# Adobe Inc. Comments on the
# U.S. Copyright Office Notice of Inquiry and Request for Comments on
# Artificial Intelligence and Copyright

**October 30, 2023**

Adobe Inc. ("Adobe") welcomes this opportunity to provide comments to the United States Copyright Office ("US Copyright Office" or "USCO") in response to the Notice of Inquiry and Request for Comments ("RFC") on Artificial Intelligence ("AI") and Copyright published on August 30, 2023.

## I.      Background

Adobe has a long history of pioneering innovation. We are the company that brought the world Acrobat and PDF to enable you to create and share digital documents everywhere. We have a digital experience business that powers small, medium, and large businesses' websites and e-commerce experiences. And we have a collection of products that enable all types of creative expression—including Adobe Illustrator for graphics, Adobe Premiere for professional video editing, and of course, Photoshop, our leading image editing solution. Since our founding in 1982, we've continued to invest in transformative technologies that allow our customers to unleash their creativity, perfect their craft, and power their businesses in a digital world.

At Adobe, our customers represent a vast array of creators—ranging from aspiring digital artists, war-time photographers, fashion designers, marketing and advertising specialists, and the world's leading enterprises. Our millions of customers use Adobe's tools to create the billboards you see in Times Square, Academy Award-winning movies, the cover art of albums, illustrations in children's books, the photographs you see in magazines and newspapers, and even the flyers on the windows at your local coffee shop. Throughout our history, we have worked together with our creative community as we harness the power of emerging technologies in the tools we provide to support them.

The newest advance is artificial intelligence. Adobe has been incorporating AI into our tools for more than a decade to help creators realize their potential. In content creation, AI will usher in a new age. Billions of new creative voices will be able to express themselves through these new easy-to-use tools. For professionals, AI will provide an amazing first step in their creative process and an opportunity to design entirely new experiences bounded only by their imagination and allow them to be more productive than ever before. This is the power and promise of AI.

As we, and the industry at large, continue to harness the power of this new technology, it's important that we be responsible about the impact it can have on society. When it comes to creators and intellectual property rights in the age of AI, there are important questions raised throughout the AI generation process—from training an AI model to outputs of that AI system—that highlight profound policy concerns.

At Adobe, we considered these questions and developed a comprehensive approach to AI—from training our AI models to economic solutions for artists—that aims to empower creators and enable AI

innovation to thrive. Adobe recognized the various unanswered legal questions around access to data in designing our first family of generative AI models, [Adobe Firefly](), which we launched in March 2023. We trained our first Firefly model on licensed images from our own Adobe Stock photography collection, openly licensed content, and public domain images. To help ensure copyrighted or branded materials are not created as part of Firefly's output, we have a content moderation team that performs extra filtering on the images before they become part of the Firefly dataset. Yet as AI continues to evolve, it's important that we work together to address the many still-unanswered legal and policy questions around this technology.

## II.    Training

### 1.    Access to Data is Key

AI is only as good as the data on which it is trained, which is why the question of data access is important to companies and organizations building foundation AI models. Like the human brain, AI learns from the experiences or information you give it. And like the human brain, the more information you give it, the better it will perform. If, for example, you've never been taught what a car is, it would be hard for you to accurately depict one or answer a question about what a car is or what it does. Likewise for AI, these models to need learn from a large dataset representing the universe of possible answers and diverse set of facts to produce accurate results.

That means an AI system trained on a small dataset is at greater risk of producing wrong or unsatisfactory results or reproducing harmful biases that exist within the dataset. Say, for example, you've only ever seen male lawyers. You are then more likely to conjure up an image in your head of a man when someone is talking about a lawyer, even though over half of the graduates of law schools are women. In the same way for AI, training on a larger dataset can help ensure you capture a broader set of perspectives in the training data itself. That way, when you prompt an AI model about "lawyer," you'll see a set of results that reflects the society in which you live. Given those technical realities, governments need to support access to data to ensure that AI innovation can flourish both accurately and responsibly.

In addition to inclusivity in training sets, governments need to support inclusivity in the AI ecosystem more broadly when designing AI policies. Because today's AI models require vast amounts of data to train on, smaller AI developers like startups need ready access to such datasets to avoid industry consolidation. And to ensure that progress in the AI field is not subject to industry capture, academics and other non-profit research organizations with finite resources also need ready access to the large datasets their projects require.

### 2.    AI Training & Fair Use

One important issue arising from AI's need for data access is its implication on creators and copyright. Top of mind for the creative community is how training an AI model on copyrighted materials might impact creators' rights.

Copyright is a critical intellectual property protection that has been part of our legal system since the founding of our nation. It is enshrined in our Constitution as an important lever to "promote the

progress of science and useful arts" by giving creators a right to protect their interests by protecting against unauthorized copying of their work.

But that protection "has negative features," such as "sometimes stand[ing] in the way of others exercising their own creative powers."[1] The fair use doctrine has therefore played an equally important role in the application of copyright law, originating in the courts as an "equitable rule of reason" and codified in the 1976 Copyright Act as a statutory exception to copyright. It is designed to allow copyrighted works to be used for purposes such as parody, research, and other transformative uses. Fair use ensures that copyright law does not "stifle the very creativity which [it] is designed to foster"[2] by balancing "relevant circumstances, including 'significant changes in technology.'"[3]

Training a generative AI model is a multistep process. Text-to-image models, for example, are trained on a large dataset—currently up to hundreds of millions, or billions—of images, along with their associated captions ("inputs"). The images are used as part of a feedback loop to create "weights" (mathematical constructs that are built into the AI model), which help the AI model learn a mathematical function that can produce an image, given a caption or "prompt." Through the learning process, information from the dataset is distilled into the weights such that the AI produces the appropriate visual attributes in response to a text prompt. The weights are iteratively adjusted until the model generates a satisfactory output based on the prompt. Once the model has been trained, a user can submit a prompt and the trained model will generate an output based on its weights, or its learned understanding of what visual attributes it thinks would be most likely associated with the material in the text prompt.

### a) Relevant Case Law

Fair use precedent dealing with "significant changes in technology" make clear that use of copyrighted works for purposes like training AI models is transformative. In *Sega v. Accolade*,[4] the Ninth Circuit held that intermediate copying of Sega's software was fair use. The defendant made copies while reverse engineering to discover the functional requirements—unprotected information—for making games compatible with Sega's gaming console.[5] Such intermediate copying also benefited the public: it led to an increase in the number of independently designed video games (which contain a mix of functional and creative aspects) available for Sega's console. This growth in creative expression was precisely what the Copyright Act was intended to promote.[6]

*Sony Computer Entertainment, Inc. v. Connectix Corp.*[7] similarly applied fair use to intermediate copying necessary to reverse engineer access to unprotected functional elements within a program. The defendant's use of Sony's software was transformative because it allowed users to play PlayStation games in a new computing environment (personal computers) and created "a wholly new product, notwithstanding the similarity of uses and functions" between the PlayStation gaming console and the defendant's program.[8]

---

[1] Google LLC v. Oracle Am., Inc., 141 S. Ct. 1183, 1195 (2021).
[2] Campbell v. Acuff-Rose Music, 510 U.S. 569, 577 (1994) (quotation omitted).
[3] Google, 141 S. Ct. at 1197 (quoting Sony Corp. of America v. Universal City Studios, Inc., 464 U.S. 417, 430 (1984)).
[4] 977 F.2d 1510 (9th Cir. 1992).
[5] Id. at 1522.
[6] Id. at 1523.
[7] 203 F.3d 596 (9th Cir. 2000).
[8] Id at. 606.

So too with AI model training. Inputs are temporarily accessed for the unprotected ideas, concepts, and styles contained in the dataset—say, the number of fingers a human hand has, or what cars look like—to help the AI model learn facts about the world. As in *Sega* and *Sony*, such intermediate copying benefits creators by enabling a new platform for creativity and fostering new works at a broader scale.

Other analogous case law show that the process of an AI model's training on inputs—the iterative creation of mathematical "weights" that help an AI model learn how to produce a satisfactory output— is a purpose different from that of any expressive content in those training sets and therefore constitutes fair use. In *Kelly v. Arriba Soft Corp.*, for example, the Ninth Circuit held that Arriba's creation of thumbnails of copyrighted images was a transformative use because the thumbnails in Arriba's search engine "function[] as a tool to help index and improve access to images on the internet," instead of serving a purpose "to inform and to engage the viewer in an aesthetic experience."[9] Similarly, the Second Circuit held in *Authors Guild v. Google, Inc.* that "Google's copying of the original copyrighted books" to "make available significant information *about those books*" in its search database was transformative use.[10]

*Andy Warhol Foundation v. Goldsmith*[11] does not change the impact of relevant fair use precedent. As one court recently observed in *Thomson Reuters Enter. Ctr. Gmbh v. Ross Intelligence Inc.*,[12] compared to the facts in *Andy Warhol*, AI model training involves "a technological context much more like" in *Google v. Oracle America*—a case that also addressed new technology that reshaped a market (for the smartphone space) and where the use was commercial but transformative nonetheless. And as in *Google v. Oracle America*, AI models offer users "a highly creative and innovative tool"; their training on input content is therefore "consistent with the creative progress that is the basic constitutional objective of copyright itself."[13]

### 3.    Opt-Out Mechanism: "Do Not Train" tags

As we developed Adobe Firefly, we also believed there were important steps industry could take as a whole to empower creators in the age of AI, including giving creators greater control over how their content is used in AI model training.

At Adobe, we developed a technology called Content Credentials based on our work with the Adobe-led Content Authenticity Initiative. Content Credentials allow creators to securely attach a "Do Not Train" tag in the metadata of their work to indicate a preference to opt out of AI training. This credential will travel with the content wherever it goes, and a web-crawler scraping the web to build a dataset will recognize the Do Not Train tag. Creators can thus use this credential to prevent AI developers from training on their works. This technology is part of the open standard behind the Content Authenticity Initiative, and anyone can join to implement this credential. With government support, we can make this standard industry-wide so that everyone can enable and respect these credentials and therefore empower an artist to opt out of training if that's their preference.

---

[9] 336 F.3d 811, 819 (9th Cir. 2003).

[10] 804 F.3d 202, 214, 217 (2d Cir. 2015) (emphasis in original). *See also* Authors Guild, Inc. v. HathiTrust, 755 F.3d 87 (2d Cir. 2014) (holding that the creation of a full-text searchable database resulted in a "word search [that] is different in purpose, character, expression, meaning, and  message" from the original books and therefore fair use).

[11] 598 U.S. 508 (2023).

[12] 1:20-cv-613-SB, 17 (D. Del. Sep. 25, 2023).

[13] 141 S. Ct. 1183, 1203 (quotations omitted).

### 4. Unlearning

While "unlearning" methods are not yet ready to be implemented in a commercial setting, as part of Adobe's Responsible AI efforts, our researchers are actively exploring unlearning techniques that could benefit both AI developers (by saving retraining costs) and copyright owners (by giving a more feasible way to get their content out of a training set) generally. Adobe encourages the Copyright Office to design policies that will support research concerning this important question.

### 5. International Harmonization

It is important to ensure the responsible development and use of AI is governed by a unified, global framework. We are encouraged to see the Copyright Office taking a comprehensive approach to understand AI, the application of current copyright law, and potential new policies and safeguards that can be put into place to help AI realize its full potential to benefit society.

The reach of AI technology, however, is borderless. International harmonization of copyright protections will thus be key to AI development. Given the numerous factors AI developers must already navigate in this quickly evolving landscape and the steep costs and resources required for developing AI technology, AI developers need a consistent copyright framework to operate within. The Copyright Office's collaboration across jurisdictions to achieve harmonized policies and ensure the United States stays at the forefront of AI development is an important step to achieving this.

## III. Transparency & Recordkeeping

Transparency is one of Adobe's AI Ethics principles and key to our overarching approach to AI. At Adobe, we inform our customers of when and how AI is used in our tools and products. We are also transparent about the types of datasets we train our models on. For example, we trained our first Firefly model on licensed images from our own Adobe Stock photography collection, openly licensed content, and public domain images.

Companies should disclose general information about their dataset sources, as we did with our first Firefly family of models. However, AI developers may regard specific information regarding dataset sources as proprietary information, as the selection of curated training materials may give a competitive advantage. There may also be other concerns such as confidentiality obligations when licensing datasets, for example, that could hinder an AI developer from disclosing those specific datasets. Transparency must therefore be balanced with these concerns.

## IV. Copyrightability

Another key question creators are grappling with is whether the output of an AI model can earn a copyright. Global copyright law says that you can't copyright an idea. What you *can* copyright is the expression of an idea. So, for example, while you can't copyright the word "spaceship"—an idea— you can copyright a painting of a spaceship—the expression of the idea. When applying this principle to AI image generation, that means a prompt is not copyrightable because the prompt represents the idea. When you submit a   prompt (or idea), you then receive an output based solely on the AI's interpretation of that prompt. In other words, AI is creating the expression of the idea, not the human. Copyright law

exists to protect the expression created by human creators; therefore an AI's expression of an idea is not copyrightable.[14]

But what about the creator who ideates and brainstorms in a generative AI tool and then goes on to add their own style and flare? Take our spaceship example. Maybe the artist starts with the AI-generated output but then uses other tools to change the spaceship's color, draw more planets, and add in astronauts inside the spaceship. Or maybe they already created an image of a spaceship using non-AI tools and they used AI to simply swap out the sky for a more star-filled horizon. Additionally, as AI becomes more sophisticated, prompts may become multimodal, where an input could be comprised of not just text, but also a creator's own starting image, which would be considered copyrightable. It's important to make sure that as creators leverage new technologies, they have ways to get ownership and credit for the human creativity and expression they bring to their work.

If a creator wants to use generative AI as a starting point and later register their work with the Copyright Office, they will likely need to disclaim which portions were generated by AI. In the event the Copyright Office (or courts) need to validate this information, Content Credentials give creators a way to show what they have done to an image after AI was used to demonstrate the level of human expression. As with so many aspects of copyright law, courts will need to determine the appropriate level of human expression needed in order to obtain a copyright.

Knowing how a piece of digital content was created—whether with human direction, AI assistance, or AI generation—is more important than ever. The Copyright Office has an important leadership role to play in ensuring creators can get ownership and credit for their work and helping the public understand where that work came from and who has the rights to it. Adobe will continue to explore ways to empower creators in this era of new technology.

## V.      Labeling or Identification

### 1.      Transparency in Digital Content

As AI becomes more powerful and prevalent, it's increasingly important for people to have a way to tell how a piece of content was created. AI-generating tools make it easy to create realistic synthetic imagery in the span of seconds. In the hands of bad actors, this technology raises valid concerns over the ability to produce and spread false content at mass scale. Without a way to tell where a piece of content came from, people are more prone to believe lies. And once they realize there's misinformation out there, it becomes harder to know what's true. To restore trust in the age of AI, it's essential that we implement tools to help people navigate an AI-powered world by providing more transparency around digital content.

This is the approach behind Content Credentials—an open-source technology that lets you see exactly where a piece of content came from, who created it, and what edits were made to it along the way. This approach is backed by nearly 2,000 members of the Adobe-led Content Authenticity Initiative and built upon an open standard developed by the Coalition for Content Provenance and Authenticity (C2PA). Members range from technology companies like Adobe and NVIDIA to gen AI developers like Stability AI to news organizations like the New York Times and the Wall Street Journal to camera companies like

---

[14] *Thaler v. Perlmutter*, 1:22-cv-01564-BAH (ECF #24), D.D.C. (Aug. 18, 2023).

Nikon and Leica to academic organizations, non-profits, and more, and we are already starting to see this technology implemented across member company tools and platforms.

2. **Content Credentials: Transparency Systems to Effectively Restore Trust in AI and Digital Content**

Some approaches have suggested labeling AI-generated content with visible watermarks to combat misinformation. Aside from wholly generated images for which this may be helpful, merely labeling something as having AI in it does not provide consumers with enough information or context to be useful in most cases. AI labels do not account for a host of nuanced ways AI could be used to generate content. Creators are already using this technology to bring precision and ease to their workflows; conjure up beautiful imagery; design marketing materials, album covers, social media posts, birthday invitations; and more. Beyond creative use cases, there are many good reasons to use AI in newsworthy content. For example, a journalist may want to use AI to quickly remove the visible license plate numbers in the background of an image to protect people's privacy. AI is a positive tool in the hands of the vast majority of users, and soon we will find ourselves in a world where almost every piece of content has AI-generated pixels in it. That's why, beyond knowing that AI was used, it's more impactful to know *how* it was used. For example, was AI used to quickly remove a blemish or change the color of the sky? Or did someone misuse it to manipulate an image to try to deceive people? These are very different applications of AI that mere labeling would not explain. Having more context about how AI was used in each scenario is an important factor in determining whether to trust the content.

Technologies like Content Credentials provide this important level of context. By giving good actors—like those who want to deliver trustworthy content—a way to be believed when using AI to add quality and clarity to their work more easily and efficiently they enable people to make an informed decision about whether to trust any AI-edited content they are consuming.

3. **Government's Role**

We believe government has an important role to play in restoring trust in digital content. In order for solutions like Content Credentials to work, we need them everywhere. Adobe is committed to working together across the public and private sector to drive broad adoption of this technology. CAI members have already begun embedding Content Credentials into their tools and platforms and people are already attaching them to bring more transparency to their work.

Government can help support this solution by becoming early adopters of this technology. For example, any U.S. federal agency or department with a website and official images can implement Content Credentials. And to the extent content contains Content Credentials already, government should also require that these credentials stay attached to the content and not be stripped away. By requiring platforms to carry and display Content Credentials metadata, users can digest content in an informed way regardless of where it's found on the Internet or on which platform it is hosted. To add weight to this requirement, government can consider implementing penalties for those who strip out this important metadata.

**VI.     New Legal Rights in Age of AI**

As the Copyright Office considers these important questions raised by generative AI, we believe government should consider new regulation that can provide specific protections for creators that copyright law currently cannot cover.

One of the core questions raised by Adobe's customers is, what happens when someone uses an AI model to replicate their style, in direct economic competition with their original work?  Adobe believes the intentional misuse of AI tools in this way could pose a legitimate issue for our creative community.

Adobe has proposed that Congress establish a new Federal Anti-Impersonation Right (the "FAIR" Act) to address this type of economic harm. Such a law should provide a right of action to an artist against those who are intentionally and commercially impersonating their work and style through AI tools. These protections would provide a new mechanism for artists to protect their livelihood from people misusing this new technology without having to rely solely on mismatched laws around copyright and fair use. The premise of the proposed FAIR Act is simple: intentional impersonation using AI tools for commercial gain isn't fair.

Here are a few key points about this approach:

1) This right applies specifically to generative AI work (and does not extend to any existing rights in the physical world).

2) This right creates liability for the misuser of the AI tool, allowing the artist to go after the misuser directly.

3) The right requires intent to impersonate. If an AI generates work that is accidentally similar in style, no liability is created. Additionally, if the generative AI user had no knowledge of the original artist's work, no liability is created (just as in copyright today, independent creation is a defense).

4) This right should be enacted at the federal level to avoid multiple conflicting state laws.

5) This right should include statutory damages that award a preset fee for every harm, to minimize the burden on the artist to prove actual economic damages.

The FAIR Act is drafted narrowly to specifically focus on intentional impersonation for commercial gain while allowing style innovation to flourish.

**VII.     Conclusion**

AI is a transformative technology that has the power to unleash human creativity in new ways, raise human ingenuity to new levels, and lift our society to unimagined heights. The ways creative professionals have already begun to leverage this technology have been impactful and inspiring. As with all technologies, it is important to consider thoughtfully how AI will transform industries and society at

large, and to work together to implement the right safeguards and frameworks to bring this technology to the world responsibly.

Adobe is committed to continuing to advocate for policies governing the development and use of AI that empower the creative community we serve. Copyright and creators' rights in the age of AI is a critical area we must focus on and we applaud the Copyright Office in taking a leadership role in seeking public input on this important topic. Adobe looks forward to ongoing collaboration with government, industry, creators, and the public to ensure AI is developed and implemented in the right way for everyone.