

Exhibit A

Counsel for Amici Curiae Copyright Law Professors Robert Brauneis, Sandra Aistars, Mary Catherine Amerine, Samuel F. Ernst, Vicenç Feliú, Jon Garon, Thomas Haley, Justin Hughes, Deirdré A. Keller, Yolanda M. King, Michael Landau, Yvette Joy Liebesman, Adam Mossoff, Loren E. Mulraine, Christopher Newman, Seán O'Connor, Dr. David W. Opderbeck, Dr. Julia Powles, Trevor G. Reed, Marketa Trimble, and Bhamati Viswanathan

TABLE OF CONTENTS

	<u>Page</u>
TABLE OF AUTHORITIES	ii
STATEMENT OF INTEREST OF AMICI CURIAE	1
SUMMARY OF ARGUMENT	1
ARGUMENT	3
I. DEFENDANTS’ USE OF UNAUTHORIZED COPIES OF PLAINTIFFS’ EXPRESSIVE WORKS TO TRAIN COMPUTER SYSTEMS TO OUTPERFORM HUMAN AUTHORS INFRINGES COPYRIGHT IN THOSE WORKS	3
A. Protected Expression is Inextricably Intertwined with Facts and Ideas in Plaintiffs’ Works	3
B. The Core Purpose of U.S. Copyright Law—to Advance Theoretical and Practical Knowledge—Requires Protection of Plaintiffs’ Works	6
II. DEFENDANTS’ COPYING OF PLAINTIFFS’ WORKS TO GAIN KNOWLEDGE AND SKILLS TO OUTPERFORM HUMAN AUTHORS IS NOT A FAIR USE.....	8
A. Defendant’s Copying of the Entirety of Plaintiffs’ Works to Impart Capabilities to its Large Language Models is Minimally Transformative, if at All	8
B. Defendants’ Purpose in Training Large Language Models and in Operating Retrieval Augmented Generation Systems is Commercial to the Core.....	14
C. Defendants’ Use of Plaintiffs’ Works is a Relevant Market-Impacting Use.....	15
D. Defendants are not Engaging in the Type of “Intermediate Copying” that has Been Held to be Fair Use.....	16
E. Defendants’ Use is not an Allowable “Non-Expressive” Use	18
CONCLUSION.....	19
APPENDIX.....	21

TABLE OF AUTHORITIES

	Page(s)
Cases	
<i>A & M Records, Inc. v. Napster, Inc.</i> , 239 F.3d 1004 (9th Cir. 2001)	15
<i>American Geophysical Union v. Texaco Inc.</i> , 60 F.3d 913 (2d Cir. 1994).....	4
<i>Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith</i> , 598 U.S. 508 (2023).....	8, 11, 14
<i>Authors Guild, Inc. v. Google Inc.</i> , 954 F. Supp. 2d 282 (S.D.N.Y. 2013).....	11
<i>Authors Guild, Inc. v. HathiTrust</i> , 755 F.3d 87 (2d Cir. 2014).....	9
<i>Authors Guild v. Google, Inc.</i> , 804 F.3d 202 (2d Cir. 2015).....	9, 11
<i>Baker v. Selden</i> , 101 U.S. 99 (1880).....	5
<i>Bartz v. Anthropic Pbc</i> , 787 F.Supp.3d 1007 (N.D. Cal. 2025)	7, 9, 16
<i>Campbell v. Acuff-Rose Music, Inc.</i> , 510 U.S. 569 (1994).....	8, 14, 15
<i>Feist Publications, Inc. v. Rural Telephone Service Co.</i> , 499 U.S. 340 (1991).....	5
<i>Kadrey v. Meta</i> , 788 F.Supp.3d 1026 (N.D. Cal. 2025)	passim
<i>Sega Enterprises Ltd. v. Accolade, Inc.</i> , 977 F.2d 1510 (9th Cir. 1992)	16, 17
<i>Sony Computer Entertainment, Inc. v. Connectix Corp.</i> , 203 F.3d 596 (9th Cir. 2000)	17
<i>Sony Corp. of America v. Universal City Studios, Inc.</i> , 464 U.S. 417 (1984).....	14

Page(s)

<i>Thomson Reuters Enter. Centre v. Ross Intelligence</i> , 2026 WL 2932557 (3d Cir. Sept. 29, 2026)	17
<i>UMG Recordings, Inc v. MP3.COM, Inc.</i> , 2000 WL 710056 (S.D.N.Y. 2000).....	9
<i>UMG Recordings, Inc. v. MP3. Com, Inc.</i> , 92 F. Supp. 2d 349 (S.D.N.Y. 2000).....	9
<i>Wall Data v. Los Angeles County Sheriff’s Dept.</i> , 447 F.3d 769 (9th Cir. 2006)	15
Statutes	
17 U.S.C. § 101.....	6
17 U.S.C. § 102.....	7
17 U.S.C. § 107.....	8, 14
17 U.S.C. § 504.....	7
Act of May 31, 1790, ch. 15, § 1, 1 Stat. 124, 124	6
Other Authorities	
David Atkinson, <i>Unfair Learning: GenAI Exceptionalism and Copyright Law</i> , 25 Chi.-Kent J. Intell. Prop. 43 (2026)	13
Bikram Pratim Bhuyan et al., <i>Neuro-Symbolic Artificial Intelligence: A Survey</i> , 36 Neural Comput. & Applic. 12809 (2024)	13
Rishi Bommasani et al., <i>On the Opportunities and Risks of Foundation Models</i> 21–52 (July 12, 2022), http://arxiv.org/abs/2108.07258	12
Robert Brauneis, <i>Copyright and the Training of Human Authors and Generative Machines</i> , 48 Colum. J. L. & Arts 1 (2025)	6, 18
Tom B. Brown et al., <i>Language Models Are Few-Shot Learners</i> (July 22, 2020), http://arxiv.org/abs/2005.14165	12
Comment of OpenAI, LP on U.S. Patent and Trademark Office’s Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation, Docket No. PTO-C-2019-0038.....	10
Mark A. Lemley, <i>The Fruit of the Poisonous Tree in IP Law</i> , 103 Iowa L. Rev. 245 (2017).....	7

Page(s)

Adam Nagourney and James Estrin, <i>Dudamel Lands in New York, With Big Ambitions and a Packed Schedule</i> , N.Y. Times (Sept. 10, 2026), available at https://www.nytimes.com/2026/09/10/arts/music/gustavo-dudamel-starts-job-at-new-york-philharmonic.html	3
<i>OpenAI Charter</i> , https://openai.com/charter/	2, 12, 14
<i>OpenAI Partnership</i> , Bloomberg (May 13, 2026), https://www.bloomberg.com/news/articles/2026-05-13/microsoft-spent-over-100-billion-on-openai-partnership-to-date	10
Yam Peleg, archived Tweet, https://archive.is/2RQ8X	6
Pierre Perruchet & Sebastien Pacton, <i>Implicit learning and statistical learning: one phenomenon, two approaches</i> , 5 Trends in Cognitive Sciences 233 (2006)	12
Arthur S. Reber, <i>Implicit learning and tacit knowledge</i> , 118 J. Experimental Psychology: General 218 (1989)	12
Reuters, <i>ChatGPT app hits 1 billion monthly active users in record time, data shows</i> (June 2, 2026), https://www.reuters.com/technology/chatgpt-app-hits-1-billion-monthly-active-users-record-time-data-shows-2026-06-02/	16
Benjamin L.W. Sobel, <i>Artificial Intelligence's Fair Use Crisis</i> , 41 Colum. J.L. & Arts 45 (2017)	5
U.S. Const., Art. I, Sec. 8, Cl. 8	3, 6

STATEMENT OF INTEREST OF AMICI CURIAE

Amici, listed in the Appendix, are professors who teach and write about copyright law and related subjects. Our sole interest in this case is in the development of copyright law in a way that serves the public interest.¹

SUMMARY OF ARGUMENT

Defendants’ unauthorized copying of plaintiffs’ copyrighted works to train large language models and to power retrieval augmented generation systems constitutes direct copyright infringement, and fair use does not excuse it. Defendants cannot evade liability by asserting that they copy only unprotected facts and ideas; in plaintiffs’ works, expression is inextricably intertwined with those facts and ideas. Copyright law permits the free reuse of facts and ideas once a work is lawfully acquired, but it has never permitted the wholesale, unauthorized copying of expressive works to gain access to underlying knowledge. Sanctioning such copying would deny human authors compensation, while forcing them to pay to learn from works that defendants take for free, and would thus dismantle the core incentive structure established “To promote the Progress of Science and useful Arts.”

Defendants’ systematic and wholesale ingestion of plaintiffs’ works is not protected by the fair use doctrine. The technological novelty of generative AI does not render its operational purpose transformative. The dominant original purpose of plaintiffs’ works and the purpose of defendants’ training use is the same: imparting knowledge, skills, and expressive capabilities. Defendant OpenAI has put its purpose plainly: its corporate mission is to create “highly

¹ No party or counsel for any party authored this brief in whole or in part, and no person other than Amici and undersigned counsel made a monetary contribution intended to fund the preparation or submission of this brief. The views expressed herein are those of the listed professors and do not represent the position of any university or other institution.

autonomous systems that outperform humans at most economically valuable work.”²

Defendants’ conduct is also commercial to the core and inflicts profound market harm. Defendants rely on unauthorized copies to build multibillion-dollar commercial enterprises serving over a billion monthly active users. This widespread, unauthorized copying inflicts direct, cognizable market harm by depriving copyright owners of revenue and replacing authorized sales or subscriptions for the same kind of use for which human users typically pay, directly or indirectly.

Defendants’ use of plaintiffs’ works is neither permissible intermediate copying nor non-expressive use. Unlike permissible intermediate copying—which extracts purely functional software elements to facilitate independent human creation and interoperability—defendants harvest human expression itself to train automated systems designed to outperform human authors. And unlike permissible non-expressive use, human expression is at the heart of what defendants use in order to teach their models to sound like human authors.

Defendants essentially ask this Court to create an expansive fair use *exemption* for generative AI developers. Doing so would establish a dangerous form of AI exceptionalism, stripping human creators of their statutory rights and frustrating copyright law’s constitutional purpose.

² *OpenAI Charter*, <https://openai.com/charter/> (last visited October 29, 2026) [archived at <https://perma.cc/9HQU-GNVK>].

ARGUMENT

I.

DEFENDANTS’ USE OF UNAUTHORIZED COPIES OF PLAINTIFFS’ EXPRESSIVE WORKS TO TRAIN COMPUTER SYSTEMS TO OUTPERFORM HUMAN AUTHORS INFRINGES COPYRIGHT IN THOSE WORKS

Defendants’ copying of New York Times articles directly implicates the reproduction right granted in Section 106(1) of the Copyright Act. Defendants do not take only facts and ideas from plaintiffs’ works, they copy articles in their entirety, taking the protected expression that is inextricably intertwined with the facts and ideas. That act undermines the core constitutional purpose of copyright protection, “To promote the Progress of Science and useful Arts.” U.S. Const., Art. I, Sec. 8, Cl. 8.

A. Protected Expression is Inextricably Intertwined with Facts and Ideas in Plaintiffs’ Works.

Defendants admit that they copied plaintiffs’ works to train their AI models. Defendants claim they use only the unprotected facts and ideas contained in New York Times articles. But those facts and ideas are almost always intertwined with copyrightable expression. For example, a New York Times article that presents the “fact” that Gustavo Dudamel has become the Music Director of the New York Philharmonic does so in the context of extensive copyrightable expression, such as the vivid description of Dudamel “[w]earing bright white sneakers and carrying a New York Philharmonic tote bag,” or the explanation of his “meticulously choreographed arrival” that “signals that Dudamel is moving swiftly to put his imprint on the New York Philharmonic and classical music in general.”³ When defendants use such an article,

³ Adam Nagourney and James Estrin, *Dudamel Lands in New York, With Big Ambitions and a Packed Schedule*, N.Y. Times (Sept. 10, 2026), available at <https://www.nytimes.com/2026/09/10/arts/music/gustavo-dudamel-starts-job-at-new-york-philharmonic.html>.

they copy all of it, facts and expression alike.

Defendants’ use of copyrightable expression does not stop there. When “pretraining” large language models, machine learning algorithms extract patterns from the full text without segregating facts or ideas from protectable expression. Those algorithms cannot, and do not, distinguish facts from expression. Similarly, in retrieval augmented generation operations—like OpenAI’s Browse or Microsoft’s Copilot—defendants copy entire articles and instruct the model to synthesize answers from that full context. Either way, defendants’ wholesale copying sweeps in plaintiffs’ protected expression, not merely facts or ideas.

Copyright law permits the free use of facts and ideas found in copyrighted works, but it does not permit the unauthorized copying of entire expressive works for the purpose of gaining access to those facts and ideas. Doing so would turn an essential freedom for users into the destruction of an essential right for authors. Newspaper readers read newspapers primarily for the news—for facts about current events—and for the ideas in the commentary and editorials. Students read textbooks primarily to learn about physics, economics, history, or international affairs. Teachers assign novels to high school students primarily to acquaint them with good writing—in style, theme, and the development of plots and characters. Writers use word-processing computer programs primarily for their function as word processors. Scientists read scientific journal articles primarily for their scientific insights. In none of these cases does copyright law permit the users to make unauthorized copies of newspapers, textbooks, novels, computer programs, and scientific journal articles simply because they are learning facts, ideas, and skills from those works. *Cf. American Geophysical Union v. Texaco Inc.*, 60 F.3d 913 (2d Cir. 1994) (photocopying of scientific journal articles to aid in chemical engineering research was not fair use). Put simply: “[n]o human can rebut an infringement claim merely by showing

that he has learned by consuming the works he copied, even if he puts this new knowledge to productive use later on.” Benjamin L.W. Sobel, *Artificial Intelligence’s Fair Use Crisis*, 41 Colum. J.L. & Arts 45, 73 (2017). If such rebuttal were possible, copyright would vanish.

The Supreme Court made exactly this point in the landmark decision of *Baker v. Selden*, 101 U.S. 99 (1880). The book at issue in *Baker* contained both blank forms for a particular bookkeeping system and narrative text explaining how to use that system. The Supreme Court held that the forms were unprotected because they were necessary to conduct that kind of bookkeeping, and copyright cannot protect a method of bookkeeping. *Id.* at 104–05. However, the Supreme Court affirmed copyright in the explanatory text:

There is no doubt that a work on the subject of bookkeeping . . . may be the subject of a copyright. Such a book may be explanatory either of old systems, or of an entirely new system; and . . . conveying information on the subject of book-keeping, and containing detailed explanations of the art, it may be a very valuable acquisition to the practical knowledge of the community.

Id. at 101–02.

Had defendants here copied only works with little or no creative expression, like the blank forms in *Baker* or white-pages telephone books in *Feist Publications, Inc. v. Rural Telephone Service Co.*, 499 U.S. 340 (1991), they could avoid claims of copyright infringement. But such works are a negligible fraction of what defendants copied. Overwhelmingly, defendants copied works containing substantial copyrightable expression intertwined with whatever facts or ideas may be presented in them. Wholesale copying of such expressive works, even to access the embedded facts and ideas, is a fundamental act of copyright infringement.⁴

⁴ Another argument that has been raised in favor of AI companies—though not asserted here—is that computers and algorithms are really just “reading” the plaintiffs’ works, an act that when it is performed by human beings does not implicate any of the exclusive rights granted to copyright owners. However, defendants’ computers have abilities, in consistency of processing and scale of memory and deployment, that far exceed those of human beings. For example, OpenAI’s GPT-4 was apparently trained using works that contained about 13 trillion tokens of text, which

B. The Core Purpose of U.S. Copyright Law—to Advance Theoretical and Practical Knowledge—Requires Protection of Plaintiffs’ Works.

At issue in this case is the core purpose of U.S. copyright law. Copyright law grants exclusive rights in works in which expression is intertwined with facts and ideas not merely to encourage personal expression, but to incentivize the public dissemination of valuable knowledge. That is a subtle point, but a crucial one to understand, lest copyright be mistakenly considered solely as a vehicle to promote personal literary or artistic visions. The constitutional goal has never been “to Promote Personal Expression,” but rather “[t]o promote the Progress of Science and useful Arts,” U.S. Const., Art. I, sec. 8, cl. 8—in other words, to advance both theoretical and practical knowledge. In the very first copyright act—the Copyright Act of 1790—Congress granted copyright protection to maps and charts. *See* Act of May 31, 1790, ch. 15, § 1, 1 Stat. 124, 124. It did so not to encourage mapmakers and chartmakers to express their personalities, but to incentivize them to accurately survey and graphically represent features of land and water, in order to provide aids to navigation. Maps and charts have remained statutory copyrightable subject matter from 1790 to today, *see* 17 U.S.C. § 101 (defining “pictorial, graphic, and sculptural works” as including “maps, globes, and charts”), valued primarily for their informational utility rather than as purely artistic expression.

works out to about 10 trillion words. *See* Yam Peleg, archived Tweet, <https://archive.is/2RQ8X>. At a reading rate of 250 words per minute, a human being would take about 76,000 years to read 10 trillion words. The basic copyright distinction between fixed copies and mere displays and performances that are “read” or experienced only makes sense in the context of human limitations. If humans remembered perfectly everything they experienced only once in a display or performance, there would be little or no need for reproductions. *See* Robert Brauneis, *Copyright and the Training of Human Authors and Generative Machines*, 48 Colum. J. L. & Arts 1, 45–53 (2025). Since defendants are working far outside of human limitations, enabled by electronic copies that they have made, this Court should reject any argument that they are merely “reading” plaintiffs’ works.

To be sure, copyright does not extend to “any idea, . . . process, . . . or discovery” described in a work. 17 U.S.C. § 102(b). That provision introduces two very important limitations on copyright protection. First, once someone obtains an authorized copy of a work—whether through purchase, or borrowing, or a gift—they can freely reuse its underlying facts and ideas. Second, even if someone makes an unauthorized copy, the actual damages of the owner of copyright in the work are limited to the market value of that work, as represented by either the plaintiff’s lost profits, or any profit the defendant made from selling the unauthorized copy. *See* 17 U.S.C. § 504(b). Actual damages do not capture the downstream value of extracted ideas or facts. *See, e.g.,* Mark A. Lemley, *The Fruit of the Poisonous Tree in IP Law*, 103 Iowa L. Rev. 245 (2017). An accountant who learns the trade from an infringing copy of Selden’s explanatory text does not owe copyright damages on subsequent professional income.

Thus, the *Bartz* court is correct when it states that copyright law does not and should not “make anyone pay specifically for the use of a book . . . each time they later draw upon it when writing new things in new ways.” *Bartz v. Anthropic Pbc*, 787 F.Supp.3d 1007, 1021 (N.D. Cal. 2025). Similarly, the United States is correct when it states that copyright law should not make Joan Didion pay Ernest Hemingway every time she publishes something because she admits that she learned how to write from Hemingway novels. *See* Statement of Interest of the United States, Dkt. 1682 at 16–17. Yet the inference that *Bartz* and the United States then draw—that the only way to avoid these results is to declare defendants’ copying to be fair use or otherwise not actionable—is spectacularly unwarranted. While the remedial limitations that are built into copyright law preserve vital freedoms for users, they do not grant license to make unauthorized copies of expressive works simply because the copier is only interested in the underlying ideas.

That would destroy copyright law altogether. Because virtually all of plaintiffs' works contain substantial expression, defendants' unauthorized copying is infringing.

Defendants' unauthorized copying of plaintiffs' works puts the human authors of those works at a double disadvantage. First, defendants deny those human authors any income from the defendants' profitable uses of those works, thus diminishing their incentive to write. Second, defendants instill knowledge and skills in their generative models by training them with copyrighted works without consent or compensation, while the human authors of plaintiffs' works, as well as all human authors, must still pay something for most of the works from which they learn and gain skills, as copyright law justifiably provides. Copyright law does not and should not condone such unfair and incentive-eroding treatment.

II.

DEFENDANTS' COPYING OF PLAINTIFFS' WORKS TO GAIN KNOWLEDGE AND SKILLS TO OUTPERFORM HUMAN AUTHORS IS NOT A FAIR USE

Defendants' technology is new, but what they are doing is not. The copying of protected expression for commercial profit is not fair use, regardless of whether that copying is done by humans or by computers. None of defendants' arguments fits within established fair use doctrine, and none justifies the creation of a new exception for AI models.

A. Defendant's Copying of the Entirety of Plaintiffs' Works to Impart Capabilities to its Large Language Models is Minimally Transformative, if at All.

The first fair use factor in § 107 of the Copyright Act considers "the purpose and character of the use, including whether such use is of a commercial nature." 17 U.S.C. § 107(1). A dominant consideration under this factor is the degree to which the copier's use is "transformative." *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 579 (1994). A use is transformative to the extent that it has "a further purpose or different character" than the allegedly infringed work, *Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith*, 598

U.S. 508, 525 (2023), or in other words, to the extent that it “serves a new and different function.” *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 96 (2d Cir. 2014); *Authors Guild v. Google, Inc.*, 804 F.3d 202, 217 (2d Cir. 2015). Thus, “transformativeness” in the fair use sense requires a comparison of purposes—the original purpose of plaintiffs’ works, and the purpose of defendants, the alleged infringers.

In this context, “transformativeness” does not refer to how innovative the copier’s technology is. The question is not whether the copier’s *technology* is transformative, it is whether the copier’s *use* is transformative.⁵ Generative AI is undoubtedly “transformative” in the sense that it is highly innovative and will have broad impacts—positive and negative—on our economy and society. But the fact that the technology is dazzling does not make the unauthorized use of works to train generative AI models “transformative” in the fair use sense. A quarter-century ago, infringement defendants similarly argued that internet technology was innovative and thus transformative. As the Southern District of New York responded, “the fact that [copyright infringement] was clothed in the exotic webbing of the Internet does not disguise its illegality.” *UMG Recordings, Inc v. MP3.COM, Inc.*, 2000 WL 710056, at *1 (S.D.N.Y. 2000); *see also UMG Recordings, Inc. v. MP3. Com, Inc.*, 92 F. Supp. 2d 349, 351 (S.D.N.Y. 2000) (rejecting defendant’s claim of transformativeness in more detail). The proper question is whether the relevant original purpose of the copied works substantially differs from defendants’ purpose.

Defendants do not meet that test. A core original purpose of all of the plaintiffs’ works is

⁵ The Northern District of California misapplied this factor in holding that another generative AI developer, Anthropic, made fair use of works in part because “[t]he technology at issue was among the most transformative many of us will see in our lifetimes.” *Bartz*, 787 F.Supp.3d at 1033.

to impart knowledge, expression, and skills to human beings that are both useful and enriching in our economy and society. Defendants copy those works to pursue OpenAI’s corporate mission, supported by over \$100 billion from Microsoft, to create “highly autonomous systems that outperform humans at most economically valuable work.”⁶ Because the “economically valuable work” at issue is performed by providing written text, the relevant mission and purpose here is to outperform and outcompete human authors and human-authored works.

In other words, OpenAI, supported by Microsoft, extracts knowledge, skills, experience, and expression from plaintiffs’ works to impart those capabilities to computer systems to outcompete human authors. That plaintiffs’ works were not created specifically to train computer systems does not make that use transformative. Neither copyright law nor fair use doctrine has the goal of privileging computers over human beings as recipients of capabilities. OpenAI’s use of plaintiffs’ works is minimally transformative, if at all.

Defendants try to avoid this conclusion by mischaracterizing the purpose for which plaintiffs’ works were created. OpenAI previously claimed that works like plaintiffs’ were created for their “standalone entertainment value.” Comment of OpenAI, LP on U.S. Patent and Trademark Office’s Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation, Docket No. PTO–C–2019–0038, at 5. Defendants now veil their claim in slightly more moderate language, arguing that “News Plaintiffs wrote their works to engage human readers.” OpenAI News Brief at 7. The false implication remains: because computers cannot be “entertained” or “engaged,” defendants’ use must be different.

⁶ OpenAI Charter, *supra* note 2; see Rachel Metz, Brody Ford, and Madlin Mekelburg, *Microsoft Spent Over \$100 Billion on OpenAI Partnership*, Bloomberg (May 13, 2026), <https://www.bloomberg.com/news/articles/2026-05-13/microsoft-spent-over-100-billion-on-openai-partnership-to-date>.

We human beings are always “on” as learners—we are always updating our understanding of the world through our interaction with works of authorship, as well as through other experiences. *See, e.g.*, National Academies of Sciences et al., *How People Learn II: Learners, Contexts, and Cultures* 3–4 (2018). Because humans always learn from works, the imputed objective purpose of those works centrally includes developing knowledge and skills. That objective purpose, stemming from recognized, normal use, governs copyright and fair use analysis. *See Warhol*, 598 U.S. at 545 (fair use requires “an objective inquiry into what use was made,” and “cannot turn merely on the stated or perceived intent of the artist.”) (quotation omitted).

While mischaracterizing plaintiffs’ purpose, defendants present a radically incomplete and misleading view of their own. Defendants state that their purpose is “to distill broad patterns of human language and facts about the world from an enormous text corpus.” OpenAI News Brief at 7. Unfortunately, in *Kadrey v. Meta*, the court accepted a similar assertion, and described generative model training as reducible to “ingest[ing] text to learn ‘statistical patterns’ of how words are used together in different contexts.” *Kadrey v. Meta*, 788 F.Supp.3d 1026, 1044 (N.D. Cal. 2025). The references to statistical patterns may be intended to bring to mind the “ngrams” tool mentioned over a decade ago as supporting Google’s fair use claim for its Google Library Project. That was a tool that used full-text scans of books to “furnis[h] statistical information . . . about the frequency of word and phrase usage over centuries,” such as “‘track[ing] the frequency of references to the United States as a single entity (‘the United States is’) versus references to the United States in the plural (‘the United States are’) and how that usage has changed over time.’” *Google*, 804 F.3d at 209 (quoting *Authors Guild, Inc. v. Google Inc.*, 954 F. Supp. 2d 282, 287 (S.D.N.Y. 2013)).

Defendants’ project is stunningly more ambitious, and their implicit comparison to Google’s ngrams tool is entirely inappropriate. Quantitative reports of patterns of word use across multiple copied works provide additional historical insights without substituting in any way for those works, or for the human skills that created them. By contrast, Large Language Models sit at the core of defendants’ strategy to “outperform humans at most economically valuable work.” OpenAI Charter, *supra* note 2. That is because so much of the economically valuable work that defendants seek to surpass humans at is taught, explained, and performed with language. Human beings create works of authorship in which they solve math problems, translate from one language to another, explain scientific concepts, edit prose, compose poetry, debate philosophy, argue about legal issues, and do many, many other things. OpenAI’s release of its GPT-3 model led its researchers and others to understand how exposure to a large number of works enables large language models to perform many tasks without specialized fine tuning.⁷

The fact that OpenAI models are gaining these abilities through analyzing patterns of words in works of authorship is no reason to hold that this use of plaintiffs’ works is transformative fair use. First, much of human learning also consists of automatic recognition of patterns across many works. There is an entire field of psychology devoted to “implicit learning”—to how we human beings learn languages and other subjects by automatically and unconsciously updating our understandings of words and phrases and syntax as we are exposed to texts. *See, e.g.,* Arthur S. Reber, *Implicit learning and tacit knowledge*, 118 J. Experimental Psychology: General 218 (1989); Pierre Perruchet & Sebastien Pacton, *Implicit learning and*

⁷ *See, e.g.,* Tom B. Brown et al., *Language Models Are Few-Shot Learners* (July 22, 2020), <http://arxiv.org/abs/2005.14165> (reporting GPT-3’s superior performance on a variety of standard tasks); Rishi Bommasani et al., *On the Opportunities and Risks of Foundation Models* 21–52 (July 12, 2022), <http://arxiv.org/abs/2108.07258> (explaining the capabilities of large language models and related image and video models).

statistical learning: one phenomenon, two approaches, 5 Trends in Cognitive Sciences 233

(2006). When the court in *Kadrey* asserted that pattern detection “is not how a human reads a book,” *Kadrey*, 788 F.Supp.3d at 1044, it unfortunately ignored a major component of human learning.

Second, differences between human and computer modes of learning should not matter for application of copyright law, just as differences between diverging computer technologies should not matter. Suppose that two companies train computers to engage in dialogue about accounting, and to perform accounting tasks, by ingesting copyrighted accounting textbooks without authorization. One builds a model that predicts words in sequence (like OpenAI). Another builds a model that derives distinct concepts and rules from the texts, following a well-known branch of artificial intelligence sometimes known as “symbolic AI.” See, e.g., Bikram Pratim Bhuyan et al., *Neuro-Symbolic Artificial Intelligence: A Survey*, 36 Neural Comput. & Applic. 12809 (2024). In spite of their disparate internal characters, both models are able to write texts and engage in dialogue about accounting, and both are able to do so only because they ingested unauthorized copies of copyrighted accounting textbooks. Holding that constructing one model is fair use while constructing the other one is not would be arbitrary, because both acquire valuable capabilities from copyrighted works of authorship. It would be equally arbitrary to hold that computer acquisition of capabilities is fair use while human acquisition is not. See David Atkinson, *Unfair Learning: GenAI Exceptionalism and Copyright Law*, 25 Chi.-Kent J. Intell. Prop. 43, 61-65 (2026) (demonstrating that four-factor fair use arguments for machine learning apply equally to human learners).

B. Defendants’ Purpose in Training Large Language Models and in Operating Retrieval Augmented Generation Systems is Commercial to the Core.

In addition to transformativeness, the first fair use factor considers whether the copier’s use is of a commercial nature, with commercial character weighing against fair use. *See Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417, 448–50 (1984); *Campbell*, 510 U.S. at 583–84; *Warhol*, 598 U.S. at 531. Defendants OpenAI and Microsoft are both commercial entities operating services such as ChatGPT and Browse for profit. As OpenAI proudly proclaimed in March of this year:

Within a year of launching ChatGPT, we reached \$1B in revenue. By the end of 2024 we were generating \$1B per quarter. We are now generating \$2B in revenue per month. At this stage, we are growing revenue four times faster than the companies who defined the Internet and mobile eras, including Alphabet and Meta.⁸

It is this kind of unprecedented revenue growth that led the court in *Kadrey v. Meta* to reject claims that generative AI developers could not afford to license the works on which they trained their models:

[T]he suggestion that adverse copyright rulings would stop this technology in its tracks is ridiculous. These products are expected to generate billions, even trillions, of dollars for the companies that are developing them. If using copyrighted works to train the models is as necessary as the companies say, they will figure out a way to compensate copyright holders for it.

Kadrey, 788 F.Supp.3d at 1036. While commerciality matters less when a use is “transformative,” *see Campbell*, 510 U.S. at 579, defendants’ use is both not transformative, and commercial to the core. Their AI systems are designed to outperform humans and human authors, not necessarily in all respects, but “at most *economically valuable* work,” OpenAI Charter, *supra* note 2 (emphasis added)—at the work that will make defendants a profit.

⁸ OpenAI, *OpenAI raises \$122 billion to accelerate the next phase of AI* (March 31, 2026), <https://openai.com/index/accelerating-the-next-phase-ai/>.

C. Defendants’ Use of Plaintiffs’ Works is a Relevant Market-Impacting Use.

The fourth fair use factor in § 107 of the Copyright Act considers “the effect of the use upon the potential market for or value of the copyrighted work.” 17 U.S.C. § 107(4). Because defendants’ use of plaintiffs’ works is comparable to common uses by human readers, defendants’ unauthorized copying has at least (1) the market impact of the plaintiffs’ loss of one sale of, or one subscription to, their works; or (2) the defendants’ gain from having avoided having to pay the market price for the works. Suppose that a human being testified that they had made an unauthorized copy of an accounting textbook in order to learn everything that they could from it—to learn how to talk about accounting, how to do accounting, and more broadly, how the authors of the textbook write in English—so that they could provide informal consulting services about accounting. The lost sale would unquestionably be a cognizable market impact in any inquiry about whether that unauthorized use was a fair use. *Cf. A & M Records, Inc. v. Napster, Inc.*, 239 F.3d 1004, 1015 (9th Cir. 2001) (no fair use when “repeated and exploitative unauthorized copies of copyrighted works were made to save the expense of purchasing authorized copies”); *Wall Data v. Los Angeles County Sheriff’s Dept.*, 447 F.3d 769, 779 (9th Cir. 2006) (same). Alternatively, suppose that a human being testified that they had made an unauthorized copy of an accounting textbook so that they could keep it on hand and consult it as a reference work when others asked them questions, even though they would of course answer the questions in their own words, as OpenAI’s Browse or Microsoft’s Copilot would. Again, the resulting lost sale would be a cognizable market impact in a fair use inquiry.

The market impact here is doubly magnified. First, fair use inquiry considers “whether unrestricted and widespread conduct of the sort engaged in by the defendant . . . would result in a substantially adverse impact on the potential market’ for the original.” *Campbell*, 510 U.S. at 590 (quoting Nimmer on Copyright § 13.05[A][4], p. 13-102.61 (footnote omitted)). From the

plethora of copyright actions brought against many AI developers, we know that it has already become an industry practice to make unauthorized copies of human works to train generative AI models and to use them as reference or “grounding” context for retrieval augmented generation operations. Second, the scale of defendants’ operations alone is vast, and displaces an incalculable number of human beings who might otherwise be directly reading or consulting a newspaper or accounting textbook. In June of this year, ChatGPT reached one billion active monthly users.⁹ When one unauthorized copy of a work is used to build a model that is serving a billion people, there is no doubt that the making of that copy could have a substantial impact on the market value of that work.

In this respect, cognizable fourth-factor market impact in this case should not be limited either to AI output that is substantially similar to a particular training work, or to the broader “market dilution” of a particular training work by that AI output.¹⁰ Defendants are using plaintiffs’ works to gain knowledge and abilities that are comparable to those gained by human beings from those works. Because they are doing so with unauthorized copies of those works, the market impact of the plaintiffs’ lost sales is properly considered in a fair use analysis.

D. Defendants are not Engaging in the Type of “Intermediate Copying” that has Been Held to be Fair Use.

Defendants and their supporters also argue that their copying is fair because it is only “intermediate copying” that does not appear in the outputs of their AI systems. However, *Sega Enterprises Ltd. v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992), the primary case cited for this argument, explicitly and thoroughly repudiates it. In *Sega*, the defendant argued that

⁹ See Reuters, *ChatGPT app hits 1 billion monthly active users in record time, data shows* (June 2, 2026), <https://www.reuters.com/technology/chatgpt-app-hits-1-billion-monthly-active-users-record-time-data-shows-2026-06-02/>.

¹⁰ Compare *Bartz*, 787 F.Supp.3d at 1033, with *Kadrey*, 788 F.Supp.3d at 1044.

“intermediate copying does not infringe the exclusive rights granted to copyright owners in section 106 of the Copyright Act unless the end product of the copying is substantially similar to the copyrighted work.” *Id.* at 1517. The Ninth Circuit flatly rejected that argument: “[T]he Copyright Act does not distinguish between unauthorized copies of a copyrighted work on the basis of what stage of the alleged infringer's work the unauthorized copies represent. The intermediate copying done by [the defendant] therefore falls squarely within the category of acts that are prohibited by the statute.” *Id.* at 1518.

Both *Sega* and the subsequent Ninth Circuit decision in *Sony Computer Entertainment, Inc. v. Connectix Corp.*, 203 F.3d 596 (9th Cir. 2000) (“*Connectix*”), stand for an important but much narrower principle: intermediate copying is permissible when done solely to identify functional elements in software, enabling human creators to use their own creativity to develop new software that is interoperable with the analyzed software. As the Third Circuit recently held in *Thomson Reuters Enter. Centre v. Ross Intelligence*, 2026 WL 2932557, at *7 (3d Cir. Sept. 29, 2026), “intermediate copying” outside of that context, such as copying to train an AI legal research tool, does not qualify as a fair use.

Here, defendants do not train on plaintiffs’ works to extract functional software elements so that defendants or their employees can then provide all of the creative expression that defendants’ models will use to produce their outputs. Defendants provide none of the expressiveness their models rely on; they harvest that expression from works created by others. In the case of the works at issue here, they harvest that expression without authorization from works still under copyright. That falls squarely outside of the fair use principle developed in *Sega* and *Connectix*, and within the kind of “intermediate copying” that *Sega* firmly establishes as infringement and that *Ross Intelligence* confirms is undeserving of fair use treatment.

E. Defendants’ Use is not an Allowable “Non-Expressive” Use.

Defendants have also argued that their use is not infringing because it is a “non-expressive” use. If by that they mean that defendants have copied and used only facts from plaintiffs’ works, and not the expressive choices of the authors of those works, that is simply not true. To serve their mission of outperforming human authors, defendants do not just want a model that “knows English”; rather, they want a model that contains patterns taken from the way human authors actually express themselves. Indeed, defendants want a model that contains patterns taken from the way human authors express themselves *currently*, which is why they are not satisfied with using works already in the public domain as training works. *See* Robert Brauneis, *Copyright and the Training of Human Authors and Generative Machines*, 48 Colum. J. L. & Arts 1, 22–26, 30–35 (2025) (distinguishing between different meanings of the term “expression,” and noting that large language models and human authors both use the creative choices in the works from which they learn). As the court in *Kadrey v. Meta* recognized, if generative AI models “are to generate high-quality text, [the developers] need coherent, reasonably high-quality training data. In other words, they need high-quality expression.” *Kadrey*, 788 F.Supp.3d at 1049.

If defendants mean that the outputs of their models rarely if ever infringe the works on which the models were trained, that is the same as the argument that “intermediate copying” is never infringing, an argument that the Ninth Circuit definitively rejected—twice—as we have discussed above. Finally, if defendants mean that their use is not infringing because their computers have no aesthetic or emotional reactions to the training works, that is irrelevant. Just as the principal purpose of U.S. copyright law is not to provide an incentive for personal artistic expression, neither is it to provide aesthetic or emotional satisfaction to beings capable of experiencing such satisfaction. Rather, it is to provide an incentive to advance theoretical and

practical knowledge of all kinds. Defendants’ use is imparting valuable capabilities and knowledge to their generative models, and the expressive choices of the authors of plaintiffs’ works are key to that value. Thus, defendants’ use is not “non-expressive” in any of the three possible meanings of that term.

CONCLUSION

Defendants’ unauthorized, wholesale copying of plaintiffs’ expressive works to train commercial models that directly compete with human authors is basic infringement, and it exceeds any recognized boundary of fair use. What defendants seek in this case is the creation of a new AI model exemption from long established copyright doctrine. Doing so would establish a dangerous form of AI exceptionalism, stripping human creators of their statutory incentives and undermining the constitutional mandate to promote the progress of science and useful arts. Amici copyright law professors respectfully submit that this Court should preserve the integrity of copyright law by granting plaintiffs’ motion for summary judgment and denying defendants’ cross-motion.

Dated: New York, New York
October 30, 2026

Respectfully Submitted,

FRIEDMAN KAPLAN SEILER
ADELMAN & ROBBINS LLP

/s/ Katherine L. Pringle

Katherine L. Pringle
Kyle C. Bailey (admission pending)
7 Times Square
New York, New York 10036-6516
(212) 833-1100
kpringle@fklaw.com
kbailey@fklaw.com

Counsel for Amici Curiae Copyright Law

Professors Robert Brauneis, Sandra Aistars, Mary Catherine Amerine, Samuel F. Ernst, Vicenç Feliú, Jon Garon, Thomas Haley, Justin Hughes, Deirdré A. Keller, Yolanda M. King, Michael Landau, Yvette Joy Liebesman, Adam Mossoff, Loren E. Mulraine, Christopher Newman, Seán O'Connor, Dr. David W. Opderbeck, Dr. Julia Powles, Trevor G. Reed, Marketa Trimble, and Bhamati Viswanathan

APPENDIX¹¹

Robert Brauneis
Michael J. McKeon Professor of Intellectual Property Law
The George Washington University Law School

Sandra Aistars
Distinguished Counselor in Residence & Professor of Law
IPPI: The IP Policy Institute, The University of Akron School of Law

Mary Catherine Amerine
Associate Professor of Law
Loyola Marymount University - Loyola Law School

Samuel F. Ernst
Professor of Law
Western State College of Law

Vicenç Feliú
Associate Dean for Library Services & Professor of Law
Nova Southeastern University - Shepard Broad College of Law

Jon Garon
Associate Dean for Technology and Innovation and Professor of Law
Nova Southeastern University - Shepard Broad College of Law

Thomas Haley
Associate Professor
University of Florida Levin College of Law

Justin Hughes
Hon. William Matthew Byrne Jr. Professor of Law - Loyola Law School, Loyola Marymount
University
Visiting Professor of Law - Oxford University

Deirdré A. Keller
Visiting Professor of Law
Ohio Northern University Claude W. Pettit College of Law

Yolanda M. King
Associate Professor of Law and Associate Dean of Student Affairs and Strategy
University of Illinois Chicago School of Law

¹¹ Institutional affiliations are listed for identification purposes only.

Michael Landau
Professor Emeritus
Georgia State University College of Law

Yvette Joy Liebesman
Professor of Law
Saint Louis University School of Law

Adam Mossoff
Professor of Law
George Mason University Antonin Scalia Law School

Loren E. Mulraine
Professor of Law and Director of Music and Entertainment Law Studies
Belmont University - College of Law

Christopher Newman
Associate Professor of Law
George Mason University Antonin Scalia Law School

Seán O'Connor
Allen Grubman Chair in Media & Entertainment Law
Brooklyn Law School

Dr. David W. Opderbeck
Professor of Law
Seton Hall University School of Law

Dr. Julia Powles
Professor and Executive Director, UCLA Institute for Technology, Law & Policy
University of California Los Angeles School of Law

Trevor G. Reed
Professor of Law
University of California Irvine School of Law

Marketa Trimble
Samuel S. Lionel Professor of Intellectual Property Law
William S. Boyd School of Law, University of Nevada, Las Vegas

Bhamati Viswanathan
Visiting Assistant Professor
Suffolk University Law School

CERTIFICATION OF WORD COUNT

This brief complies with the word-count limitation ordered by the Court, Dkt. 1663, because it contains 5,861 words (based on the Microsoft Word word-count function), excluding the parts of the brief not counted against such word limit under Local Civil Rule 7.1.

Dated: New York, New York
October 30, 2026

Respectfully Submitted,

FRIEDMAN KAPLAN SEILER
ADELMAN & ROBBINS LLP

/s/ Katherine L. Pringle

Katherine L. Pringle (kpringle@fklaw.com)
7 Times Square
New York, New York 10036-6516
(212) 833-1100

*Counsel for Amici Curiae Copyright Law
Professors Robert Brauneis, Sandra Aistars, Mary
Catherine Amerine, Samuel F. Ernst, Vicenç Feliú,
Jon Garon, Thomas Haley, Justin Hughes, Deirdré
A. Keller, Yolanda M. King, Michael Landau, Yvette
Joy Liebesman, Adam Mossoff, Loren E. Mulraine,
Christopher Newman, Seán O'Connor, Dr. David
W. Opderbeck, Dr. Julia Powles, Trevor G. Reed,
Marketa Trimble, and Bhamati Viswanathan*