Congressional Research Service
Informing the legislative debate since 1914

Legal Sidebar

Generative Artificial Intelligence and Copyright Law

Updated November 17, 2026

Generative artificial intelligence (AI) programs, which generate new text, images, music, or other content ("outputs") in response to user inputs ("prompts"), have raised several questions in the field of copyright law. These questions include whether content created using generative AI may receive copyright protection as well as whether training and using generative AI programs may infringe existing copyrights.

This Legal Sidebar provides an overview of how U.S. courts and stakeholders have sought to apply the copyright laws enacted by Congress to issues raised by generative AI, including summaries of key court decisions and actions by the U.S. Copyright Office. It also provides a brief introduction to the copyright law concepts of authorship, infringement, and fair use underpinning these issues.

Other CRS Legal Sidebars explore questions AI raises in the intellectual property fields of patent law and the right of publicity.

Copyright in Works Created with Generative AI

Do Copyrighted Works Require a Human Author?

The question of whether copyright protection may be afforded to generative AI outputs hinges largely on the legal concept of "authorship." Article I, Section 8, Clause 8 of the U.S. Constitution, often referred to as the Intellectual Property (IP) Clause, empowers Congress to "secur[e] for limited Times to Authors . . . the exclusive Right to their . . . Writings." Based on this authority, the Copyright Act affords copyright protection to "original works of authorship." While the Constitution and Copyright Act do not explicitly define who (or what) may be an "author," U.S. courts have not recognized copyright in works that lack a human author, including works created autonomously by AI systems.

Before the proliferation of generative AI, a series of U.S. court decisions rejected copyright protection for other kinds of nonhuman authors, holding that a monkey who took photos of himself lacked standing to sue under the Copyright Act, that human authorship was required to copyright a book purportedly inspired by celestial beings, and that a living garden could not be copyrighted. The U.S. Copyright Office has also long maintained that copyrighted works must be "created by a human being" and therefore

Congressional Research Service
https://crsreports.congress.gov
LSB10922

refused to register works that are "produced by a machine or mere mechanical process that operates randomly or automatically without any creative input or intervention from a human author."

One recent lawsuit has unsuccessfully challenged the human-authorship requirement in the AI context. In June 2022, Stephen Thaler sued the Copyright Office for denying his application to register a visual artwork that he claimed was authored "autonomously" by an AI program. Dr. Thaler argued that human authorship is not required by the Copyright Act. In August 2023, a U.S. district court granted summary judgment in favor of the Copyright Office. The court held that "human authorship is an essential part of a valid copyright claim," reasoning that only human authors need copyright as a creative incentive.

In March 2025, the U.S. Court of Appeals for the D.C. Circuit affirmed the district court's decision in *Thaler v. Perlmutter*, holding that the Copyright Act "requires all eligible work to be authored in the first instance by a human being." The court reasoned that several provisions of the Copyright Act imply that the word "author" refers only to human beings, including provisions that refer variously to authors' "property, . . . lifespans, family members, domiciles, nationalities . . . or signatures." In addition, the D.C. Circuit observed that the Copyright Office had adopted the human-authorship requirement before Congress enacted the current Copyright Act. The court thus inferred that Congress meant to enshrine the human-authorship requirement when it enacted the law. Based on its holding that the Copyright Act requires human authorship, the court found it unnecessary to consider the Copyright Office's alternative argument that the Constitution's IP Clause requires human authorship.

In May 2025, the D.C. Circuit denied Dr. Thaler's petition to rehear the case en banc (i.e., by all of the court's judges). In March 2026, the U.S. Supreme Court denied Dr. Thaler's petition to review the case.

May Humans Copyright Works That They Create Using AI?

The human-authorship requirement does not necessarily preclude copyright protection for works created with the assistance of generative AI, depending on the nature of human involvement in the creative process. The Copyright Office, as the agency responsible for evaluating applications to register copyrights, has been faced with articulating how the human-authorship requirement may be satisfied for works made using generative AI. (Although copyrights do not need to be registered, registration provides certain benefits to copyright owners and is a prerequisite for filing a lawsuit for copyright infringement.)

In March 2023, the Copyright Office released Copyright Registration Guidance for "works containing material generated by [AI]" (the AI Guidance). Granting that human authors may use AI in their creative process, the AI Guidance states that "what matters is the extent to which the human had creative control over the work's expression." The AI Guidance explains that when AI "determines the expressive elements of its output, the generated material is not the product of human authorship." Thus, it states, human authorship is not present "when an AI technology receives solely a prompt from a human."

Nonetheless, the AI Guidance states that works containing AI-generated material may be copyrighted under some circumstances, such as "sufficiently creative" human arrangements or modifications of AI-generated material, or works that combine AI-generated and human-authored material. The AI Guidance states that authors may claim copyright protection only "for their own contributions" to such works, and they must identify and disclaim AI-generated parts of the works when applying to register their copyright.

To illustrate these principles, the Copyright Office has highlighted three of its decisions denying registration for works made with generative AI:

- *Zarya of the Dawn*: A February 2023 decision that AI-generated illustrations for a graphic novel were not copyrightable, although the human-authored text of the novel and overall selection and arrangement of the images and text in the novel could be copyrighted.
- *Théâtre D'opéra Spatial*: A September 2023 decision that an artwork generated by AI and then modified by the applicant could not be copyrighted, since the applicant failed to identify and disclaim the AI-generated portions as required by the AI Guidance.
- *SURYAST*: A December 2023 decision that an artwork generated by an AI system combining a "base image" (an original photo taken by the applicant) and a "style image" the applicant selected (Vincent van Gogh's *The Starry Night*) could not be copyrighted, since the AI system was "responsible for determining how to interpolate [i.e., combine] the base and style images."

While the foregoing decisions indicate that it may not be possible to obtain copyright protection for many AI-generated works, the issue remains unsettled, partly because the Copyright Act allows applicants to challenge the Copyright Office's final refusal to register a copyright in federal court. The putative author of *Théâtre D'opéra Spatial*, for instance, has sued the Copyright Office for declining to register that work. Both the plaintiff and Copyright Office have filed motions for summary judgment in that case (*Allen v. Perlmutter*), expressing conflicting views regarding whether works created by generative AI in response to user prompts may be copyrighted. While the AI Guidance observes that courts may give some weight to the Copyright Office's expertise, courts are not legally bound to adopt the office's interpretation of the Copyright Act, such as how it applies the human-authorship requirement to AI-assisted works.

In January 2025, the Copyright Office published an installment of its multipart report *Copyright and Artificial Intelligence* addressing the copyrightability of AI-generated works. Reinforcing the AI Guidance's emphasis on "creative control," the report concludes that, "given current generally available technology, prompts alone do not provide sufficient human control to make users of an AI system the authors of the output." The report contends that the Copyright Act's distinction between copyrightable "works" and noncopyrightable "ideas" precludes copyrightability for works generated by AI in response to user prompts. Specifically, the report argues, "[p]rompts essentially function as instructions that convey unprotectible ideas" and "do not control how the AI system processes them in generating the output."

Some commentators assert that some AI-generated works should receive copyright protection, comparing AI programs to other tools human authors have used to create copyrighted works. For example, the Supreme Court held in the 1884 case *Burrow-Giles Lithographic Co. v. Sarony* that photographs can be entitled to copyright protection where the photographer makes decisions regarding creative elements such as composition, arrangement, and lighting. Some copyright applicants argue that generative AI programs may function as tools analogous to cameras. The Copyright Office disputes this analogy, arguing that AI users do not exercise sufficient creative control to characterize generative AI as a tool. Instead, the Copyright Office has analogized an AI user to "a client who hires an artist" and provides only "general directions." Others argue that the Copyright Act does not require such exacting control.

Regarding works that combine human-authored and AI-generated material, the Copyright Office reported in January 2025 that it "ha[d] registered hundreds of works that incorporate AI-generated material, with the registration covering the human author's contribution to the work." The office took the position that new legislation regarding "the copyrightability of AI-generated material" is currently not needed, indicating that courts "will provide further guidance on the human authorship requirement as it applies to specific uses of AI" and that, since each work must be analyzed individually, "greater clarity would be difficult to achieve" through legislation.

Copyright Infringement by Generative AI Programs

Does the AI Training Process Infringe Copyrights in Other Works?

Generative AI systems are trained to create literary, visual, and other works by exposing them to large amounts of data, which may include text, images, sound recordings, or other works obtained by AI companies. This process often involves making digital copies of existing works that may be copyrighted, without the copyright owners' permission. According to one survey, copyright owners have filed over 150 lawsuits against AI companies in the United States, generally claiming that the AI training process infringes their exclusive right to make reproductions (or copies).

Many AI companies, some legal scholars, and the White House assert that using copyrighted works to train AI systems constitutes fair use and is therefore noninfringing. In September 2026, for example, the U.S. Department of Justice filed a brief supporting OpenAI's defense against copyright infringement litigation, arguing that "the United States has a strong interest in . . . rejecting any argument that training [large language models] on copyrighted texts violates copyright law."

Whether or not unauthorized copying constitutes fair use depends on four nonexclusive factors that Congress codified in the Copyright Act:

1. the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
2. the nature of the copyrighted work;
3. the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
4. the effect of the use upon the potential market for or value of the copyrighted work.

The Supreme Court has stated that fair use is a "flexible" doctrine, and "its application may well vary depending upon context." As to the first factor, the Court has held that some uses with a "transformative" purpose (such as parodies) may be fair, although in 2023 it cautioned that transformativeness is "a matter of degree." The Court has described the fourth factor as the "most important" one.

In June 2025, the U.S. District Court for the Northern District of California issued summary judgment decisions in two lawsuits concerning generative AI and fair use. In one case, *Bartz v. Anthropic PBC*, the defendant trained an AI system (Claude) on books obtained through a combination of buying print copies and downloading digital copies from "pirate sites." The district court held that it was fair use to copy the books to train Claude, reasoning that generative AI is "quintessentially transformative" (factor 1), that copying entire books was "reasonably necessary" to train Claude (factor 3), and that Claude does not create copyright-infringing outputs that would "displace demand" for the books (factor 4). The court found that only factor 2 tilted against fair use, since the books were used for their "expressive qualities." The court also held that it was fair use for Anthropic to convert print books into digital versions. But, the court held, it was not fair use for Anthropic to download pirated books and maintain them in a "central library," only parts of which were used for AI training. In July 2026, the court approved a class action settlement by which Anthropic agreed to pay $1.5 billion to copyright owners and their attorneys.

In the other case, *Kadrey v. Meta Platforms, Inc.*, a different judge ruled that it was fair use for Meta to train a generative AI system (Llama) on books it downloaded from pirated "shadow libraries." As in *Bartz*, the court in *Kadrey* held that this use was "highly transformative" and that factor 1 as well as factor 3 favored fair use, while factor 2 did not. Regarding factor 4, the court criticized plaintiffs for failing to develop evidence of market harm, potentially forfeiting a winning argument. Disagreeing with the judge in *Bartz* about the potential scope of factor 4, the court in *Kadrey* reasoned that market competition from noninfringing AI outputs (for example, AI-generated books about the same topics as the

copyrighted books) may weigh against fair use no less than competition from infringing outputs. The court opined in dicta that factor 4 might weigh "decisively" against fair use in other generative AI cases. The *Bartz* and *Kadrey* judges' disagreement reflects a wider debate over whether "market dilution" from noninfringing AI outputs is a cognizable harm under factor 4. The Copyright Office, for instance, has argued that market dilution may weigh against fair use, while some scholars have criticized this view.

The *Bartz* and *Kadrey* decisions also took different approaches to the initial downloading of plaintiffs' books without the authors' permission. Unlike in *Bartz*, the court in *Kadrey* did not conduct a separate fair-use analysis of these downloads. The court reasoned that Meta's purpose in downloading the books was to train Llama, and "[b]ecause Meta's ultimate use of the plaintiffs' books was transformative, so too was Meta's downloading . . . ." By contrast, the court in *Bartz* found that Anthropic had impermissibly obtained "a central library of works to be available for any number of further uses." Yet in dicta, the *Bartz* court doubted that it would ever be fair use to download books from pirate sites.

In light of these unsettled legal debates and the fact-specific nature of fair use, courts in other lawsuits may reach different decisions. In May 2025, the Copyright Office released a prepublication version of the Generative AI Training part of its *Copyright and Artificial Intelligence* report, which concluded that "some uses of copyrighted works for generative AI training will qualify as fair use, and some will not."

In September 2026, the U.S. Court of Appeals for the Third Circuit ruled that it was not fair use for a company (ROSS Intelligence) to copy case summaries from Westlaw, a legal research platform, to train an AI program to quote passages from legal opinions in response to user questions. This decision may have limited importance to copyright litigation regarding generative AI, since, as the court observed, "ROSS's AI was not a generative AI, meaning it would not create any new expression."

Do AI Outputs Infringe Copyrights in Other Works?

Some outputs of a generative AI program might infringe copyrights in works that were used to train the program. Copyright owners may be able to establish that such outputs infringe their copyrights if the AI program both (1) copied their works and (2) created "substantially similar" outputs. First, to establish copyright infringement, a plaintiff must prove the infringer "actually copied" the underlying work. This element is sometimes proven circumstantially by evidence that the infringer "had access to the work." For AI outputs, access might be shown by evidence that the AI program was trained on the underlying work.

Second, a plaintiff must prove that the new work is "substantially similar" to the copyrighted work to establish infringement. The substantial similarity test is difficult to define. Courts have variously described the test as requiring, for example, that the works have "a substantially similar total concept and feel" or "overall look and feel" or that "the ordinary reasonable person would fail to differentiate between the two works." Leading cases have also stated that this determination considers both "the qualitative and quantitative significance of the copied portion in relation to the plaintiff's work as a whole." For AI-generated outputs, no less than for traditional works, the "substantial similarity" analysis may require courts to make such comparisons between the output and the plaintiff's work.

Two kinds of AI outputs may raise special concerns. First, some AI programs may create outputs that include existing fictional characters. These outputs may run a heightened risk of infringement, since characters sometimes enjoy copyright protection distinct from the specific works in which they appear. In one pending lawsuit, for example, copyright owners Disney, Universal, and their affiliates claim that Midjourney's AI service creates copyright-infringing visual outputs with fictional characters from the *Star Wars, Frozen*, and *Minions* films, among other works.

Second, some AI programs may be prompted to create works "in the style of" a particular artist or author, although some AI programs may now be designed to "decline" such prompts. These outputs are not necessarily infringing, as copyright law generally protects only against the copying of specific works, not

an artist's general aesthetic style. For example, a song generated by AI in the style and simulated voice of a human performer might not infringe any copyright, because it does not copy the music or lyrics of any particular song. (Voice simulations might nonetheless violate some state right-of-publicity laws or federal trademark laws.) As discussed above, there is a separate debate regarding whether market dilution from stylistically similar outputs weighs against a fair-use defense for training AI.

Considerations for Congress

Congress may consider whether to resolve any of the copyright law questions raised by generative AI programs through legislation. Congress may, for example, consider legislation clarifying whether AI-generated works are copyrightable or under what circumstances the process of training generative AI programs may constitute fair use. Some bills and proposals in the 119th Congress would impose civil liability for training AI on some or all copyrighted material, such as the COPIED Media Act (S. 1396), the AI Accountability and Personal Data Protection Act (S. 2367), and the TRUMP AMERICA AI Act (discussion draft). Others would authorize subpoenas for copyright owners to determine if their works were used to train AI programs (the TRAIN Act, S. 2455 and H.R. 7209) or require disclosure to the Copyright Office of copyrighted works used to train AI programs (the CLEAR Act, S. 3813).

Alternatively, Congress may take a wait-and-see approach to give courts an opportunity to decide more cases involving generative AI. The National Policy Framework for Artificial Intelligence published by the White House in March 2026 states that, "[a]lthough the [Trump] Administration believes that training of AI models on copyrighted material does not violate copyright laws, it acknowledges arguments to the contrary exist and therefore supports allowing the Courts to resolve this issue."

Congress may also consider the practical implications of requiring AI companies to seek licenses from copyright owners in the event that training generative AI systems is not a fair use. Commentators debate whether it is practicable for AI companies to identify and pay owners of the large number of works needed to train AI systems. One scholar argues that it may be feasible to create markets for AI training data via content aggregation (e.g., TV streaming services), collective management organizations (or CMOs, such as those that manage rights to musical works), compulsory licenses (which exist for certain uses of sound recordings), or technological measures (e.g., giving rightsholders a means to opt out of AI training). Another scholar questions the practicality of CMOs for AI training data while observing that some "large rights holders" have individually contracted with AI companies for the use of their works.

In the Generative AI Training part of its *Copyright and Artificial Intelligence* report, the Copyright Office contends that voluntary licensing by individual rightsholders or CMOs is sometimes (though perhaps not always) feasible, noting that "[s]ome AI systems have now been trained exclusively on licensed or public domain works." The report expresses normative and practical reservations about nonvoluntary approaches to licensing training data, such as compulsory licenses or requiring copyright holders to "opt out" if they do not consent to the use of their works to train AI. Thus, the report "recommends allowing the licensing market to continue to develop without government intervention" for now. Congress may consider these recommendations as well as differing perspectives such as those surveyed in the report.

Author Information

Christopher T. Zirpoli
Legislative Attorney

Disclaimer

This document was prepared by the Congressional Research Service (CRS). CRS serves as nonpartisan shared staff to congressional committees and Members of Congress. It operates solely at the behest of and under the direction of Congress. Information in a CRS Report should not be relied upon for purposes other than public understanding of information that has been provided by CRS to Members of Congress in connection with CRS's institutional role. CRS Reports, as a work of the United States Government, are not subject to copyright protection in the United States. Any CRS Report may be reproduced and distributed in its entirety without permission from CRS. However, as a CRS Report may include copyrighted images or material from a third party, you may need to obtain the permission of the copyright holder if you wish to copy or otherwise use copyrighted material.