FREE STATE FSF FOUNDATION

Perspectives from FSF Scholars
November 12, 2026
Vol. 21, No. 48

Government Presents Distorted View of Fair Use in OpenAI Copyright Case

by

Seth L. Cooper *

I. Introduction and Summary

On November 20, 2026, reply briefs are due on competing motions for summary judgment in *New York Times Company v. OpenAI*, potentially one of the most consequential copyright infringement cases involving generative artificial intelligence (AI). The case presents the issue of whether making and using unauthorized copies of copyrighted written works to train generative AI models constitutes "fair use" under copyright law. Although the court has yet to rule on the fair use question, the United States' Statement of Interest fails to recognize the full scope of copyright owners' exclusive rights, including their rights to control and license their works for AI training.

The government's apparent position that copyright law allows unauthorized copying of protected works to train AI in all cases reflects a misguided interpretation that, if adopted, would convert the narrow, case-by-case affirmative defense of fair use effectively into a zero-royalty licensing regime. Rather than examining specific acts of unauthorized reproduction under the statutory fair use factors, the government abstracts the underlying conduct simply as "AI training" to shortcut legal analysis. The government also elevates broad concerns about innovation, economic

competitiveness, and national security above the specific rights and protections Congress established in copyright law. The government's claim that licensing would raise AI developers' costs only underscores the value of the copyrighted property they seek to use – and the need to safeguard copyright owners' exclusive rights, rather than grant developers blanket permission to appropriate that property without consent or compensation. Moreover, the government's lopsided position on fair use would deny copyright owners opportunities to develop and participate in emerging AI-training licensing markets, rendering the Trump Administration's stated support for voluntary licensing arrangements effectively empty. The court should reject the government's distortion of fair use and safeguard copyright owners' rights to pursue AI-training licensing opportunities.

The Constitution's Article I, Section 8 Copyright Clause empowers Congress to promote progress by securing authors' property rights, while fair use is a case-by-case defense to infringement. Yet the Statement stakes out a stunningly absolutist position, urging the court to reject "any argument that training LLMs on copyrighted texts violates copyright law." Functionally, this would turn fair use into an industry-wide compulsory license with a royalty rate of zero. The government's broad appeals to American AI leadership, technological innovation, economic competitiveness, and national security do not justify disregarding the specific protections Congress enacted in the Copyright Act. Copyright law itself, following the Constitution's Copyright Clause, promotes progress by securing authors' rights and providing opportunities for voluntary licensing and exchange.

Emblematic of the government's failure to stand up for copyright owners is its inattention to OpenAI's undisputed downloading of pirated books from Libgen to develop GPT-3 and GPT-3.5 training datasets. Nowhere does the government confront this conduct or suggest that it should factor into the court's fair use analysis. Instead, its seemingly categorical position on the lawfulness of "AI training" obscures the specific acts of unauthorized copying at issue. Courts should examine what generative AI developers actually did, rather than subsume everything under the general label of "training." The government should have an interest in ensuring that American AI systems are developed using lawfully acquired rather than pirated content.

The government's simplistic argument that the expense of obtaining permission to use copyrighted works would increase AI model input costs and hamper AI service competition is no justification for making reproduction and use of valuable copyrighted works free for AI developers. All businesses face costs in obtaining inputs for its products and services, but those costs do not entitle them to free use of another's valuable property.

The government's sweeping fair use position threatens the exclusive rights on which existing and emerging AI-training licensing markets depend. Copyright owners, not AI developers, should decide whether and on what terms to license their works. The fourth statutory fair use factor requires courts to consider harm to potential as well as existing markets. By declaring a developing category of AI uses free at the outset, the government's position in the OpenAI litigation would eliminate incentives for AI developers to obtain licenses, extinguishing licensing opportunities before they take hold.

If the development of generative AI necessitates a special exception permitting mass reproduction of copyrighted works, especially for purposes of national security, Congress is the proper institution to consider the matter and strike the right balance. Hopefully, the court will reject the government's blanket support for mass unauthorized copying to train AI models, apply the statutory fair use factors to the specific acts of copying at issue, and safeguard copyright owners' rights and market opportunities.

II. The OpenAI Copyright Infringement Case

In re OpenAI, Inc., Copyright Infringement Litigation[1] is a consolidated multidistrict proceeding involving copyright infringement lawsuits brought by The New York Times Company, other news organizations and publishers, the Authors Guild, and several best-selling authors against OpenAI and Microsoft Corporation. OpenAI operates a for-profit business that develops and sells access to a series of large language models (LLMs) through its ChatGPT service and application programming interface (API), as well as through products offered by Microsoft. Microsoft is a major investor in OpenAI, and the two companies are development and collaboration partners.

The plaintiffs allege, among other things, that the defendants unlawfully reproduced and used their copyrighted works to develop and train generative AI models. It is undisputed in the litigation that, in 2018, an OpenAI employee downloaded pirated copies of copyrighted books from the Library Genesis (Libgen) website and that OpenAI used the downloaded books to develop datasets for training its GPT-3 and GPT-3.5 models. The latter model provided the foundation for ChatGPT's public launch in November 2022.

In September 2026, the parties filed competing motions for summary judgment in the litigation, including on the questions of whether OpenAI's and Microsoft's reproduction and use of plaintiffs' copyrighted works to develop and train AI models constitute copyright infringement and whether such conduct qualifies as fair use. Earlier that month, the Department of Justice filed a Statement of Interest on behalf of the United States in the litigation.[2]

Notably, the developing case law has not yet resolved the fair use question presented in the OpenAI litigation. In separate 2025 decisions, two judges in the Northern District of California held that using copyrighted books to train generative AI models constituted fair use, although their rulings were limited to the circumstances before them. Both decisions, *Bartz v. Anthropic*[3] and *Kadrey v. Meta*,[4] were analyzed in my April 2026 *Perspectives from FSF Scholars* paper, "Assessing the Impact of Trump AI Policy Framework on Intellectual Property."[5] However, the settlement in *Bartz* received final approval in July 2026, while *Kadrey* remains pending at the district court level. Thus, no federal appellate court has yet decided whether unauthorized copying of protected works to train generative AI models constitutes fair use. In September of this year, the Third Circuit rejected an AI-training fair use defense in *Thomson Reuters v. ROSS Intelligence*,[6] but that case involved a non-generative AI system. The decision was analyzed in my October 2026 *Perspectives from FSF Scholars* paper, "In Key Decision, Court Affirms Copyright Protection in New AI Licensing Markets."[7]

Against this unsettled legal backdrop, the OpenAI litigation could have far-reaching implications for copyright owners' exclusive rights and emerging markets for licensing works to train generative AI. With major copyright owners and leading AI developers on opposing sides, the court could help determine whether mass copying of protected works for AI training constitutes fair use. Yet the Justice Department has urged the court in *In re OpenAI* to adopt an expansive interpretation of fair use for generative AI training.

III. The Government's Distortion of Fair Use

The Statement of Interest filed by the Department of Justice advances a misguided interpretation of copyright law that fails to recognize the full scope of copyright owners' exclusive rights, including their rights to control and license the reproduction of their works for AI model training. The Constitution's Article I, Section 8 Copyright Clause empowers Congress "[t]o 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."[8] Certainly, the Constitution makes promoting progress in knowledge and the arts for the public an end goal, but it does so through the means of securing the property rights of authors.

Accordingly, Section 106 of the Copyright Act sets forth copyright owners' exclusive rights to reproduce their works, prepare derivative works, distribute copies, and publicly perform or display their works.[9] Taken together, these rights ensure the marketability, use, and value of copyrighted property. But unauthorized exercise of these rights by third parties constitutes infringement unless permitted by an applicable statutory limitation or defense. Codified in Section 107,[10] fair use is an affirmative defense to infringement, evaluated on a case-by-case basis under four non-exhaustive statutory factors: (1) the purpose and character of the use; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used; and (4) the effect of the use upon the potential market for or value of the copyrighted work.

The Statement of Interest of the United States invokes the Trump Administration's March 2026 National Policy Framework for Artificial Intelligence, which states that the Administration believes training AI models on copyrighted material "does not violate copyright laws," while acknowledging contrary arguments and supporting licensing frameworks and collective rights systems.[11] Yet the Statement stakes out a seemingly categorical position that "the United States has a strong interest in this Court **rejecting any argument** that training LLMs on copyrighted texts violates copyright law."[12] This categorical declaration effectively calls on courts to recognize a blanket fair use exemption for reproducing and using protected expressive content to train generative AI models.

This distortion of fair use, if accepted by courts, would transform a narrow, case-specific affirmative defense into a standing authorization to reproduce vast quantities of copyrighted works for commercial purposes. AI developers could use protected expression without consent, negotiated terms, or compensation. Functionally, this turns fair use into an industry-wide compulsory license with a royalty rate of zero.

Moreover, the government's broad appeals in the Statement of Interest to American AI leadership, technological innovation, economic competitiveness, and national security do not

justify disregarding the specific protections Congress enacted in the Copyright Act. Generalized assertions about the public benefits of AI cannot substitute for examining the particular acts of unauthorized copying at issue, the statutory requirements of fair use, and the resulting harms to copyright owners.

The Third Circuit's September 2026 decision in *Thomson Reuters v. ROSS Intelligence*[13] rejected broad claims by a private AI developer that copyright enforcement would impede technological development and harm national security. Although the United States did not appear in the *ROSS Intelligence* case, the Third Circuit expressly referenced DOJ's Statement of Interest in the OpenAI litigation. As the Third Circuit explained, "While some AI technology may implicate national security concerns, this does not give ROSS carte blanche to violate copyright law merely because it incorporates AI."[14] As discussed in my October 2026 *Perspectives from FSF Scholars* paper, the decision in *ROSS Intelligence* reinforces the principle that generalized claims about AI innovation and national security cannot substitute for examining specific acts of copying, the statutory protections afforded to copyright owners, and the resulting harms.

IV. Government Overlooks and Obscures Specific Acts of Infringement

Emblematic of the government's failure to stand up for copyright owners is its inattention to the undisputed downloading of pirated books from the Library Genesis (Libgen) website, which OpenAI used to develop datasets for training its GPT-3 and GPT-3.5 models. Nowhere in its Statement of Interest does the government confront this conduct or suggest that it should factor into the court's fair use analysis. Instead, the government's absolutist position on the lawfulness of "AI training" obscures the specific acts of unauthorized copying at issue.

Contrary to the government's approach, courts should examine what generative AI developers actually did, rather than subsume everything under the general label of "training." As the Supreme Court emphasized in *Andy Warhol Foundation v. Goldsmith* (2023),[15] "[t]he fair use provision, and the first factor in particular, requires an analysis of the specific 'use' of a copyrighted work that is alleged to be 'an infringement.'"

The government's failure to acknowledge the unlawful Libgen copying reflects its broader pattern of overlooking specific acts of potentially infringing conduct by appealing to the generalized purpose of "AI training." That same pattern is evident in the government's reliance on *Bartz v. Anthropic* for its AI-training fair use holding while ignoring the court's conclusion concerning Anthropic's separate acquisition and retention of millions of pirated books: "The downloaded pirated copies used to build a central library were not justified by a fair use."[16]

Surely, the United States has a substantial interest in combating digital piracy and protecting copyrighted works against large-scale unauthorized copying. Federal law expressly prohibits[17] certain forms of willful copyright infringement, and the Justice Department is responsible for enforcing those protections. The government should also have an interest in ensuring that American AI systems are developed using lawfully acquired rather than pirated content.

V. Licensing Costs Do Not Justify Free Use

The government's argument that requiring AI developers to obtain licenses would raise costs and "hamper competition in the market for LLMs"[18] turns copyright's property principle backwards. The expense of obtaining permission to use copyrighted works does not entitle third parties to use them for free. Indeed, the potentially substantial cost of licensing books, news articles, and other creative works for AI training demonstrates that these works are valuable copyrighted property. Yet the government's position would permit AI developers to exploit that property for commercial profit without compensating copyright owners, denying them opportunities to license their works. This amounts to a free ride at the copyright owners' expense.

The government contends that licensing fees would disproportionately benefit legacy media outlets and that LLMs can level the playing field for smaller publishers and authors. Indeed, the Statement observes that "[a]uthors with limited resources can use LLMs to compete."[19] The government's appeal to the plight of small AI startups is also decidedly unpersuasive in a case involving OpenAI and Microsoft, two major technology companies with substantial financial resources. Moreover, the government's argument overlooks the interests of individual authors, journalists, artists, and small publishers whose copyrighted works are used to train AI models. Their property rights and opportunities to earn licensing revenues deserve protection, too. Copyright law should not favor small AI developers by transferring the value of creative works belonging to others without compensation.

Nor does copyright law deprive AI developers of access to knowledge and useful material for training their models. Under Section 102 of the Copyright Act, facts, ideas, procedures, systems, methods, and discoveries are not protected by copyright.[20] Public-domain works and federal government publications are also generally available for unrestricted use.[21] Developers can license protected works, commission new content, obtain voluntarily contributed material, or invest their own labor in creating training datasets. Copyright protects original expression, not knowledge itself.

The Third Circuit's September 2026 *Thomson Reuters v. ROSS Intelligence* decision reinforces this point. ROSS could have developed its training materials using freely available judicial opinions, but instead copied Westlaw's copyrighted headnotes because doing so was easier. As the court explained, "Unlike necessity, ease is not a justification for copying."[22] The Third Circuit also distinguished the intermediate-copying cases *Google v. Oracle* (2021),[23] *Sega v. Accolade* (1992),[24] and *Sony v. Connectix* (2000),[25] in which copying computer code was necessary to access otherwise inaccessible, unprotected functional elements. Here, by contrast, OpenAI seeks to justify copying protected expressive content from books and articles to train its generative AI models, not to access unprotected functional elements of computer code. In short, copyright law does not guarantee AI developers the cheapest or easiest means of developing their products, much less an entitlement to free use of valuable copyrighted property.

VI. Government's Position on Fair Use Threatens AI-Training Licensing Markets

The government's sweeping fair use position threatens to undermine the exclusive rights on which existing and emerging AI-training licensing markets depend. Copyright law already

provides the foundation for voluntary licensing markets: copyright owners possess the exclusive right to determine whether and on what terms to use and license their works to third parties. Yet the government's position would effectively deprive copyright owners of those choices by permitting widespread commercial copying without permission or compensation.

The government's fair use position also disregards the importance of protecting existing and emerging markets for licensing copyrighted works to train AI models. Under the fourth statutory fair use factor, courts must consider the effect of unauthorized copying of works on potential as well as existing markets. As the Supreme Court explained in *Campbell v. Acuff-Rose Music* (1994): "The market for potential derivative uses includes only those that creators of original works would in general develop or license others to develop."[26] Copyright owners need not already have fully established such a licensing market to receive protection against unauthorized copying that would undermine its development.

As the Supreme Court recognized in *Harper & Row, Publishers, Inc. v. Nation Enterprises* (1985) and *Campbell v. Acuff-Rose Music*, the fourth fair use factor also requires considering the market consequences if conduct of the kind challenged becomes unrestricted and widespread.[27] Here, widespread copying is not hypothetical. The government's proposed interpretation could govern the reproduction and use of millions or billions of protected works across the generative AI industry. A doctrine designed as a limitation on exclusive rights should not become the prevailing economic rule governing access to copyrighted works for AI training.

The Third Circuit recently applied these principles in *Thomson Reuters v. ROSS Intelligence*. The court recognized that the market for licensing copyrighted legal headnotes as text to train AI was "rapidly developing," even though Thomson Reuters had not previously licensed its headnotes to third parties for that purpose.[28] By copying the headnotes without permission, ROSS "usurped Thomson Reuters's opportunity to enter that derivative market" and license its works for AI training.[29] As discussed more fully in my October 2026 *Perspectives from FSF Scholars* paper, "In Key Decision, Court Affirms Copyright Protection in New AI Licensing Markets," the Third Circuit's decision affirms that emerging AI-training markets deserve consideration under established copyright principles.

According to the Statement of Interest: "The United States takes no position on whether a licensing regime would be financially or logistically feasible given the breadth of works at issue, the liability that developers would face, and coordination difficulties."[30] Here again, the government's blanket endorsement of unauthorized copying to train AI models as a fair use renders the Trump Administration's professed support for voluntary licensing arrangements effectively empty. By substantially eliminating the incentive for AI developers to obtain licenses, the government's approach would deny copyright owners opportunities to develop and participate in emerging AI-training licensing markets.

As I explained in my October 2026 *Perspectives* paper, declaring a developing category of AI uses free at the outset "risks effectively extinguishing copyright owners' opportunity to enter a new market."[31] Copyright owners should retain the discretion to develop their own products, license their works for AI training, or decline to participate in those markets. Courts should not

impose a zero-royalty fair use regime that extinguishes those opportunities before voluntary licensing markets have a chance to take hold.

VII. Conclusion: Major AI Copyright Exceptions Are for Congress

If the development of generative AI requires a special exception permitting mass copying of copyrighted works, especially for national security purposes, Congress is the constitutionally designated institution to consider and enact one. Congress can weigh the competing interests and determine whether a statutory exception, collective licensing mechanism, or compensated compulsory license is warranted. Courts should not stretch Section 107's case-specific fair use defense into the practical equivalent of an industry-wide compulsory license with a royalty rate of zero. Having professed support for congressional measures to promote voluntary licensing arrangements, the Trump Administration should look to Congress rather than advocate sweeping, effectively blanket fair use rules through its Statement of Interest that would deny copyright owners their rights and opportunities to license their works for AI training.

Hopefully, the court will reject the government's misguided interpretation of copyright law and its seemingly categorical position that copying works to train AI models constitutes "fair use." Instead, the Copyright Act's exclusive rights and case-specific fair use requirements should apply to the actual acts of copying alleged in the litigation, including the acquisition of pirated books. The court should not permit generalized claims about AI innovation, economic competitiveness, national security, or transformative purpose to displace examination of the defendants' conduct and its consequences for copyright owners. Nor should the expense of licensing protected works become a justification for obtaining those valuable inputs without payment when lawful alternatives remain available. Above all, the court should protect copyright owners' rights and opportunities in AI-training licensing markets.

* Seth L. Cooper is an Adjunct Senior Fellow of the Free State Foundation, a free market-oriented think tank in Potomac, MD. The views expressed in this *Perspectives* do not necessarily reflect the views of others on the staff of the Free State Foundation or those affiliated with it.

Further Readings

Seth L. Cooper, "In Key Decision, Court Affirms Copyright Protection in New AI Licensing Markets," *Perspectives from FSF Scholars*, Vol. 21, No. 43 (October 15, 2026).

Seth L. Cooper, "The Declaration of Independence and the Constitution: Securing Rights and the Example of Copyright," *Perspectives from FSF Scholars*, Vol. 21, No. 37 (September 17, 2026).

Seth L. Cooper, "AI Infringement Case Against Google Tests Limits of Platform Service Terms," *Perspectives from FSF Scholars*, Vol. 21, No. 27 (July 2, 2026).

Randolph J. May and Seth L. Cooper, "World IP Day 2026 and America's 250th Birthday: Follow the Declaration's Principles for Securing Intellectual Property Rights," *Perspectives from FSF Scholars*, Vol. 21, No. 19 (April 24, 2026).

Seth L. Cooper, "Assessing the Impact of Trump AI Policy Framework on Intellectual Property," *Perspectives from FSF Scholars*, Vol. 21, No. 17 (April 15, 2026).

Seth L. Cooper, "No Need for Supreme Court Review of Case Rejecting Copyright for Autonomous AI," *Perspectives from FSF Scholars*, Vol. 21, No. 10 (February 24, 2026).

Seth L. Cooper, "Use of Copyrighted Content to Train AI Models Requires Owners' Consent," *Perspectives from FSF Scholars*, Vol. 20, No. 24 (June 3, 2025).

Seth L. Cooper, "Copyright Office Report Confirms Copyrightability of AI-Generated Works," *Perspectives from FSF Scholars*, Vol. 20, No. 9 (February 19, 2025).

---

[1] *In re OpenAI, Inc., Copyright Infringement Litigation*, MDL No. 25-md-3143 (S.D.N.Y.); *The New York Times Co. v. Microsoft Corp*., No. 23-cv-11195 (S.D.N.Y.).
[2] Statement of Interest of the United States, *In re OpenAI, Inc., Copyright Infringement Litigation*.
[3] *Bartz v. Anthropic PBC*, 787 F. Supp. 3d 1007, 1025–27 (N.D. Cal. 2025).
[4] *Kadrey v. Meta Platforms, Inc*., 788 F. Supp. 3d 1026 (N.D. Cal. 2025).
[5] Seth L. Cooper, "Assessing the Impact of Trump AI Policy Framework on Intellectual Property," *Perspectives from FSF Scholars*, Vol. 21, No. 17 (Apr. 15, 2026).
[6] *Thomson Reuters Enterprise Centre GmbH v. ROSS Intelligence Inc.*, No. 25-2153, slip op. at 1–2, 24–27 (3d Cir. Sept. 29, 2026).
[7] Seth L. Cooper, "In Key Decision, Court Affirms Copyright Protection in New AI Licensing Markets," *Perspectives from FSF Scholars*, Vol. 21, No. 43 (Oct. 15, 2026).
[8] U.S. Const. art. I, § 8, cl. 8 (Copyright Clause).
[9] 17 U.S.C. § 106 (exclusive rights).
[10] 17 U.S.C. § 107 (fair use).
[11] The White House, A National Policy Framework for Artificial Intelligence (Mar. 20, 2026).
[12] Statement of Interest of the United States, *supra*, at 4.
[13] *ROSS Intelligence*, No. 25-2153, slip op. at 26–27 & n.11 (3d Cir. Sept. 29, 2026).
[14] *Id.* at 27.
[15] *Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 529, 533 (2023).
[16] *Bartz v. Anthropic PBC*, 787 F. Supp. 3d 1007, 1033 (N.D. Cal. 2025).
[17] *See* 17 U.S.C. § 506(a); 18 U.S.C. § 2319 (criminal copyright infringement provisions).
[18] Statement of Interest of the United States, *supra*, at 4.
[19] *Id.* at 3–4.
[20] 17 U.S.C. § 102(b).
[21] 17 U.S.C. § 105 (copyright protection generally unavailable for U.S. government works).
[22] *ROSS Intelligence*, No. 25-2153, slip op. at 21.
[23] *Google LLC v. Oracle America, Inc*., 593 U.S. 1 (2021).
[24] *Sega Enterprises Ltd. v. Accolade, Inc*., 977 F.2d 1510 (9th Cir. 1992).
[25] *Sony Computer Entertainment, Inc. v. Connectix Corp*., 203 F.3d 596 (9th Cir. 2000).
[26] *Campbell v. Acuff-Rose Music, Inc*., 510 U.S. 569, 592 (1994).
[27] *Harper & Row, Publishers, Inc. v. Nation Enterprises*, 471 U.S. 539, 568 (1985); *Campbell v. Acuff-Rose Music, Inc*., 510 U.S. 569, 590 (1994).
[28] *ROSS Intelligence*, No. 25-2153, slip op. at 26.
[29] *Id.*
[30] Statement of Interest of the United States, *supra*, at 4 n.13.
[31] Cooper, "In Key Decision, Court Affirms Copyright Protection in New AI Licensing Markets," *supra*, at 5.