as "masking" or "dropout." Failure to hide or remove information during training makes models unusable.

The danger in the law student analogy is that it makes the AI application easier to anthropomorphize. The process of training is not creative or selective. The AI models do not "think" or analogize. All the "learning" that occurs within the AI application is simply the rote construction or use of massively complex probability functions.

The Training Process

With these analogies in mind, we can examine how the same steps–receive, predict, check, adjust–are used to train an AI model. However, while the steps are conceptually similar to the mental processes of the art inspector and the law student, they are turned into a mechanical process that can be repeatedly performed by a computer.

Creating a “Brain”: the Architecture of Machine Learning Models

Because computers do not have brains and senses like humans, the first step is to create a logical “brain”–a structure that can receive and process input. This structure is sometimes referred to as the “architecture” of the AI application.

To build a model, a data scientist begins by defining a logical structure for processing inputs to create outputs. Each part of the training process corresponds to a different part of the structure. These structures–initially inspired by the interconnections between brain cells–are called "neural networks." There are many different types of neural networks, but they share three general structures: an input layer, one or more "hidden" layers, and an output layer. These layers are made up of “nodes”–logical structures where values are temporarily stored and computation can occur. These nodes are highly interconnected by logical paths to other nodes. A stylized illustration is shown in Fig. 1.

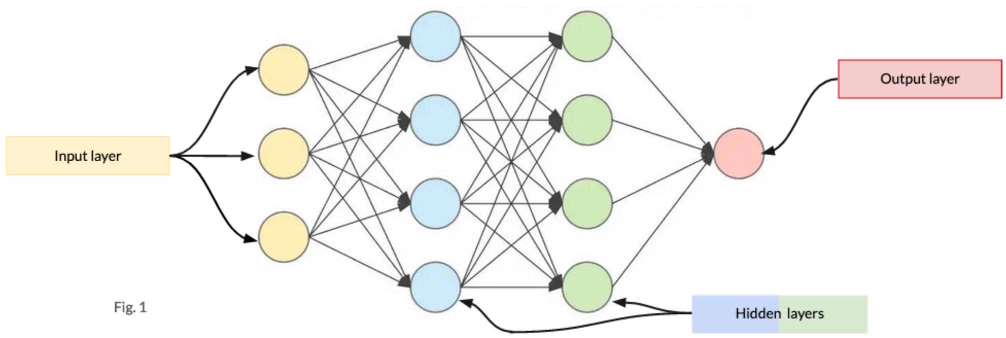

Fig. 1

This stylized figure has three nodes in the input layer, eight nodes in two hidden layers, and one node in the output layer. Different AI applications can have different numbers of nodes in each layer and can have many different types of interconnections.

The Input Layer (Receive)

The input layer of a neural network is where the data is provided to the model. It is similar to the art inspector viewing a painting or the law student reading a case.

Unlike humans, who can process whole pictures or cases at a time, computers are more limited. Each node in the input layer has a memory designed to receive a single element of the input data. The goal of the input layer is to provide a uniform representation of the raw data that the model will use to make predictions.

As humans, we might think about these inputs as representing the pixels in an image or words on a page. However, from the computer's point of view, the input is just a list of numerical values called a "vector." For example, in a model that processes images, each node in the input layer just gets a number. Depending on the application, the number could represent the brightness of one part of an individual pixel. In a model that processes text, each node in the input layer might receive a value that represents a word or character. See Fig. 2.

The Hidden Layers (Predict)

The hidden layers in a neural network are where the majority of the processing occurs in an AI application. These layers are called "hidden" because the data that is processed within them is not directly observable from the inputs or outputs of the model. The hidden layers contain a series of interconnected nodes, each of which performs a mathematical calculation on the inputs received from the previous layer. After performing the calculation, the node can pass forward the same value, a changed value, or nothing at all.

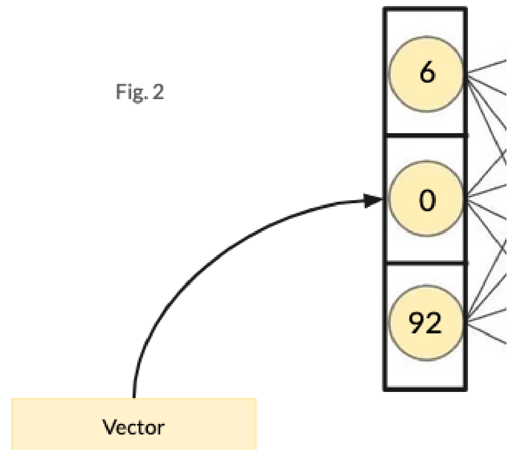

Fig. 2

Each hidden node has an associated "weight" that changes the probability that a value will be changed or passed on. The weight corresponds to the model's "best guess" as to how the inputs should be used. See Fig. 3.

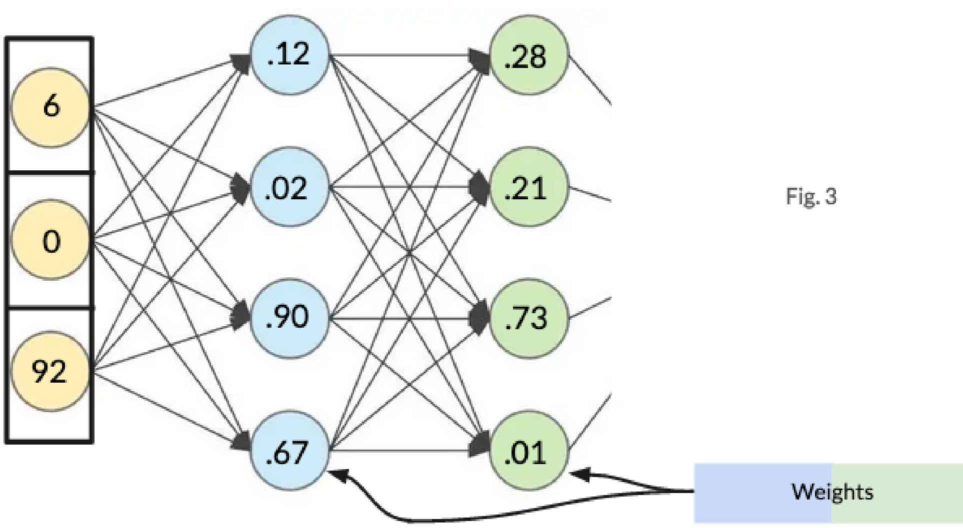

Fig. 3

Similar to the art inspector, scientists have no idea what information will end up being important for the evolution of the neural network. In the past, data scientists tried to identify specific "features" of the input data that they would provide to the neural network. However, isolating the right features took time, was error-prone, and didn't work as well. The current trend is simply to provide *all* of the data to the neural network and let the computer identify which correlations are useful. As a result, the correlations developed during the training process can be unexpected.

Despite the identification of these correlations, neither the “style” nor the content of an image is saved as part of the model during training. Like the law student that extracts principles taught in court cases, the model extracts correlations that, to humans, resemble certain artistic styles.

Applying this more concretely to the functioning of a real AI application, a common use of some AI applications is to render pictures as if they had the unique brushwork of Vincent van Gogh. This is called “Neural Style Transfer.” However, to perform this function, it would be counterproductive to save the brushstroke pattern for any existing van Gogh painting–none of the brushstrokes would fit a different image. Instead, it appears that one or more layers of the neural style transfer model contain a function that spreads out, moves, or changes the input values associated with each pixel in a way that creates a result that, to humans, resembles the brushwork style of van Gogh.

For humans, we might describe this process of creating a painting with van Gogh-style brushwork as “applying each color in a thick, contoured slab of paint.” But for a model, this becomes an instruction like “for every pixel of blue, adjust the blue values of all surrounding pixels by an amount corresponding to this equation.” It involves none of the creativity and judgment that a human artist would need to use to apply the technique. Instead, it is “just” an evolved probability function that is literally inhuman in its complexity. Stable Diffusion, a modern model for processing and generating images, has 890 million parameters. GPT3, a model for understanding natural language, has 175 billion. Each parameter can be thought of as a conditional probability associated with one possible state of the neural network.

The Output Layer (Check)

The output of the hidden layers are then passed to the output layer, where the final result is provided. Just as with the input layers, the AI application doesn’t “know” what the output represents. Like the input, it is just a vector of numbers. See Fig. 5.

How the result is interpreted is up to the human user. For example, the example value .82 in Fig. 5 could be interpreted as a classification (“there is an 82% chance this email is spam”), a recommendation (“if you liked this show, there is an 82% chance you might also like…”), or in the case of generative AI, one part of a pixel value of a new picture (change the blue in this output pixel to 82% of

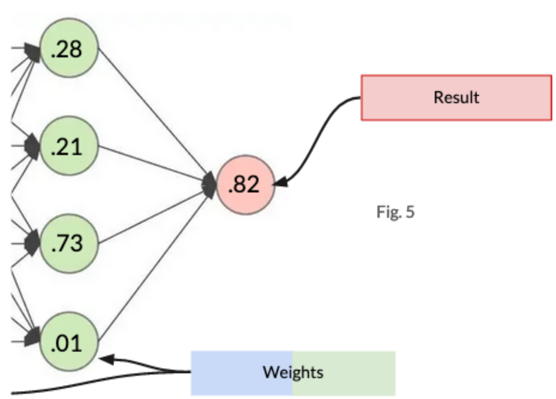

Fig. 5

its maximum value) or the next word of a new sentence (the number .82 corresponds to the word "elephant").

Update the Weights (Adjust)

During training, every input has a known correct output (or possible output, if there are multiple correct possibilities). The result is compared with the correct output and the weights in the neural network are adjusted a tiny bit so that the next time the model receives a similar input, it will be more likely to provide a similar answer.

Try Again (Repeat)

This same process is then repeated. After processing and recording the predicted probabilities over the millions of provided examples, the application builds a comprehensive statistical picture of the range of possible answers for any given input and the likelihood of each answer. In many cases, the same inputs are re-used in different rounds of training to see if there are any further statistical correlations that can be learned from each example.

A portion (10-20%) of the input examples are never used as training inputs but are instead saved as a "testing" set. The testing set is never used as part of the training. The application receives the training input, predicts the output, and checks it against a known correct answer. However, the differences between the prediction and the correct answer are never used to adjust the weights in the model. Instead, the performance of the model is evaluated by using these never-used inputs as a barometer for how good the model's predictions have become.

At some point, the model's predictions stop improving. At that point, training is complete. The only way to further improve the model's predictions further is to provide more example inputs for training. As more examples are used in the training process, the resulting model has a better, more coherent picture of the interrelated probabilities required to process the inputs "correctly."

Question 7.1. *How are training materials used and/or reproduced when training an AI model? Please include your understanding of the nature and duration of any reproduction of works that occur during the training process, as well as your views on the extent to which these activities implicate the exclusive rights of copyright owners.*

Training materials are reproduced during training to the extent that they are "read" or "viewed" by the AI system. This reproduction can have multiple steps, such as making a copy of a work onto a computer performing training, and then making a temporary copy of the work (sometimes called a "RAM copy") into the working memory of the computer so that it can be used as input for the AI training system. These reproductions are necessary steps for the use of any work by (or using) a computer system.

The incidental reproduction necessary to "read" a work should not implicate any of a copyright owner's exclusive rights for two reasons: 1) the reproduction is limited to "reading" the work, and 2) section 117 of the Copyright Act excludes this reproduction from copyright owners' rights under section 106.

AI Training Only "Reads" the Work

The exclusive rights granted to copyright owners only address reproduction and distribution. "Reading" a work is not one of the reserved rights under copyright, even if someone (or some*thing*) is reading the work via a computer.[13] Allowing the incidental copies necessary for the functioning of any computer is tantamount to giving a copyright owner exclusive control over any computer-assisted "reading" of a work. The Copyright Act was never meant to grant such control.

By way of comparison, the reproduction that occurs as part of AI training is exactly the same as the reproduction that occurs when a person views a work over the Internet. When someone uses a web browser, a copy is made on that person's local computer. Then the web browser "reads" the local copy of the text or image file in order to show it to the human using the web browser.

Comparing AI training to web browsing is not an idle comparison. Many recent datasets (such as LAION 5B) only consist of links to pictures and pages on the Internet, not the images themselves.[14] Systems using the LAION datasets to train an AI use an automated "browser" to request (and receive) copies of the works they use to train the AI system, exactly the same way a human would receive the work.

[13] See Jessica Litman, The Exclusive Right to Read, 13 CARDOZO ARTS & ENT. L.J. 29, 34 (1994) ("Ninety years later, the U.S. copyright law is even more technical, inconsistent and difficult to understand; more importantly, it touches everyone and everything....Most of us can no longer spend even an hour without colliding with the copyright law. Reading one's mail or picking up one's telephone messages these days requires many of us to commit acts that the government's Information Infrastructure Task Force now tells us ought to be viewed as unauthorized reproductions or transmissions.")); Jessica Litman, *Readers' Copyright*, 58 J. COPYRIGHT SOC'Y U.S.A. 325 (2010) ("Copyright gives no exclusive rights to control private performance or display.80 What you do with a book, movie, or sound recording in your living room is not copyright infringement, even if your copy is pirated. Private performance and display is simply off limits. (That isn't because copyright owners didn't ask for private performance and display rights - they did. But nobody took those demands seriously, I think, because at some level everyone understood that the freedom to read and enjoy material without the copyright police looking over your shoulder is an interest that copyright law has respected and should protect.").

[14] See the answer to question 6.1 for more detail.

No one would argue that a human viewer is infringing a copyright simply because the computer temporarily reproduces the work in the process of allowing it to be perceived. The law should treat the incidental reproductions used during the AI training process equivalently.

Limitations on the Copyright Owner's Exclusive Rights under 17 U.S.C. 117(a)

A second reason that AI training should not implicate the exclusive rights of copyright owners is because Section 17 U.S.C. 117(a)(1) provides an exception that covers this use. It states in relevant part:

> **(a) Making of Additional Copy or Adaptation by Owner of Copy**.—Notwithstanding the provisions of section 106, it is not an infringement for the owner of a copy of a computer program to make or authorize the making of another copy or adaptation of that computer program provided:
>
> (1) that such a new copy or adaptation is created as an essential step in the utilization of the computer program in conjunction with a machine and that it is used in no other manner....[15]

A "computer program" as defined in section 101 is "a set of statements or instructions to be used directly or indirectly in a computer in order to bring about a certain result."[16] While it may be surprising to some people unfamiliar with the underlying functioning of the computer, this definition includes digital media files.

When we think about an image or a song on our computer, we usually don't pay attention to the distinction between how a file is encoded and stored on the computer and the copyrighted work itself. However, just as a physical book is not the same as the copyrighted content of the book, the files we use on our computer are not the same as the copyrighted work itself.

Using images as an example, when a person views an image on a computer, the image they are viewing consists of individually controlled pixels on the computer screen that create the pattern of light and color we recognize as a picture. This is the "certain result" envisioned by the statute. Most people are familiar with many of the common file extensions, such as .jpg and .png. Two image files with different extensions may result in identical images being placed on the screen, but the process followed by the computer is different for every type of file, because each type of file has a different type of instructions used to create the image we see.

---

[15] 17 U.S.C. §117(a)(1).
[16] 17 U.S.C. §101.

These instructions are usually converted to a binary format for ease of use on a computer–but not always.[17] However, each digital media file is "a set of statements or instructions" used by a computer to bring about "a certain result" of allowing a human to perceive the copyrighted work.

If a person training an AI system has lawfully come into possession of a copy of a work then they are the "owners" of that copy of the work. They should be able to use it pursuant to the exception granted by Congress.

Question 7.2. *How are inferences gained from the training process stored or represented within an AI model?*

The result of the AI training procedure is an AI "model." The model is the combination of the neural network design and the "weights" – a set of numbers that encode statistical probabilities about the inputs that have been processed during training.

If a person were to save a well-trained model to disk and examine it, what that person would see would be a gigantic matrix of numbers–the learned weights associated with the nodes in the hidden layers. An AI model does not directly contain any copies of its source inputs, even in compressed form.

It is hard to say exactly what any single probability within the model represents. It doesn't record knowledge about any single input, and no single input is represented in the weights more than an infinitesimal amount. Collectively, however, the weights provide a detailed statistical picture of the collective experience of humanity when it comes to the inputs. These statistics encode knowledge about what makes an image a "picture" as opposed to a bunch of noise, or what makes a bunch of words a "story" or an "article" instead of a bunch of gibberish. Or more specifically, what makes a song a "pop song from the 80s" as opposed to an "aria." Because humanity has embedded so much implicit knowledge in words, pictures, and songs, these models can appear quite "intelligent." However, these AI models should be considered an advanced form of autocomplete.

The knowledge embedded within the model is called "latent" knowledge, in the sense that it is information that is hidden but unexpressed. The logical landscape of this hidden knowledge–what concepts are "close together" or "far apart," and in which ways, is sometimes called the "latent domain."

In copyright terms, the latent domain is a statistical expression of the ideas underlying every expression that it was trained on. It does not contain "The Hobbit." Instead it contains a representation of what a fantasy story *is*, how it is

[17] One type of image format (.ps, for "Postscript") consists only of human-readable instructions that tell the computer how to draw the desired image on the screen.

structured, and what sort of elements fantasy stories might include (the scènes à faire).

Question 7.3. *Is it possible for an AI model to "unlearn" inferences it gained from training on a particular piece of training material? If so, is it economically feasible? In addition to retraining a model, are there other ways to "unlearn" inferences from training?*

Unlearning is not currently possible but can be simulated in some circumstances.

We have a very limited understanding of how each input affects the model. Right now it is beyond the state of the art to accurately and directly measure the influence of a particular training point on the final weights. However, we can counter-train a model to selectively "forget" an input by further training the model to respond as if the input had never been provided. However, these forgetting procedures are themselves probabilistic.

See for example Golatkar et al., "*Eternal Sunshine of the Spotless Net: Selective Forgetting in Deep Networks*," arXiv:1911.04933v5 [cs.LG] (31 Mar 2020) or Bourtoule et al., "*Machine Unlearning*," 2021 IEEE Symposium on Security and Privacy (SP) (2021).

Question 7.4. *Absent access to the underlying dataset, is it possible to identify whether an AI model was trained on a particular piece of training material?*

This is known as "membership inference" or a "membership inference attack." It is possible to predict whether a particular piece of training material was used. In practice there is a tradeoff between the accuracy of the prediction and the amount of resources needed to make the prediction. In addition, the larger the model (i.e., the more works were used as training material), the harder it is to perform.

See for example R. Shokri, M. Stronati, C. Song and V. Shmatikov, "*Membership Inference Attacks Against Machine Learning Models*," *2*017 IEEE Symposium on Security and Privacy (SP), San Jose, CA, USA, 2017, pp. 3-18, doi: 10.1109/SP.2017.41.

Question 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.*

In the context of fair use, it is essential to evaluate the AI training procedure separately from the output of the procedure (the model), separate from any the output of the procedure (the model). Under Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith et al., ("Warhol")[18] some uses may be infringing, whereas other uses may be fair use.[19]

The determination of whether a particular use is considered fair use is decided on a case-by-case basis and is a combination of law and fact.[20] There is no automatic assumption of fair use. Fair use is an affirmative defense, with the defendant bearing the burden of proof.[21]

Fair use is evaluated based on four factors: 1) the purpose and character of the use; 2) the nature of the copyrighted work; 3) the amount or substantiality of the portion used; and 4) the effect of the use on the potential market or value of the work.[22] Courts have also taken into account whether a particular use advances the public purpose of encouraging the creation of new works.[23]

Factor 1: The Purpose and Character of the Use

Regarding the first factor of fair use, the purpose and character of the use, several aspects of AI training are directly relevant to the inquiry. These are: 1) the use is transformative; 2) the use is limited to "reading" the work; and 3) the work is only used for making measurements and recording facts about the content.

The Use in AI Training is Transformative

A primary consideration relevant to the character and use of the work is whether the use of the work is transformative.

The concept of transformative use comes from the 1994 Supreme Court decision in *Campbell v. Acuff-Rose Music*.[24] In *Campbell*, the Supreme Court described "transformative" use as being the key element underlying the first fair use factor:

---

[18] Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith et al., ______ U.S. ______, ______ S. Ct._____ (2023), slip. op available at https://www.supremecourt.gov/opinions/22pdf/21-869_87ad.pdf.

[19] *Id*. at 21 (Only [AWF's commercial licensing of Orange Prince to Condé Nast] . . . is alleged to be infringing. We limit our analysis accordingly. In particular, the Court expresses no opinion as to the creation, display, or sale of any of the original Prince Series works"). *See also* fn. 10 ("Congress has directed courts to examine the purpose and character of the challenged 'use.' . . . Had AWF's use been solely for teaching purposes, that clearly would affect the analysis").

[20] Harper & Row, Publrs. v. Nation Enters., 471 U.S. 539, 560 (1985).

[21] Am. Geophysical Union v. Texaco Inc, 60 F. 3d 913, 918 (2d Cir. 1994).

[22] Authors Guild, Inc. v. HathiTrust, 755 F.3d 87, 96 (2d Cir. 2014).

[23] Id. at 94 (quoting Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 575 (1994)) ("for fair use of copyrighted materials has been thought necessary to fulfill copyright's very purpose, "to promote the Progress of Science and useful Arts...."); see also U.S. Const., Art. I, § 8, cl. 8. ("To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.").

[24] Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 586, 114 S. Ct. 1164, 1175 (1994).

> The central purpose of this investigation is to see, in Justice Story's words, whether the new work merely "supersede[s] the objects'" of the original creation, or instead adds something new, with a further purpose or different character, altering the first with new expression, meaning, or message; it asks, in other words, whether and to what extent the new work is "transformative."[25]

Even if a particular output is found to infringe a work, the AI model should be evaluated on its own. To put it in the context of *Sony Corp. of Am. v. Universal City Studios, Inc.*[26] (*Sony*), a VCR can be used to make an infringing copy of a movie. But the VCR itself is not the original input movie or an infringing copy of the movie. It is a technical device that can in some circumstances be used to make infringing outputs.

The cases most similar to AI training are *Authors Guild, Inc. v. HathiTrust*[27] (*HathiTrust*) and the related case *Authors Guild v. Google*[28] (*Google*). The factual background of these two cases is similar: the Authors Guild sued HathiTrust and Google for copyright infringement because of the defendants' mass digitization of books. Hathitrust, a digital library consortium, created its digital copies for preservation, for accessibility for visually impaired users, and to create a search index. Google created Google Book search to facilitate researchers in locating relevant information. Users could search across books for specific words and phrases and then see a "snippet view" with the search result highlighted.

In these cases, there were two accused processes: 1) the creation of a digital copy of the books for the purposes of creating a search index, and 2) the distribution of whole or partial copies to users. Since the process of training an AI model does not result in distribution, the relevant portion for our purposes is the creation of the digital copy for indexing.

In the Authors Guild cases, digitization was accomplished by "mak[ing] a digital scan of each book, extract[ing] a machine-readable text, and creat[ing] an index of the machine-readable text of each book."[29] The end result was a search index enabling users to find content within the books more effectively as well as research new types of questions.[30]

With regard to the creation of the search index, the Hathitrust court said:

---

[25] *Id*. at 579.
[26] Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417 (1984).
[27] Authors Guild, Inc. v. HathiTrust, 755 F.3d 87, 97 (2d Cir. 2014)
[28] Authors Guild, Inc. v. Google, Inc., 804 F.3d 202, 217 (2d Cir. 2015)
[29] *Id.* at 208.
[30] *Id*.

> [We] conclude that the creation of a full-text searchable database is a quintessentially transformative use.... the result of a word search is different in purpose, character, expression, meaning, and message from the page (and the book) from which it is drawn. Indeed, we can discern little or no resemblance between the original text and the results of the [defendant's] full-text search.[31]

The *Google* court further elaborated on the transformative nature of the search index by highlighting the new statistical research tools that it made possible:

> [The] purpose of Google's copying of the original copyrighted books is to make available significant information about those books, permitting a searcher to identify those that contain a word or term of interest, as well as those that do not include reference to it. In addition, through the ngrams tool, Google allows readers to learn the frequency of usage of selected words in the aggregate corpus of published books in different historical periods. We have no doubt that the purpose of this copying is the sort of transformative purpose described in Campbell as strongly favoring satisfaction of the first factor.[32]

The Authors Guild cases are not alone. Many other courts looking at similar fact patterns have found the same. For example, in *A.V. v. iParadigms, LLC*,the court found that the copying and archiving of student papers is permissible when aimed at detecting and preventing plagiarism rather than capturing expressive content.[33] In *Perfect 10 v. Amazon.com, Inc*., the court ruled that Google's copying of Internet content to make it searchable was considered transformative as it turned the image into a "pointer" directing the user to a source of information.[34] Similarly, in *Kelly v. Arriba Soft Corp*., the court ruled that copying to produce exact replicas of artistic works displayed in thumbnail form on the internet was transformative as it was unrelated to any aesthetic purpose and was aimed at facilitating searches.[35]

Just like the building of a search index is a "quintessentially transformative use," so too is the building of an AI model. The result of the machine-based processing is a product with wholly different purposes, capabilities, and uses. There is no way in which an AI model could be mistaken for any of its training inputs. The mass of statistical probabilities that make up a generative AI model

---

[31] *HathiTrust* at 97.
[32] Authors Guild v. Google, Inc., 804 F.3d 202, 217 (2d Cir. 2015)
[33] AV Ex Rel. Vanderhye v. iParadigms, LLc, 562 F. 3d (4th Cir. 2009)
[34] Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146 (9th Cir. 2007)
[35] Kelly v. Arriba Soft Corp., 336 F.3d 811 (9th Cir. 2003)

are so different from the training material that there is no question it is "different in purpose, character, expression, meaning, and message"[36] from any (or all) of the works that were used as input.

The Use in AI Training is Limited to "Reading" the Work by a Computer

One persistent misunderstanding is the perception that the AI training process makes repeated or derivative copies of each work used as input.[37] There is also the perception that the model somehow "stores" the works used for training within the model. Both of these perceptions are incorrect.

As described above relative to questions 7. And 7.1, there is only one copy of the work needed for training: the initial copying of the work into the input layer of the AI model. This is the process by which the model "reads" the input in order to perform the training process. Reading (or viewing) a work, including on or by a computer, has been repeatedly found to be fair use.[38]

Sony Corp. of Am. v. Universal City Studios, Inc.[39] is instructive. The case centered around Sony's Betamax video cassette recorder (VCR), which allowed users to record television programs for later viewing, a practice known as "time-shifting."

The Supreme Court ruled in favor of Sony, stating that noncommercial home use of the VCR to record television programs for later viewing is fair use. Although time-shifting required making a copy of a copyrighted work, it was non-infringing because the purpose was to allow the users to receive the work at the time of their choosing—which was not distribution, publishing or

---

[36] Authors Guild, Inc. v. HathiTrust, 755 F.3d 87, 97 (2d Cir. 2014)

[37] *See* Benjamin L.W. Sobel, Artificial Intelligence's Fair Use Crisis, 41 COLUM. J.L. & ARTS 45, 48 (2017) (These "training data" often comprise thousands of unauthorized copies of copyrighted works, which are reduplicated and modified countless more times throughout the training process**.");** *id.* at 62 ("Once an input dataset has been compiled, it may be copied, emulated, and re-copied thousands of times during the learning process."); U.S. Pat. & Trademark Off., Public Views on Artificial Intelligence and Intellectual Property Policy (2020) (saying that AI "functions by ingesting copyright works" which results in "mass digitization.").

Courts do not seem to have an incorrect notion of ML, rather they often have no notion at all. *See Carpenter v. McDonald'S Corp.,* 580 F. Supp. 3d 512, 516 (N.D. Ill. 2022) (describing AI simply as "a form of artificial intelligence."); Performance Pricing, Inc. v. Google, Inc., No. 2:07cv432, 2009 U.S. Dist. LEXIS 77538, at *5 fn. 3 (E.D. Tex. Aug. 28, 2009) (describing AI as "a type of computational algorithm which is derived by other algorithms."). *But see Ocean Tomo, LLC v. Patentratings, LLC,* 375 F. Supp. 3d 915, 956 (N.D. Ill. 2019) ("At a high level, machine learning tools attempt to discern patterns within data, but with no pre-conceived concepts or requirements as to the structure of these data. Machine learning uses an iterative process, in which the system initially forecasts an outcome based on combinations of input variables. The system then determines the errors of its forecasts, and adjusts accordingly, iterating until these error terms are minimized.").

[38] *Supra* at 7 And 7.1.

[39] Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417 (1984).

performance.[40]

It is also significant that the only "reader" in an AI training scenario is a computer, not a human. As highlighted by Professor Grimmelmann in "Copyright for Literate Robots," courts have almost uniformly found that machine "reading" is fair use.[41]

AI Model Training is Limited to Making Measurements and Recording Facts

Also relevant to the first fair use factor is that AI training is limited to making measurements and recording facts. Because the outputs of AI applications *seem* so expressive, people mistakenly assume that AI applications copy creative expression from inputs and use the copied expression to generate derivative outputs.[42]

Rather than copying any expression, however, the model training process records facts about the work. Think of the analogy of the art inspector from question 7. In that analogy, the art inspector examines paintings by taking and recording every measurement possible–brushstrokes per square inch, correlations between colors six inches apart, and the number of syllables in the artist's name. No one would argue that the art inspector's list of measurements is a derivative work of the measured paintings – and in fact they would recognize that the list of measurements was not a copyrightable object at all. Facts *about* a work cannot be copyrighted and are not part of the expressive content of a work.[43]

---

[40] Id. at 449 ("time-shifting merely enables a viewer to see such a work which he had been invited to witness in its entirety free of charge, the fact that the entire work is reproduced, does not have its ordinary effect of militating against a finding of fair use."). The defendants in Sony brought substantial evidence that the copyright holders wanted users to make copies, if necessary to view their works. Id. at 445 (Fred Rogers, copyright holder of Mister Rogers' neighborhood testified that "he had absolutely no objection to home taping for noncommercial use and expressed the opinion that it is a real service to families to be able to record children's programs and to show them at appropriate times.").

[41] Courts have almost uniformly found that machine "reading" is fair use. *See generally* James Grimmelmann, *Copyright for Literate Robots*, 101 Iowa L. Rev. 657, 659 (2016).

[42] In Doe v. GitHub Inc., the plaintiffs accuse the defendant's of "distributing" the input code "to Copilot users as if it were created by Copilot." No. 3:22-cv-06823, at *6 (N.D. Cal. filed Nov. 3, 2022). Although the plaintiff's admitted that "Codex and Copilot do not retain copies of the materials they are trained on," they argue that "[i]n practice, however, the Output is often a near-identical reproduction of code from the training data." *Id.* at *15. Likewise, in Anderson v. Stability AI LTD., the plaintiffs argue that "[t]hese 'new' images are based entirely on the Training Images and are derivative works of the particular images Stable Diffusion draws from when assembling a given output." No. 3:23-cv-00201, at *3 (N.D. Cal. filed Jan. 13, 2023). This leads them to conclude that the AI tool "is merely a complex collage tool." *Id.*

[43] "[C]opyright's idea/expression dichotomy strikes a definitional balance between the First Amendment and the Copyright Act by permitting free communication of facts while still protecting an author's expression. No author may copyright his ideas or the facts he narrates." Harper & Row, Publrs. v. Nation Enters., 471 U.S. 539, 556 (1985) (internal citations omitted).

This distinction between factual and expressive content was made clear by *Feist Publications, Inc. v. Rural Telephone Service Co*.[44] In *Feist,* Rural Telephone Service Co. (Rural) created a telephone directory that included listings for its customers, while Feist Publications, Inc. (Feist) was a company that specialized in producing area-wide telephone directories. Feist used Rural's listings without permission in its own directory, resulting in Rural suing Feist for copyright infringement.

The key issue in the case was whether Rural's telephone directory was eligible for copyright protection. The Supreme Court held that the directory was not protected by copyright because it lacked the necessary originality and creativity. The *Feist* court emphasized that facts, such as names, addresses, and phone numbers, are not copyrightable because they are discovered rather than created, and thus do not meet the originality requirement. The court noted that *compilations* of facts may be copyrightable, but only those compilations that show sufficient human creativity.

Though similar to the phone books in *Feist,* AI training is even further from infringement in that the factual content recorded in the model is generated by the training process. In *Feist*, the factual content was directly copied from Rural's phone book. But in AI training, the statistical probabilities associated with each input are not part of the work at all. They are generated in response to the "predict" and "adjust" phases of the training process. This would be like if Feist recorded measurements such as the number of businesses associated with each letter, the correlation of phone numbers with names, and other similar facts, and then published those facts without publishing any part of the phone book itself. If the court was unwilling to find that straightforward facts violated copyright, then the argument for abstract and newly generated facts like the facts recorded in AI models is even stronger. Ultimately, whether abstract or straightforward, the rule remains that facts are not copyrightable.

Even viewing the model as a whole, the statistical measurements within a model are not selected due to any human creativity or judgment. It is the computer process that identifies correlations and records the facts. Moreover, humans are currently unable to even understand the connection between any particular weight in the model and any fact observed during training.

One case that moves towards addressing "abstract facts", is *New York Mercantile Exchange, Inc. v. IntercontinentalExchange, Inc* which focused on the issue of copyright protection for market settlement prices.[45]

New York Mercantile Exchange, Inc. (NYMEX) sued IntercontinentalExchange, Inc. (ICE), claiming that ICE had infringed on NYMEX's copyrights by republishing its market settlement prices without authorization. NYMEX argued

---

[44] 499 U.S. 340 (1991).
[45] 497 F.3d 109 (2d Cir. 2007).

that the settlement prices were original and creative works deserving of copyright protection.

The Second Circuit disagreed with NYMEX's argument. The court held that the settlement prices were not copyrightable because they were factual information and not original expressions. The court reasoned that the prices were determined by an objective process involving the exchange of bids and offers and thus did not possess the requisite level of creativity and originality required for copyright protection. Like the measurements taken as part of the AI training process, the bid and offer prices in *New York Mercantile* were independent facts, albeit difficult to extract, available for anyone to use.

As the *Feist* court noted:

> [F]acts do not owe their origin to an act of authorship. The distinction is one between creation and discovery: The first person to find and report a particular fact has not created the fact; he or she has merely discovered its existence. To borrow from Burrow-Giles, one who discovers a fact is not its "maker" or "originator." The discoverer merely finds and records. Census takers, for example, do not "create" the population figures that emerge from their efforts; in a sense, they copy these figures from the world around them.[46]

Just like the census taker in the example given by the Supreme Court, the AI model records only facts in the form of statistical probabilities. These facts are available for anyone, or any process, to copy and use. And just as there is no copyrightable expression in a mechanistic set of measurements about a work, there is no expression copied *from* the work to make such a set of facts.

The First Factor Weighs in Favor of Fair Use

For the reasons discussed above relative to question 8, the “purpose and character” of AI training leans heavily in favor of fair use. Especially important is the fact that AI training is *completely* transformative to the point that the AI model does not contain any part of the works used for training at all.

Factor 2: The Nature of the Copyrighted Work

The second fair use factor is the nature of the copyrighted work. When evaluated against AI training as a whole, this factor does not influence the fair use analysis either way. None of the types of works that could be used for training receive any special favor or analysis. To the AI application, the exact type of content used for training is irrelevant; the model only sees a series of numbers. All types of works, copyrightable or not, highly original or not, are all treated equivalently. Thus this factor does not bear any weight in the analysis.

[46] Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 347 (1991) (internal citations omitted).

Factor 3: The Amount or Substantiality of the Portion Used

The third factor in the fair use analysis is whether the amount of copying exceeded what was necessary and if it was excessive. There are no strict rules on how much of a copyrighted work can be copied while still being considered fair use.[47] The permissible extent of copying depends on the purpose and character of the use. Excessive copying is copying anything "more" than what is reasonably "necessary."[48] In some cases, copying the entire work might be necessary, and in such instances, this factor doesn't weigh against a finding of fair use.

In the case of AI model training, the entire work is typically reproduced as part of the "reading" process. However, *none* of the input works are reproduced within the AI model. Depending on whether the focus is on intermediate, ephemeral copies or how much copyrightable expression is included in the ultimate output, the amount of the work used can vary from "all of it" to "none of it."

Examining the purpose of the use, however, the intent of the AI training procedure is not to copy any particular input. Instead, it is designed to identify the non-copyrightable similarities that undergird all the different expressions used during training. To accurately identify the non-copyrightable commonalities, the only reasonable method is to examine all the inputs possible, including copying entire works. Since using the entire works is reasonably necessary to enable AI model training, the copying is not excessive. Accordingly, the third fair use factor should not incline the fair use analysis either way.

Factor 4: The Effect on the Market

The last factor in the fair use analysis is how the use affects the market for the original work.

Part of the widespread concern regarding AI has to do with the possibility that AI systems may compete in the market against human creators. This is a reasonable concern, but the fourth fair use factor is not about competitiveness or "the market" generally. Rather, the "effect on the market" is the extent to which the result of the copying serves as a substitute for the original work.[49]

---

[47] *Maxtone-Graham v. Burtchaell*, 803 F.2d 1253, 1263 (2d Cir. 1986). "[T]he extent of permissible copying varies with the purpose and character of the use." *Campbell*, 510 U.S. at 586-87. "The crux of the inquiry is whether "no more was taken than necessary." *Id*. at 589.

[48] *See* Harper & Row v. Nation Enterprises, 471 U.S. 539.

[49] "Even when an entire copyrighted work was recorded, the District Court regarded the copying as fair use because there is no accompanying reduction in the market for plaintiff's original work." Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 425-26, 104 S. Ct. 774, 780 (1984), fair use depends on "the likelihood that the parody may serve as a market substitute for the original" Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 586, 114 S. Ct. 1164, 1175 (1994).

As stated by the Supreme Court in *Campbell v. Acuff-Rose Music*: "[T]he only harm to derivatives that need concern us, as discussed above, is the harm of market substitution."[50]

Trained AI models and AI applications are wholly different types of goods than the inputs that they trained on. There is no possible market substitution between the AI model and any particular input it is trained on. The AI model is useless as an artwork, song, poem, or as any other type of creative work. As described above, anyone looking at an AI model would only see a "gigantic matrix of numbers" inscrutable to any process but the AI application itself.

Using Copyrighted Works to Train an AI Model is Fair Use

From the analysis above, it becomes clear that AI training is "quintessential" fair use. When AI model training is examined with the correct factual background, the strength of its legality surpasses that of even the most obvious and well-known cases. Existing case law convincingly makes the case that AI model training is fair use.

It is undisputed that copyrighted works are necessary for many types of AI training. But as stated by the *Feist* court, "[t]he primary objective of copyright is not to reward the labor of authors, but to promote the Progress of Science and useful Arts. To this end, copyright assures authors the right to their original expression, but encourages others to build freely upon the ideas and information conveyed by a work."[51] This is what AI models enable people to do.

Question 8.1. I*n 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 distinction between *Warhol* and *Google v. Oracle*[52] is the extent to which the use "superseded the objects . . . of the original work."[53] In *Warhol*, the court found that the exact use examined by the Court (licensing Orange Prince for a magazine cover) directly substituted for Goldsmith's original photo. In contrast, in *Google v. Oracle*, the overriding factor in that case was the way in which it drove the creation of new, independent works. "To the extent that Google used parts of the Sun Java API to create a new platform that could be readily

[50] Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 593, 114 S. Ct. 1164, 1178 (1994)
[51] *Feist Publ'ns Inc.*, 499 U.S. at 349-50 (internal quotation marks and citations omitted).
[52] *Id.*
[53] *Folsom v. Marsh*, 9. F.Cas. 342 (C.C.D. Mass. 1841)

used by programmers, its use was consistent with that creative 'progress' that is the basic constitutional objective of copyright itself."[54]

Applying this distinction to AI training, the question is whether AI training (either the process, or the resulting model) substitutes for the original work, or whether it fulfills the purpose of copyright by encouraging the creation of new works.

From this perspective, the purpose and character of AI training, at every phase, is *much more* weighted toward fair use than the copied source code at issue in *Google v. Oracle*. It is much more transformative than search indexes at issue in the *Authors Guild* cases, or any of the other cases that found fair use. People can and do use generative AI models to facilitate the almost unlimited generation of *new* works. Generative AI models make artistic creation accessible to a broad portion of the population–and it is evident that new works are being created every hour of every day. This fulfills the "basic constitutional purpose" of copyright to an unprecedented degree.

Question 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?*

Entities that collect and distribute copyrighted material for training should be evaluated individually according to the fair use factors. Is the entity not for profit? Is the dataset created for the purpose of research? Is the training data provided in a way that can easily be misused for clearly infringing uses? These are fact-specific questions that cannot be answered in the abstract.

Question 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?[45] Does it make a difference if funding for these noncommercial or research uses is provided by for-profit developers of AI systems?*

The "purpose and character" of a use and the importance of transformativeness is a matter of degree that must be weighed against other considerations, like the commercial nature of the copying.[55]

Much AI training is commercial. On its own, commerciality weighs against a finding of fair use. Balancing the commerciality of some AI training, however, is the "degree of difference" between the input works used for training and the output of the training–the model itself.[56] As expressed by the Warhol court, "[t]he fair use provision, and the first factor in particular, requires an analysis of

---

[54] *Id.*
[55] *Warhol*, slip. op. at 12.
[56] *See id.*

the specific 'use' of a copyrighted work."[57] The court found that only "AWF's commercial licensing of Orange Prince" was unjustified, and the Court "expresse[d] no opinion" as to any other uses of the work.[58]

In contrast, the differences between AI models and the works they are trained on are so stark that there is no reasonable comparison between them. The Warhol court found that both Orange Prince and Goldsmith's original photograph were licensed for magazine covers, showing that the two works were substitutes in that market. However, an AI model is not consumable in the same way as the works used to train the model. The weights are just a set of numbers, completely unintelligible to humans. There is no way to perceive any particular input within the model weights, let alone all the works used during training.

For example, Goldsmith's photo could be an input to an AI training procedure, but the AI model trained in part on Goldsmith's photo could not be licensed to replace the original photo in any circumstance. It is conceivable that an image later generated using the model could possibly be infringing, but the model itself is distinct.

Put another way, AI training *completely* transforms the training materials into a set of statistical measurements. Accordingly, the high degree of transformativeness makes the commercial nature of any AI training less relevant for fair use purposes.

Question 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?*

Developers have found that the greater the amount of training data used, the better the AI model. This is because the goal of training is to develop a statistical model of how to generate new works, not how to copy existing works.

This leads to the somewhat counterintuitive result that the larger the amount of material used to train a particular AI system, the more likely it should be considered a fair use. This is because there is an inverse relationship between the quantity of training materials used to train a particular AI system and the possible effect that any one work has on the AI model.

For example, imagine a model trained on a single work. The only information contained within the model would come from a single source. In that case any outputs from the AI model would necessarily be based on a single known

[57] *Id*. at 20.
[58] *Id*. at 21.

source (even though the model itself would still just be a set of statistical measurements).

As the number and variety of training material grows, however, the AI model starts to truly reflect the ideas and structures behind all of the expressions. The model, and any outputs of the model, stop reflecting any individual training input in more than an infinitesimal, *de minimis* way.

Question 8.5. *Under the fourth factor of the fair use analysis, how should the effect on the potential market for or value of a copyrighted work used to train an AI model be measured? Should the inquiry be whether the outputs of the AI system incorporating the model compete with a particular copyrighted work, the body of works of the same author, or the market for that general class of works?*

The case law and policy are aligned on this question: the only relevant market is the market for a *particular registered work*, not for the works by a particular author or works in general.[59] As the court in *Dave Grossman Designs v. Bortin* [60] stated:

> The law of copyright is clear that only specific expressions of an idea may be copyrighted, that other parties may copy that idea, but that other parties may not copy that specific expression of the idea or portions thereof. For example, Picasso may be entitled to a copyright on his portrait of three women painted in his Cubist motif. Any artist, however, may paint a picture of any subject in the Cubist motif, including a portrait of three women, and not violate Picasso's copyright so long as the second artist does not substantially copy Picasso's specific expression of his idea.[61]

As described in *Dave Grossman Designs*, copyright law does not prohibit the copying of an artist's distinctive style in the context of a new work.[62] Looking at other types of works, doing things "in the style of" another artist is even more attenuated. There is no copyrightable interest in the written style of a particular author, nor of the general style of a musical artist. There must always be specific copied expression.

---

[59] "[W]hen a commercial use amounts to mere duplication of the entirety of an original, it clearly "supersede[s] the objects," Folsom v. Marsh, supra, at 348, of the original and serves as a market replacement for it, making it likely that cognizable market harm to the original will occur." Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 591, 114 S. Ct. 1164, 1177 (1994).

[60] Dave Grossman Designs, Inc. v. Bortin, 347 F. Supp. 1150 (N.D. Ill. 1972).

[61] *Id*. at 1156.

[62] An artist or author may have causes of action other than copyright, including trademark liability or potential right of publicity claims, or claims under a new law such as the "No Fakes" Act.

More generally, AI systems are just tools that will be used by authors to help them create new works. These works will not be spontaneously generated; they will be the result of deliberate creative effort. Artists and authors competing in the market using AI tools is just… artists and authors competing in the market.

Question 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)?*

There may be market pressures that encourage developers to use opt-in or opt-out mechanics when developing AI systems. However, opt-in vs. opt-out is only an issue if a work is legally required to be licensed. As stated by the Supreme Court in *Campbell v. Acuff-Rose Music, Inc.*: “If the use is otherwise fair, then no permission need be sought or granted.”[63]

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

All consent systems become increasingly unwieldy when considering the number of works used to train modern AI systems. As discussed relative to question 6.1, the overwhelming majority of training materials are unregistered materials posted on the Internet. Both opt-in and opt-out regimes cannot effectively handle the terabytes of existing datasets created over the past twenty years for research purposes.

Further, “commercial” vs. “non-commercial” activity can be difficult to define ahead of time. In one recent case, the actions of a nonprofit entity that did not charge any users still were considered “commercial” because the activity still attracted people to its website and bolstered its standing in the community.[64]

Question 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?*

There are putative efforts toward an opt-out approach, such as a “do not train” indicator usable by websites. However, just like the comparable robots.txt “do not index” indicator, compliance is ultimately voluntary.

---

[63] Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 586 (1994)
[64] Hachette Book Group, Inc. et al. v. Internet Archive et al., No. 20-cv-4160, 2023 WL 2623787, 2023 U.S. Dist. LEXIS 50749.

Question 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?*

It is not reasonably feasible to get consent in advance. For all existing datasets, no advance consent is possible. This would exclude the past decades worth of research datasets from use.

Going forward, most content on the Internet is unregistered, posted by (possibly anonymous) individuals. Individually contacting these individuals is not possible. In practice, advance consent would just become a line in an ignored terms of service page.

Question 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?*

If a court ultimately finds that training AI systems results in infringement, then we have existing remedies for copyright infringement.

Question 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?*

The United States has consistently rejected the concept of moral rights, except in the case where artistic works are being defaced, destroyed, or where individuals are misattributing authorship.

None of these situations apply to AI training, and there is no need to extend moral rights to be able to exclude AI training.

It is reasonable to allow authors and artists to stop others from falsely claiming that an AI-generated work is their work.

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

If owner consent was required, the only feasible mechanism would be a compulsory licensing regime. But without a mechanism to provide licenses for unregistered content, any owners' consent regime would be drastically incomplete and would act as a wealth transfer from the poor to those privileged enough to be able to register their works.

Compulsory licensing is not feasible given that the majority of AI training inputs are (and will likely continue to be) anonymous, pseudonymous, and unregistered.

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

The only sector where direct voluntary licensing might be feasible is in the music sector, due to the existing compulsory license structures and the fact that a large percentage of musical works are registered. However, even in the music sector there are thousands or millions of unregistered musical works posted on the Internet.

Question 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?*

People can engage in voluntary collective licensing without any action by Congress. There is value in better-curated, cleaner, enriched, or more up-to-date data. Many companies (such as Lexis-Nexis and Westlaw) take free, public-domain data and process it to create valuable databases–including for AI system use. No changes to the law would be required to enable this business model.

Question 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?*

No, Congress should not establish a compulsory licensing regime unless it was also royalty free. No licensing regime, even if compulsory, could reasonably deal with royalties for anonymous, pseudonymous, or unregistered content.

Unregistered, "common person" content is extremely valuable for helping systems understand common concepts in language, art, and music. Given that people create new copyrighted works constantly, this unregistered content would always be the majority of content, making royalties impossible to administer.

Question 10.5. *Should licensing regimes vary based on the type of work at issue?*

No. As well as not having any basis in law, AI systems are increasingly "multimodal," meaning that they are trained on many different types of content concurrently. It would not make sense to have a single project that required different licensing for works that were subject to the same process.

Question 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 only possible reasonable entity to arrange training licenses would be a dataset curator. Only the curator would know where the data came from and would have a reasonable way to aggregate licensing terms for AI training.

Developers who train an AI model are dealing with bulk records. The volume of records included in modern AI system precludes engagement with any individual records.

Companies employing AI systems do not know (and in most cases, have no ability to know) which specific inputs were included in a dataset or what role they played in training.

In practice, dataset licensing usually results in an inability to use the dataset. For example, a recent study found that conflicting licenses made the use of most data sources in one field commercially unavailable.[65] Another study found that "showing you this map of aggregated bullfrog occurrences would be illegal."[66]

Question 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.*

Right now it is beyond the state of the art to identify the degree to which one particular input contributes to a particular output in general. While it might be possible to make guesses, to find particular "memorized" inputs, or to construct toy AI models that have this property, It is not possible for any current system that is commercially reasonable or academically interesting.

---

[65] Gopi Krishnan Rajbahadur, Erika Tuck, Li Zi, Dayi Lin, Boyuan Chen, Zhen Ming (Jack) Jiang, and Daniel M. German. 2022. "*Can I use this publicly available dataset to build commercial AI software?-A Case Study on Publicly Available Image Datasets"* ACM Trans. Softw. Eng. Methodol. , (August 2022), 19 pages.

[66] Desmet, Peter, "*Showing you this map of aggregated bullfrog occurrences would be illegal*," peterdesmet.com, Oct. 17, 2013. *Available at* https://peterdesmet.com/posts/illegal-bullfrogs.html. *See generally* Meeker, Heather, "*Beyond Open Data: The Only Good License Is No License*," PLI Chronicle, Apr. 1 2022. *Available at* https://plus.pli.edu/Details/Details?fq=id:(352066-ATL2).

Intuitively, modern AI systems use *so much* data that the connection between any particular input and any output is necessarily attenuated. But even more fundamental are the mathematics behind AI training that force systems to generalize away from particular inputs and instead identify the underlying "latent" structures that represent the ideas and patterns behind *all* the inputs.

See the technical discussion of AI system training relative to questions 7-7.4 for more detail.

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

A licensing requirement for AI systems would stop most AI research and development in the United States. The vast number of training inputs would make licensing infeasible for anyone but the largest entities.

In particular, a large portion of all AI system development has taken place in the open source community. Essentially all AI tooling is open source, and open source developers are the cause of many fundamental advances in AI development and deployment.

But these open source developers do not usually have institutions supporting them or rights clearance offices. They are individuals that develop and contribute source code for personal reasons. If open source developers needed to engage in rights clearances for every AI training project, all AI development would abruptly become too expensive and difficult for this large and essential development population.

Question 14. *Please describe any other factors you believe are relevant with respect to potential copyright liability for training AI models.*

When evaluating potential liability for AI training, current law also asks whether a product that might have some infringing capabilities also has "significant noninfringing uses" and whether the product is marketed to users as a means to infringe copyright.

The "significant noninfringing uses" doctrine originated from Sony Corp. v. Universal City Studios.[67] In *Sony*, the accused product was Sony's Betamax video cassette recorder (VCR). It was undisputed by both sides in Sony that Sony's VCRs were capable of making infringing copies of movies or other material. However, the court found that if a product had substantial noninfringing uses, the manufacturer would not be liable for contributory copyright infringement. In the case of the VCR, making copies for private,

[67] 464 U.S. 417 (1984).

non-commercial use was sufficient.[68] The Court reasoned that stifling the distribution of products with legitimate uses would impede technological progress and undermine the goals of copyright law–to promote the progress of science and useful arts.

In the context of generative AI, the primary purpose (and the primary use) is the generation of new works. Compare with copy machines, which have significant noninfringing uses even though they are designed to make it easy to copy things.

In contrast, generative AI is defined by its ability to *generate* new things; it is a poor copyist. While it is possible to generate infringing works using such applications, the overwhelming majority of users generate original art, original text, or original code. This is not just a "significant noninfringing purpose," it is in furtherance of the purposes of copyright.

Question 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?*

Collecting and disclosing this information may be considered a good practice, but doing so at a level that would identify individual works would be so burdensome as to stop most AI training. Many common datasets comprise billions of individual works, most of which were gathered in an automated fashion. In many–even most–cases, the copyright "records" simply do not exist.

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

If AI training is a fair use, there should be no obligation. As stated by the Supreme Court in *Campbell v. Acuff-Rose Music, Inc.*: "If the use is otherwise fair, then no permission need be sought or granted."[69]

---

[68] "The question is thus whether the Betamax is capable of commercially significant non-infringing uses. In order to resolve that question, we need not explore all the different potential uses of the machine and determine whether or not they would constitute infringement. Rather, we need only consider whether on the basis of the facts as found by the District Court a significant number of them would be non-infringing. "Moreover, in order to resolve this case we need not give precise content to the question of how much use is commercially significant.", Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 442, 104 S. Ct. 774, 789 (1984).
[69] Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 586 (1994)

Question 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?*

Users of generative AI systems should be able to be able to claim copyright in the works that they create using AI tools. The current Office practice is based on incorrect factual assumptions, and is inconsistent with case law and ordinary practice.

Describing Generative AI

Generative AI applications are usually designed to produce outputs of the same type as the inputs. For instance, a generative model trained on text data may be used to generate new text, such as sentences, paragraphs, or even entire articles. A model trained on images may be used to create new images. However, there is no inherent restriction forcing AI systems to generate outputs of the same type as their inputs. For example, some generative systems, like the one in the *Zarya* case, can take an image as an input and return a textual description of the input, whereas other generative systems can take a textual description of a scene and return an image providing a rendering of the scene described by the user.

Referring back to the description of AI systems given relative to question 7, the generation process relies on the statistical patterns learned from the training data to create a predicted output. This output is returned as the result (or as part of the result). This is analogous to the law student's ability to create a new, coherent legal analysis based on the principles and lessons she derived from her case studies.

The difference between the law student and the AI application, however, is that the law student uses her intelligence and creativity to generate her answers, whereas an AI system has neither intelligence nor creativity. What the AI system has instead is *context* and *randomness*.

Context

Taking the example of text, scientists have known since the 1960s that it was possible to construct sentences by analyzing a bunch of writing, finding which words tend to follow each other, and then repeatedly picking out the next word with the highest probability.

Humans instinctively perform this kind of analysis. For example, if someone was asked to predict the next word in the sentence "It was a dark and stormy _________," almost everyone would respond with the word "night." Sometimes there are a number of possible "next words," such as in the sentence "The wizard raised his ___________." Some people might predict the next word might be "wand," "staff," or "hand."

When scientists tried to get computers to imitate humans, however, they quickly figured out that just choosing the most probable next word resulted in sentences that were trite, ungrammatical, and repetitive. The difference was that the computer was only considering the single preceding word. Humans take into account all the words in the sentence, as well as the millions of words of context accumulated throughout our lives.

The logical way to improve the quality of the sentences created was to use more context when determining the most probable next word. Instead of only looking at the immediately preceding word, the computer could look at the two preceding words, the three immediately preceding words, or more. Nevertheless, using more than about five words of context usually resulted in systems that were too big to run on the computers of the time.

In the past fifteen years, however, the storage and processing capabilities of computers and networked computer systems have grown exponentially. As of the writing of this article, state-of-the-art text generation systems are able to take in about fifty typewritten pages of context when determining what word to generate next.[70] Scientists have also identified methods (called "attention") of helping the model adaptively use different parts of its provided context to improve generation.

Randomness

The second ingredient in generation is randomness. Scientists have discovered that one ingredient that makes humans creative is the element of surprise. Humans don't always use the highest probability outputs. We vary how we express ourselves in order to produce different effects on readers or viewers.

To emulate this tendency in humans, data scientists building generative AI systems include a parameter (frequently called "temperature") that is

[70] GPT-4 Technical Report, https://arxiv.org/abs/2303.08774, What is the difference between the GPT-4 models? https://help.openai.com/en/articles/7127966-what-is-the-difference-between-the-gpt-4-models, OpenAI's GPT-4 is a safer and more useful ChatGPT that understands images, https://the-decoder.com/open-ai-gpt-4-announcement/ ("The context length of GPT-4 is limited to about 8,000 tokens, or about 25,000 words. There is also a version that can handle up to 32,000 tokens, or about 50 pages….")

interpreted as a probability that the model should choose a slightly lower-probability path for a part of its output. For example, a temperature of 0.7 could mean that there is a 70% chance that the highest probability path will be used when generating an output, and there is a 30% chance that one of the lower probability paths will be taken instead.

The "temperature" used in an application does not correspond to any physical or logical law. It is a heuristic, derived over time and observation, that causes AI systems to seem more "human" in their outputs. Many AI systems, including Midjourney, allow users to control the temperature used for a particular generation. This allows a human using the AI system the ability to guide the course of generation by using or constraining the level of randomness affecting the output.

Incorrect Factual Assumptions Informing the Office's Current Policy

The current reasoning of the Office is laid out in its decision regarding *Zarya of the Dawn*.[71] In the *Zarya* case, the Office took administrative notice of some documentation regarding how the generative AI system used by the registrant Kashtanova worked. However, the Office expressed an incorrect understanding of the technology, leading to incorrect results.

The crux of the Office's reasoning in *Zarya* is that Kashtanova had only limited control over the picture that was generated using the AI system. They recognized that the author has some control over what comes out of Midjourney but not enough: "the process is not controlled by the user because it is not possible to predict what Midjourney will create ahead of time" (p. 8) or "Rather than a tool that Ms. Kashtanova controlled and guided to reach her desired image, Midjourney generates images in an unpredictable way." (p. 9)

There are a number of errors with the Office's arguments, some legal and some factual. However, they all seem to stem from a core factual misunderstanding of the role that randomness plays in AI system generation.

The Office reasoned that the outputs of AI systems are almost totally "random" and "unpredictable," so whatever the artist puts into the AI system does not matter. At most it's a "suggestion" that can be ignored. (p, 10)

The Legal Standard for Copyrightability

The most fundamental error with the Office's expressed position is that it used the wrong legal standard. The standard is whether there is a "modicum of creativity," not whether Kashtanova could "predict what Midjourney [the AI

---

[71] U.S. Copyright Office, Cancellation Decision re: Zarya of the Dawn (VAu001480196) at 1 (Feb. 21, 2023), https://www.copyright.gov/docs/zarya-of-the-dawn.pdf

system] [would] create ahead of time." In other words, the Office incorrectly focused on the output of the tool rather than the input from the human.

Jackson Pollack famously couldn't predict how the paint he used would drip onto the canvas. Pollack designed his paintings - he knew what he wanted the end result to be - but he used a process involving random dripping and flicking of paint to make his art.

In photography, the closest analogous art, there are famous photographs that captured animals, people, or humorous situations entirely by mistake. Nevertheless, the output is still copyrightable, because a human had at least a "modicum" of input into the shot.

When examined from the correct legal standard - did Kashtanova provide a "modicum" of input - the Office's answer seems inconsistent. The Office recognizes that Kashtanova personally authored the prompts and other inputs. It just doesn't think Kashtanova did enough. But a "modicum" is the merest sliver of originality and creativity - "a very modest quantum of originality will suffice" (1 Nimmer on Copyright § 2.08). Courts have found that almost any decision that goes into a photograph will do - even decisions like which camera to use, choosing a brand of film, and taking several shots and picking one. Kashtanova made exactly analogous decisions and had the additional element of the personally composed prompts instead of just a simple "button press" as would occur on a camera.

Controlling the Generation of AI Systems

One factual error is encapsulated in the Office's statement that as "[The AI system's] specific output cannot be predicted by users… makes [the tool] different for copyright purposes than other tools used by artists." (p.9) The problem is that if this statement is construed narrowly enough to make it true, then it promulgates an incorrect standard. If this statement is construed more broadly, then it is factually false.

Randomness is part of the generative process–but *the output of an AI model is not random*. A human using the AI system typically describes what should be generated and/or provides other inputs that are used to initialize and guide the generative process. These inputs are usually referred to as the "prompt."

The AI system takes the prompt and analyzes it as if it were an input. It then uses the analyzed prompt to identify a place in the latent domain to focus on when running the generative process.

The practice of developing a prompt that will give the desired output is sometimes referred to as "prompt engineering." Prompt engineering is actually an exploration through the latent space of the model–the probabilistic landscape of ideas and meanings–to match the generated expression to the author's or artist's conception. The goal of the author is to develop the exact set of inputs–images, words, and options–that will lead to the generation of the desired output.

If the Office's statement is interpreted narrowly, it is true: under most circumstances, the *exact* output is not predictable. However, having 100% control over the output is also not the correct standard. The standard is a modicum of creative input leading to the output.

If the Office's statement is interpreted broadly, it is factually false. When a user provides "cute purple dinosaur" into an image generator, the application returns images of a cute purple dinosaur, not a motorcycle or a cloud. A truly random output would not be so controlled. Further, the more information that is given within the prompt, the more control is exerted over the output.

Examining the images in the *Zarya* case, they were created with consistent characters and a consistent style, both corresponding to the artistic vision provided by Kashtanova. Any generation that was essentially uncontrollable would not be capable of being bent to illustrate a consistent artistic vision.

Finally, there is a subtle factual error that appeared to influence the Office's thinking. The Office quoted from Midjourney's documentation describing diffusion models generally, and based on that reading, came to the conclusion that "Because [the AI system] starts with randomly generated noise that evolves into a final image, there is no guarantee that a particular prompt will generate any particular visual output." The Office compares the prompt to a mere "suggestion" that may or may not be followed by the tool."

The subtle error comes in a misunderstanding about the "randomly generated noise." Generative AI systems have two layers: a semantic "latent" layer associated with meaning and a rendering layer that generates the output. When a person inputs a prompt into the tool, the effect is to focus the attention of the tool on one or more specific spots in the latent domain, places that are statistically associated with particular meanings. There is random noise, but the visual layer evolves the final image from the noise **based upon the latent "meaning" in the prompt provided by the AI tool user**.

So when the Office compares the prompt to a "suggestion" like a patron might give to a painter, it is anthropomorphizing the tool and coming to an invalid conclusion. AI tools can't take "suggestions." They can only do exactly as they are programmed to do and pull from an *artist-chosen* place in its massive table of probabilities to drive the generation of an output.

Generative AI and Photography.

We have had all of these arguments before, about 150 years ago, relative to photographs. When photography was first invented, it was not considered art. In the 1844 publication "The Pencil of Nature," William Henry Fox Talbot wrote:

> [T]he the plates of this work have been obtained by the mere action of Light upon sensitive paper. They have been formed or depicted by optical and chemical means alone, and without the aid of any one acquainted with the art of drawing. It is needless, therefore, to say that they differ in all respects, and as widely us possible, in their origin, from plates of the ordinary kind, which owe their existence to the united skill of the Artist and the Engraver.[72]

Burrow-Giles Lithographic Co. v. Sarony[73] was first Supreme Court case to declare that photographers were authors and that their photos were protectable, despite photography's mechanized process.

The dispute in Sarony concerned the lithographic company's unauthorized reproduction and selling of 85,000 copies of photographer Napoleon Sarony's portrait of Oscar Wilde. The defendant made two arguments: First, that a photograph is not a "writing" by an "author" as the Constitution requires. Instead, Burrow-Giles argued that a photograph is mere output of a machine and thus was not the proper subject of copyright, which protects authors, not automated processes. Second, the defendant argued that photographs are by definition unoriginal, removing them from the scope of copyright protection.

The Supreme Court found that Sarony's photograph was copyrightable, but its explanation is instructive. It wasn't that the photograph was in itself copyrightable – it was *all the other things* that were done by the photographer that made it so:

> [P]laintiff made the [photo] . . . entirely from his own original mental conception, to which he gave visible form by posing the said Oscar Wilde in front of the camera, selecting and arranging the costume, draperies, and other various accessories in said photograph, arranging the subject so as to present graceful outlines, arranging and disposing the light and shade, suggesting and evoking the desired expression, and from such disposition, arrangement, or representation, made entirely by plaintiff, he produced the picture in suit.[74]

---

[72] Talbot, William. "*The Pencil of Nature*" Longmans, London 1844.
[73] 111 U.S. 53, 57–59 (1884).
[74] *Id*. at 110-111.

That left open the question as to whether all photographs were copyrightable. Perhaps if Sarony hadn't done as much posing and arranging, the photo would not bear his artistic stamp.

The next milestone was the 1903 case of *Bleistein v. Donaldson*.[75] In Bleistein, Justice Holmes found that an illustrated circus advertisement could enjoy copyright. The case is often remembered for the holding that the commercial character of a work does not deprive it of copyright. However, the decision also gives some important insight into originality. While Sarony speaks of the need for creativity, however modest, Bleistein only refers to uniqueness. "The copy," speaking of the poster, says Holmes:

> is the personal reaction of an individual upon nature. Personality always contains something unique. It expresses its singularity even in handwriting, and a very modest grade of art has something irreducible, which is one man's alone.[76]

In Holmes' view, originality depends on the author's original contribution, the personal stamp, not on intellectual labor. For Holmes, making artistic judgments was not a suitable approach because it risked leaving unappreciated creative works unprotected.

This view of copyright – that all that is needed is the bare minimum of originality, which came naturally from the artist's interaction with the tool – informed the next century of photographic copyright cases. Notable is the 1922 case J*ewelers' Circular Pub. Co v. Keystone Pub. Co.*.[77] *Jeweler's Circular* was about a directory showing photographs of different Jewelers' trademarks. In part, *Jewelers Circular* deals with the sweat of the brow doctrine, which was decisively rejected in *Feist*. But *Jewelers' Circular* also included commentary on the copyrightability of photographs by District Judge Learned Hand:

> Burrow-Giles...left open an intimation that some photographs might not be protected...I think that, even as to these, Bleistein...rules, because no photograph however simple, can be unaffected by the personal influence of the author, and no two will be absolutely alike. Moreover, this all seems to me quite beside the point, because under § 5(j) photographs are protected, without regard to the degree of "personality" which enters into them. At least there has been no case since 1909 in which that has been held to be a condition. The suggestion that

[75] Bleistein v. Donaldson Lithographing Co., 188 U.S. 239 (1903)
[76] *Id*. at 188.
[77] 281 F. 83. (2d Cir. 1922), cert. denied, 259 U.S. 581 (1922),

> the Constitution might not include all photographs seems to me overstrained.[78]

This view – that the copyright law extends to all photographs – was described as the law in *Time Inc. v. Bernard Geis Associates*,[79] the case about the Zapruder film. According to the judge's reading of the then current Copyright Act, "Congress has expressly made photographs the subject of copyright, without any limitation" because no photograph "can be unaffected by the personal influence of the author."[80]

The court then quoted *Nimmer on Copyright*, page 99, endorsing the reasoning in *Jewelers Circular* as "the prevailing view." The court further quoted Nimmer saying:

> * * * any (or as will be indicated below, almost any) photograph may claim the necessary originality to support a copyright merely by virtue of the photographers' personal choice of subject matter, angle of photograph, lighting and determination of the precise time when the photograph is to be taken.[81]

The *Zarya* decision from the Office only quoted Sarony–and accordingly came to the *Saronyl*-like decision that AI-assisted art is only copyrightable because all the additional things that a human does around the generation by the AI tool. But *Sarony* was not the last word in the copyrightability of machine-mediated art.

The only consistent outcome is to treat AI art as we did photography: that as long as a human had some minimal input into the generation of the work, that work should be by default copyrightable. The Office should not be trying to evaluate what came from the human and what came from the machine. As it says in Feist:

> Original, as the term is used in copyright, means only that the work was independently created by the author (as opposed to copied from other works), and that it possesses at least some minimal degree of creativity. To be sure, the requisite level of creativity is extremely low; even a slight amount will suffice. The vast majority of works make the grade quite easily, as they possess some creative spark, "no matter how crude, humble or obvious" it might be.

[78] *Id*. at 934-
[79] Time Inc. v. Bernard Geis Assocs., 293 F. Supp. 130.
[80] *Id.* at 141.
[81] *Id,* at 142.

Question 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 changes are needed to the Copyright Act to clarify the human authorship requirements. Only humans are eligible for copyright. However, the Office should in most cases not be evaluating the extent of the human input as part of the registration process. Photographs are considered *prima facie* copyrightable, unless there is additional information that suggests another outcome. The Office should treat AI-assisted art equally.

Question 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?*

Legal protection of AI-assisted art is the most urgent need from the Copyright Office. AI tools are entering every creative endeavor. They are enabling thousands of new artists to create new works. But under the Office's current rules, none of those works are protectable.

Sincerely,

Van Lindberg
Taylor English Duma, LLP