JIPEL

NYU Journal of Intellectual Property
& Entertainment Law

VOLUME 14 NUMBER 2

Statement of Purpose

Consistent with its unique development, the New York University Journal of Intellectual Property & Entertainment Law (JIPEL) is a nonpartisan periodical specializing in the analysis of timely and cutting-edge topics in the world of intellectual property and entertainment law. As NYU's first online-only journal, JIPEL also provides an opportunity for discourse through comments from all of its readers. There are no subscriptions or subscription fees; in keeping with the open-access and free discourse goals of the students responsible for JIPEL's existence, the content is available for free to anyone interested in intellectual property and entertainment law.

The New York University Journal of Intellectual Property & Entertainment Law is published up to three times per year at the New York University School of Law, 139 MacDougal Street, New York, New York, 10012. In keeping with the Journal's open access and free discourse goals, subscriptions are free of charge and can be accessed via www.jipel.law.nyu.edu. Inquiries may be made via telephone (212-998-6101) or e-mail (submissions.jipel@gmail.com).

The Journal invites authors to submit pieces for publication consideration. Footnotes and citations should follow the rules set forth in the latest edition of The Bluebook: A Uniform System of Citation. All pieces submitted become the property of the Journal. We review submissions through Scholastica (scholasticahq.com) and through e-mail (submissions.jipel@gmail.com).

The Journal is also available on WESTLAW, LEXIS-NEXIS and HeinOnline.

NEW YORK UNIVERSITY

JOURNAL OF INTELLECTUAL PROPERTY AND ENTERTAINMENT LAW

VOL. 14 BOARD OF EDITORS – ACADEMIC YEAR 2024-2025

NEW YORK UNIVERSITY

JOURNAL OF INTELLECTUAL PROPERTY AND ENTERTAINMENT LAW

VOLUME 14 SPRING 2025 NUMBER 2

THE AI INPUT CLASS: CONSTITUTIONAL URGENCY AND FAIR LICENSING IN AI COPYRIGHT CLASS ACTIONS

XUAN-THAO NGUYEN* & ELIZABETH PORTER**

The humanities have long been under attack; now Big Tech is eating them for breakfast. Artificial Intelligence is undermining the cultural and constitutional values of human creativity; it is also threatening the livelihoods of the creative working class. Because neither Congress nor regulators can keep pace with AI's pace of change, class actions have stepped into the breach. In the first wave of such litigation, authors and artists in courts on both coasts claim that AI companies violate intellectual property law when they "train" their systems on copyrighted works.

This Article analyzes what we call "input" AI claims—that is, proposed class action copyright claims against AI companies for inputting creative works into their models. This action is variously described as copying, scraping, feeding, or training. We argue that input copyright class action claims have both merit and power. Substantively, the creators' claims of direct copyright infringement are supported by almost three decades of caselaw applying the concept of "copying" to code and other innovative technologies. Procedurally, input claims based on AI "training" fall into the heartland of the letter and spirit of Rule 23, which governs federal class actions. Defendants' claims of fair use, while seductive, would swallow copyright law and crush human-centered creativity.

*Pendleton Miller Chair in Law, University of Washington School of Law. Special thanks to my coauthor for a wonderful collaboration on several projects overseas and at UW Law School. Thank you to my friend Robert Gomulkiewicz for his comments on the manuscript over lunch at the usual spot. Anna Johansen, UW Law Class of 2024, provided valuable research assistance and Cindy Fester added her professional touch to this manuscript.

**James W. Mifflin Professor of Law, University of Washington School of Law. My sincere thanks to my wonderful research assistant Robert Colton.

Class action settlements—which combine private ordering with judicial oversight—may offer the best opportunity to establish a licensing regime that protects creators from future infringement. In lieu of fair use, we urge a different doctrine: "fair licensing." Licensing poses its own risks to creativity; it risks creating moats around knowledge and paying creators symbolic amounts for giving up control of their works. But AI class actions can and should play an influential role in spurring the creation of fair AI licensing regimes. Courts and litigants are wise to act expeditiously in doing so: there is a constitutional urgency to protecting the essence of human creativity.

Introduction

Generative artificial intelligence models are fast becoming a normal feature of the digital landscape.[1] So quickly, in fact, that it has proven impossible for lawmakers and regulators to fully conceive of, must less respond to, the host of issues raised by this new but artificially natural-feeling technology.[2] Indeed,

[1] OpenAI, the leading AI company, has released successive versions of its GPT foundation model for users to generate text, images, and voices. *See Hello GPT-4o*, OpenAI (May 13, 2024), https://openai.com/index/hello-gpt-4o/ [https://perma.cc/B8NH-EBTM] ("We're announcing GPT-4o, our new flagship model that can reason across audio, vision, and text in real time."); Jennifer Elias, *Google Rolls Out Its Most Powerful AI Models as Competition from OpenAI Heats Up*, CNBC (May 14, 2024), https://www.cnbc.com/2024/05/14/google-announces-lightweight-ai-model-gemini-flash-1point5-at-google-i/o.html [https://perma.cc/GK8U-EYX6]; *Introducing the Next Generation of Claude*, Anthropic (Mar. 4, 2024), https://www.anthropic.com/news/claude-3-family [https://perma.cc/A2WD-BUTG] ("Today, we're announcing the Claude 3 model family, which sets new industry benchmarks across a wide range of cognitive tasks. The family includes three state-of-the-art models in ascending order of capability: Claude 3 Haiku, Claude 3 Sonnet, and Claude 3 Opus.").

[2] *Senator Schumer to Propose Framework for AI Legislation in Coming Weeks*, Reuters (May 9, 2024), https://www.reuters.com/technology/legislative-roadmap-ai-is-coming-weeks-schumer-says-2024-05-08/ [https://perma.cc/J2RP-8L4B] ("Senate Majority Leader Chuck Schumer said on Wednesday he would provide a framework for addressing the risks and potential benefits of artificial intelligence in the coming weeks, which will then be translated into piecemeal legislation."); Brian Fung, *Chuck Schumer and Bipartisan Group of Senators Unveil Plan to Control AI–While Investing Billions of Dollars In It*, CNN (May 15, 2024), https://www.cnn.com/2024/05/15/tech/schumer-ai-framework/index.html [https://perma.cc/FY2X-U27W] (reporting that the policy action plan "instructs multiple Senate committees to come up with guardrails for AI to address some of its biggest risks, such as AI-enabled discrimination, job displacement and election interference").

AI's creators are struggling with the impact of their own work.[3] The rapidity with which generative AI is spreading, combined with lawmakers' deep-rooted commitment—embodied in the Communications Decency Act—to minimizing regulation of the Internet, means that litigation, rather than regulation, will play the lead role in establishing legal limits on 21st-century AI systems in the U.S.[4]

Moreover, the sheer magnitude of people and entities affected by AI foundation models, and the depersonalized, algorithmic mechanisms that propel those models, mean that many legal claims against AI companies will only provide effective relief if they are brought as class actions.[5] Without the benefits of aggregation, there is no viable way for the creative working class—from journalists, novelists, playwrights, poets, nonfiction writers, screenwriters, composers, lyricists, musicians, videogame developers, and cartoonists, to illustrators, painters, and artists—who allege unlawful activity by AI corporations to pursue compensation and/or injunctive relief against further harm. Even with class action lawsuits, a favorable outcome is uncertain.

We anticipate waves of class litigation by parties impacted by generative AI as they seek to shelter their businesses, livelihoods, and identities from the rising

---

[3] *See, e.g.*, Ed Newton-Rex, @ednewtonrex, X (Mar. 4, 2024), https://x.com/ednewtonrex/status/1764667915075543256 [https://perma.cc/K9DH-PQBR]; Ed Newton-Rex, @ednewtonrex, X, (Nov. 15, 2023), https://x.com/ednewtonrex/status/1724902327151452486?s=20&mx=2 [https://perma.cc/5MFT-3GVP] ("I've resigned my role leading the Audio team at Stability AI, because I don't agree with the company's opinion that training generative AI models on copyrighted works is 'fair use'."); Gareth Vipers et al., *Elon Musk Sues OpenAI, Sam Altman, Saying They Abandoned Founding Mission*, Wall St. J. (Mar. 1, 2024), https://www.wsj.com/tech/ai/elon-musk-sues-openai-sam-altman-for-breach-of-contract-0864979d [https://perma.cc/A5LA-AVFT] (describing Musk's lawsuit as saying OpenAI was "giving priority to profit over the benefits of humanity").

[4] Nowhere does the May 2024 Senate's roadmap for AI policy address copyright AI concerns. Chuck Schumer, Mike Rounds, Martin Heinrich & Todd Young, Driving U.S. Innovation in Artificial Intelligence: A Roadmap for Artificial Intelligence Policy in the United States Senate (May 2024), https://www.schumer.senate.gov/imo/media/doc/Roadmap_Electronic1.32pm.pdf [https://perma.cc/B424-HPHK]. In a law that went into effect on August 1, 2024, European regulators have taken a stance that is more protective of creative rights and individual privacy. *See* The EU Artificial Intelligence Act, Future of Life Inst., https://artificialintelligenceact.eu/ [https://perma.cc/2KL2-RUCW] (last visited Oct. 9, 2024). For a description of international AI copyright cases, see Aklovya Panwar, *Generative AI & Copyright Issues Globally: ANI Media v, OpenAI*, Tech Pol'y Press (Jan. 8, 2025), https://www.techpolicy.press/generative-ai-and-copyright-issues-globally-ani-media-v-openai/ [https://perma.cc/4LVZ-PK5L].

[5] *See* Xiyin Tang, *The Class Action as Licensing and Reform Device*, 122 Colum. L. Rev. 1627, 1632 (arguing that class actions "hav[e] a part to play in making substantive copyright law") [hereinafter *The Class Action*].

influence of artificial intelligence in everyday life.[6] This Article addresses the first wave—intellectual property disputes against the AI companies themselves—which is already here.

Copyright—the protection for "the creative powers of the mind"[7]—is enshrined in the U.S. Constitution.[8] To replicate natural-seeming language, AI models are dependent on complex, rich narratives—narratives that can be found in novels, short stories, academic articles, poetry, songs, and other copyright-protected creative output. So AI companies copied that output without slowing down to seek permission or offer compensation. The companies decided that it was better to ask forgiveness than permission.[9] At a time when the humanities are already under attack, technology companies are eating them for breakfast.[10]

Over the past year, authors, artists, coders, publishers, and organizations representing them have filed suits arguing that AI foundation model companies, from OpenAI to Anthropic—backed by Big Tech money—are violating their copyrights by "training" their large language models on tens of thousands of

[6] In addition to the copyright class actions discussed here, other lawsuits are investigating the role of AI companies in scraping individuals' data, see, e.g., Class Action Complaint at 1, P.M. v. OpenAI LP, No. 23-CV-03199 (N.D. Cal. June 28, 2023) (alleging scraping of publicly available information violated right to privacy), and using AI to reduce insurance payouts, Class Action Complaint at 1, Kisting-Leung v. Cigna Corp., No. 23-CV-01477 (E.D. Cal. July 24, 2023) (alleging AI algorithm wrongly denied insurance claims).

[7] *In re* Trade-Mark Cases, 100 U.S. 82, 94 (1879); *see generally* Julie E. Cohen, *Creativity & Culture in Copyright Theory*, 40 U.C. Davis L. Rev. 1151 (2007) ("Creativity is universally agreed to be a good that copyright law should seek to promote, yet copyright scholarship and policymaking have proceeded largely on the basis of assumptions about what it actually is.").

[8] U.S. Const. art. I, § 8, cl. 8 (Congress shall have power "To promote the Progress of Science and Useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries").

[9] Alexei Oreskovic, *Can We Afford to Let AI Companies Ask for Forgiveness Instead of Permission?*, Yahoo Fin. (May 31, 2024), https://finance.yahoo.com/news/afford-let-ai-companies-ask-190305079.html [https://perma.cc/QKJ6-FK4R] ("Asking for forgiveness, rather than permission, is Silicon Valley's favorite business model—from Uber's early days entering cities without seeking approval from local officials to the social networking companies' loose treatment of user data. With the AI market booming, the forgiveness cycle is kicking into high gear once again."); Torshie Torto, *AI May Not Replace You But It Will Steal From You*, Medium (Sept. 29, 2023), https://medium.com/@naatorsh/generative-ai-is-theft-d99643bbfcf4 [https://perma.cc/R8V3-98EF].

[10] *See, e.g.*, Ignacio M. Sanchez Prado, *The Humanities Are Worth Fighting For*, L.A. Rev. of Books (July 14, 2023), https://lareviewofbooks.org/article/the-humanities-are-worth-fighting-for/ [https://perma.cc/427S-SV5T] (noting that "the humanities were already being dismantled" as early as 2001, but pointing to the 2008 financial crisis as a tipping point).

copyrighted creative works without even the veneer of an attempt to obtain consent or provide compensation.[11]

Despite the enormous influence of a handful of class suits on the development of copyright law, there remains very little scholarship on the intersection of class actions and copyright.[12] In this Article, we seek to advance the study of copyright class actions in the field of artificial intelligence. To that end, we argue that class action plaintiffs' input claims are correct on the merits. Beginning in the 1990s with the Ninth Circuit's decisions in the so-called *MAI* Trio, courts have consistently and adeptly used existing copyright law to preserve authorial rights against incursions by new duplicating technologies. Under that precedent, AI "training" constitutes unlawful copying. This same set of precedents also undermines AI companies' fair use defense.

Procedurally, this Article argues that the Rule 23 class action device is an effective and appropriate tool with which to impose legal guardrails for direct infringement on AI companies. Examining earlier copyright class actions, we show that the AI plaintiffs' direct copyright infringement claims—which we call "input" claims because they arise out of material that was input into AI engines—fall squarely within the letter of Rule 23. AI class actions also embody Rule 23's spirit. Class actions are appropriate when a defendant's conduct impacts a group of people in functionally the same way. This maps on precisely to the facts of what AI companies did. Moreover, the class action device can achieve a remedy for large numbers of people who otherwise would be without the resources or knowledge to

[11] Alex Reisner, *Generative AI is Challenging a 234-Year-Old Law*, The Atl. (Feb. 29, 2024), https://www.theatlantic.com/technology/archive/2024/02/generative-ai-lawsuits-copyright-fair-use/677595/ [https://perma.cc/B4CW-QNXP].

[12] The preeminent article in this area is Tang, *The Class Action*, *supra* note 5. A few other works have focused on particular copyright cases. *See, e.g.*, James Grimmelman, *Future Conduct and the Limits of Class Action Settlements*, 91 N.C. L. Rev. 387, 419–20 (2013) (discussing the *Google Book Project* case); Pamela Samuelson, *The Google Book Settlement as Copyright Reform*, 2011 Wis. L. Rev. 479, 560 (2011) (same) [hereinafter *The Google Book Settlement*]; Renee G. Stern, Note, *Taming the "Frankenstein Monster": Copyright Claim Compatibility with The Class Action Mechanism*, 38 Colum. J.L. & Arts 549 (2015) (discussing Football Ass'n Premier League Ltd. v. YouTube, Inc., 297 F.R.D. 64 (S.D.N.Y. 2013)). Professor Tang has also argued that licensing regimes—such as we propose here—carry the risk of entrenching tech oligopolies that will later evade licensing regimes. *See* Xiyin Tang, *Copyright's Techno-Pessimist Creep*, 90 Fordham L. Rev. 1151, 1184–87 (2021). We discuss our pragmatic response to this. *See infra* Part VII.

vindicate their rights. That is also true here: Only as a group do creators have the clout to negotiate with companies valued in the billions of dollars.[13]

Importantly, input claims are only one facet of creators' challenges to AI. Plaintiffs in AI cases have also alleged various downstream, post-"training" infringement claims, which we refer to as "output" claims.[14] For example, plaintiffs have argued that AI engine outputs may themselves be infringing.[15] Our focus here is on input claims, but two observations are in order. First, while class certification may be appropriate for adjudicating some output claims, requirements for individualized proof and/or individualized assessment of defenses will render class certification more challenging for such claims. Second, if (as we think likely) the parties reach a settlement in the pending cases, any such agreement is likely to encompass both input and output claims, which arise from the same basic factual predicate. Any release of future output claims will thus hover in the air during resolution of the input claims. In assessing the fairness of a settlement, the key question will be whether the agreement preserves authorial voice and independence into the future in a sustainable way.

With regard to both input and output claims, the class action device serves a vital forward-focused role. Xiyin Tang has argued that copyright class actions spur the creation of licensing schemes and generate blueprints for legislative gap-filling.[16] That is already happening in the AI context. Ironically, at the very same time that AI companies are defending their "fair learning" justification for stealing from the creative working class, they are already entering into licensing deals with news publishers as well as user platforms such as Reddit.[17] The difference in the

[13] Cade Metz & Tripp Mickle, *OpenAI Completes Deal That Values the Company at $80 Billion*, N.Y. Times (Feb. 16, 2024), https://www.nytimes.com/2024/02/16/technology/openai-artificial-intelligence-deal-valuation.html [https://perma.cc/T5LX-M6L3z].

[14] *See* Class Action Complaint at 22, Andersen v. Stability AI Ltd., No. 23-CV-00201 (N.D. Cal. Jan. 13, 2023).

[15] *See id.*; First Amended Complaint at 2, Doe v. GitHub, Inc., No. 22-CV-06823-JST (N.D. Cal. July 21, 2023); Complaint at 4, Getty Images (US), Inc. v. Stability AI, Inc., No. 23-CV-00135-UNA (D. Del. Feb. 3, 2023); Complaint at 6, Concord Music Grp., Inc. v. Anthropic PBC, No. 23-CV-01092 (M.D. Tenn. Oct. 18, 2023); Complaint at 32, N.Y. Times Co. v. Microsoft Corp., No. 23-CV-11195 (S.D.N.Y. Dec. 27, 2023).

[16] Tang, *The Class Action*, *supra* note 5, at 1644 ("[A]ggregate litigation and its ensuing settlements act as both blanket licenses, where individualized licensing negotiations proved impossible, *and* as the blueprint for gap-filling legislation.").

[17] Todd Spangler, *OpenAI Inks Licensing Deals to Bring Vox Media, The Atlantic Content to ChatGPT*, Variety (May 29, 2024), https://variety.com/2024/digital/news/

class action context is that a court will need to approve any settlement; an important element of that judicial review will be ensuring that working class artists, and not only best-selling authors, receive fair compensation for any licensing agreement that results.

We conclude by centering the risk to human creativity that is at the heart of the pending class actions. Artificial intelligence inevitably has the buzz, and the buzzwords, of cutting-edge post-modern life.[18] But generative AI raises questions at least as old and as important as the Constitution. And the Constitution's far-sighted protection of intellectual property rights should not be forgotten in the heady rush of innovation. In fundamental ways, AI copyright class actions are not solely about the right of this or that author, or the economic impact of copyright

---

openai-vox-media-atlantic-chatgpt-licensing-deals-1236018547/ [https://perma.cc/4TTC-29NZ]; Katie Robertson, *OpenAI Strikes a Deal to License News Corp Content*, N.Y. Times (May 22, 2024), https://www.nytimes.com/2024/05/22/business/media/openai-news-corp-content-deal.html [https://perma.cc/YFQ3-7KBF] ("The deal gives OpenAI's chatbots access to new and archived material from The Wall Street Journal, The New York Post, MarketWatch and Barron's, among others."); Emilia David, *OpenAI Strikes Licensing Deal With the Magazine Giant Behind People*, The Verge (May 7, 2024), https://www.theverge.com/2024/5/7/24151171/openai-dotdash-meredith-people-instyle-licensing [https://perma.cc/H8NJ-X6XU] (reporting Dotdash Meredith, publisher of *People, Better Homes & Gardens, Investopedia*, *Food & Wine*, and *InStyle*, will license its content to OpenAI to "train AI models with its articles"); Angela Christy et al., *OpenAI to Use FT Content for Training AI Models in Latest Media Tie-up*, Reuters (Apr. 30, 2024), https://www.reuters.com/technology/financial-times-openai-sign-content-licensing-partnership-2024-04-29/ [https://perma.cc/5HD6-F6GM]; A.W. Ohlheiser, *A Poster's Guide to Who's Selling Your Data to Train AI*, VOX (Feb. 29, 2024), https://www.vox.com/technology/24086039/reddit-tumblr-wordpress-whos-selling-your-data-to-train-ai [https://perma.cc/TCJ9-P9NJ] ("The Associated Press has licensed part of its archives to OpenAI. Shutterstock, the stock photo archive, has signed a six-year deal with OpenAI to provide training data, which includes access to its photo, video, and music databases."); *see also id.* ("Reddit and Google entered into a $60 million deal that would give Google access to Reddit's API in order to, among other things, train its generative AI models."); Chris Morris, *Here Are the Companies OpenAI Has Made Deals With to Train ChatGPT*, Fast Co. (May 29, 2024), https://www.fastcompany.com/91130785/companies-reddit-news-corp-deals-openai-train-chatgpt-partnerships [https://perma.cc/6SXL-R8DV] ("In the past month, OpenAI has partnered with Reddit, News Corp., Vox Media, and the Atlantic as it looks for data to train ChatGPT.").

[18] Garielle Olya, *OpenAI Engineers Earn $900K Per Year: How Does That Compare To Salaries at Other Top Tech Companies?*, Yahoo Fin. (Apr. 25, 2024), https://finance.yahoo.com/news/openai-engineers-earn-900k-per-200010721.html [https://perma.cc/6QMG-HNHD]; *Writer Salary*, ZipRecruiter, https://www.ziprecruiter.com/Salaries/Writer-Salary [https://perma.cc/G4WR-VKD7] (last visited Mar. 19, 2025) ("As of Mar 12, 2025, the average annual pay for a Writer in the United States is $28,500 a year. Just in case you need a simple salary calculator, that works out to be approximately $13.70 an hour. This is the equivalent of $548/week or $2,375/month.").

infringement on individuals; rather, these claims are about a threat to the concept of authorship as a human endeavor, and to the vision of America's founders that copyright will protect and incentivize human creativity.

Since the Founding, copyright law has withstood the challenges of successive new technologies. But in a world where 90% of online content will be synthetically generated by 2026, human creativity is endangered.[19] Thus, we conclude by arguing that there is a constitutional—and not merely a statutory or economic—urgency to protecting the human role, and human agency, in the process of creation. Class actions are a feasible, viable way to respect this constitutional urgency. We recognize that licensing agreements are the most likely, and perhaps the only, way to resolve the impasse between creators and AI companies.

But licensing regimes pose their own risks to the livelihood and dignity of artists. Fair licensing must do more than give authors a smidgeon of cash in exchange for their creative souls; it must address the constitutional urgency, and the unknowable future, brought on by generative AI. To qualify as far, we argue that a licensing regime must bear three hallmarks: 1) it must allow creators to opt out; 2) it must not allow licensees' unilateral expansion of the scope of the license; and 3) it must provide a mechanism for attribution to human authors.

Part I of this Article illustrates how AI transforms, and taints, human creativity.

Part II briefly describes the pending intellectual property class actions currently pending against AI companies and describes their central copyright claims.

Part III traces how, beginning in the 1990s, courts have successfully applied existing copyright law to successive waves of technological disruption. It shows

[19] *AI Experts Predict by 2026, 90% of Online Content Will Be Generated by Artificial Intelligence*, IDCA Press (Sept. 27, 2022), https://idc-a.org/news/industry/AI-Experts-Predict-By-2026-90-Of-Online-Content/127ab0c0-34ba-4c03-8bad-1e4f21923f31 [https://perma.cc/43JB-AZXL]; Alexandra Garfinkle, *90% of Online Content Could Be Generated by AI by 2025, Expert Says*, Yahoo Fin. (Jan. 13, 2023), https://finance.yahoo.com/news/90-of-online-content-could-be-generated-by-ai-by-2025-expert-says-201023872.html [https://perma.cc/MX7Q-UKNB].

that, under these precedents, AI companies are liable for feeding copyright-protected creative works into their engines.

Part IV analyzes courts' application of Rule 23 in pre-AI copyright class actions. Current AI class claims reflect the same fundamental pressure points as did earlier suits.

Part V argues that, just as existing copyright law can meet this moment, so too can existing class action doctrine. Using two current AI cases as examples—one from the Second Circuit and one from the Ninth—we argue that AI plaintiffs' input claims fall in the heartland of what Rule 23 is designed to accomplish. We also recognize that class certification is likely to prove more difficult (though not impossible) for "output" claims—claims for derivative infringement and other downstream claims that arise out of the original unlawful copying.

Part VI contests the AI companies' fair use defense, which comes down to one philosophy: "trust us."

Part VII concludes by urging courts and Congress to reject fair learning in favor of fair licensing. It proposes three terms that can ensure that a licensing regime stays faithful to the constitutional essence of copyright law.

I
AI and the Risk of Post-Human Creativity

In a rock band, four or five members each play a role or more: a lead guitar, rhythm guitar, bass guitar, drum, keyboard, and lead vocal.[20] A symphony requires many musicians with talent in their respective instruments to collaborate under a conductor's baton perfecting a beautiful movement.[21] Jazz musicians are known for their partnership merging their different styles and elevating their art or seizing inspiration from listening to other jazz musicians.[22] The "intimate exchange of the creative pair, such as John Lennon and Paul McCartney" illuminates the value of human connectivity for creative process and achievements.[23] Recording an album needs more than just the band but the producer, the studio, the engineer, the mixing engineer, and the mastering engineer.[24] Connecting humans for collaboration is essential for creativity.

Michelangelo is known for the breathtaking painting of the Sistine Chapel ceiling. But he could not accomplish the feat alone. He "worked with a team of assistants" to achieve the results.[25] Picasso made Cubism together with Georges Braque and rivaled with Henri Matisse leading to adversarial collaboration.[26]

---

[20] *Rock of Ages: A Closer Look*, 5th Ave. Theater, https://www.5thavenue.org/media/h1hofn0t/1819_roa_study_guide.pdf [https://perma.cc/A8Z9-TRYC] (last visited Oct. 11, 2024).

[21] *What is the Difference Between a Symphony and an Orchestra?*, Colo. Symphony (Sept. 2, 2022), https://coloradosymphony.org/symphony-vs-orchestra/ [https://perma.cc/DP85-FBV9] ("A symphony is an orchestra's large-scale musical composition, usually requiring 50–80 instruments. Symphonies refer to the music, not the musicians performing it, and similar to stage plays, they may have several movements or acts of, often complex, elaborate classical music.").

[22] *The Most Essential Jazz Collaborations*, Spotify (Nov. 26, 2013), https://open.spotify.com/album/4681YDZS6qj3wCsVP6ndmG [https://perma.cc/EZW9-3MMZ] (listing 15 songs of famous collaborations).

[23] Joshua Wolf Shenk, *The End of 'Genius'*, N.Y. Times (July 19, 2014), https://www.nytimes.com/2014/07/20/opinion/sunday/the-end-of-genius.html [https://perma.cc/6SDS-P37Z].

[24] Ari Herstand, *What to Know Before You Record Your Album*, ASCAP (July 24, 2013), https://www.ascap.com/help/career-development/how-to-record-your-album [https://perma.cc/52RV-EW4C].

[25] Fernando Teixeira & Izabela Cardozo, *The 'Lone Genius' Myth: Why Even Great Minds Collaborate*, BBC (Mar. 12, 2021), https://www.bbc.com/worklife/article/20210308-the-lone-genius-myth-why-even-great-minds-collaborate [https://perma.cc/63CX-TLJT].

[26] *Cubism*, Tate, https://www.tate.org.uk/art/art-terms/c/cubism [https://perma.cc/8XG6-MTY9] (last visited Oct. 11, 2024) ("Cubism was a revolutionary new approach to representing reality invented in around 1907–08 by artists Pablo Picasso and Georges Braque. They brought different views of subjects (usually objects or figures) together in the same picture, resulting in paintings that appear fragmented and abstracted."); Paul Trachtman, *Matisse & Picasso*, Smithsonian Mag., Feb. 2003, at 62, 62 ("[T]hese

Albert Einstein accumulated inventive knowledge from reading and analyzing other people's inventions while he was working as a patent examiner at the Swiss Patent Office,[27] and testing his idea for his theory of relativity in conversation with the engineer Michele Besson, "the best sounding board in Europe."[28] Creativity is demonstratively the product of human connectivity and collaboration.

With generative AI, creativity may become *post*-human. A new rock song, a newly improvised jazz piece, a new symphony, among other myriad new music products, can be synthetically generated.[29] A known musician's style can be imitated through AI generation.[30] A famed vocalist's unique sound can be generated because GPT-4o "can respond to audio inputs in as little as 232 milliseconds."[31] Connectivity and collaboration are not invited. All the models need is prompts.

Do you want to write in the style of Tolstoy's novels—perhaps *War and Peace*, or *Anna Karenina*? AI models are standing by to assist in generating deepfake Tolstoy-like new text through a series of prompts: "1. Directly ask the AI to write like the writer you have in mind[,] 2. Describe what the writer's voice is like, and ask the AI to write like that (preferably with examples)[, and] 3. Fine-tune the AI

friends—and rivals—spurred each other to change the course of 20th-century art[.]"); John Richardson, *Between Picasso and Matisse*, Vanity Fair, May 18, 2009, at 144, 147 ("[T]he two greatest painters and sculptors of the 20th century bounced off each other, outdid each other, honored each other, and occasionally ignored each other, in ways that were sometimes calculated, sometimes instinctive, and sometimes fortuitous.").

[27] Teixeira & Cardozo, *supra* note 25.

[28] Shenk, *supra* note 23.

[29] There are many AI music generators available today. For example, see Loudly, https://www.loudly.com/ai-music-generator [https://perma.cc/738W-QZLL] (last visited Jan. 29, 2025) ("Studio quality music … Make AI-generated music in seconds and add to your videos, social channels, advertising media, podcasts, apps and more."); Soundraw, https://soundraw.io/ [https://perma.cc/Z4NY-QPHL] (last visited Jan. 29, 2025) ("Create tracks with AI for background music, making your own songs, distributing songs, [and] selling songs[.]"); Soundful, https://soundful.com/en-us/ [https://perma.cc/5MUH-UEHH] (last visited Jan. 29, 2025) ("The future of music is here with Soundful's AI Music Generator. Leverage the power of AI to generate royalty free background music at the click of a button for your videos, streams, podcasts and much more.").

[30] If you are inclined to create a Beethoven-like symphony, see Today, *AI Scientists Use Computer to Create Beethoven-Like Symphony*, YouTube (Oct. 24, 2021), https://www.youtube.com/watch?v=0AU10opANTE [https://perma.cc/32US-4LK7].

[31] OpenAI, *supra* note 1.

on a specific writer[.]"[32] Do you wish to draft a sitcom dialogue mimicking Jerry and Kramer in a new Seinfeld episode?[33] In minutes, AI models trained on novels and scripts written by human authors present you with synthetic literary text of the original authors' style.

Writing pitches for sequels to popular movies in an instant? AI models' performance in the movie sequels task garnered praises like "surprising," "good," "good enough that you could see someone buying a ticket to the film," and "brilliant."[34] For instance, in response to *Casablanca* in the prompt, AI wrote: "One possible plot for a sequel to *Casablanca* could explore what happened to Rick Blaine and Ilsa Lund after they parted ways in Casablanca. The sequel could pick up several years after the events of the original film, with Rick and Ilsa living separate lives in different parts of the world. Perhaps circumstances conspire to bring them back together, and they must grapple with the complicated feelings they have for each other and the choices they made in the past."[35] Having ingested all the scripts and movies, AI models produce new pitches deriving from the originals.

AI can generate a painting that helps you win a jury prize.[36] Prompts, again, are what you input at the computer.[37] Learning how to paint from teachers, imitating styles from accomplished artists, studying their techniques, and listening to critiques from others, none of these steps are necessary. AI models trained on existing images from all sources are ready at your prompt command to generate

[32] Dan Shipper, *How to Make AI Write Like Your Favorite Author, a Step-by-Step Guide from Prompting to Fine-Tuning*, Every (Mar. 3, 2023), https://every.to/chain-of-thought/how-to-make-ai-write-like-your-favorite-author [perma.cc/MYL3-5JNK].

[33] *Id.* (featuring an AI-generated dialogue between Jerry and Kramer to illustrate the differences between using different AI models for generating text).

[34] Matt Singer, *We Asked an Artificial Intelligence to Write Sequels to Beloved Movies*, Screen Crush (Mar 2, 2023), https://screencrush.com/artificial-intelligence-writes-movie-sequels/ [https://perma.cc/E62C-6JCE].

[35] *Id.*

[36] Kevin Roose, *An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy*, N.Y. Times (Sept. 2, 2022), https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html [https://perma.cc/YE98-JGCB] (reporting Jason M. Allen won the Colorado State Fair's annual art competition for his AI-generated piece).

[37] *Id.* (reporting Jason M. Allen typed prompts into Midjourney AI to generate "Théâtre D'opéra Spatial").

whatever you desire and whenever you wish.[38] All you need is to type in the prompts, and generative AI programs will do the rest in seconds.[39]

Learning to speak or sing beautifully in a foreign language typically takes years of learning from teachers and coaches. In many cases, the students may travel abroad to submerge themselves in living in foreign countries for the opportunities to connect with the people and culture while improving their language abilities. In the AI age, just ignore human connectivity and creativity regarding learning foreign languages. For instance, AI can generate an English-speaking rock band vocalist (who is real) rocking away fluently and with nuanced flair in Japanese and Korean languages (synthetically, of course).[40]

Albums of new songs by different artists can be synthetically generated without ever stepping into a studio.[41] Google's Dream Track AI tool excites performers to generate new content in their replicas.[42] The days of collaboration

[38] For "best" prompts to help generating paintings and images, see *Best Midjourney Prompts*, PromptHero, https://prompthero.com/midjourney-prompts [https://perma.cc/84MB-XE2E] (last visited Oct. 10, 2024) ("Search the best Midjourney prompts and get millions of ideas for your next prompt engineering session. Showing only good prompts for Midjourney, ranked by users' upvotes and popularity."); Hiba Amin, *How to Master Midjourney Prompts (Best Prompts in 2024)*, Superside (Oct. 5, 2024), https://www.superside.com/blog/midjourney-prompts [https://perma.cc/GT49-PG23].

[39] *Prompts*, Midjourney, https://docs.midjourney.com/docs/prompts [https://perma.cc/82GW-VUA5] (last visited Oct. 10, 2024) ("A prompt is a short text phrase that the Midjourney Bot interprets to produce an image. The Midjourney Bot breaks down the words and phrases in a prompt into smaller pieces, called tokens, that are compared to its training data and then used to generate an image. A well-crafted prompt can help make unique and exciting images. . . . A basic prompt can be as simple as a single word, phrase or emoji. The Midjourney Bot works best with simple, short phrases that describe what you want to see. Avoid long lists of requests and instructions. Instead of: Show me a picture of lots of blooming California poppies, make them bright, vibrant orange, and draw them in an illustrated style with colored pencils Try: Bright orange California poppies drawn with colored pencils.").

[40] FKA Twigs, whose real name is Tahliah Debrett Barnett, can replicate herself singing in different languages. *See* Prithvi Iyer, *Transcript of Senate Hearing on NO FAKES ACT*, Tech Pol'y Press (May 20, 2024), https://www.techpolicy.press/transcript-us-senate-judiciary-subcommittee-hearing-on-the-no-fakes-act/ [https://perma.cc/JC5Q-CBBW] ("In the past year, I have developed my own deepfake version of myself that is not only trained in my personality but can also use my exact tone of voice to speak many languages.").

[41] Chloe Veltman, *Google's Latest AI Music Tool Creates Tracks Using Famous Singers' Voice Clones*, NPR (Nov. 17, 2023), https://www.npr.org/2023/11/17/1213551049/googles-latest-ai-music-tool-creates-tracks-using-famous-singers-voice-clones [https://perma.cc/T9UP-MQCH].

[42] *Id.* ("Whoever gets paid for [their voice clone] will be happy because they'll be able to sit at home and not have to go to a recording session.").

in the studio are numbered. Obviously, for now, those famous few get paid by Google's Deep Mind AI for having their voices cloned as part of new synthetic content creation without collaborating with others in the studio. For many others with ordinary voices, there is nothing to collect.[43]

Generating a haiku in seconds? Ditto. AI models ingest haiku and are trained to generate haiku, poems known for having 17 syllables, and apply the world's shortest and clearest rules to compose new poems and evaluate their aesthetic. AI can generate haiku quicker than humans, and humans fail to distinguish which haikus are synthetic and which are human-created.[44] John Grisham took years to research and develop characters and stories for his books. However, AI can generate sequel chapters mimicking John Grisham's style in an instant.

AI poses an existential threat to the creative working class, and also to the very concept of human creativity.[45] If we do not act soon, we will no longer be able to distinguish—or perhaps even remember—what human creativity is.

II
Copying and Fair Use in Disruptive Technologies

The Constitution's Copyright Clause was drafted during the age of lithography—more than 50 years before the rotary press spurred modern print media.[46] Yet its language, and its mission, have remained relevant.[47] Copyright law has outlasted cycles of disruptive technological advancements. When personal computers first began to enter businesses and households, critics lamented

[43] Xuan-Thao Nguyen, *Fakes and the New Property Right?*, 53 Fla. St. U. L. Rev. (forthcoming 2026).

[44] Jimpei Hitsuwari et al., *Does Human-AI Collaboration Lead to More Creative Art? Aesthetic Evaluation of Human-Made and AI-Generated Haiku Poetry*, 139 Comput. in Hum. Behav., Feb. 2023, at 1, 1, https://www.sciencedirect.com/science/article/abs/pii/S0747563222003223 [https://perma.cc/EU4T-DX7R].

[45] Emma Saunders, *John Grisham: Threat from AI Cannot Be Truly Appreciated*, BBC (Oct. 17, 2023), https://www.bbc.com/news/entertainment-arts-67134595 [https://perma.cc/W67V-WMNR]; Susan Hornik, *Hollywood Writers Fear Losing Work to AI*, BBC (July 26, 2023), https://www.bbc.com/news/business-66289583 [https://perma.cc/L7UA-H42S].

[46] The rotary printing press, invented by Richard Hoe, was first available in 1847. *See* Princeton Univ., *Hoe's Eight-Cylinder Printing Press*, Graphic Arts Collection (July 23, 2013), https://graphicarts.princeton.edu/2013/07/23/hoes-eight-cylinder-printing-press/ [https://perma.cc/EHM2-4LVF].

[47] U.S. Const. art. I, § 8, cl. 8 (granting Congress the power "To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries").

that "as people and businesses often find the transition to computers difficult, integrating computer software into the realm of copyright law has proven to be a daunting task."[48] Ultimately, however, both courts and Congress rose to the challenge, and copyright law has remained a relevant, practical tool for protecting authorial rights from technology-enabled infringement. Beneath each technological innovation, the heart of the issue has remained: the protection of human creative output. And in each instance, the Constitution's Copyright Clause has served as a compass, guiding courts to encourage technological innovation while simultaneously protecting human creativity.

This Part recounts how courts have thoughtfully applied copyright doctrine—both affirmative claims, as well as the defense of fair use—to waves of disruptive digital technology. The precedents described below provide a firm basis for holding AI companies liable for their unauthorized "training" of AI models on copyrighted creative works, and for denying a "fair learning" defense. Simultaneously, however, this part recounts the Supreme Court's somewhat expansive use of the fair use doctrine to protect some code-copying in its 2021 decision in *Google LLC v. Oracle America, Inc.*[49]

A. The MAI *Trio: Code Copying*

The Copyright Act explicitly contemplates the evolution of technological methods of copying creative output. It defines "copies" as "material objects … in which a word is *fixed by any method now known or later developed*."[50] In the early 1990s, the Ninth Circuit penned three opinions later referred to as the *MAI* Trio, defining the meaning of "copying" under the Copyright Act during the rise of software and personal computers. These decisions, which granted property rights—and copyright protections—to the output of the software age, apply directly in the age of artificial intelligence.[51]

---

[48] Jule L. Sigall, Comment, *Copyright Infringement Was Never This Easy: RAM Copies and Their Impact on the Scope of Copyright Protection for Computer Programs*, 45 CATH. U. L. REV. 181, 181 (1995); Pamela Samuelson et al., *A Manifesto Concerning the Legal Protection of Computer Programs*, 94 COLUM. L. REV. 2308, 2310 (1994) (observing that copyright law fails to protect computer programs); Comput. Assocs., Int'l, Inc. v. Altai, Inc., 982 F.2d 693, 712 (2d Cir. 1992).

[49] Google LLC v. Oracle Am., Inc., 593 U.S. 1 (2021).

[50] Copyright Act of 1976, 17 U.S.C. § 101 (emphasis added).

[51] *See* Danielle D'Onfro, *Contract-Wrapped Property*, 137 HARV. L. REV. 1058, 1072–73 (2024) (*MAI Systems Corp.* "arguably chang[ed] the shape of modern property ownership").

1. Code Is a "Copy"

The first of the Trio, *MAI Systems Corp. v. Peak Computer*, established that even temporary copies of software code are protected by copyright law.[52] In so doing, *MAI Systems* defined the concept of "copying" for the digital age.

The case concerned a dispute between MAI—a computer manufacturer and software developer—and Peak, a computer maintenance company.[53] When Peak provided maintenance and repair services to those of its customers using MAI computers, it necessarily ran MAI software.[54] MAI argued that this unlicensed use of its software constituted "copying" for purposes of the Copyright Act; it sought and obtained a permanent injunction against Peak.[55] The Court framed the question as whether "copying" occurs when a computer program is transferred from a permanent storage device to a computer's random-access memory (RAM).[56] It answered that question in the affirmative.

To begin, the Court noted that it was "generally accepted that the loading of software into a computer constitutes the creation of a copy." It applied this same logic to RAM copies.[57] A computer executing any program must load that program into RAM.[58] Peak argued that RAM copies were not "fixed" under copyright law. But the Ninth Circuit held that a RAM copy is "sufficiently permanent or stable to permit it to be perceived, reproduced, or otherwise communicated for a period of

---

[52] MAI Sys. Corp. v. Peak Comput., 991 F.2d 511 (9th Cir. 1993).

[53] *Id.*

[54] *Id.* at 513.

[55] *Id.* at 514–15.

[56] *Id.*

[57] *Id.* at 519 (citing Vault Corp. v. Quaid Software Ltd., 847 F.2d 255, 260 (5th Cir.1988)); 2 Melville B. Nimmer & David Nimmer, Nimmer on Copyright § 8.08 (1983); Nat'l Comm'n on the New Tech. Uses of Copyrighted Works, Final Report of the National Commission on the New Technological Uses of Copyrighted Works 13 (1978)).

[58] *See What Is Computer and Laptop RAM?*, Intel, https://www.intel.com/content/www/us/en/tech-tips-and-tricks/computer-ram.html [https://perma.cc/M3E3-42UK] (last visited Oct. 10, 2024) ("RAM is a common computing acronym that stands for random-access memory. Sometimes it's called PC memory or just memory. In essence, RAM is your computer or laptop's short-term memory. It's where the data is stored that your computer processor needs to run your applications and open your files. Inside your computer, RAM typically comes in the form of a rectangular flat circuit board with memory chips attached, also referred to as a memory module. Computers typically come with at least two RAM modules with room to add more, if needed. These RAM modules are critical components that work hand in hand with your computer's central processing unit (CPU) and must be working optimally for you to have a good experience.").

more than transitory duration."[59] Therefore, it held that the "loading of copyrighted computer software from a storage medium (hard disk, floppy disk, or read-only memory) into the memory of a central processing unit "CPU") causes a copy to be made."[60]

The immediate impact of the case was that MAI could prevent Peak from using its software, thereby preventing Peak from offering its services as a competitor. The lasting impact was even broader. *MAI* was innovative, clear, and—given the lack of Supreme Court precedent in copyright—"far-reaching."[61] It was also problematic. Scholars have criticized *MAI* on textual and policy grounds.[62] As Danielle D'Onfro recently observed, the effect of *MAI* was to entrench restrictive intellectual property licenses into anything related to software—which is fast becoming everything.[63] Indeed, the Ninth Circuit now appears to offer a more nuanced take on software copies.[64]

Nevertheless, what is good for the goose should be good for the gander. *MAI*'s broad conception of a fixed copy—which has tended to benefit corporate software developers—provides authors and other copyright owners with a strong shield against non-consensual copying by AI companies.

2. Limiting Fair Use for Code Copying

The second case in the *MAI* Trio, *Triad Systems Corp. v. Southeastern Exp. Co.*, which reached the Ninth Circuit in 1995, narrowly construed fair use in

---

[59] *MAI Sys. Corp.*, 991 F.2d at 518.

[60] *Id.*

[61] *See* Stephen Yelderman, *The Supreme Court's Fragile Copyright Law*, 50 Fla. St. U. L. Rev. 335, 345–46 (2023) (noting that, given the rarity of Supreme Court review of copyright decisions, circuit law has broad influence).

[62] Aaron Perzanowski, *Fixing RAM Copies*, 104 Nw. U. L. Rev. 1067, 1075–80 (2010).

[63] D'Onfro, *supra* note 51, at 1073.

[64] CDK Glob. LLC v. Brnovich, 16 F.4th 1266, 1276 (9th Cir. 2021) ("Loading software into a computer's memory satisfies the embodiment requirement because a computer's memory is a medium from which software 'can be perceived, reproduced, or otherwise communicated.' But embodiment alone does not result in the creation of a copy; the embodiment must also persist for a period of more than transitory duration. We have not previously considered what is required for a copy to persist for more than a transitory period. . . . *MAI* and the cases following it establish only that 'loading a program into a computer's memory can result in copying that program,' not that 'loading a program into a form of memory always results in copying.'") (citations omitted).

the copying of software.[65] Triad manufactured computers for use by automotive parts stores.[66] Triad also developed software for its service technicians to use.[67] Southeastern competed against Triad for the business of servicing and maintaining Triad computers.[68] Upon discovering that Southeastern had been making RAM copies of Triad software when providing services to customers, Triad sued.[69] Following *MAI*, Triad moved for summary judgment on its copyright infringement claim; Southeastern countered with a fair use defense.[70] The district court denied both motions, the case proceeded to trial, and the jury found that Southeastern had violated Triad's copyrights.[71] Following the verdict, the district court granted Triad's motion for a preliminary injunction, pending the resolution of other issues in the case.[72] Southeastern appealed.

Applying *MAI*, the Ninth Circuit expressed "no doubt" that Southeastern had made infringing copies of Triad's software.[73] The Court also noted that, for purposes of injunctive relief, there is a presumption of irreparable harm from copyright infringement, notwithstanding that a copyright plaintiff can theoretically be made whole through money damages.[74]

---

[65] *See* Triad Sys. Corp. v. Se. Exp. Co., 64 F.3d 1330, 1335–36 (9th Cir. 1995), *superseded by statute on other grounds*, 17 U.S.C. § 117(c), *as stated in* 4Die4Kourt v. Hillair Cap. Mgmt., LLC, 692 F. App'x. 366, 369 (9th Cir. 2017). 17 U.S.C. § 117(c) provides:

> Machine Maintenance or Repair.—Notwithstanding the provisions of section 106, it is not an infringement for the owner or lessee of a machine to make or authorize the making of a copy of a computer program if such copy is made solely by virtue of the activation of a machine that lawfully contains an authorized copy of the computer program, for purposes only of maintenance or repair of that machine, if—
> (1) such new copy is used in no other manner and is destroyed immediately after the maintenance or repair is completed; and
> (2) with respect to any computer program or part thereof that is not necessary for that machine to be activated, such program or part thereof is not accessed or used other than to make such new copy by virtue of the activation of the machine.

[66] *Triad Sys. Corp.*, 64 F.3d at 1333.
[67] *Id.*
[68] *Id.*
[69] *Id.*
[70] *Id.* at 1334.
[71] *Id.*
[72] *Id.*
[73] *Id.* at 1335.
[74] *Id.*

Most significantly, the Ninth Circuit rejected Southeastern's fair use defense.[75] Southeastern, the Court held, "has invented nothing of its own; its use of Triad's software is … neither creative nor transformative, and does not provide the marketplace with new creative works."[76] Southeastern was "simply commandeering its customers' software and using it for the very purpose for which, and in precisely the manner in which, it was designed to be used."[77] The Court also found that Southeastern's conduct yielded no appreciable public benefit.[78] Accordingly, the Ninth Circuit rejected the fair use defense in RAM copy infringement cases.

In 1998—after *MAI* and *Triad*—Congress amended Section 117 of the Digital Millenium Copyright Act to prohibit software developers from using the RAM copy doctrine to eliminate competition from computer service companies.[79] But Congress did not eliminate the RAM copy doctrine itself; it merely provided an exception to it.[80] Congress also did not overturn *MAI*'s holding that license agreements define the limits of licensees' rights with regard to software.[81]

3. The Age of Licenses

A decade later, the Ninth Circuit decided the final case in the *MAI* Trio, *Wall Data Inc. v. LA County Sheriff's Dept.*[82] Wall Data developed RUMBA software and the Los Angeles County Sheriff's Department purchased licenses for 3,663 computers.[83] But the Sheriff's Department installed the software on 6,007

---

[75] *Id.* at 1336.

[76] *Id.*

[77] *Id.* at 1337.

[78] *Id.*

[79] *See* 17 U.S.C. § 117(c).

[80] *See* Melissa A. Bogden, *Fixing Fixation: The RAM Copy Doctrine*, 43 Ariz. St. L. J. 181, 197 (2011) ("By codifying an exception for RAM copies created during computer maintenance and repair, Congress implicitly approved MAI's conclusion that a RAM copy could be fixed.").

[81] *See* Vernor v. Autodesk, Inc., 621 F.3d 1102, 1111 (9th Cir. 2010) ("In response to *MAI*, Congress amended § 117 to permit a *computer owner* to copy software for maintenance or repair purposes. *See* 17 U.S.C. § 117(c); *see also* H.R. Rep. No. 105-551, pt. 1, at 27 (1998). However, Congress did not disturb MAI's holding that licensees are not entitled to the essential step defense.").

[82] Wall Data Inc. v. L.A. Cnty Sheriff's Dep't, 447 F.3d 769 (9th Cir. 2006).

[83] *Id.* at 773.

computers.[84] Wall Data brought a copyright infringement suit against the Sheriff's Department, which raised "fair use" and "essential step" defenses.[85]

The Court rejected fair use. It ruled that making verbatim copies of software, for the same purpose, was not transformative, nor was it an industry norm.[86] It also held that, were the practice of duplicating software in excess of one's license to become widespread, it "could seriously impact" the market.[87]

The Court also rejected an "essential step" defense under 17 U.S.C. § 117.[88] "Essential step" means the making of another copy of a software program in the utilization of the computer program in conjunction with a machine. The copy is used in no other manner. The Ninth Circuit held that the "essential step" defense only applies to an "owner" of a program, not to a licensee such as the LA Sherriff's Department.[89] The Ninth Circuit cited *MAI* in support of its ruling that the Sheriff's Department's licensee is bound by the license terms, and as a licensee, the Sheriff's Department did not have the statutory right to make extra copies of the software.[90]

---

[84] *Id.* at 776.

[85] *Id.*

[86] *Id.* at 776–82.

[87] *Id.* at 781–82.

[88] 17 U.S.C. § 117 provides:

> (a) Making of Additional Copy or Adaptation by Owner of Copy.—Notwithstanding the provisions of §106, it is not an infringement for the owner of a copy of a computer program to make or authorize the making of another copy or adaptation of that computer program provided:
> (1) that such a new copy or adaptation is created as an essential step in the utilization of the computer program in conjunction with a machine and that it is used in no other manner, or
> (2) that such new copy or adaptation is for archival purposes only and that all archival copies are destroyed in the event that continued possession of the computer program should cease to be rightful.

[89] 447 F.3d at 784–85 ("We conclude that the Sheriff's Department received licenses to the RUMBA software. Generally, if the copyright owner makes it clear that she or he is granting only a license to the copy of software and imposes significant restrictions on the purchaser's ability to redistribute or transfer that copy, the purchaser is considered a licensee, not an owner, of the software.").

[90] *Id.* at 785 ("In this case, as in *MAI*, the licensing agreement imposed severe restrictions on the Sheriff's Department's rights with respect to the software. Such restrictions would not be imposed on a party who owned the software. . . . The Sheriff's Department's use of and rights to the RUMBA software products were restricted under the terms of the click-through and volume booklet licenses. These restrictions were sufficient to classify the transaction as a grant of license to Wall Data's software, and not a sale of Wall Data's software. For these reasons, under *MAI*, the Sheriff's Department is not the "owner" of copies of Wall Data's software for purposes of § 117.").

The *MAI* Trio applied long-standing copyright law to burgeoning digital technologies. The Trio established that digital copying of copyrighted works in the computer software context is unlawful infringement, barring permission or a statutory exemption. It also empowered and protected software creators by allowing them to define the terms of use for their product through the (now ubiquitous) use of licenses. Notably, courts—not Congress—took the lead in analyzing the application of the Copyright Act to new technologies.[91] Although Congress eventually weighed in, amending § 117, it did so only after the issues of copying and fair use of software had been explicated through an adversary process between those parties most directly impacted.

B. Cached Contents and Fair Use: Field v. Google

The explosion of networked computers in the early 2000s forced copyright law to update copyright doctrine once again. Companies such as Google, Yahoo, AltaVista, and Lycos developed technologies to capture the emerging search sector of the digital economy. To do so, they sought to copy, analyze, and cache every existing web page. Google's automatic program for finding and cataloging web pages was called the "Googlebot."[92] Search companies did not seek permission from the copyright owners of any of these web pages. Litigation soon followed, pitting copyright owners against the search companies.

In this battle, the search companies held their own.

When a user conducts a search at Google, the search engine displays its search results, comprising the title of a webpage, short snippets from the webpage, the URL for the page, and another link labeled "cached."[93] Clicking on the "cached" link "directs an Internet user to the archival copy of a Web page stored in Google's system cache, rather than to the original Web site for that page."[94] A disclaimer explains that the page is a snapshot from the Google cache, not the original page.[95]

---

[91] *See also* Bogden, *supra* note 80, at 186 (defending the role of courts in elucidating copyright doctrine of emerging technologies).

[92] *Googlebot*, GOOGLE FOR DEVS., https://developers.google.com/search/docs/crawling-indexing/googlebot [https://perma.cc/YYM9-W5EN] (last visited Jan. 20, 2024).

[93] Field v. Google, Inc., 412 F. Supp. 2d 1106 (D. Nev. 2006).

[94] *Id.* at 1111.

[95] *Id.*

The snapshot is therefore not current if the webpage has been updated by the owner.[96]

The "cached" links served several purposes, according to Google, including: (1) providing archival copies that allowed users to access the pages in the event the website became inaccessible, (2) providing webpage comparisons, enabling users to determine how a particular webpage changed over time, and (3) identifying search query terms, thus allowing users to determine why a particular page was more responsive to their query.[97] Also, according to Google, the vastness of the internet prevented it from contacting website owners to ascertain whether they wanted their pages listed in search results or accessible through "cached" links.[98]

In *Field*, the plaintiff—an attorney and amateur author—brought a copyright infringement action against Google for copying his website, on which he had posted 50 creative works, and storing these copies in a cache without authorization.[99] The plaintiff argued that Google directly infringed his copyrights when a user clicked on a "cached" link to the web pages containing the plaintiff's copyrighted works and downloaded a copy of those pages from Google's cache.[100] That meant, according to the plaintiff, that Google itself had created and distributed copies of the plaintiff's works.[101]

Field did not prevail. The district court held that the *user*, not *Google*, creates and downloads a copy of the cached web page when the user requests a web page contained in the Google cache.[102] Without a user's request, no copy is created or distributed; even with such a request, the "automated, non-volitional conduct by Google in response to a user's request does not constitute direct infringement" under copyright law.[103] The district court also sustained Google's implied license defense. At the time of the lawsuit, Field had an easy way to protect his digital

[96] *Id.* at 1112.
[97] *Id.* at 1111–12.
[98] *Id.* at 1112.
[99] *Id.* at 1109.
[100] *Id.* at 1115.
[101] *Id.*
[102] *Id.*
[103] *Id.* The Court also analyzed the other three factors, nature of copyrighted works, the amount and substantiality of the use, and the effect of the use upon the potential market for or value of the copyrighted work. The district court ruled for Google for each of the three remaining factors. *Id.* at 1120–22.

material from being cached: he could have used a "no-archive" meta-tag, which would have informed Google not to display "cached" links to his page. Field did not do this.[104] Instead, the plaintiff "made a conscious decision to permit" Google to make available the cached links.[105]

Most significantly, the district court found that Google's copying and distributing of the plaintiff's copyrighted works by allowing users access to them through cached links constituted fair use.[106] Under the first fair use factor—the purpose and character of Google's use—the court observed that Google's system cache serves a different purpose from that of the plaintiff's original works. The plaintiff intended his copyrighted works to serve an artistic function.[107] But Google's cache enabled users to access content when the original page is inaccessible, detect changes in a particular website, and allow users to understand why a page was or was not responsive to their query.[108] "Because Google serves different and socially important purposes … and does not merely supersede the objectives of the original creations," the district court concluded that Google's alleged copying and distribution of the plaintiff's web pages was transformative.[109]

Assessing the three remaining fair use factors, the district court found that Google operated its cache system in good faith; that it provided a disclaimer and explanation; and that it followed industry protocols by not displaying "cached" links if the owner of a page opted out.[110]

Field established a blueprint for other copyright infringement actions relating to cached contents.[111] For instance, the district court in *Parker v. Yahoo!* adopted much of *Field*'s reasoning and held that Yahoo!, a search engine with similar caching practices, had an implied license to display a cached copy of the plaintiff's

---

[104] *Id.* at 1116.

[105] *Id.*

[106] *Id.* at 1118–23.

[107] *Id.* at 1118.

[108] *Id.* at 1118–19.

[109] *Id.* at 1119.

[110] *Id.* at 1123. The district court also held that Section 512(b) of the Digital Millenium Copyright Act—which provides a safe harbor to Internet service providers for "caching" activities—immunized Google from liability in damages. *Id.* at 1123–25.

[111] *See, e.g.*, Parker v. Google, Inc., 422 F. Supp. 2d 492 (E.D. Pa. 2006), *aff'd*, 242 F. App'x 833 (3d Cir. 2007), *cert. denied*, 552 U.S. 1156 (2008); Parker v. Yahoo!, Inc., No. Civ.A.07-2757, 2008 WL 4410095 (E.D. Penn. Sept. 25, 2008).

copyrighted works.[112] AI companies will likely invoke *Field* to argue that their "training" is simply caching by another name. As we discuss further below, however, AI large language models are a far cry from Google's caches.[113]

C. Licensing and Fair Use: American Geophysical Union v. Texaco, Inc.[114]

As courts expanded the power of licensors in copyright, they also incorporated licensing schemes into their concept of fair use. This trend is best embodied in *American Geophysical Union v. Texaco, Inc.*, where the Second Circuit expressly considered the availability of a licensing mechanism in rejecting a fair use defense.[115] In *Texaco*, academic researchers sued oil giant Texaco, alleging that the company routinely photocopied academic research publications for its scientists without permission from or compensation to the publishers.[116] The court's rejection of fair use in *Texaco* provides an important check on *Field*'s fair use approach.

In *Texaco*, the court questioned whether fair use could even apply to mechanical copying of an entire document.[117] Even under the traditional four-factor fair use test, however, the court rejected the defense. It found that Texaco's practice of making copies for individual scientists was "part of a systematic process of encouraging employee researchers to copy articles so as to multiply available copies while avoiding payment."[118] While the second factor—the nature of the copyrighted work—favored Texaco, the third and fourth factors tipped in favor of the plaintiff authors. The Court was particularly persuaded by the fact that Texaco copied each copyrighted work in its entirety. Texaco had attempted to reframe the question by noting that only a small fraction of any particular *volume* of the academic journal was copied—but the Court refused to take a macro view, given that "each article enjoys independent copyright protection."[119]

---

[112] *Parker*, 2008 WL 4410095, at *3–4.

[113] *See infra* Part VI.

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

[115] *Id.*

[116] *Id.* at 915 (noting that "all or most" of Texaco's 400–500 researchers "presumably photocopy" research materials). Although *Texaco* was brought as a class action, the court did not address class certification because the parties stipulated to a set of facts for a pre-certification determination of whether the photocopying constituted fair use. *Id.*

[117] *Id.* at 917.

[118] *Id.* at 920.

[119] *Id.* at 926.

Most importantly, the Court found in favor of the plaintiffs on the fourth factor—purportedly the "most important" fair use element—because of the availability of the Copyright Clearance Center, a private licensing organization that could negotiate a royalty in exchange for additional copies.[120] The Court noted that existence of "traditional, reasonable, or likely to be developed" licensing markets is an important factor in assessing the effect of an infringing activity on the market; if paying for a use is relatively easy, free use is, simply, "less fair."[121]

Some scholars have criticized *Texaco* for allowing the existence of a private licensing scheme to warp the traditional fair use analysis.[122] As discussed below, however, the doctrine of fair use must respond to current copyright culture. And within that culture, licensing organizations and agreements have become a routine and important way—perhaps the only feasible way—to mediate between copyright holders and digital innovators.

D. Digital Copying of Creative Works

Beginning early in the 2000s, technology enabled companies to duplicate creative works with increasing efficiency, opening the door to mass-scale digital duplication, resulting in individual as well as class suits seeking redress for copyright infringement by digital disruptors.

1. Digital Copying: New York Times Co. v. Tasini[123]

In *New York Times Co. v. Tasini*, the Supreme Court confronted the first wave of large-scale digital replication of copyrighted works. The plaintiffs in *Tasini* had published their works in print periodicals. The periodicals' publishers had, without consent, submitted the authors' pieces to an electronic database, making them retrievable by users as individual works, detached from the volume in which they first appeared. The defendant publishers argued that this was a "revision" of a work

[120] *Id.* at 930 (finding that the Copyright Clearance Center provided a "workable market for institutional users to obtain licenses").

[121] *Id.* at 931.

[122] *See, e.g.*, Mark A. Lemley, *Should a Licensing Market Require Licensing?*, 70 L. & CONTEMP. PROBS. 185, 190 (2007) (criticizing the "well-known circularity" of the relationship between a licensing market and the fourth fair use factor); Pamela Samuelson, *Unbundling Fair Uses*, 77 FORDHAM L. REV. 2537, 2585–86 (2009).

[123] N.Y. Times Co. v. Tasini, 533 U.S. 483 (2001).

authorized by copyright law. The Court rejected that argument.[124] Detaching these works from their context and dropping them into a massive database, the Court held, "effectively overrides the Authors' exclusive right to control the individual reproduction and distribution" of their work.[125]

The publishers offered several other defenses that recur in digital copying cases, none of which prevailed (and many of which were raised anew by subsequent digital disruptors). First, they claimed that the decontextualization of the copyrighted works could be solved technologically by enabling users to "manipulate the Databases" to retrieve all the works contained within the original volume.[126] The Court was unimpressed. Second, defendants argued that, at most, they could be contributory infringers; end users, rather than the Databases, would be committing the infringing acts. Again, the Court disagreed: "It is the copies themselves, without any manipulation by users," that violate the Copyright Act.[127] Finally, the defendants argued that disallowing inclusion of these works in electronic databases would, as the Court phrased it, "punch gaping holes in the electronic record of history."[128] The Court remained unmoved.[129] This, according to the Court, is what licensing agreements are made for.[130]

2. *Fair Use:* Google Book Project[131]

Although *Tasini* was a Supreme Court decision, it is overshadowed in cultural memory by the Second Circuit's adjudication of Google's book project. Google began with an ambitious (and ultimately unrealized) plan to create a comprehensive database of books—a quasi-mythical universal library.[132] In 2004, announcing the "Google Books Library Project," Google secured agreements with the New York Public Library, the Library of Congress, and several university libraries to digitally

---

[124] *Id.* at 502 (rejecting the defendants' claims that inclusion in the database was similar to the conversion of a work from newsprint to microfilm).

[125] *Id.* at 503–04.

[126] *Id.* at 504.

[127] *Id.* at 505.

[128] *Id.*

[129] *Id.* at 505–06 (declining to rule based on "speculation of future harms").

[130] *Id.*

[131] Authors Guild, Inc. v. Google, Inc., 804 F.3d 202 (2d Cir. 2015) [hereinafter *Google Book Project*].

[132] Scott Rosenberg, *How Google Book Search Got Lost*, WIRED (Apr. 11, 2017), https://www.wired.com/2017/04/how-google-book-search-got-lost/ [https://perma.cc/78JT-8LM9] ("In its youth, Google Books inspired the world with a vision of a library of utopia.").

copy their collections, and it scanned more than twenty million volumes in its quest to create a massive digital book repository.[133] Left out of this vision was a process for seeking permission from, or paying compensation to, copyright holders included in the scanned book collection.

The Authors Guild, Inc., an organization of published authors, as well as several individual authors, challenged the project in a class action. Google asserted that its unauthorized activities—copying, storing, distributing, and displaying the copyrighted book—were protected by the doctrine of fair use.[134] The parties participated in extensive negotiations and entered into a proposed settlement.[135] That settlement would have granted Google broad (though not exclusive) rights over the books in its electronic database, including the right to sell subscriptions to an e-book database and the right to sell online access to individual books, in exchange for paying 63% of the revenues to copyright holders.[136] Objectors contested the settlement on the ground that it did not comply with Rule 23; they also alleged that it violated copyright, privacy, and antitrust law, among others.[137] The district court rejected the proposed settlement.[138]

Google reduced the scope of its project so that users would only be able to see "snippets" of the copied texts in the public domain; what was originally a digital bookstore became a vast public index instead. The Second Circuit found that Google's now-much-narrowed search and snippet functions satisfy fair use.[139] It held that the first factor—the purpose and character of the use—favored Google.[140] Google's use of copyrighted works "highly transformative" because its digitizing technology enables "a search for identification of books containing a term of interest to the searcher" and permits "a searcher to identify those that contain a word or term of interest".[141] With the snippet views, searchers can only "read snippets from the book searched" but gain "important value to the

[133] *Google Book Project*, 954 F. Supp. 2d 282, 286 (S.D.N.Y. 2013), *aff'd*, 804 F.3d 202 (2d Cir. 2015).
[134] *Id.* at 288 (stating that from "the outset" Google's principal defense was fair use).
[135] *Id.*
[136] *Google Book Project*, 770 F. Supp. 2d 666, 671 (S.D.N.Y. 2011).
[137] *Id.* at 673–74.
[138] *Google Book Project*, 954 F. Supp. 2d at 288 (citing 770 F. Supp. 2d 666).
[139] *Google Book Project*, 804 F.3d at 218 (noting that "[s]nippet view thus adds importantly to the highly transformative purpose of identifying books of interest to the searcher").
[140] *Id.* (finding that the first factor "favors a finding of fair use").
[141] *Id.* at 217.

basic transformative search function, which tells only whether and how often the searched term appears in the book."[142] Google's tiny snippets show "the searcher just enough context surrounding the searched term to help her evaluate whether the book falls within the scope of her interest (without revealing so much as to threaten the author's copyright interests)."[143] In other words, Google's transformative use of the copyrighted works did not directly compete as substitutes for the plaintiffs' books.[144]

On the second factor, the nature of copyrighted works, the Second Circuit noted that "each of three" books scanned by Google is "factual" and concluded that nonfiction or fiction books are "not dispositive" in finding fair use.[145] The Circuit observed that "[n]othing in this case influences us one way or the other with respect to the second factor considered in isolation."[146] Considering *both* the first factor and second factor of fair use doctrine together, the Circuit stated, "the second factor favors fair use not because Plaintiffs' works are factual, but because the secondary use transformatively provides valuable information about the original, rather than replicating protected expression in a manner that provides a meaningful substitute for the original."[147] The Circuit, however, did not provide extensive explanation. This seems to suggest that authors of published nonfiction books may face a conclusory finding of fair use under the second factor, or that the second factor is no longer relevant in fair use analysis.

For the third factor, the amount and substantiality of the portion used, the Second Circuit found that "[w]hile Google makes an unauthorized digital copy of the entire book, it does not reveal that digital copy to the public. The copy is made to enable the search functions to reveal limited, important information about the books."[148] Accordingly, with respect to the search function, Google satisfied the third factor. As to the snippet view function, Google constructed the snippet feature in a "manner that substantially protects against its serving as an

[142] *Id.*
[143] *Id.* at 218.
[144] *Id.*
[145] *Id.* at 220.
[146] *Id.*
[147] *Id.*
[148] *Id.* at 221–22.

effectively competing substitute for Plaintiffs' books".[149] In fact, Google included many limitations on the snippet function:

> These include the small size of the snippets (normally one eighth of a page), the blacklisting of one snippet per page and of one page in every ten, the fact that no more than three snippets are shown—and no more than one per page—for each term searched, and the fact that the same snippets are shown for a searched term no matter how many times, or from how many different computers, the term is searched. In addition, Google does not provide snippet view for types of books, such as dictionaries and cookbooks, for which viewing a small segment is likely to satisfy the searcher's need. The result of these restrictions is, . . . , that a searcher cannot succeed, even after long extended effort to multiply what can be revealed, in revealing through a snippet search what could usefully serve as a competing substitute for the original.[150]
>
> . . . .
>
> The other restrictions built into the program work together to ensure that, even after protracted effort over a substantial period of time, only small and randomly scattered portions of a book will be accessible.[151]

With such limitations and restrictions imposed by Google in creating its search and snippet functions, Google satisfied the third factor for both the search and snippet functions.

Under the fourth factor, the effect of the use upon potential market or value, the Second Circuit ruled for Google because Google did not sell scanned books in competition with the plaintiffs. Specifically, the snippet views in the aggregate amount to "no more than 16% of a book."[152] The Circuit also found that a loss of sale of a book occurs "in relation to interests that are not protected by the copyright" because "a searcher's need for access to a copyrighted book will at times be because the snippet conveys a historical fact that the searcher needs to ascertain."[153] In

[149] *Id.* at 222.
[150] *Id.*
[151] *Id.*
[152] *Id.* at 224.
[153] *Id.*

summary, Google's making of a complete digital copy of the copyrighted works for the narrow and unchanging purpose of providing the public with its search and snippet view functions did not constitute copyright infringement.

Google, the defendant in both *Field* and *Google Book Project*, prevailed in its fair use defense in the second suit because both cached content and the snippet index involved limited, unchanging, and transformative use of existing copyrights. In neither case did the search company's output interfere with copyright holders' potential market of selling their books or negatively impact the value of copyrighted works. In the context of AI input class actions, by contrast, there are no firm guardrails on the use of the copied copyrighted works. There is also a strong profit motive to resist the formation of such guardrails. AI companies rely on existing literary text, images, and sounds to create AI models. Without human creative works to train AI models, there can be no AI models.[154] That creativity is the essence of what copyright law protects.

3. Fair Use of Code: Google LLC v. Oracle America, Inc.

In setting up its open-source Android platform, Google engineers wrote millions of lines of new code. But they also copied 11,500 lines of code from Java SE, a program ultimately owned by Oracle.[155] By using this code—which included tasks that were essential to using Java coding language at all—Google enabled "millions of programmers, familiar with Java, to be able easily to work with its new Android platform."[156] Oracle sued Google, arguing among other things that Google's use of the code violated copyright. The question before the Supreme Court was whether this copied code is subject to copyright protections (the trial court had found otherwise[157]), and—if yes—whether Google's actions were shielded as fair use.

---

[154] In fact, some are already warning that AI models are running up against limits on available content on which to train. *See, e.g.*, Nicola Jones, *The AI revolution Is Running out of Data. What Can Researchers Do?*, 636 NATURE 290 (2024), https://www.nature.com/articles/d41586-024-03990-2 [https://perma.cc/P7ND-JAUN].

[155] Google LLC v. Oracle Am., Inc., 593 U.S. 1 (2021).

[156] *Id.* at 9; *see also id.* at 9–14 (explaining the nature and use of the copied code).

[157] *Id.* at 14–15.

Reasonable minds could and did disagree about these issues. One jury deadlocked on the question of fair use;[158] following an appeal, a subsequent jury found Google's use of the code did constitute fair use.[159] On appeal again, the Federal Circuit reversed, finding "nothing fair about taking a copyrighted work verbatim and using it for the same purpose and function as the original in a competing platform."[160]

The Supreme Court, in an archetypal opinion by Justice Stephen Breyer, reversed the reversal.[161] It assumed, without deciding, that the code at issue was protected by copyright.[162] But it concluded that Google's use of the 11,500 lines of code *was* protected by fair use.[163] The purpose of the judge-made, equitable doctrine of fair use, the Court noted, is to "permit courts to avoid rigid application of the copyright statute when, on occasion, it would stifle the very creativity which that law is designed to foster."[164] In the Court's view, Google's use of the Java code merited that protection.

To a large extent, the Court's fair use holding reflects underlying skepticism toward the concept of granting copyright protection to computer code. Although it assumed for purposes of argument that code is copyrightable, it strongly suggested that—to the extent it is protected at all—code should enjoy much thinner protection than fiction, and even than less-protected non-fiction.[165] In the Court's view, the code at issue was a type of functional organizer—the equivalent of cabinets, drawers, and files.[166] On this basis, when analyzing the first fair use factor, the

[158] *Id.* at 14.

[159] *Id.* at 16.

[160] *Id.*

[161] *Id.* at 16 ("And the Circuit again reversed the District Court. … Google then filed a petition for certiorari[.]"); *id.* at 20 ("The second question asks us to determine whether Google's use of the API was a 'fair use.' … Unlike the Federal Circuit, we conclude that it was.").

[162] *Id.* at 20 ("We shall assume, but purely for argument's sake, that the entire Sun Java API falls within the definition of that which can be copyrighted. We shall ask instead whether Google's use of part of that API was a 'fair use.'").

[163] *Id.* at 33, 40.

[164] *Id.* at 18 (quoting Stewart v. Abend, 495 U.S. 207, 236 (1990)).

[165] *Id.* at 20–21; *see also id.* at 21 (discussing computer programs and stating that "[g]enerally speaking, computer programs differ from books, films, and many other 'literary works' in that such programs almost always serve functional purposes").

[166] *Id.* at 27.

nature of the copyrighted work, the Court ruled that the code was, "if copyrightable at all, further than are most computer programs … from the core of copyright."[167]

The Court also indicated strong support for Google's argument that its use of the Java code lines was transformative, thus tipping the second fair use factor in Google's favor. In the Court's view, Google's use of the code furthered the development of computer systems in the new and important realm of smartphone platforms.[168] The Court found that the third factor—the amount and substantiality of the portion used—also favored Google: the Court found that the copied code was only 0.4 percent of the total set of Java's code.[169] Finally, the Court found that Google's use of the code lines posed little threat to Oracle's market, because Oracle was "poorly positioned to succeed in the mobile market."[170] In other words, Google was not developing a market substitute for Oracle's products.[171] The Court also saw Oracle's attempt to shield this code as a potentially monopolistic move that would thwart innovation.[172] It ruled in favor of Google on the ground of fair use.

***

The cases involving RAM copies, cached contents, digital libraries of scanned books, and lines of computer code, illustrate the ways that copyright law has adapted in response to the computer revolution in the last five decades. These precedents show that AI companies, in "feeding" copyrighted works to their systems, were engaging in what copyright law defines as unlawful infringement. Yet that is not the end of the matter. In the era of networked computers,

[167] *Id.* at 29. The Court also cited with approval an amicus brief by the American Antitrust Institute warning that "Copyright on largely functional elements of software that [have] become an industry standard gives a copyright holder anti-competitive power." *Id.* at 32 (quoting Brief for Am. Antitrust Inst. as Amicus Curiae at 7).

[168] *Id.* at 31–32. In so holding, the Court relied on two Ninth Circuit cases finding fair use in situations where companies had reverse-engineered or copied code as an intermediate step in developing a new product. *Id.* at 22 (citing with approval Sony Comput. Ent., Inc. v. Connectix Corp., 203 F.3d 596, 603–08 (9th Cir. 2000) and Sega Enters. Ltd. v. Accolade Inc., 977 F.2d 1510, 1521–27 (9th Cir. 1992)). For a thoughtful analysis of *Oracle*, *Connectix* and *Sega*, see Justin Hughes, *The Sub Rosa Rules of Copyright Fair Use*, 64 Ariz. L. Rev. 1, 27–28 (2022).

[169] *Oracle*, 593 U.S. at 28.

[170] *Id.* at 31.

[171] *Id.* at 37.

[172] *Id.* at 39.

courts—squeamish about squelching innovation—appear to have become more open to findings of fair use. Where Big Tech enjoyed fair use protection, no asking for forgiveness was necessary. But the scope of fair use doctrine remains ambiguous, and fair use decisions such as *Field* and *Oracle* apply imperfectly, if at all, to the wholesale copying of thousands of copyrighted works of fiction and non-fiction.

This uncertainty lays the foundation for the current wave of class actions by creators against AI companies. The technology is new, but the questions are familiar: what constitutes infringement? What infringement is protected by fair use? With one notable recent exception, courts in the many pending cases have yet to answer.[173]

III
Generative AI and the First Wave of Litigation

The unexpected arrival of generative AI in 2022 as the world resumed normalcy after the COVID-19 pandemic sent a shockwave across the globe. Generative AI dominated headlines in late 2022 and early 2023. This new technology was lauded—and also criticized for trampling on copyright and privacy laws. This Section briefly describes the three primary types of cases in the current, first wave of litigation arising from generative AI.[174]

A. Copyright Infringement Actions

Generative AI systems rely on data to train models for the creation and generation of new works that mirror the training data. The quality of that data matters.[175] The phrase *garbage in, garbage out*, first uttered by IBM programmer

[173] Thomson Reuters Enter. Ctr. GmbH v. Ross Intel. Inc., No. 20-CV-00613-SB, 2025 WL 458520, *7 (D. Del. Feb. 11, 2025) (finding that defendant Ross Intelligence, an AI company, violated Thompson Reuter's copyright when Ross copied the headnotes from Thompson Reuter's Westlaw research platform for the purpose of training Ross's AI-driven legal research platform).

[174] For a helpful visual of American AI copyright cases, see Kate Knibbs, *Every AI Copyright Lawsuit in the U.S., Visualized*, Wired (Dec. 19, 2024), https://www.wired.com/story/ai-copyright-case-tracker/ [https://perma.cc/7EBF-BGYS].

[175] "In computer science, garbage in, garbage out (GIGO) is the concept that flawed, biased or poor quality ('garbage') information or input produces a result or output of similar ('garbage') quality." *Garbage in, garbage out*, Wikipedia, https://en.wikipedia.org/wiki/Garbage_in,_garbage_out [https://perma.cc/J4ZY-N3CF] (last visited Oct. 14, 2024).

and instructor George Fueschel,[176] applies squarely to generative AI.[177] Indeed, AI developers actively seek to avoid poor data in training their models. This has led AI companies to use copyrighted content to train models, without taking the time to seek permission from, or offer compensation to, copyright holders. This type of infringement falls in the "input" claims category.

In general, copyright refers to a bundle of rights, namely, the right to make copies, the right to prepare derivative works, the right to distribute the works, the right to publicly perform the works, and the right to publicly display the works.[178] Before the arrival of generative AI, most copyright infringement cases centered on unauthorized copying and distribution. In *American Geophysical Union v. Texaco, Inc.*, for example, the defendant's employees used a photocopy machine to make unauthorized copies of journal articles.[179]

The age of networked computers spurred the creation of new technologies for ease of copying and distribution, as seen in *A&M Records, Inc. v. Napster* where the defendant was accused of providing a central server that facilitated music file uploading and sharing among users without permission from copyright holders.[180] As a result, fewer copyright disputes rested on violation of the right to prepare derivative work, and if they did, the focus was on situations where a subsequent work was allegedly derived from the original.[181] Moreover, the development of well-recognized IP structures, such as the licensing model and movie right options,

[176] *Id.*; Rahul Awati, *Garbage in, Garbage out (GIGO)*, TechTarget, https://www.techtarget.com/searchsoftwarequality/definition/garbage-in-garbage-out [https://perma.cc/G6K2-F83P] (last visited Oct. 31, 2023).

[177] *E.g.*, Heather Rodriguez, *Garbage In, Garbage Out: The Potential Pitfalls of Artificial Intelligence*, Tex. A&M Univ. Coll. of Arts & Sci. (Jan. 19, 2023), https://artsci.tamu.edu/news/2023/01/garbage-in-garbage-out-the-potential-pitfalls-of-artificial-intelligence.html [https://perma.cc/ZT4K-HJFL]; Brooks Hanson et al., *Garbage in Garbage Out: Mitigating Risks and Maximizing Benefits of AI in Research*, 623 Nature 28 (2023), https://www.nature.com/articles/d41586-023-03316-8 [https://perma.cc/K5DP-LGG2].

[178] 17 U.S.C. § 106.

[179] Am. Geophysical Union v. Texaco, Inc., 60 F.3d 913, 914 (2d Cir. 1994) (holding that copying journal articles for private corporate library use is infringement, not fair use).

[180] A&M Recs., Inc. v. Napster, Inc., 239 F.3d 1004, 1011 (9th Cir. 2001).

[181] Brett Snider, *Is It Illegal to Remix Music*, FindLaw (Aug. 31, 2013), https://www.findlaw.com/legalblogs/law-and-life/is-it-illegal-to-remix-music/ [https://perma.cc/3JE8-Y6UC] ("The main legal issue with remixes is that they are derivative works, meaning that they are derived mostly from other artists' songs. In theory, this isn't an issue. But almost all music released for consumers has copyright protection, preventing remixers from making their own version of certain songs without permission."); Jay T. Westermeier,

helped guard against infringing derivative works. For instance, J.K. Rowling sold both movie and merchandise rights based on her Harry Potter series.[182]

In the AI era, copyright disputes have taken a darker turn. Pending copyright class actions zero in on the violation of both the right to make copies and the right to prepare derivative works. In *Getty Images v. Stability AI*, for example, Getty Images alleges that Stability AI copied more than 12 million photographs without permission for compensation, with the purpose of "build[ing] a competing business."[183] In the course of so doing, Getty Images alleges, Stability AI also either removed or altered Getty Images' copyright management information and infringed on its trademarks.[184] Now stuffed to the gills with rich and complex images created by human artists, Stability AI dazzles the public with its generative AI models, with which users can generate stunning AI images in seconds. In sum, unlike the copyright infringement actions of prior decades, AI copyright suits seek to protect creators' rights to authorial creative integrity against a pending torrent of copycat creativity.

Two unsettled doctrinal questions dominate the pending copyright suits. The first is whether the AI companies' actions constitute "copying" for purposes of copyright law. We argue above that, in light of a series of cases applying copyright law to emerging technologies, it does.[185]

The second question is whether AI developers' use of copyrighted works is protected by fair use. Courts, including the Supreme Court, have struggled to apply the traditional four fair use factors in this new context.[186] In *Oracle*, the Supreme

Understanding the Importance of Derivative Works, FINNEGAN (Mar. 2009), https://www.finnegan.com/en/insights/articles/understanding-the-importance-of-derivative-works.html [https://perma.cc/9QBG-PE29].

[182] Michael Cieply, *Warner and J.K. Rowling Reach Wide-Ranging Deal*, N.Y. TIMES (Sept. 12, 2013), https://www.nytimes.com/2013/09/13/business/media/warner-jk-rowling-partnership-will-include-new-wizardry-film.html [https://perma.cc/5CR4-X7AR]; Nathan Reiff, *NBCUniversal Buys Harry Potter Rights in Enormous Franchise Deal (NBCU,TWX,DWA)*, INVESTOPEDIA (June 25, 2019), https://www.investopedia.com/news/nbcuniversal-buys-harry-potter-rights-enormous-franchise-deal-nbcutwxdwa/ [https://perma.cc/KS7E-5K2T].

[183] Amended Complaint at 1, Getty Images (US), Inc. v. Stability AI, Inc., No. 23-CV-00135-GBW (D. Del. Mar. 29, 2023), https://storage.courtlistener.com/recap/gov.uscourts.ded.81407/gov.uscourts.ded.81407.13.0.pdf [https://perma.cc/B4YG-6DHD].

[184] *Id.*

[185] *See supra* Part II.

[186] The fair use factors are:

Court held that fair use protected Google's use of 11,500 lines of Oracle's Java code.[187] But neither *Oracle* nor other cases address fair use when creative works are being copied and made available to the public wholesale, nor when the purpose of this copying is to enable the instantaneous creation of limitless derivative works that will directly compete against the works that were copied.

The pending AI copyright suits, which are seeking substantial money damages for past infringement,[188] leverage the utilitarian and incentivization theories underlying copyright infringement claims.[189] These theories, however, understate the core value of copyright to society.[190] The heart of the copyright violations by AI companies is the complete disregard for creators' originality, manifested as unique voices, styles, signatures, manners, and genres.[191] Rich, diverse works of authorship enhance the progress of humankind. Encouraging

---

(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.

17 U.S.C. § 107.

[187] Google LLC v. Oracle Am. Inc., 593 U.S. 1 (2021).

[188] *See* Class Action Complaint at 47, Authors Guild v. OpenAI, Inc., No. 23-CV-08292 (S.D.N.Y. Sept. 19, 2023); Class Action Complaint at 43, Andersen v. Stability AI Ltd., No. 23-CV-00201 (N.D. Cal. Jan. 13, 2023); First Amended Complaint at 64–65, Doe v. GitHub, Inc., No. 22-CV-06823-JST (N.D. Cal. July 21, 2023); Complaint at 35–36, Getty Images (US), Inc. v. Stability AI, Inc., No. 23-CV-00135-UNA (D. Del. Feb. 3, 2023); Complaint at 58, Concord Music Grp., Inc. v. Anthropic PBC, No. 23-CV-01092 (M.D. Tenn. Oct. 18, 2023); Complaint at 67, N.Y. Times Co. v. Microsoft Corp., No. 23-CV-11195 (S.D.N.Y. Dec. 27, 2023). These complaints note how lucrative generative AI deals have delivered for the defendants, in particular, for OpenAI.

[189] *See, e.g.*, Mei-lan Stark, *5 Ways Copyright Laws Encourage Personal Expression and Creativity*, U.S. Chamber of Com. (Apr. 25, 2022), https://www.uschamber.com/intellectual-property/five-ways-copyright-laws-encourage-personal-expression-and-creativity [https://perma.cc/XFB7-X8KX]; Kristelia Garcia, *Monetizing Infringement*, 54 U.C. Davis L. Rev. 265 (2020); Shyamkrishna Balganesh, *Foreseeability and Copyright Incentives*, 122 Harv. L. Rev. 1569 (2009).

[190] U.S. Const. art. I, § 8, cl. 8 ("[t]o promote the Progress of Science and useful Arts"); *see also* Deepak Somaya & Lav R. Varshney, *Ownership Dilemmas in an Age of Creative Machines*, 36 Issues 79, 79–80 (2020), https://issues.org/ownership-dilemmas-in-an-age-of-creative-machines/ [https://perma.cc/4EXQ-MPLV].

[191] *See* Judy Estrin, *The Case Against AI Everything, Everywhere, All at Once*, Time (Aug. 11, 2023), https://time.com/6302761/ai-risks-autonomy/ [https://perma.cc/6CX5-3W9P] ("Artificial Intelligence is not just chat bots, but a broad field of study. One implementation capturing today's attention, machine learning, has expanded beyond *predicting* our behavior to generating content—called Generative AI. The awe of machines

that progress is an explicit purpose of the Intellectual Property Clause of the Constitution.

B. Claims Under the Digital Millenium Copyright Act

Open-source software developers are also in the first wave of generative AI class actions. In *Doe v. GitHub, Inc.*,[192] anonymous members of a proposed class of developers allege that AI companies have taken billions of lines of open-source code, without attribution or permission, to train their generative models, thus violating traditional copyright law as well as the Digital Millenium Copyright Act of 1998 (DMCA).[193] This type of claims is also in the "input" claims category.

Computer software developers spend significant time and creativity drafting software code.[194] Coders have adopted open-source licensing regimes to facilitate collaborative efficiency and creativity within the copyright law framework. For more than three decades, coders worldwide have followed the requirements of open-source license agreements, including GPL, MIT's Open-Source License.[195]

---

wielding the power of language is seductive, but *Performative* AI might be a more appropriate name, as it leans toward production and mimicry—and sometimes fakery—over deep creativity, accuracy, or empathy.").

[192] First Amended Complaint at 21–23, Doe v. GitHub, Inc., No. 22-CV-06823-JST (N.D. Cal. July 21, 2023).

[193] Digital Millenium Copyright Act, Pub. L. No. 105-304, 112 Stat. 2860 (1998).

[194] Christian Owens, *Move Over Actors and Artists, Software Developers Are True Creative Heroes*, Forbes (Feb. 14, 2019), https://www.forbes.com/sites/christianowens/2019/02/14/move-over-actors-and-artists-software-developers-are-true-creative-heroes/ [https://perma.cc/DY3X-NRBD]; Ralph D. Clifford et al., *Answering Question One in* Google v. Oracle: *The Creativity of Computer Programmers*, 70 J. Copyright Soc'y 127 (2023), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4267334 [https://perma.cc/4N7R-L447].

[195] For instance, MIT Open Source License Provides:

> Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:
>
> The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.
>
> THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY,

In these agreements, the copyright management information (CMI) such as the copyrighted work's title, copyright registration number, the copyright owner's name, the creator's name, and terms and conditions for use of the work, are included as part of the licensed materials.[196] Under these agreements, billions of lines of code have become publicly accessible. Developers post code in so-called Git depositaries—hence the name GitHub. The open-source software movement fueled the rapid development of the World Wide Web, among countless other projects.

Big Tech companies, notably Microsoft, in teaming up with startup OpenAI, are releasing generative AI systems that are trained on billions of lines of code taken from these publicly accessible code depositaries. But coders allege that the companies have exploited this code without complying with the open-source license terms.[197] In fact, according to the class action complaint brought by coders, Big Tech companies removed or altered the copyright management information (CMI) from the licensed codes, cleansing them of copyright identity and protection.[198] The cleansed codes are used as training data for generative AI models. The output of the training data is new codes that are either copies of the original codes or derivatives thereof. Though AI companies have not made their codes public, the output exposes the scope and the damages to coders and authors alike. The output mirroring the training data demonstrates that when it comes to your children, you can immediately identify them because you were the one who gave birth to them. The coders seek to bring their grievances as a class.

WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

The MIT License, Open Source Initiative, https://opensource.org/license/mit/ [https://perma.cc/2JUV-NACW] (last visited Oct. 14, 2024).

[196] *Copyright Management Information (CMI)*, Copyright All., https://copyrightalliance.org/education/copyright-law-explained/the-digital-millennium-copyright-act-dmca/copyright-management-information/ [https://perma.cc/R5D2-76FK] (last visited Oct. 14, 2024) ("The DMCA includes provisions that protect the integrity of copyright management information. Copyright management information, or CMI, is information about a copyrighted work, its creator, its owner, or use of the work that is conveyed in connection with a copyrighted work. For example, CMI would include the copyrighted work's title, ISBN number or copyright registration number; the copyright owner's name; the creator's name; and terms and conditions for use of the work.").

[197] First Amended Complaint at 54–55, Doe v. GitHub, Inc., No. 22-CV-06823-JST (N.D. Cal. July 21, 2023).

[198] *Id.* at 34–35.

C. Data Privacy Violations

The third type of AI class action challenges the unlawful taking of private information, including personally identifiable information, from hundreds of millions of internet users, including children of different age groups, across platforms and sites, without their informed consent or knowledge. The allegations in *J.L. v. Google LLC* and *A.T. v. OpenAI* exhibit this "move fast and break things" attitude.[199] Indeed, Google, the defendant in one of the two data privacy class actions, has repeatedly engaged in the strategy of "Do First, Ask Forgiveness Later" that entangled the company in legal troubles with both regulators and rights holders.[200] The data privacy claims, however, is not in the copyright's "input" claims category.

In *J.L.*, eight plaintiffs filed a proposed class action against Google regarding its generative AI system Bard.[201] The plaintiffs alleged that Bard scraped personal data and property for its training data. Google moved to dismiss the complaint based on a wide range of procedural and substantive defects.[202] Google asserted that the use of the data is necessary to train Bard and the lawsuit would "take a sledgehammer not just to Google's services but to the very idea of generative AI."[203] Google's response reflects the current arms race among Big Tech companies in their race to develop and train what will become the dominant generative AI models.

---

[199] Dana Kanze, Mark A. Conley & E. Tory Higgins, *Research: Organizations That Move Fast Really Do Break Things*, Harv. Bus. Rev. (Feb. 12, 2020), https://hbr.org/2020/02/research-organizations-that-move-fast-really-do-break-things [https://perma.cc/XZ8V-HFS8] (observing that "Silicon Valley has long been known for its 'ask forgiveness, not permission' and 'move fast and break things' attitudes, but lately it's had to reckon with the consequences of that mindset. Examples from Uber to WeWork suggest that this modus operandi exposes fast-growing firms to significant risks.").

[200] Erik Sherman, *Google Struggles with Its "Do First, Ask Forgiveness Later" Strategy*, CBS News (Mar. 12, 2010), https://www.cbsnews.com/news/google-struggles-with-its-do-first-ask-forgiveness-later-strategy/ [https://perma.cc/8SPA-6BFY] (naming many examples of Google undertaking business initiatives without asking permission involving copyrights and private data throughout its existence that got the company in trouble repeatedly).

[201] Class Action Complaint, J.L. v. Alphabet Inc., No. 23-CV-03440-LB (N.D. Cal. July 11, 2023), https://www.classaction.org/media/jl-et-al-v-alphabet-inc-et-al.pdf [https://perma.cc/CG9J-5UZH].

[202] Blake Brittain, *Google Says Data-Scraping Lawsuit Would Take 'Sledgehammer' to Generative AI*, Reuters (Oct. 17, 2023), https://www.reuters.com/legal/litigation/google-says-data-scraping-lawsuit-would-take-sledgehammer-generative-ai-2023-10-17/ [https://perma.cc/8CCB-26MC].

[203] *Id.*

In another class action, *A.T. v. OpenAI*, the complaint alleges that OpenAI, together with Microsoft, has continually released new versions of ChatGPT, Dall-E, and Vall-E since the first version was released, jolting and entrancing the public with "human-like" generative AI capability.[204] Subsequently, with ChatGPT 4.0, for instance, OpenAI charges a subscription fee of $20 per month per user. OpenAI is amassing billions of dollars from both its alliance with Microsoft and charging high user fees. According to the complaint filed in the class action, to develop ChatGPT, Dall-E, and Vall-E or generative AI products, OpenAI turns to the internet to steal private information, including personal data from millions of unsuspecting consumers worldwide. This stolen data, the complaint alleges, includes private information and private conversations, medical data, information about children, and other type of data OpenAI can scrape off the internet to feed its large language models and deep language algorithms.

In addition, the complaint asserts, OpenAI takes private and personal information from its 100 million registered users, without disclosing that to them. This information includes all input details, account information users enter for registration, names, contact details, login credentials, emails, payment information for paid users, transaction records, identifying data harvested from users' devices and browsers, social media information, chat log data, usage data, analytics, cookies, keystrokes, typed searches, and other online activity data.[205] Most significantly, OpenAI scrapes users' digital footprints accumulated in real-time and as far back as fifteen years ago.

The allegations paint a stark picture that the massive personal data gives OpenAI sufficient information for its generative AI products to be created and trained to have the ability to "create our digital clones, including the ability to replicate our voice and likeness and predict and manipulate our next move," "misappropriate our skill sets and encourage our own professional obsolescence."[206]

***

[204] Class Action Complaint at 2–3, A.T. v. OpenAI LP, No. 23-CV-04557 (N.D. Cal. Sept. 5, 2023), https://fingfx.thomsonreuters.com/gfx/legaldocs/xmpjlnldzpr/OPENAI%20PRIVACY%20LAWSUIT%20complaint.pdf [https://perma.cc/S8HD-X64F].

[205] *Id.* at 6.

[206] *Id.* at 7.

AI models pose a threat to privacy and human creativity. In October 2023, President Biden signed an executive order directing agencies to develop "New Standards for AI Safety and Security."[207] But that directive is unlikely to provide compensation to those creators whose copyrights have already been disrespected, or to those people whose privacy has been violated. Especially given the pace and scale of AI companies' efforts, litigation is the best option for copyright holders—including everyday people—to shelter their creative works and personal data from this new force. The first wave of AI class actions seeks to do just that.

IV

Class Certification in Pre-AI Copyright Suits

Class actions, governed in federal courts by Rule 23, are a well-established (if sometimes controversial) vehicle for groups of individuals alleging similar harm to seek redress for that harm in a single suit led by representative plaintiffs.[208] The quintessential class action is for so-called "negative-value" claims, where—in the absence of a class—individuals would not have the incentive or resources to vindicate their legal rights in court.[209] AI class actions brought by creators whose works have been used to "train" large language models fit this framework perfectly. Yet class actions are high-stakes, uncertain endeavors. The very things that make class actions so valuable—their flexibility, their potentially broad preclusive effect, and their ability to generate productive settlement discussions—also make them risky for everyone involved.

---

[207] *Fact Sheet: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence*, White House (Oct. 30, 2023), https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/ [https://perma.cc/X28T-Q78K].

[208] *See* 7A Charles Alan Wright & Arthur R. Miller, Federal Practice and Procedure § 1751 (4th ed. 2024) ("The obvious advantage of the representative suit was that it was far cheaper and more convenient to maintain a single proceeding in equity than to adjudicate the controversy in piecemeal fashion by multiple actions at law.").

[209] *See, e.g.*, Linda Sandstrom Simard, *A View from Within the Fortune 500: An Empirical Study of Negative Value Class Actions*, 47 Ind. L. Rev. 739, 740 (2023) ("By aggregating groups of small value claims together, the cost of litigation is shared . . . , thus making litigation more feasible for claims that would otherwise never see the light of a courtroom.").

Although group litigation has ancient roots, the modern class action dates to 1966, when rulemakers promulgated what is now Rule 23.[210] From the outset, many proponents of Rule 23 conceived of it as a quasi-regulatory device, resolving concrete disputes in a way that could provide closure to large numbers of people, many of whom would lack the resources to seek individual vindication of their rights.[211] Today, civil rights, consumer, and tort litigants argue that class actions are "an important substitute for, or addition to, public administration."[212] In contrast to this quasi-regulatory conception, the defense bar has sought to depict Rule 23 as a mere joinder device, albeit one that, in their view, is being exploited by rapacious plaintiffs' lawyers. David Marcus refers to this formalist conception as an "adjectival" view of Rule 23—one in which the procedure is subordinate to, and should not distort, the substantive law.[213]

Inevitably, these theoretical tensions have played out in copyright class actions.[214] Even prior to the Roberts Court's procedural revival[215]—which has created heightened barriers to class certification—courts evinced skepticism toward copyright class actions. One court described a proposed copyright class unflatteringly as a "Frankenstein monster."[216] Perhaps as a result, until recently, there were relatively few such suits, with the *Google Book Project* and *Spotify* cases likely the only ones to receive mainstream media coverage.[217]

---

[210] David Marcus, *The History of the Modern Class Action, Part I:* Sturm und Drang, *1953–1980*, 90 Wash. U. L. Rev. 587, 588 (2013).

[211] *See, e.g.*, *id.* at 590.

[212] *Id.*; *see also* Robert L. Carter, *The Federal Rules of Civil Procedure as a Vindicator of Civil Rights*, 137 U. Penn. L. Rev. 2179, 2185 (1989) (class actions are "closely associated with the figure of the private attorney general").

[213] *Id.*

[214] *See, e.g.*, Samuelson, *The Google Book Settlement*, *supra* note 12 (arguing, based on the *Google Book Project* case, that "courts should engage in heightened scrutiny of the certifiability of a settlement class when the settlement would, in effect, achieve legislative outcomes"); C. Scott Hemphill, *Collusive and Exclusive Settlements of Intellectual Property Litigation*, 2010 Colum. Bus. L. Rev. 685, 690–91 (2010) (arguing that the *Google Book Project* settlement "made brilliant use of the class action mechanism").

[215] *See* Howard M. Wasserman, *The Roberts Court and the Civil Procedure Revival*, 31 Rev. Litig. 311, 312–13 (2012).

[216] Football Ass'n Premier League, Ltd. v. YouTube, Inc., 297 F.R.D. 64, 64 (S.D.N.Y. 2013) (quoting Eisen v. Carlisle & Jacquelin, 391 F.2d 555, 572 (2d Cir. 1968)).

[217] *See* Tang, *The Class Action*, *supra* note 5, at 1645 (noting that "few copyright class actions were filed between 1938, when Rule 23 was promulgated, and 1990"). The earliest known copyright class was certified in 1938. *See* Buck v. Russo, 25. F. Supp. 317 (D. Mass. 1938) (suit brought by President of the American Society of Composers, Authors & Publishers alleging infringement of copyrighted musical compositions).

Yet over time many groups of creators did seek to bring class actions to vindicate their copyright claims. As described below, plaintiffs in earlier suits invoked different categories of Rule 23, sought different forms of relief, and met with varying degrees of success. Among those that failed, some founded on substantive grounds (such as a fair use defense) unrelated to class certification, while others were deemed not to have met the requirements of Rule 23. Many of these decisions were unpublished, which limited their precedential value.[218]

Nevertheless, these pre-AI copyright class actions provide insight into the applicability of Rule 23 in today's AI suits. In general, when courts concluded that the heart of a claim was a defendant's allegedly infringing conduct, they tended to find Rule 23 satisfied. Conversely, courts focusing on the individualized nature of copyrights, or on individualized defenses such as notice or fair use, tended to find class resolution inappropriate.[219] Predictably, then, plaintiffs' theory of liability was rooted in the defendants' conduct: the class argued that the defendant acted according to a uniform policy or practice, often with a dollop (or more) of bad faith. Conversely, copyright class defendants emphasized the scattered nature of the plaintiff copyright holders; they told a story of uniqueness, messy layers of contracts, and the need for individualized determinations of copyright ownership and fair use.[220] Details of Rule 23 aside, certification appears to depend on which of these conflicting narratives—and class action theories—prevails.

This Section describes how courts have applied the certification requirements of Rule 23(a) and (b) in pre-AI copyright class actions. It also explains the specific way that courts have applied Rule 23(e)—which governs approval of class action settlements—in the context of copyright. To approve a class action settlement, a court must find that the proposed class satisfies all of these provisions.

A. Copyright and Class Action Rule 23(a)

To obtain class certification—in order to litigate, or to obtain judicial approval of a settlement—a party bears the burden to satisfy the Rule 23 criteria. First,

---

[218] *See infra* Part IV.

[219] *See, e.g.*, Ryan v. CARL Corp., No. C97-3873, 1999 WL 16320, at *6 (N.D. Cal. Jan. 13, 1999) ("[T]he trend appears to be to deny certification if the plaintiffs are suing a 'habitual infringer' but have little else in common.").

[220] *See, e.g.*, Football Ass'n Premier League, Ltd. v. YouTube, Inc., 297 F.R.D. 64, 65 (S.D.N.Y. 2013) ("Generally speaking, copyright claims are poor candidates for class-action treatment.").

the proposed class representative must demonstrate that the claims meet the four requirements of Rule 23(a), which are referred to in shorthand as numerosity, commonality, typicality, and adequacy of representation.[221] Taken collectively, these four factors are intended to ensure that a proposed class is sufficiently large and yet sufficiently cohesive to litigate as a collective, through the proxies of class representatives.

Prior to the Roberts Court's heightened scrutiny of class actions, 23(a) was not considered particularly onerous. This shoot-from-the-hip view of Rule 23(a) applied in early copyright class actions as it did more generally. For example, in *Bernstein v. Universal Pictures, Inc.*, the district court provisionally certified a class of lyricists and composers in a suit alleging that movie and television producers used adhesive contracts to constrain plaintiffs from entering the industry.[222] The Second Circuit ultimately rejected the suit on its merits, but the class certification appears to have been decided in a cursory preliminary order. Another court certified a class of composers alleging infringement by cable company Showtime/The Movie Channel in five breezy paragraphs (at least, breezy by civil procedure standards).[223]

Later decisions, however, gave more attention to 23(a) factors:

23(a)(1): To meet the numerosity requirement, a party seeking class certification must show that the class "is so numerous that joinder of all members is impracticable."[224] There is no numerical cutoff for a class; this finding is within the discretion of the court. Class sizes smaller than 50 people are unusual, but not unheard of.[225] Copyright class actions have not typically struggled to meet the numerosity threshold, although in one early case, a court denied certification in a one-page order because the plaintiffs had failed to demonstrate the existence of a

[221] *See* Fed. R. Civ. P. 23(a); *see, e.g.*, Wal-Mart Stores, Inc. v. Dukes, 564 U.S. 338, 345 (2011) (listing Rule 23(a) factors).

[222] Bernstein v. Universal Pictures, Inc., 379 F. Supp. 933, 934 (S.D.N.Y. 1974), *rev'd*, 517 F.2d 976 (2d Cir. 1975).

[223] David v. Showtime/The Movie Channel, Inc., 697 F. Supp. 752, 756–57 (S.D.N.Y. 1988).

[224] Fed. R. Civ. P. 23(a)(1).

[225] *See* Wright & Miller, *supra* note 208, § 1762; *see also* 28 U.S.C. § 1332(d)(5)(B) (granting diversity jurisdiction over certain class actions with a class size of at least 100).

sufficient number of freelance photographers affected by the defendant's alleged infringement.[226]

23(a)(2): This factor, which requires a showing of "questions of law or fact common to the class,"[227] was long considered a minimal burden—a view that is reflected in early copyright class actions.[228] In *Ryan v. CARL Corporation*, for example, the plaintiff class argued that the defendant's document retrieval service, which functioned "like a private interlibrary loan service," had provided over 700,000 copies of plaintiffs' copyrighted publications to users of its service, making payments for only 461 of them.[229] In affirming commonality, the court noted that "plaintiffs are essentially challenging a standard business procedure" with a "common core of salient facts."[230] The same logic applied in *American Geophysical Union v. Texaco, Inc.*, although that case settled without a judicial analysis of Rule 23, after courts in the Second Circuit ruled against Texaco's fiar use defense,[231] Texaco's activity of making copies of academic articles for its employees was uniform in its purpose and effect.

Since 2011, however, commonality has been a higher hurdle. That year, the Supreme Court decided *Wal-Mart v. Dukes*, a Title VII class action in which over a million employees claimed gender discrimination, the Court emphasized that not any common questions would do; claims must depend on a "common contention . . . that is capable of classwide resolution—which means that

---

[226] *See* Resnick v. Copyright Clearance Ctr., Inc., No. CIV.A.01-11520-RWZ, 2003 WL 22176619, at *1 (D. Mass. Sept. 22, 2003).

[227] FED. R. CIV. P. 23(a)(2).

[228] *See, e.g.*, Utopia Ent., Inc. v. Claiborne Par., No. CIV.A.03-1355, 2006 WL 8435006, at *4 (W.D. La. Jan. 10, 2006) (noting that defendants conceded commonality), *R. & R. adopted*, 2006 WL 548476 (W.D. La. Mar. 6, 2006); *Showtime*, 697 F. Supp. at 757 (finding that "[i]ncidental differences do not defeat commonality").

[229] Ryan v. CARL Corp., No. C97-3873, 1999 WL 16320, at *1 (N.D. Cal. Jan. 13, 1999).

[230] *Id.* at *6–7.

[231] Am. Geophysical Union v. Texaco, Inc. 802 F. Supp. 1, 4 (S.D.N.Y. 1992) (describing "[t]his class action"), *aff'd*, 60 F.3d 916 (2d Cir. 1994), *cert. dismissed*, 516 U.S. 1005 (1995). Although there is no Rule 23 certification analysis in the published opinions, the settlement agreement reached required approval by the court, a requirement for all class actions, suggesting that the case was treated as a certified class. *See Texaco, Publishers Agree to Settle Copyright Case (ARL 180)*, STANFORD LIBRS., https://fairuse.stanford.edu/texaco/settlement-arl-180/ [https://perma.cc/9SXF-5F6Q] (last visited Mar. 30, 2025) (noting that "the settlement is subject to the approval of the entire group of publishers and the court").

determination of its truth or falsity will resolve an issue that is central to the validity of each one of the claims in one stroke."[232]

Unsurprisingly, *Wal-Mart*'s more stringent approach affected proposed copyright classes. For example, in *Wu v. Pearson Education, Inc.*, the district court initially certified a class of copyright owners who alleged that Pearson had printed educational materials containing the owners' work product in numbers in excess of the amount for which Pearson had paid.[233] Subsequently the case was transferred to a different district court judge, who decertified the class, in part based on a finding that "layers of individually negotiated contractual arrangements" defeated commonality.[234] Similarly, the district court in *Football Association Premier League, Ltd. v. YouTube* found no commonality in a suit by a "worldwide class" of copyright owners against YouTube, given the need for individual determinations of the validity of a copyright, notice to YouTube of infringement, fair use, damages, and other specific defenses.[235] Commonality continues to pose a significant challenge to class plaintiffs, including those in pending copyright class actions.

23(a)(3): This provision asks whether "the claims or defenses of the representative parties are typical of the claims or defenses of the class."[236] Although there is some overlap between this provision and both (a)(2) and (a)(4),[237] the focus of the typicality factor is on the congruency between the claim of the named plaintiff(s) and those absent class members: it requires that the claims of the class representatives are fundamentally aligned with those of the absentees.[238]

Analysis of typicality in copyright class actions has often been cursory. In one early suit by a proposed *plaintiff* class of music publishers against a proposed *defendant* class of producers and distributors, the court held that the single named plaintiff lacked Article III standing to make claims against defendants other than

[232] Wal-Mart Stores, Inc. v. Dukes, 564 U.S. 338, 350 (2011).
[233] Wu v. Pearson Educ., Inc., No. 09 CIV. 6557, 2012 WL 6681701, at *1 (S.D.N.Y. Dec. 21, 2012).
[234] *Id.* at *7.
[235] Football Ass'n Premier League, Ltd. v. YouTube, Inc., 297 F.R.D. 64, 66 (S.D.N.Y. 2013).
[236] Fed. R. Civ. P. 23(a)(3).
[237] *See, e.g.*, Brown v. Kelly, 609 F.3d 467, 475 (2d Cir. 2010) (noting that "the commonality and typicality requirements often tend to merge into one another").
[238] *See* Wright & Miller, *supra* note 208, § 1764.

the one that had allegedly reproduced his copyrighted music.[239] For that reason, the court held that the plaintiff also failed to show typicality.[240]

But typicality may also be a sticking point in plaintiff-side class actions. For example, the court in *Football Association* minced no words in rejecting certification based on typicality as well as commonality.[241] The class claims in that case were dauntingly broad. The proposed class included copyright holders worldwide who fell into either of two categories: (1) YouTube had previously blocked infringement of their copyrights on its site, but there were subsequent infringing uploads; and/or (2) YouTube had allowed copyrighted musical compositions to be used without authorization.[242] The court noted that, to establish liability, each class member of this proposed worldwide class would be required to show that YouTube had legal knowledge of the unlawful conduct.[243] As to each allegedly unlawful download, moreover, YouTube would have a potential fair use defense.[244] "The unique nature of each work and of its infringement," the court stated, "cannot be obliterated by its inclusion in a sea of other claims."[245] Other courts have also evinced skepticism of typicality in broadly framed class claims.[246]

23(a)(4): The final, and most important, element of 23(a) is adequacy of representation. This provision encapsulates the need to ensure due process before allowing representative parties to litigate the rights of those who are absent.[247]

---

[239] Angel Music, Inc. v. ABC Sports, Inc., 112 F.R.D. 70, 73 (S.D.N.Y. 1986).

[240] *Id.*; *see also* Estate of Berlin v. Stash Recs., Inc., No. 95 Civ. 6575, 1996 WL 374176, at *2 (S.D.N.Y. July 2, 1996) (finding that "other class members have not been injured by the same course of conduct").

[241] Football Ass'n Premier League, Ltd. v. YouTube, Inc., 297 F.R.D. 64, 64 (S.D.N.Y. 2013) (calling the case "a Frankenstein monster posing as a class action").

[242] *Id.* at 65.

[243] *Id.*

[244] *Id.*

[245] *Id.*

[246] *See, e.g.*, Schneider v. YouTube, Inc., 674 F. Supp. 3d 704, 727 (N.D. Cal. 2023) (describing typicality as "questionable" where "[e]ach work-in-suit will be the subject of individualized proof of ownership and infringement"); *see also* Blackman v. Teespring, Inc., No. 19-CV-01494-RS, 2019 WL 7832600, at *2 (N.D. Cal. Jul. 12, 2019) (granting motion to sever joined copyright claims under Rule 21 because claims would need to be resolved upon individualized facts).

[247] WRIGHT & MILLER, *supra* note 208, § 1765 (noting that adequacy of representation is mandatory because "[t]he binding effect of all class-action decrees raises substantial due-process questions").

Adequacy contemplates two separate inquiries. The first—and the one most pressing in copyright class actions—is whether the named class representatives will fairly and adequately represent the class. To protect this fundamental requirement, courts must ensure there are no significant conflicts of interest between the named representatives and absent class members. In *Amchem Products, Inc. v. Windsor*, the Supreme Court rejected a massive asbestos-related class settlement on the ground that the class representatives, who had current asbestos-related injuries, were structurally inadequate to represent class members whose injuries might manifest in a more-distant future.[248] Although the Court pointed to specific examples that gave rise to its concerns—such as the settlement's failure to adjust for future inflation—its holding was premised not only on the terms of the settlement, but the structural conflict of interests between representatives and absent class members.

Courts have found adequacy to be an important consideration in copyright class settlements involving future licensing agreements. For example, the district court rejected a proposed 2011 settlement in *Authors Guild v. Google* (*Google Book Project*) in part because the "interest and values" of certain absent class members differed from those of the class representatives.[249] Under such circumstances, the court found it unacceptable that absent class members might be "deemed—by their silence—to have granted to Google a license to future use of the copyrighted works."[250] Similarly, the Second Circuit reversed approval of a proposed settlement in *In re Literary Works in Electronic Database Copyright Litigation* because the Court agreed with objectors that the named plaintiffs had a structural incentive to favor class members with registered copyrights over class members without any registered copyrights—a bias that was reflected in the structure of the proposed settlement.[251]

[248] *See* Amchem Prods., Inc. v. Windsor, 521 U.S. 591, 625 (1997) (Rule 23(a)(4) "serves to uncover conflicts of interest between named parties and the class they seek to represent.").

[249] *Google Book Project*, 770 F. Supp. 2d 666, 679 (S.D.N.Y. 2011).

[250] *Id.* at 680.

[251] *In re* Literary Works in Elec. Database Copyright Litig., 654 F.3d 242, 254 (2d Cir. 2011); *see also* Utopia Ent., Inc. v. Claiborne Par., No. CIV.A.03-1355, 2006 WL 8435006, at *1 (W.D. La. Jan. 10, 2006) (no adequacy of representation when one of the named representatives might be viewed as having condoned the infringement, thus undermining the class claims).

In both *Google Book Project* and *In re Literary Works*, the plaintiffs were subsequently able to satisfy the adequacy requirement, at least at the district court level. In *Google Book Project*, Judge Chin in 2012 approved Authors Guild's motion for class certification notwithstanding Google's claim that the class representatives were inadequate because a large percentage of absent class members in fact supported Google's book project and/or believed they benefited from it.[252] In dicta, the Second Circuit noted its view that Google's adequacy argument "may carry some force," but it vacated certification without analysis to prioritize resolution of the fair use question.[253]

The class in *In re Literary Works* also ultimately satisfied the adequacy requirement. The Second Circuit mandated the creation of sub-classes to ensure adequate representation of authors with only non-registered copyrights.[254] Subsequently the district court approved a revised settlement agreement that included such a sub-class.[255]

The second facet of adequacy addresses whether class counsel has the requisite experience and resources to prosecute a complex action on behalf of a class. As of 2003, Rule 23 mandates that courts adjudicating class actions appoint class counsel.[256] Since that time, it is unusual for a court to deny class certification on the ground that counsel is inadequate. But in one case—coincidentally decided in 2003—the court did deny certification in a copyright case based in part on poor representation.[257]

---

[252] *Google Book Project*, 282 F.R.D. 384, 394 (S.D.N.Y. 2012), *vacated on other grounds*, 721 F.3d 132 (2d Cir. 2013).

[253] *Google Book Project*, 721 F.3d 132, 134 (2d Cir. 2013).

[254] *See* FED. R. CIV. P. 23(c)(5) ("When appropriate, a class may be divided into subclasses that are each treated as a class under this rule.").

[255] Order Granting Final Approval of Revised Proposed Settlement and Final Judgment, *In re* Literary Works in Elec. Databases Copyright Litig., No. 00-MD-01379-GBD (S.D.N.Y. June 10, 2014), https://storage.courtlistener.com/recap/gov.uscourts.nysd.410272.51.0.pdf [https://perma.cc/URJ4-XSET].

[256] FED. R. CIV. P. 23(g)(1).

[257] Auscape Int'l v. Nat'l Geographic Soc'y, No. 02 Civ. 6441, 2003 WL 23531750, at *4–8 (S.D.N.Y. July 25, 2003). The plaintiff represented a proposed class arguing that National Geographic had violated their copyrights by reproducing their articles (which had previously appeared in the print version in the magazine) on microfiche and CD Rom. The district court found class counsel inadequate in a 6-page description of ineffectiveness, sanctions, and prior misconduct. The court also found that the class failed to satisfy Rule 23(b)(3). *Id.* at *17.

B. Copyright and Class Action Rule 23(b)

In addition to Rule 23(a), a proposed class must also comport with one provision of Rule 23(b), which delineates three major forms of the class action. Given the small number of proposed copyright classes that survived Rule 23(a), there is minimal caselaw analyzing Rule 23(b) in the context of copyright claims. That said, copyright plaintiffs have sought certification under all three Rule 23(b) class subtypes—a pattern that is replicated in pending AI class actions. As described below, the various Rule 23(b) categories differ in ways that are significant for copyright cases. For example, some Rule 23(b) categories prohibit class members from opting out; one category mandates such an opt-out right. Some categories allow money damages; others bar them. As described below, the AI copyright plaintiffs have pleaded certification under all these potential categories, leaving their options open.

Rule 23(b)(1): Although it comes first chronologically, Rule 23(b)(1) is a lesser known, somewhat obscure provision of the class action rule. It contains two subparts.[258] Referred to colloquially as the "prejudice class" provision, (b)(1) authorizes class resolution in situations where individual suits reaching inconsistent results might prejudice either the party opposing the class (typically the defendant), or members of the class (typically the plaintiffs). Rule 23(b)(1) does not permit putative class members to opt out of a certified class; this means that a class certified under (b)(1) will bind—for good or for ill—all who are encompassed within the class definition. In addition, there is no requirement that class members receive individualized notice in (b)(1) classes, although Rule 23 does provide that a court "may direct appropriate notice to the class."[259]

(b)(1)(A): This sub-provision of (b)(1) allows certification where individual suits might "establish incompatible standards of conduct."[260] On its face, this language seems to contemplate class certification under a very broad set of

---

[258] FED. R. CIV. P. 23(b)(1).

[259] FED. R. CIV. P. 23(c)(2)(A).

[260] FED. R. CIV. P. 23(b)(1)(A) ("A class action may be maintained if Rule 23(a) is satisfied and if:

> (1) prosecuting separate actions by or against individual class members would create a risk of:
> (A) inconsistent or varying adjudications with respect to individual class members that would establish incompatible standards of conduct for the party opposing the class").

circumstances. That is not true in practice. Most importantly, this provision does not make certification appropriate simply due to the risk of inconsistent jury verdicts. Construed that way, (b)(1)(A) would make most cases certifiable.[261] In practice, this Rule 23(b) category typically applies in contexts where a ruling against the defendant as to one party would force the court to take identical steps with respect to others similarly situated, such as in a case brought by one of multiple riparian landowners.[262] In modern day, this subtype is most commonly used in ERISA cases, where a plaintiff's action is brought on behalf of a retirement plan.[263]

The court in *Showtime/The Movie Channel* granted certification in a copyright suit under (b)(1)(A), based on its finding that individual suits by copyright owners against the defendant cable company might create contrary, seemingly arbitrary, results.[264] Reflecting an earlier, more naïve approach to class action analysis, however, the court's analysis was barely a paragraph long.[265]

(b)(1)(B): This sub-provision allows certification where individual judgments "as a practical matter, would be dispositive of the interests" of those not before the court.[266] Just as with (b)(1)(A), *stare decisis*—or the possibility that different courts might reach different results—is not enough to justify certification under this sub-provision. As with Rule 23(b)(1)(A), absent class members do not have a right to opt out of a (b)(1)(B) class.

This is not a strong basis for class certification in the AI context. The paradigmatic use of (b)(1)(B) is—or was—where a defendant's resources are inadequate to meet the substantial legal claims against it; in the absence of a class action, the first plaintiffs to reach judgment will deplete the available funds at the expense of those later in line. But forcing all affected parties to sue as part of a class raises due process concerns, especially given the lack of an opt-out right. In

[261] 2 William Rubenstein, Alba Conte & Herbert B. Newberg, Newberg and Rubenstein on Class Actions § 4:7 (6th ed. 2024) ("This would cast too broad a net.").

[262] Fed. R. Civ. P. 23(b)(1)(A) advisory committee's note to 1966 amendment.

[263] Rubenstein et al., *supra* note 261, § 4:12 (noting that courts "regularly certify ERISA cases under Rule 23(b)(1)(A)").

[264] David v. Showtime/The Movie Channel, Inc., 697 F. Supp. 752, 757 (S.D.N.Y. 1988).

[265] *Id.*

[266] Fed. R. Civ. P. 23(b)(1)(B) (allowing a class action where adjudications with respect to individual class members that, as a practical matter, would be dispositive of the interests of the other members not parties to the individual adjudications or would substantially impair or impede their ability to protect their interests).

Ortiz v. Fibreboard, the Supreme Court cast a shadow on this type of class action. It rejected a settlement of an asbestos class action that would have been confined to the available insurance proceeds of the defendant (and not every available dollar); and it warned against "adventurous application" of Rule 23(b)(1)(B).[267] As a result, courts rarely certify limited fund classes; like its counterpart (b)(1)(A), this subtype is predominantly used in ERISA suits.[268]

Courts have not certified copyright class actions under Rule 23(b)(1). In *WB Music Corp. v. Rykodisk, Inc.*, the court refused such certification, noting that because each class member's infringement claim was particular to them, resolution of one would not impair the rights of others.[269] Similarly, the First Circuit rejected an attempt by a graphic artist to sue a defendant class made up of wholesalers and retailers who infringed the copyright to her wallpaper design.[270] In essence, the court found, the plaintiff was arguing, incorrectly, that *stare decisis* was sufficient to qualify as a prejudice requiring class resolution; the court was unpersuaded.[271]

23(b)(2): This class action category, referred to as the "civil rights" provision, applies where "the party opposing the class has acted or refused to act on grounds that apply generally to the class, so that final injunctive relief or corresponding declaratory relief is appropriate respecting the class as a whole."[272] This provision is a vehicle for obtaining declaratory and injunctive relief, rather than money damages.[273] As with 23(a), there is no right to opt out of a certified (b)(2) class: a decision is binding on all those who fall within the class.

Rule 23(b)(2) may provide relief to groups of copyright owners in two ways. First, (b)(2) certification can provide the basis for class-wide injunctive relief—that is, it may be invoked to prevent future infringement. Second, parties and courts may

---

[267] Ortiz v. Fibreboard Corp., 527 U.S. 815, 845 (1999).

[268] *See* 1 Joseph M. McLaughlin, McLaughlin on Class Actions § 5:14 (21st ed. 2024) (noting that "classes seeking to remedy alleged breaches of fiduciary duties under ERISA are generally certified under Rule 23(b)(1)(B) or, alternatively, under Rule 23(b)(1)(A) or (b)(2)").

[269] WB Music Corp. v. Rykodisk, Inc., No. CIV.A.94-2902, 1995 WL 631690, at *4 (E.D. Pa. Oct. 26, 1995).

[270] Tilley v. TJX Co., 345 F.3d 34, 40–43 (1st Cir. 2003).

[271] *Id.* at 41.

[272] Fed. R. Civ. P. 23(b)(2).

[273] Fed. R. Civ. P. 23(b)(2) advisory committee's note to 1966 amendment (Rule 23(b)(2) does not apply where "final relief relates exclusively or predominantly to money damages"); Suzette Malveaux, *Class Actions at the Crossroads: An Answer to Wal-Mart v. Dukes*, 5 Harv. L. & Pol'y Rev. 375, 391–96 (2011).

invoke (b)(2) in so-called bifurcated proceedings. Under Rule 23(c)(4), a court may certify a class only for certain issues. Using this issue-class provision, a court could certify a class under (b)(2) for purposes of determining liability, while leaving damages or other downstream claims to be resolved through a different, perhaps individual, mechanism.[274]

Two early copyright class actions illustrate this. In one of the cases (b)(2) certification was not appropriate given the plaintiffs' request for substantial damages awards.[275] But the plaintiffs in *Ryan v. CARL Corp.* were successful in obtaining certification under (b)(2), despite potential for money damages. The *Ryan* class requested that the court bifurcated issues into liability—for which (b)(2) was appropriate—and damages, which could be adjudicated later, if necessary, either in the form of a (b)(3) class or in individual proceedings.[276] The court in *Ryan* agreed: it approved (b)(2) certification given the "uniformity" of the defendant's allegedly infringing practice and stayed consideration of a damages class for another day.[277]

Yet courts have been skeptical of (b)(2) copyright claims where plaintiffs cannot show a high degree of uniformity in the defendant's relationship with class members. In *Kihn v. Bill Graham Archives LLC*, the Ninth Circuit reversed a district court decision certifying a class of musicians alleging copyright violations by a company that allows users access to recordings of live concerts from the 1950s to 1990s.[278] The district court had certified the class under both (b)(2) and (b)(3); the appellate court rejected both.[279] As to (b)(2), the Court of Appeals found that the result of certification—an injunction requiring the defendant to remove a whole category of recordings from its website—failed to recognize significant distinctions among class members.[280] For example, some of the musicians involved might have reached agreements with the defendant, and others may wish to do so in the

[274] Wright & Miller, *supra* note 208, § 1790; *see also, e.g.*, McReynolds v. Merrill Lynch, Pierce, Fenner & Smith, Inc., 672 F.3d 482, 491–92 (7th Cir. 2012) (invoking Rule 23(c)(4) and allowing certification of Rule 23(b)(2) class for purposes of determining liability in Title VII case).

[275] WB Music Corp. v. Rykodisk, Inc., No. CIV.A.94-2902, 1995 WL 631690, at *4 (E.D. Pa. Oct. 26, 1995).

[276] Rule 23's flexible judicial management provisions allow such bifurcation. Fed. R. Civ. P. 23(c)(4)(A).

[277] Ryan v. CARL Corp., No. C97-3873, 1999 WL 16320, at *11 (N.D. Cal. Jan. 13, 1999).

[278] Kihn v. Bill Graham Archives LLC, No. 20-17397, 2022 WL 18935, at *1 (9th Cir. Jan. 3, 2022).

[279] *Id.*

[280] *Id.* at *3.

future.[281] Under (b)(2), those musicians would have no opportunity to opt out of the class.[282]

23(b)(3): This final class action category, added in 1966, allows certification of a class for purposes of seeking money damages for individual class members.[283] The pending AI copyright class actions all seek some form of money damages, which means that they will very likely need to contend with the thorny requirements of (b)(3). The Supreme Court has labeled (b)(3) "the most adventuresome innovation" of the modern Rule 23.[284] As one scholar stated two decades ago, the (b)(3) standard is fuzzy at best; the rules for interpreting it "have received virtually no clarification from the Supreme Court, have bewildered lower courts, and have not attracted substantial scholarly scrutiny."[285] Notwithstanding the Roberts Court's strong procedural bent, that remains largely true.

Because of inherent tensions between individuals' right to their own day in court and the monolithic nature of a class determination, plaintiffs seeking (b)(3) certification must make additional showings to demonstrate that a representative action—rather than individual suits—is appropriate. Specifically, parties seeking (b)(3) certification must prove that common issues in the case predominate over individual issues, and also that a class action is superior to individual suits or other

---

[281] *Id.*

[282] *Id.*

[283] Rule 23(b)(3) provides:

> (3) the court finds that the questions of law or fact common to class members predominate over any questions affecting only individual members, and that a class action is superior to other available methods for fairly and efficiently adjudicating the controversy. The matters pertinent to these findings include:
> (A) the class members' interests in individually controlling the prosecution or defense of separate actions;
> (B) the extent and nature of any litigation concerning the controversy already begun by or against class members;
> (C) the desirability or undesirability of concentrating the litigation of the claims in the particular forum; and
> (D) the likely difficulties in managing a class action.

FED. R. CIV. P. 23(b)(3).

[284] Amchem Prods., Inc. v. Windsor, 521 U.S. 591, 592 (1997).

[285] Allan Erbsen, *From Predominance to "Resolvability": A New Approach for Regulating Class Actions*, 58 VAN. L. REV. 995, 1001 (2005).

dispute resolution formats.[286] Finally, a court certifying a class under Rule (b)(3) must also find that administering the class is feasible,[287] although the Supreme Court has clarified that a court need not address trial manageability if it is certifying a class for purposes of settlement.[288] Unlike in the other 23(b) categories, members of (b)(3) classes have the right to opt out of membership in the class.[289]

To show predominance, a proposed class must show that the class issues are not only common, but in fact are central—that they are the driving force in the dispute. Not surprisingly, therefore, proposed class suits with questionable commonality and typicality will inevitably founder on the extra predominance requirement.[290]

In *Google Book Project*, for example, the district court used Rule 23(b)(3) to certify a class of copyright holders in their suit alleging that Google's project to digitize and create a scannable index of books was a "massive copyright infringement."[291] The court found that "[e]very potential class member's claim arises out of Google's uniform, widespread practice of copying entire books without permission of the copyright holder and displaying snippets of those books for search."[292] Google's fair use defense, the court found, could also be resolved on an aggregate basis.[293] The district court also found that a class action was a superior method of resolving the claims against Google, saving both time and money.[294] On appeal, the Second Circuit vacated the certification order.[295] The appellate court did not take specific issue with the district court's analysis, although it intimated

[286] *See* FED. R. CIV. P. 23(b)(3).
[287] FED. R. CIV. P. 23(b)(3)(D).
[288] *Amchem*, 521 U.S. at 620.
[289] *See* FED. R. CIV. P. 23(c)(2)(B)(v).
[290] *See, e.g.*, Waite v. UMG Recordings, Inc., No. 19-CV-01091, 2023 WL 1069690, at *8 (S.D.N.Y. Jan. 27, 2023) (moving to the "crux" of defendants' arguments and rejecting class certification under Rule 23(b)(3)).
[291] *Google Book Project*, 282 F.R.D. 384, 387 (S.D.N.Y. 2012), *vacated on other grounds*, 721 F.3d 132 (2d Cir. 2013).
[292] *Id.* at 395.
[293] *See id.*
[294] *See id.*
[295] *Google Book Project*, 721 F.3d 132 (2d Cir. 2013).

some skepticism.[296] Rather, it ordered the district court hold its class certification in abeyance until after it resolved the fair use question.[297]

Class actions involve competing narratives—connectivity v. entropy. Copyright certification decisions reflect this tension. For example, in opposing (b)(3) certification, Google argued that copyright ownership was too varied and complex for resolution on a class-wide basis.[298] But the district court found that those individual issues were subsidiary to the central question of the legality of Google's book project.[299] In contrast, in *Waite v. UMG Recordings*, Inc., the district court denied (b)(3) certification on the ground that individual issues—specifically, whether the musical compositions at issue were "made for hire"—prevented resolution of the class claims on common proof.[300] Resolution of these same connectivity v. entropy questions will determine class certification in the AI context.

C. Copyright and Class Action Settlements Under Rule 23(e)

Settlements of class actions are different—and more onerous—than the private settlements that are traditional in litigation. Because any settlement will bind absent class members to a judgment, Rule 23(e) requires a court to ensure that those absent class members have been adequately represented and that the proposed settlement is "fair, reasonable, and adequate."[301] Reflecting the quasi-regulatory nature of class actions, the end-product of Rule 23(e) is a hybrid creature between a court judgment and a private settlement contract.[302] Parties first bargain with each other; then they must persuade the court, over the voices of any objectors, that the settlement is worthy of a judicial stamp of approval.

This approval requires multiple steps. First, a court must determine that a proposed settlement class meets the Rule 23(a) and (b) requirements described above; the Supreme Court expressly admonished litigants and courts that

---

[296] *Id.* at 134.

[297] *Id.* at 135.

[298] *Google Book Project*, 282 F.R.D. at 395.

[299] *See id.*

[300] Waite v. UMG Recordings, Inc., No. 19-CV-01091, 2023 WL 1069690, at *5–6 (S.D.N.Y. Jan. 27, 2023).

[301] Fed. R. Civ. P. 23(e)(2).

[302] Howard M. Erichson & Ethan J. Leib, *Class Action Settlements as Contracts?*, 102 N.C. L. Rev. 73, 77 (2023).

the requirements apply—perhaps with even greater import—in the settlement context.[303] Second, Rule 23(e) mandates that courts give appropriate notice to all class members.[304] It must also hold a "fairness hearing" to evaluate the settlement, and to allow objectors to the proposal to have their say.[305] In evaluating whether a proposed settlement is adequate, a court must consider, *inter alia*: whether the representatives fairly represented the class; whether the proposal was negotiated at arms' length; and whether the relief provided for the class is adequate, taking into account factors such as the costs and risks of proceeding to trial and the method proposed for remitting the relief provided by the settlement (money damages or otherwise) to class members.[306]

Central to this fairness determination is an evaluation of whether there is a significant conflict among different subsets of the class. In *Amchem Products, Inc. v. Windsor*—the Supreme Court's master blueprint for class action settlements—the Court rejected a global asbestos settlement on the ground that the agreement impermissibly privileged class members with *current* asbestos-related injuries over those whose injuries would only manifest in the *future*.[307]

Notwithstanding the relative paucity of copyright class actions, copyright class settlements have played an influential role in establishing the parameters for technology companies' use of intellectual property. The Second Circuit's analysis of the proposed settlement in *In re Literary Works* provides an apt illustration.[308]

Following the Supreme Court's decision in *Tasini*,[309] the Second Circuit evaluated a proposed consolidated class settlement in a suit by authors whose works had been placed in electronic databases without their permission.[310] The district court had referred the case to mediation; three years later, the parties had reached an

[303] Amchem Prods., Inc. v. Windsor, 521 U.S. 591, 593 (1997) (cautioning litigants that "Rule 23(e)'s settlement prescription was designed to function as an additional requirement, not a superseding direction, to the class-qualifying criteria of Rule 23(a) and (b)").

[304] Fed. R. Civ. P. 23(e)(1).

[305] Fed. R. Civ. P. 23(e)(2), (5).

[306] Fed. R. Civ. P. 23(2)(A)–(D).

[307] *Amchem*, 521 U.S. at 626 (rejecting proposed settlement in part based on the agreement's "disparity between the currently injured and exposure-only categories of plaintiffs").

[308] *In re* Literary Works in Elec. Databases Copyright Litig., 654 F.3d 242 (2d Cir. 2011).

[309] N.Y. Times Co. v. Tasini, 533 U.S. 483, 483 (2001).

[310] *In re Literary Works*, 654 F.3d at 245.

agreement.[311] They jointly moved the district court to certify the class and approve the settlement.[312] *In re Literary Works* contains two lessons for AI class actions:

1. Class Settlements Must Not Mask Significant Intra-Class Conflicts

The Second Circuit initially rejected the proposal on the ground that it did not adequately represent certain class members.[313] Under the terms of the proposed settlement, class members were divided into three classes.[314] Category A included class members who hold registered copyrights making them eligible for statutory damages.[315] Category B included those with copyrights registered too late to render them eligible for statutory damages, and Category C included those who have not yet registered their copyrights—a group which composed 99% of the total claims.[316] The Court found that there had been inadequate representation of the Category C group.[317] Noting that the interests of Category C should have been protected by a sub-class represented by independent counsel,[318] the Court remanded with instructions for the creation of three separately represented sub-classes.[319] On remand, the district court approved a settlement with those sub-classes.[320] As discussed below, some of the pending AI suits have defined the class to include only creators who have timely registered their copyrights; other suits, however, are including both registered and non-registered copyright holders in the class, thus raising the likely need for sub-classes.[321]

2. Class Settlements May Release Defendants from Liability for Future Actions

Objectors to the proposed settlement in *In re Literary Works* also challenged the scope of the settlement release, which applied to claims that were beyond those

---

[311] *Id.*
[312] *Id.* at 247.
[313] *Id.* at 251.
[314] *Id.* at 246.
[315] *Id.*
[316] *Id.*
[317] *Id.* at 253.
[318] *Id.* at 254.
[319] *Id.* at 257–58.
[320] Order Granting Final Approval of Revised Proposed Settlement and Final Judgment at 3, *In re* Literary Works in Elec. Databases Copyright Litig., No. 00-MD-01379 (S.D.N.Y. June 10, 2014).
[321] *See* First Consolidated Class Action Complaint at 60–61, Authors Guild v. OpenAI, Inc. Nos. 23-CV-08292, 23-CV-10211 (S.D.N.Y. Feb. 6, 2024), https://admin.bakerlaw.com/wp-content/uploads/2024/02/ECF-47-First-Consolidated-Class-Action-Complaint.pdf [https://perma.cc/5F5W-Z9CZ].

that had been pleaded in the class action.[322] Under the terms of the settlement, class members who did not opt out were prohibited from barring future use of their works by the defendant publishers, including the licensing of those works to third parties.[323] Objectors argued that future infringements would be new and distinct harms, and they also objected to the provision allowing future sub-licensing.[324] But the Second Circuit overruled the objections, finding that the settlement's scope was acceptable and that the terms were reasonably within the factual predicate of the underlying suit.[325]

There may be limits to courts' tolerance of future releases in some class settlements.[326] In the proposed settlement of *Google Book Project*, for example, the district court rejected a release outlined in a 166-page settlement agreement, finding the scope of the release of future claims too broad—too unpredictably constraining on class members.[327] But the court's rejection of that settlement may be an outlier.[328] Courts have approved copyright class settlements with releases in many other contexts.

D. Administrative Rule 23 Requirements

In addition to satisfying Rule 23(a) and one of the categories of Rule 23(b), a class action must be administrable—that is, a court must find that the parties are capable of implementing a judgment in a way that is fair to absent class members. Several practical elements of Rule 23 are particularly relevant to copyright class actions. Although pre-AI copyright opinions do not expressly address all these administrative elements, they are discussed here for purposes of setting forth the main points of contention in many certification disputes.

---

[322] *In re Literary Works*, 654 F.3d at 247–48.

[323] *Id.*

[324] *Id.* at 248–49

[325] *Id.*

[326] Grimmelmann, *supra* note 12, at 418 (warning that future releases in class settlements "are harder to understand, they create unique design problems, and courts are the wrong institutions to make such decisions").

[327] *Google Book Project*, 770 F. Supp. 2d 666, 671, 677–78 (S.D.N.Y. 2011).

[328] Tang, *The Class Action*, *supra* note 5, at 1659 (describing rejection of the *Google Book Project* settlement as "the exception").

1. Defining the Class: Ascertainability and the "Fail-safe Class"

Over the past decade, some federal courts have read an "implied requirement of ascertainability" into Rule 23.[329] The basic concept is not controversial. To be certified, a class must be "sufficiently definite so that it is administratively feasible for the court to determine whether a particular individual is a member."[330] To ensure this, a class must be defined by objective criteria.[331] Sometimes this can be complicated. For example, in a consumer class action, it may be difficult to determine by an objective standard who purchased a particular food item or cosmetic.

As it is applied by some courts, ascertainability can interpose a high, seemingly arbitrary barrier to class certification.[332] The Third Circuit is the leader of this aggressive approach.[333] That Court defends its rigorous ascertainability requirement on the ground that it protects absent class members, ensures due process to the defendant, and eases the cost and administrative burden of an ill-defined class.[334] This is purportedly necessary in order to determine the preclusive scope of a judgment—"who gets the benefit of any relief and who gets the burden of any loss."[335] The term ascertainability is nowhere in the text or rulemaking commentary to Rule 23.

But the Circuits do not agree on the nature and scope of an ascertainability requirement—or even on whether it really exists—and the Supreme Court has yet to weigh in. The Second and Ninth Circuits, where most of the major AI class actions are pending, have both expressly rejected the Third Circuit's ascertainability

---

[329] *In re* Initial Pub. Offerings Secs. Litig., 471 F.3d 24, 30 (2d Cir. 2006); RUBENSTEIN ET AL., *supra* note 261, § 3:2 (noting that some courts find this requirement to be implicit in Rule 23(a), while others locate this requirement in Rule 23(c)(1)(B)).

[330] WRIGHT & MILLER, *supra* note 208, § 1760.

[331] *See, e.g.*, *In re* Petrobas Secs., 862 F.3d 250, 254 (2d Cir. 2017).

[332] *See* Zachary L. Sanders, *Administrative Feasibility Redux: a Reexamination of the Heightened Ascertainability Requirement for Class Certification*, 98 N.Y.U. L. REV. 2273, 2273 (2023) (criticizing ascertainability as "a scattershot cudgel").

[333] *See* Daniel Luks, *Ascertainability in the Third Circuit: Name That Class Member*, 82 FORDHAM L. REV. 2359 (2014) (criticizing the circuit's high ascertainability standard); Carrera v. Bayer Corp., 727 F. 3d 300, 303–04 (3d Cir. 2013).

[334] *Carrera*, 727 F.3d at 307, 310.

[335] Flo & Eddie, Inc. v. SiriusXM Radio, Inc., No. 13-5693, 2015 WL 4776932, at *6 (C.D. Cal. May 27, 2015); *see, e.g.*, Marcus v. BMW of N. Am., LLC, 687 F.3d 583 (3d Cir. 2012) (ascertainability ensures that those who will be bound by any judgment are identifiable).

requirement as grounded in policy rather than the text of Rule 23.[336] Therefore, while ascertainability—sometimes called "administrative feasibility"—may pose a challenge to class action litigants, the AI class actions are not likely to founder on this ground.

Related to ascertainability is the argument that a court should not certify a "fail-safe" class.[337] A fail-safe class is one in which membership in the class depends on whether a party has a valid claim.[338] In other words, if a class member loses their claim for relief, then by definition they were never properly a member of the class, and therefore are not bound by *res judicata*.[339] Neither the Second nor the Ninth Circuits has ruled on the existence or import of the concept of a "fail-safe" class.[340] In any case, however, as discussed below, the AI plaintiffs are not proposing such an *ipse-dixit*, preclusion-proof class. If courts in the pending AI cases certify classes of copyright holders, the members of those classes—that is, those creators who do not exercise their right to opt out[341]—may either prevail or take nothing, depending on courts' views of the merits.

2. *The Order of Operations*

Under Rule 23, district courts have wide discretion as to how they manage class actions.[342] Therefore, while Rule 23 provides that courts should assess class certification at "an early practicable time,"[343] many courts allow certification to be decided after decisions—or even judgment—on the merits.[344]

This procedural flexibility inevitably interjects another axis of strategy and uncertainty into class actions, including the pending AI copyright cases. For example, *Texaco*, discussed above, was brought as a class action. However,

---

[336] Briseno v. ConAgra Foods, Inc., 844 F.3d 1121, 1133 (9th Cir. 2017) (joining the Sixth, Seventh, and Eighth Circuits in declining to adopt a free-standing administrative feasibility requirement); *In re* Petrobas Secs., 862 F.3d 250, 264 (2d Cir. 2017) (same); *see also* Cherry v. Domestic Corp., 986 F.3d 1296, 1304 (11th Cir. 2021) (rejecting requirement).

[337] Wright & Miller, *supra* note 208, § 1760 (analyzing ascertainability and fail-safe classes).

[338] *See* Erin L. Geller, *The Fail-Safe Class as an Independent Bar to Class Certification*, 81 Fordham L. Rev. 2769, 2770–71 (2013).

[339] *See, e.g.*, Messner v. Northshore Univ. HealthSystem, 669 F.3d 802, 825 (7th Cir. 2012).

[340] *See* Makaron v. Enagic USA, Inc., 324 F.R.D. 228, 235 (C.D. Cal. 2018).

[341] *See* Fed. R. Civ. P. 23(c)(2)(v).

[342] Fed. R. Civ. P. 23(d) (allowing district court to prescribe the course of proceedings).

[343] Fed. R. Civ. P. 23(c)(1)(A).

[344] Wright & Miller, *supra* note 208, § 1785.3 (listing cases).

the parties stipulated to try the question of fair use prior to addressing class certification. Once the Second Circuit denied Texaco's fair use defense—basically ensuring that Texaco would be held liable for copying research publications for its employees—the case settled.[345]

As discussed below, litigants in the pending AI copyright class actions have taken different procedural pathways. Those decisions have pragmatic and substantive implications.

V
Certification of AI Input Classes

Against the above background, multiple content creators have filed proposed class actions in federal courts against generative AI companies, seeking relief for various forms of copyright infringement, sometimes accompanied by state law claims. These lawsuits have sprung up primarily in two locations: New York and California, although two are in Delaware.[346] The New York cases, all in the Southern District, include actions brought by fiction authors, non-fiction authors, and journalists.[347] On the west coast, several suits—including those brought by visual artists,[348] music publishers,[349] coders,[350] and additional suits by authors[351]—were brought in California; all are pending in the Northern District.

In addition to the suits seeking formal class certification, there are a handful of cases against AI companies brought by single organizations or corporations that support, aggregate, and disseminate the work of a large number of creators. These include cases by *The New York Times*, the Center for Investigative Reporting, and

[345] Sarah K. Wiant, *Settlement of* Texaco *Case*, Stanford Libs. (June 1995), https://fairuse.stanford.edu/texaco/settlement/ [https://perma.cc/GK6L-JEQG] (stating that Texaco agreed to pay over $1 million, to pay a retroactive licensing agreement with the Copyright Clearance Center, and to enter a 5-year forward looking licensing agreement).

[346] Note that two major AI suits, neither class actions, are proceeding in Delaware. Getty Images (US), Inc. v. Stability AI, Inc., No. 23-CV-00135-UNA (D. Del. filed Feb. 3, 2023); Thomson Reuters Enter. Ctr. GmbH v. Ross Intel. Inc., No. 20-CV-00613-SB, 2025 WL 458520 (D. Del. Feb. 11, 2025).

[347] *See* Huckabee v. Meta Platforms, Inc., No. 23-CV-09152-LGS (S.D.N.Y. filed Oct. 17, 2023) (non-fiction and self-help books).

[348] Andersen v. Stability AI Ltd., No. 23-CV-0021 (N.D. Cal. filed Jan. 13, 2023).

[349] Concord Music Grp., Inc. v. Anthropic PBC, No. 24-CV-03811 (N.D. Cal. filed June 26, 2024).

[350] Doe v. GitHub, Inc., No. 22-CV-06823 (N.D Cal. filed Nov. 3, 2022).

[351] Kadrey v. Meta Platforms, Inc., No. 23-CV-03417 (N.D. Cal. filed July 7, 2023); Nazemian v. NVIDIA Corp., No. 24-CV-01454 (N.D. Cal. filed Mar. 8, 2024) (authors).

Getty Images.[352] Because these suits do not implicate the technical or logistical complexities of Rule 23, courts in these suits may reach decisions on the merits, including the applicability of fair use, earlier than their class-action cousins. If that happens, those precedents will influence the default expectations—and therefore the settlement options—of AI litigants across the board.

In the class suits, the parties' litigation strategies reflect those adopted by early copyright class litigants discussed above.[353] The plaintiff creators argue that the defendants' rapacious copying of literary works is the defining conduct in the dispute, providing the basis for class certification and liability. As their complaint states, "OpenAI and Microsoft have built a business valued into the tens of billions of dollars by taking the combined works of humanity without permission."[354] In contrast, the defendants—while trying to evade liability entirely based on fair use[355]—have generally made blanket denials regarding the class-wide actions necessary to support class certification.[356] The success of these opposing strategies will largely depend on courts' views of the underlying intellectual property claims. As is always true of class actions, procedure and substantive law are intertwined. But even if the copyright claims appear to chart new ground—something we dispute—the class certification issues are, mostly, old hat.

For purposes of analyzing the likelihood of obtaining class certification, we divide AI claims into two categories. In the first category, input claims, the plaintiffs

[352] Complaint at 2, N.Y. Times Co. v. Microsoft Corp., No. 23-CV-11195 (S.D.N.Y. Dec. 27, 2023) (challenging "[d]efendants' unlawful use of The Times's work to create artificial intelligence products that compete with it"). Other cases brought by news organizations or content aggregators include: Ctr. for Investigative Reporting, Inc. v. OpenAI, Inc., No. 24-CV-04872 (S.D.N.Y. filed June 27, 2024); *see also* Daily News LP v. Microsoft Corp., No. 24-CV-03285 (S.D.N.Y. filed Apr. 30, 2024); Getty Images (US), Inc. v. Stability AI, Inc. No. 23-CV-00135-UNA (D. Del. filed Feb. 3, 2023) (photographs); and Intercept Media, Inc. v. OpenAI, Inc., No. 24-CV-01515 (S.D.N.Y. filed Feb. 28, 2024).

[353] *See generally supra* Part IV.

[354] The first of the consolidated suits, *Authors Guild v. OpenAI, Inc.*, was filed in September 2023. Class Action Complaint, Authors Guild v. OpenAI, Inc., No. 23-CV-08292 (S.D.N.Y. Sept. 19, 2023).

[355] *See, e.g.*, OpenAI Defendants' Answer to First Consolidated Class Action Complaint at 3, Authors Guild v. OpenAI, Inc., Nos. 23-CV-08292, 23-CV-10211 (S.D.N.Y. Feb. 16, 2024) (asserting that use of plaintiffs' works to teach AI models intelligence and language "would be paradigmatic transformative fair use"), https://storage.courtlistener.com/recap/gov.uscourts.nysd.606655/gov.uscourts.nysd.606655.75.0.pdf [https://perma.cc/X2JL-FH5T].

[356] *See, e.g.*, *id.* at 45–46 (responding with one word to the Complaint's class allegations: "Denied.").

have a strong claim for class certification.[357] The plaintiffs' input claim is that AI companies committed direct copyright infringement by duplicating copyrighted works, without asking permission or offering compensation, in order to "train" their AI foundation models. To be sure, these are not slam-dunk claims; the AI defendants are straining mightily for a fair use defense, and the plaintiffs must obtain evidence proving which works were copied—evidence that they claim the defendants possess. Nevertheless, we believe these input claims are not only meritorious but also that they fit solidly within the Rule 23 class action framework. By way of illustration, we describe two input class claims below.

Several of the pending AI class suits seek a remedy not only for the initial copying of their works, but also for the "market usurpation defendants have enabled by making Plaintiffs unwilling accomplices in their own replacement."[358] We refer claims based on this market usurpation as output claims.[359] Output claims, often alleging derivative infringement[360] or asserting violations of the Digital Millenium Copyright Act, allege that AI companies are using creators' content—whether computer code, plot devices, literary text, images, or even name—to instantly produce and market derivative content without permission, attribution, or compensation. Although output claims are not the focus of this Article, we note that they are likely to require more individualized evidence and be subject to more individualized analysis (for example, of fair use, or substantial

[357] *See* James Vincent, *The Lawsuit that Could Rewrite the Rules of AI Copyright*, The Verge (Nov. 8, 2022), https://www.theverge.com/2022/11/8/23446821/microsoft-openai-github-copilot-class-action-lawsuit-ai-copyright-violation-training-data [https://perma.cc/HQ5J-8YX6] (describing *Doe v. GitHub* as the "first class-action case in the U[.]S[.] challenging the training and output of AI systems").

[358] Class Action Complaint at 3, Authors Guild v. OpenAI, Inc., No. 23-CV-08292 (S.D.N.Y. Sept. 19, 2023).

[359] This type of derivative claim can take many forms. For example, author Jane Friedman accused Amazon of selling books that listed her as an author but were in fact AI-generated. Clare Duffy, *An Author Says AI is "Writing" Unauthorized Books Being Sold under Her Name on Amazon*, CNN (Aug. 10, 2023), https://www.cnn.com/2023/08/10/tech/ai-generated-books-amazon/index.html [https://perma.cc/392M-3N5Y]. Flooded with low-quality, AI-generated, "scam" books, Amazon has now implemented a cap on how many titles can be published per day. Andrew Limbong, *Authors Push Back on the Growing Number of AI "Scam" Books on Amazon*, NPR (Mar. 13, 2024), https://www.npr.org/2024/03/13/1237888126/growing-number-ai-scam-books-amazon [https://perma.cc/T5CA-3UUM].

[360] The Copyright Act grants holders exclusive rights to "prepare derivative works." 17 U.S.C. § 106(2).

similarity). This is likely to make class certification of such claims challenging.[361] In all likelihood, however, the strength of those claims will never be tested: the lingering uncertainty and risk for all parties will provide a strong incentive to reach a settlement of all outstanding claims.

Below are two illustrative examples of input claims where we believe class certification is appropriate. As these examples demonstrate, AI copyright classes are not cookie-cutter. To the contrary, parties and attorneys have adopted quite different litigation strategies in these cases.

A. The Input Class: Direct Copyright Infringement in Authors Guild v. OpenAI, Inc.

In *Authors Guild v. OpenAI*, a professional writers' organization[362] as well as two groups totaling approximately 30 plaintiffs—including best-selling fiction authors such as George R. R. Martin and Jodi Picoult, as well as non-fiction authors including Stacy Schiff and Jia Tolentino—seek to represent classes of fiction and non-fiction authors in consolidated suits against Open AI and Microsoft.[363] Among other claims, the plaintiffs allege that OpenAI and Microsoft committed direct copyright infringement when they scraped class members' works in order to "train" their AI large language models GPT-N and ChatGPT.[364]

The consolidated complaint defines the fiction class as "[a]ll natural persons in the United States who are the sole authors of, and legal or beneficial owners of Eligible Copyrights in, one or more Fiction Class Works; and all persons in the United States who are the legal or beneficial owners of Eligible Fiction Copyrights in one or more Fiction Class Works held by literary estates."[365] The non-fiction class definition is the same for non-fiction copyright owners. For purposes of

[361] Some output claims may nevertheless be strong candidates for resolution by class action. For example, author Jane Friedman accused Amazon of selling books that listed her as an author but were in fact AI-generated. Duffy, *supra* note 359.

[362] *About the Guild*, Authors Guild, https://authorsguild.org/about/ [https://perma.cc/5SCM-DTSG] (last visited Oct. 15, 2024).

[363] Three suits alleging similar claims were consolidated for pre-trial purposes. *See* Order Denying Motion to Intervene, Authors Guild v. OpenAI, Inc., Nos. 23-CV-08292, 23-CV-10211, 24-CV-00084, 23-CV-11195 (S.D.N.Y. Apr. 1, 2024). For simplicity, this Article will refer to these consolidated putative class actions simply as "*Authors Guild*."

[364] *Id.* at 17–19.

[365] *Id.* at 60.

these claims, eligible class members are limited to those who timely registered their copyrights. In their Answers, OpenAI and Microsoft repeatedly deny that any works were infringed.[366]

In their pleadings, the defendants also assert blanket denials regarding class certification.[367] Assuming that the plaintiff authors succeed in defeating a motion for summary judgment on the question of fair use, however, we believe plaintiffs' direct infringement claims merit class certification.

1. Order of Operations—Fair Use First

As an initial matter, the parties to the consolidated *Authors Guild* suit—now in the discovery process in the S.D.N.Y.—followed the blueprint established in *Texaco* and negotiated an agreement to allow the district court to first resolve the defendants' fair use defense, and only then to take up class certification if necessary.[368] In exchange for this concession by the plaintiffs, the defendants agreed to forego filing Rule 12(b)(6) motions to dismiss, and to comply with an expedited briefing and discovery schedule.[369]

However it is resolved, the *Authors Guild* court's fair use ruling, or a similar ruling in another case, will set the tone for the remainder of the case. As in *Texaco*, a rejection—even a partial one—of fair use will put the authors in a strong settlement position. A recent decision by the District of Delaware rejecting an AI company's fair use defense is one such positive development for creators. In *Thomson Reuters Enterprise Centre GmbH v. Ross Intelligence Inc.*, Thompson Reuters, owner of legal search engine Westlaw, alleged that Ross had unlawfully copied Westlaw's

[366] *See, e.g.*, OpenAI Defendants' Answer to First Consolidated Class Action Complaint at 25–38, Authors Guild v. OpenAI, Inc., Nos. 23-CV-08292, 23-CV-10211 (S.D.N.Y. Feb. 16, 2024) (repeatedly denying "that any works were infringed"). Note that this denial does not say that no works were copied. It is a legal conclusion ("infringed") couched in the format of a factual denial.

[367] Defendant Microsoft Corp.'s Answer to First Consolidated Class Action Complaint at 47–49, Authors Guild v. OpenAI, Inc., Nos. 23-CV-08292, 23-CV-10211 (S.D.N.Y. Feb. 16, 2024), https://storage.courtlistener.com/recap/gov.uscourts.nysd.606655/gov.uscourts.nysd.606655.74.0_1.pdf [https://perma.cc/JXN8-ZZ65].

[368] Author Class Plaintiffs' Opposition to Motion to Intervene and Dismiss, Stay or Transfer at 6, Authors Guild v. OpenAI, Inc., Nos. 23-CV-08292, 23-CV-10211 (S.D.N.Y. Feb. 26, 2024), https://storage.courtlistener.com/recap/gov.uscourts.nysd.606655/gov.uscourts.nysd.606655.81.0.pdf [https://perma.cc/95CH-GQ2W].

[369] *Id.*

headnotes in order to train Ross's competing product.[370] Ross had initially sought to license Westlaw's headnotes. When Westlaw refused, Ross hired a company to create "Bulk Memos" to train Ross's AI model; the Bulk Memos were created from Westlaw headnotes.[371] The district court found over 2,000 acts of direct infringement.[372]

The court rejected Ross's claim of fair use. It held that "using Thomson Reuters's headnotes as AI data to create a legal research tool to compete with Westlaw" was not transformative.[373] Further, it found that the copying could have an effect on Westlaw's market for its own current and future products.[374] The court therefore granted summary judgment to Thompson Reuters/Westlaw as to the specific infringed headnotes.[375]

While this is a promising example of a court rejecting a fair use defense by an AI company for training on copyrighted material, the decision in *Thompson Reuters* is limited by its context: legal search engine competitors.

If the court in *Authors Guild* does not reach the same conclusion, however, and it finds that the defendant's "training" constitutes fair use, either for all authors or (for example) only for the non-fiction authors, the settlement dynamics will shift in favor of the AI companies. Neither ruling is likely to destroy the incentive to reach some sort of settlement in these cases; the possibility of appellate or Supreme Court reversal will remain a threat. But the fair use determination will set the tone.

That said, no order of operations is risk-free here. Copyright class action plaintiffs on the west coast have opted for the traditional certification-first route. In *Tremblay v. OpenAI, Inc.*, the defendants have already succeeded in dismissing some of the plaintiffs' claims,[376] although not the direct infringement claims.[377]

---

[370] Thomson Reuters Enter. Ctr. GmbH v. Ross Intel. Inc., No. 20-CV-00613-SB, 2025 WL 458520, *1 (D. Del. Feb. 11, 2025).

[371] *Id.*; *see also id.* at *5 (finding that "a Bulk Memo question that looks more like a headnote than it does like the underlying judicial opinion is strong circumstantial evidence of actual copying").

[372] *Id.* at *5–6.

[373] *Id.* at *7.

[374] *Id.* at *9–10.

[375] *Id.* at *10.

[376] *See, e.g.*, Order Granting in Party and Denying in Part the Motions to Dismiss at 12, Tremblay v. OpenAI, Inc., No. 23-CV-03223 (N.D. Cal Feb. 12, 2024).

[377] Tremblay v. OpenAI, Inc., 716 F. Supp. 3d 772, 777 (N.D. Cal. 2024) (noting that defendants did not move to dismiss the plaintiffs' claims for direct infringement).

That dismissal will narrow the scope of potentially valuable discovery. Litigating class certification is complex and costly. In addition, the presence of a potential fair use defense may muddy the certification analysis. The possibility that the defendants have a valid fair use defense as to even some class members might hinder a court from finding that the plaintiffs have met their burden to show Rule 23 commonality and typicality. In contrast, the *Authors Guild* plaintiffs will have the Rule 23 wind at their back if they survive a fair use motion for summary judgment.

The decision to resolve fair use prior to certification has other collateral consequences as well, because there are competing class actions in different jurisdictions, with potential overlap in class members. Before class certification, absent class members have not yet been joined to a suit.[378] If the *Authors Guild* case suffers a serious setback prior to certification, or if certification is delayed, absent class members may migrate to *Tremblay* or a different case that reaches class certification and settlement sooner.

2. Authors Guild *Satisfies Rule 23(a) Factors*

To succeed on certification, plaintiffs must show that their class claims meet the four Rule 23(a) requirements of numerosity, commonality, typicality, and adequacy of representation.[379] As one scholar has noted, these quippy labels "mask the complexity of the case law and scholarship that has accreted over generations to give these terms their meanings."[380] Notwithstanding this complexity, and assuming that fair use is not an obstacle, the *Authors Guild* plaintiffs have strong arguments for all four 23(a) factors.

Rule 23(a)(1)—the requirement that the class be "so numerous that joinder of all members is impracticable"[381]—is satisfied. Notwithstanding the weak protests

[378] *See* Smith v. Bayer Corp., 564 U.S. 299, 315 (2011) (absent class members are not parties until a class has been certified, and therefore they are not bound).

[379] Wal-Mart Stores, Inc. v. Dukes, 564 U.S. 338, 345 (2011); FED. R. CIV. P. 23(a)(1)–(4).

[380] Jay Tidmarsh, *Diagnosis and Treatment of the "Superiority Problem,"* 69 VAND. L. REV. EN BANC 267, 267 (2016).

[381] FED. R. CIV. P. 23(a)(1).

of the defendants,[382] numerosity is not seriously in doubt.[383] Relatedly, by limiting the class membership to creators in the U.S. with registered copyrights,[384] the plaintiffs have defined the class in a way that is objective and verifiable, thus satisfying any potential ascertainability arguments. (Note: This is not necessarily true of the *Tremblay* plaintiffs in the Northern District of California, who are seeking to represent all copyright holders, whether registered or not.[385])

The *Authors Guild* plaintiffs can also likely meet their burden to show commonality under Rule 23(a)(2). Predictably, the plaintiffs' theory of the case centers on the unifying and uniform act of AI companies duplicating copyrighted works for purposes of "training" large language models. Specifically, the Consolidated Complaint alleges that "[d]efendants copied and data-mined the works of [class members], without permission or compensation."[386] In so doing, the plaintiffs allege, the registered copyrights of all class members "were infringed in materially the same way."[387]

At some level, this allegation may appear to be circular; it begs the question of what constitutes a material infringement. But the plaintiffs' claims are more straightforward than that: they allege that the AI defendants unlawfully and directly "copied Plaintiffs' works and then fed them into their 'large language models.'"[388] This is a simple, unifying claim of liability for direct infringement. Whether or not

[382] Defendant Microsoft Corp.'s Answer to First Consolidated Class Action Complaint at 48, Authors Guild v. OpenAI, Inc., Nos. 23-CV-08292, 23-CV-10211 (S.D.N.Y. Feb. 16, 2024) (Microsoft "denies that it possesses information concerning the exact number of members of Plaintiffs' putative class"). Note that Microsoft's response, denying knowledge of an "exact" number, strongly suggests that it might have an approximate number.

[383] First Consolidated Class Action Complaint at 61, Authors Guild v. OpenAI, Inc., Nos. 23-CV-08292, 23-CV-10211 (S.D.N.Y. Feb. 6, 2024) (alleging that "[b]oth Classes consist of at least tens of thousands of authors and copyright holders and thus are so numerous that joinder of all members is impractical"); *but see* Defendant Microsoft Corp.'s Answer to First Consolidated Class Action Complaint at 48, Authors Guild v. OpenAI, Inc., Nos. 23-CV-08292, 23-CV-10211 (S.D.N.Y. Feb. 16, 2024) (Microsoft "denies that it possesses information concerning the exact number of members of Plaintiffs' putative class").

[384] First Consolidated Class Action Complaint at 60–61, Authors Guild v. OpenAI, Inc., Nos. 23-CV-08292, 23-CV-10211 (S.D.N.Y. Feb. 6, 2024).

[385] First Consolidated Amended Complaint at 9, Tremblay v. OpenAI, Inc., No. 23-CV-03223 (N.D. Cal. Mar. 13, 2024).

[386] First Consolidated Class Action Complaint at 4, Authors Guild v. OpenAI, Inc., Nos. 23-CV-08292, 23-CV-10211 (S.D.N.Y. Feb. 6, 2024).

[387] *Id.* at 62.

[388] *Id.* at 3.

the defendants' actions are in fact unlawful, copying is the central question in the case for all class members.[389]

Plaintiffs' direct infringement claim also satisfies Rule 23(a)(3)'s typicality requirement, which tends to merge into the commonality analysis.[390] Typicality asks whether the claims of the class representatives align with those of the absent class members. In this case, the question is whether the fiction author representatives' claims mirror those of absent fiction authors, and whether the same dynamic is true for the non-fiction representatives vis-à-vis the absent non-fiction class members. It appears that they do. The class representatives allege that the defendant AI companies used all class members' copyrighted works in the same way and for the same general (albeit broad and evolving) purpose.[391] Beneath the technological complexity of AI, the direct infringement claims are simple and uniform—unlike the "Frankenstein" claims that the district court rejected in *Football Association*. In fact, the Second Circuit has found typicality in class actions with far more divergence among class members' experiences than exists here.[392] Moreover, unlike in *Football Association*, the classes in *Authors Guild* are defined concretely and objectively. They are limited to persons in the United States who possess (or are beneficial owners of) registered copyrights.[393] The claims are cohesive.

Finally, and for many of the same reasons, the *Authors Guild* classes satisfy Rule 23(a)(4) adequacy of representation. As just discussed, it does not appear that either the Authors Guild or the individual class representatives have interests that are misaligned with those of the absent class members. In the *Google Book Project* case, Google alleged that many plaintiff class members supported

[389] Wal-Mart Stores, Inc. v. Dukes, 564 U.S. 338, 350 (2011) ("What matters to class certification … is not the raising of 'common questions'—even in droves—but, rather the capacity of a classwide proceeding to generate common answers apt to drive the resolution of the litigation.").

[390] *See* Gen. Tel. Co. of Sw. v. Falcon, 457 U.S. 147, 157 n.13 (1982).

[391] First Consolidated Class Action Complaint at 13, Authors Guild v. OpenAI, Inc., Nos. 23-CV-08292, 23-CV-10211 (S.D.N.Y. Feb. 6, 2024) (quoting U.S. Patent & Trademark Off., Public Views on Artificial Intelligence and Intellectual Property Policy 29 (2020) for the proposition that AI training "almost by definition involve[s] the reproduction of entire works or substantial portions thereof").

[392] *See* Marisol A. v. Giuliani, 126 F.3d 372, 376–77 (2d Cir. 1997) (approving district court's finding of commonality and typicality in class of children who suffered from different aspects of child welfare system).

[393] First Consolidated Class Action Complaint at 60–61, Authors Guild v. OpenAI, Inc., Nos. 23-CV-08292, 23-CV-10211 (S.D.N.Y. Feb. 6, 2024).

Google's efforts and disagreed with the class claims, thus calling adequacy of representation into question. Although it decided the case on other grounds, the Second Circuit acknowledged that Google's argument "may carry some force."[394] As of now, however, there is no significant outcry by creators who favor giving their intellectual property to AI companies at no cost. To be sure, there are academics and policymakers who favor AI companies' free use of copyrighted material.[395] But there is as of yet no sign of a serious fissure among creators that would undermine adequacy of representation. In fact, the Library Copyright Alliance, which supports a finding of fair use for AI copying, submitted comments to the Copyright Office urging that office to defer to courts' resolution of these questions.[396] As to the second element of (a)(4) adequacy—the quality of the lawyering—there is no serious likelihood that the court or the defendants will argue that the renowned attorneys representing the class are in any way inadequate.[397]

At first glance, it may also appear that the fame and wealth of named plaintiffs like John Grisham and Jodi Picoult could undermine typicality and/or adequacy in a class composed largely of minor authors. After all, Grisham and other best-selling authors might argue that the defendants' actions resulted in vastly larger damages. But that should not prove to be a barrier. Copyright law provides for two possible forms of damages: actual and statutory.[398] It is often difficult to calculate actual damages. In lieu of that calculation, a plaintiff may elect at any time before final judgment to receive statutory damages for each act of infringement.[399] The

[394] *Google Book Project*, 721 F.3d 132, 134 (2d Cir. 2013).

[395] *See, e.g.*, Mark A. Lemley & Bryan Casey, *Fair Learning*, 99 Tex. L. Rev. 743 (2021); Andrew W. Torrance & Bill Tomlinson, *Training Is Everything: Artificial Intelligence, Copyright, and "Fair Training,"* 128 Dick. L. Rev. 233, 250 (2023) (proposing a "fair training exception" to copyright infringement).

[396] The Library Copyright Alliance and other groups submitted comments to the U.S. Copyright Office arguing that "remuneration for ingestion is neither appropriate nor feasible." Libr. Copyright All., Comments of the Library Copyright Alliance on the Inquiry Concerning Artificial Intelligence and Copyright 1 (2023), https://www.librarycopyrightalliance.org/wp-content/uploads/2023/10/CO-AI-NOI-Final.pdf [https://perma.cc/GXH3-UYZH]. However, those comments also urge the Copyright Office to defer making any policy given that the judiciary is "the most appropriate forum" for resolving creators' claims. *Id.* at 3.

[397] The plaintiff class is represented by Susman Godfrey LLP, Lieff Cabraser Heimann & Bernstein LLP, and Cowan Debaets Abrahams & Sheppard LLP. Defendants did not oppose this. First Consolidated Class Action Complaint at 67–68, Authors Guild v. OpenAI, Inc., Nos. 23-CV-08292, 23-CV-10211 (S.D.N.Y. Feb. 6, 2024).

[398] 17 U.S.C. § 504(a).

[399] 17 U.S.C. § 504(c)(1).

precise amount of a statutory damages award depends on a court's assessment of the seriousness of the infringement and the defendant's ability to pay; it does not depend on a showing of the damages suffered by the copyright owner.[400]

The *Authors Guild* Complaint notes that, should the class elect statutory damages, "the damages inquiry will likewise be common, if not identical," among all members.[401] The simplicity of statutory damages bolsters the classes' claims of commonality, typicality, and adequacy. It also raises the specter of a massive damages award.

3. Authors Guild *Satisfies Rule 23(b)*

The *Authors Guild* complaint asserts that class certification would be appropriate under all of the Rule 23(b) categories.[402] In addition, it alleges that, as an alternative to certification of the entire dispute, the court could bifurcate the proceeding into liability and damages under its power to certify a class only as to certain issues.[403] This would entail first certifying a liability-only class under Rule 23(b)(2), for the purposes of resolving the common question of whether the defendants are liable for infringement for copying class members' works,[404] while leaving the question of damages to another process on another day.

Despite the class certification laundry list in the pleadings, however, the plaintiffs' most straightforward and effective option is to obtain certification under the most rigorous of the Rule 23(b) options, Rule 23(b)(3), and they are in a strong position to do that on their direct infringement claims. This form of certification had distinct advantages. First, it will allow any objecting authors to opt out of the class and either pursue claims elsewhere or drop any potential claim—thus minimizing the force of any arguments that the class members are not united. Second, unlike

[400] Statutory damages awards may range from $750 at the low end, to $150,000 per infringement, if the infringer acted willfully. 17 U.S.C. § 504(c)(1)–(2).

[401] First Consolidated Class Action Complaint at 63, Authors Guild v. OpenAI, Inc., Nos. 23-CV-08292, 23-CV-10211 (S.D.N.Y. Feb. 6, 2024).

[402] *Id.* at 60, 62–63.

[403] Rule 23(c)(4) allows courts to certify "a class action with respect to particular issues." For a thoughtful analysis of this provision, see Myriam Gilles & Gary Friedman, *The Issue Class Revolution*, 101 B.U. L. Rev. 133, 136 (2021).

[404] First Consolidated Class Action Complaint at 62–63, Authors Guild v. OpenAI, Inc., Nos. 23-CV-08292, 23-CV-10211 (S.D.N.Y. Feb. 6, 2024).

(b)(2) classes, in which class members may only obtain declaratory or injunctive relief, (b)(3) classes may seek money damages.

As explained above, to obtain certification under Rule 23(b)(3), the *Authors Guild* class representatives will need to show what are referred to in shorthand as predominance and superiority. This they can do.

3.1. *Rule 23(b)(3) Predominance*

The predominance requirement requires plaintiffs to show that the common issues are at the heart of their case, and that any differences among plaintiffs are of only secondary importance. As discussed above, this provision is pragmatic rather than formalistic; courts have discretion as to how they interpret and apply it.[405]

But predominance should not be an insuperable barrier to certification of input copyright classes. As of now, the *Authors Guild* plaintiffs' strong arguments in favor of commonality and typicality carry over directly to this predominance argument: The plaintiffs are alleging identical conduct by the AI defendants, and they have invoked their right to seek statutory damages, which would minimize the need for thousands of fact-specific individual determinations.

3.2. *Rule 23(b)(3) Superiority*

The superiority requirement asks whether a class action is the best option for resolution of the claims at issue.[406] Rule 23(b)(3) lists five non-exclusive factors that courts should consider in evaluating superiority, including whether individual class members would have an interest in proceeding separately; the nature of any ongoing litigation in other courts; the appropriateness of the forum, and potential difficulties in managing a class suit.[407]

As one scholar has recently observed, courts apply this requirement in different and sometimes haphazard ways.[408] But the murkiness of superiority in class action doctrine writ large should not detract from its application to this case:

---

[405] *See supra* Part IV (discussing Rule 23(b)(3)).

[406] FED. R. CIV. P. 23(b)(3)(A)–(D).

[407] *Id.*

[408] Christine P. Bartholemew, *The Failed Superiority Experiment*, 69 VAND. L. REV. 1295, 1297 (2016) (noting that courts "infuse superiority with a conflicting assortment of factors, stripping away any cohesive screening function").

a class action is the superior way to resolve the case of thousands of identically situated creators and to give repose to AI defendants. The fiscal and logistical impossibility of individual small-time authors taking on AI companies makes it applicable.

Moreover, the district court in this case has already, albeit indirectly, analyzed some of the superiority factors. The west-coast plaintiffs in *Tremblay* moved to intervene in *Authors Guild* for the purpose of moving to dismiss, stay, or transfer the New York action to the Northern District of California.[409] In denying the motion, the district court stated its view that the *Authors Guild* classes had brought suit in an appropriate forum.[410] The court's order also makes clear its view that the "nature of any litigation concerning the controversy already commenced by or against members of the class is not a barrier to the New York proceedings.[411] This finding closely parallels the superiority finding required under Rule 23(b)(3)(B).[412]

B. The Input Class: Andersen v. Stability AI, Ltd.[413]

The *Authors Guild* case concerns AI models' training on text; *Anderson v. Stability AI* is about images. The named plaintiffs in this proposed class action—artists including Karla Ortiz, Gerald Brom and Jingna Zhang[414]—allege that defendants Runway, Stability AI, DeviantArt, and Midjourney downloaded billions of copyrighted images without permission for the purpose of creating and/or using an AI product called Stable Diffusion.[415]

Specifically, the plaintiffs claim that Runway worked with Stability to help train and then distribute Stable Diffusion, an AI image generator whose work

[409] Authors Guild v. OpenAI, Inc., 345 F.R.D. 585, 589 (S.D.N.Y. 2024), *appeal dismissed sub nom.* Guild v. Tremblay, No. 24-1007, 2024 WL 4564683 (2d Cir. Oct. 4, 2024), and *appeal dismissed sub nom.* Basbanes v. Microsoft Corp., No. 24-1014, 2024 WL 4564684 (2d Cir. Oct. 4, 2024).

[410] *Id.* at 592 (denying motion to intervene by west-coast plaintiffs on the ground that it would prejudice the Authors Guild plaintiffs' right to proceed in their chosen forum, and that it would "disrupt the expedited timeline agreed to by the parties").

[411] Fed. R. Civ. P. 23(b)(3)(B).

[412] *Id.* (requiring a court to evaluate "the extent and nature of any litigation concerning the controversy already begun by or against class members" prior to certifying a Rule 23(b)(3) class).

[413] Andersen v. Stability AI, Ltd., No. 23-CV-00201 (N.D. Cal. filed Jan. 13, 2023).

[414] First Amended Complaint at 1, Andersen v. Stability AI, Ltd., No. 23-CV-00201 (N.D. Cal. Nov. 29, 2023).

[415] *Id.* at 1.

products are "in the style of" the named plaintiffs and other artists.[416] The training images came from so-called LAION datasets.[417] Midjourney trained on Stable Diffusion, and DeviantArt used the model in its own image-generating product, DreamUp. In addition, plaintiffs allege that defendant Midjourney "has repeatedly promoted the use of artist names—including Plaintiffs' names—within text prompts as a means of getting better results."[418] These companies now market and sell AI image products; the plaintiffs allege that "what they're really selling is copyright infringement as a service."[419]

The named plaintiffs have alleged several causes of action against each defendant. In addition to their input claim—direct copyright infringement—they assert claims for induced copyright infringement, vicarious copyright infringement, violation of the DMCA and the Lanham Act, as well as common law claims for unjust enrichment.[420] They characterize defendants' AI image-generating products as "copyright-laundering devices, promising customers the benefits of art without the costs of artists."[421]

Unlike in *Authors Guild*, the parties in this suit did not stipulate to a non-traditional order of operations. Instead, the defendants moved to dismiss for failure to state a claim. Details aside, the plaintiffs' claims largely survived this motion. The district court found that "the plausible inferences at this juncture are that Stable Diffusion by operation by end users creates copyright infringement and was created to facilitate that infringement by design."[422]

With regard to the specific claims, none of the defendants moved to dismiss the direct copyright infringement claims based on scraping of and "training" with

---

[416] *Id.* at 73.

[417] LAION stands for "Large-Scale Artificial Intelligence Open Network." LAION is a German non-profit organization whose aim is "to make large-scale machine learning models, datasets and related code available to the general public." *About*, LAION, https://laion.ai/ [https://perma.cc/5GEC-92WD] (last visited Oct. 15, 2024). Note that in December 2023, LAION temporarily took down its datasets in response to reports that they contained child sex abuse materials. *Safety Review for LAION 5B*, LAION (Dec. 19, 2023), https://laion.ai/notes/laion-maintenance/ [https://perma.cc/6QKV-HLYK].

[418] First Amended Complaint at 3, Andersen v. Stability AI Ltd., No. 23-CV-00201 (N.D. Cal. Nov. 29, 2023).

[419] *Id.*

[420] *Id.* at 10.

[421] *Id.* at 4.

[422] Andersen v. Stability AI Ltd., 744 F. Supp. 3d 956, 969 (N.D. Cal. 2024).

copyrighted images. The court rejected defendants' motion to dismiss plaintiffs' induced copyright infringement claims.[423] For all defendants, the court dismissed the plaintiffs' DMCA claims with prejudice, and their common law unjust enrichment claims with leave to amend.[424] The court ruled variously on the other claims against each defendant:

- Stability AI: the district court granted the defendants' motion to dismiss the plaintiffs' DMCA claims as well as their unjust enrichment claims.[425] The court declined to dismiss the plaintiffs' induced infringement claim.[426]

- Runway: the district court denied Runway's motion to dismiss direct infringement claims based on plaintiffs' argument that Stable Diffusion is itself an "infringing statutory copy" of plaintiffs' works, and/or that distributing Stable Diffusion is the equivalent of distributing plaintiffs' works.[427] It also denied Runway's motion to dismiss the induced infringement claims.[428]

- Midjourney: the district court denied Midjourney's motion to dismiss plaintiffs' copyright claims; Midjourney had contended that the plaintiffs' pleadings did not plausibly allege that the artists' work was covered by registered copyrights.[429] The court also denied Midjourney's motion to dismiss plaintiffs' Lanham Act trademark claims.[430]

- DeviantArt: the district court denied DeviantArt's motion to dismiss copyright claims.[431] It granted the motion to dismiss plaintiffs' breach of contract claims and unjust enrichment claims.[432]

---

[423] *Id.* at 967–69, 975.
[424] *Id.* at 973 (Stability AI); *id.* at 975 (Runway).
[425] *Id.* at 969–71 (dismissing with prejudice claims under Sections 1202(a) and 1202(b)(1)); *id.* at 971–73 (dismissing state common law unjust enrichment claims but granting leave to amend).
[426] *Id.* at 967–69.
[427] *Id.* at 974–75.
[428] *Id.* at 975.
[429] *Id.* at 976–77
[430] *Id.* at 981.
[431] *Id.* at 984.
[432] *Id.* at 985.

The named representatives in Andersen are seeking class certification under Rules 23(b)(2) and (b)(3).[433] They break their (b)(3) class into three sub-classes, based on from which database the copyrighted images were downloaded.[434] They also seek separate certification of a (b)(3) class composed of all artists whose names were invoked in prompts by the Midjourney Image Product.[435] These multiple class claims are more nuanced than those in *Authors Guild*. But the Rule 23 analysis for direct infringement is functionally the same.

Discovery is ongoing. Based on currently available information, however, the plaintiffs can satisfy Rule 23(a) and (b) for their direct copyright claims. As with *Authors Guild*, numerosity is not controversial: the plaintiffs allege that billions of copyrighted artistic works, paired with descriptive captions, were copied without permission.[436] And as with *Authors Guild*, the driving force of the litigation—defendants' allegedly unlawful replication of plaintiffs' copyrighted works—meets the commonality and typicality requirements.[437] The First Amended Complaint states the common question as "whether Defendants violated the copyrights of Plaintiffs and the Class when they downloaded and stored copies of [class members'] Works … and … used copies of the Works to train AI Image Products."[438] This allegation of centralized and uniform conduct is the heartland of commonality and typicality. Finally, as of now there are no allegations that either the named representatives or the highly competent attorneys representing the class are inadequate to represent the rights of absent class members.

The logic of *Authors Guild* also supports certification of the Andersen plaintiffs' (b)(2) and (b)(3) classes for direct copyright infringement. Again, the centerpiece of certification under both subprovisions is the uniform conduct of the defendants—conduct that affects the plaintiff visual artists in more or less identical

[433] First Amended Complaint at 8, Andersen v. Stability AI Ltd., No. 23-CV-00201 (N.D. Cal. Nov. 29, 2023).

[434] *Id.*

[435] *Id.* at 9.

[436] *Id.* at 1.

[437] *See supra* text accompanying note 290 (defining commonality and typicality).

[438] First Amended Complaint at 10, Andersen v. Stability AI Ltd., No. 23-CV-00201 (N.D. Cal. Nov. 29, 2023).

ways. In fact, while analytically *Authors Guild* and *Andersen* are very similar, it is likely to be easier—more intuitive—to show copying of images than it is of texts.[439]

Of course, as with *Authors Guild*, the plaintiffs in *Andersen* will need to prevail against a fair use summary judgment motion. But, as explained below, we argue that fair use should not be extended to give enormous, for-profit technology companies free license to replicate copyrighted works.

VI
Stealing is Not Fair Use

In the early days of computers, Steve Jobs was known for bragging, "Good companies copy, great companies steal," while Bill Gates begged hobbyists to stop copying Microsoft software without paying. Ironically, Apple subsequently sued Microsoft for stealing the displays of Apple's Macintosh computer user interface in Windows 2.0. Perhaps both companies are great, and stealing is at the heart of Big Tech, both then and now.

Fair use, however, was not at issue in that stealing case. Stealing is and should never be deemed fair, as both great companies understand the fundamentals of good business and corporate reputation. Wholesale stealing, in particular, like what gave rise to the AI input class actions, is never fair.[440]

The defendants in the AI input class actions are riding on their fair use defense to avoid infringement liability. In addition to *Oracle*, they will rely on *Field v. Google* and *Authors Guild v. Google* for support. Unlike Google in these three cases, the AI companies are not engaging in "caching" webpages, nor indexing snippets for users to search, nor using a miniscule amount of essential code; instead, they are brazenly stealing wholesale literary text and images to train their models. AI models need copyrighted content and without the creative works of authorship

[439] *See* Elizabeth G. Porter, *Taking Images Seriously*, 114 Colum. L. Rev. 1687, 1753 (2014) (describing research showing that people perceive complex information more quickly and easily in images than in text).

[440] In addition to the "input" class cases, here are some illustrations of stealing in the "output" claims. For instance, the plots, stories, and characters in the Harry Potter books are copyrighted but any user can prompt Copilot/ChatGPT to generate multi-choice questions for the sorting hat to place someone into one of the four Hogwarts houses. By copying the entire contents of Harry Potter books, generative AI models learned about the unique characteristics of Gryffindor, Hufflepuff, Ravenclaw, and Slytherin, and generated the sorting questions. Moreover, generative AI models copied and learned about different magical creatures crafted by J.K. Rowling and produced sorting questions about Hogwarts houses and the corresponding magical creatures. *See* Appendix A.

input into the AI models, there will be no AI models. The stealing itself is not transformative. The act of copying or scraping creative works belonging to others is not transformative. The act at that moment does not create anything new to benefit anyone except AI models that have been valued at billions of dollars and enriched Big Tech and Big AI companies. Without acknowledging the existence of copyrights and without compensating creators, the stealing avoids payments to creators and disregards existing human creativity.

Furthermore, like the defendant in *American Geophysical Union v. Texaco, Inc.* who made copies of academic articles for their corporate use and faced the court's ruling that the act of copying the articles was not transformative, AI defendants copy the entirety of literary text and images for their own corporate use of building their AI models likewise is just plain copying without pay. Also, the stealing in the AI input class constitutes some exceptions under copyright law as seen in the *MAI* Trio cases, because the stealing is massive, permanent, and intentional to build powerful models of artificial intelligence for worldwide domination.

At the moment when the stealing occurs, the input claim of copyright infringement is established. That means the act at that moment of copying and scraping does not transform the literary text or images into something else except AI models, which are trained to commit the second type of copyright infringement by violating the right to prepare derivative works. AI companies will attempt to conflate the input and output claims, injecting end users as a shield for liability. At the input stage, however, there are no users to enter the prompts because the AI companies are busy stealing copyrighted works to complete the training of the AI models. Users enter the scheme only after AI companies debut their AI models long after the AI models have ingested and trained on all the copyrighted works.

The second factor of fair use is the nature of the copyrighted works. The creators will easily satisfy this factor against fair use because their literary text and nonfiction contain the level of creativity entitled to protection in fair use analysis as discussed in *Harper & Row v. Nation Enterprises.*[441] Likewise, on the third factor, the amount and substantiality of the portion used in relation to the copyrighted

[441] Harper & Row Publishers, Inc. v. Nation Enters., 471 U.S. 539, 569 (1985) (rejecting fair use defense raised by the Nation for summarizing and quoting from President Ford's forthcoming memoir of his decision to pardon Richard Nixon).

work as a whole favors the creators because the AI defendants steal the copyrights in their entirety. The last factor, the effect of the use on the potential market for the value of the copyrighted work, also favors against a finding of fair use. Indeed, as we saw in the discussion of the last factor by the court in *American Geophysical Union v. Texaco, Inc.*, the existence of a licensing model to establish royalty for creators makes stealing unfair. Also, AI defendants themselves openly admitted the licensing model for non-literary content as they paid Reddit millions for the license to use the platform's user content.

Stealing is never fair to victims, and society. Stealing is never fair when Big AI and Big Tech with enormous resources are teaming up to extract the heart and soul of the creativity class. Fair use defense is not available shielding AI defendants in the input claim. AI defendants must pay under a licensing model.

VII
Fair Licensing

The Framers enshrined the Intellectual Property Clause to "promote the Progress of Science and useful Arts" into the Constitution by securing "for limited Times" to authors "the exclusive Right" to their "Writings."[442] At the time of the Framing, "promote the Progress of Science" referred to "the creation and spread of knowledge and learning" from works of authorship.[443] The young nation needed authors and new works of authorship as expressed and embodied in literature, music, theatre, journalism, fine arts, and architecture for development and growth.[444] The works of authorship, they believed—key contributors to

[442] *See generally* Edward C. Walterscheid, *To Promote the Progress of Science and Useful Arts: The Background and Origin of the Intellectual Property Clause of the United States Constitution*, 2 J. Intell. Prop. L. 1, 1 n.1 (1994) ("[Article I, section 8, clause 8] is frequently referred to as either the Patent Clause, the Copyright Clause, or the Intellectual Property Clause, depending on the context in which it is being discussed."); Dotan Oliar, *Making Sense of the Intellectual Property Clause: Promotion of Progress as a Limitation on Congress's Intellectual Property Power*, 94 Geo. L.J. 1771, 1845 n.1 (2006).

[443] *ArtI.S8.C8.1 Overview of Congress's Power Over Intellectual Property*, Cong., https://constitution.congress.gov/browse/essay/artI-S8-C8-1/ALDE_00013060/ [https://perma.cc/B2ZA-MZFK] (last visited Mar. 23, 2025) (quoting Golan v. Holder, 565 U.S. 302, 324 (2012)).

[444] Works of authorship include eight categories:

(1) literary works;
(2) musical works, including any accompanying words;
(3) dramatic works, including any accompanying music;
(4) pantomimes and choreographic works;

humanity and to defining a new nation—must be cultivated and incentivized through uniform federal protection.[445] Without legal protection, others could freely copy authors' works, reducing the incentive to create works of authorship in the young nation and thereafter.[446] Executing this constitutional mandate without delay, the new Congress passed the first federal copyright law during its first session in 1790.[447]

Copyright law embraces human creativity, and the level of creativity necessary for a work of authorship qualified for protection is minimal. The modicum of creativity means new works of authorship are constantly in the stream of creation from the Founding time to the present. As long as the works possessing a modicum of creativity are independently created by the authors and fixed in a tangible medium of expression, the works enjoy the exclusive rights enshrined in the Constitution. An email, a tweet,[448] a poem, a song, a drawing, a painting, an essay, an investigative report, a photograph, a blueprint, a chevron design on a cheerleader uniform,[449] a musical score, a dance, a recording, a video, a game, a homemade movie, a pamphlet, and a compilation of data,[450] regardless of its aesthetic appeals, high or low culture, captivating or distasteful, are entitled to federal protection. Human creativity, in all forms, drives the nation's soul, as the Constitution's "for the progress of Science" recognizes.

---

(5) pictorial, graphic, and sculptural works;
(6) motion pictures and other audiovisual works;
(7) sound recordings; and
(8) architectural works.

17 U.S.C. § 102.

[445] Bonito Boats, Inc. v. Thunder Craft Boats, Inc., 489 U.S. 141, 162 (1989) ("One of the fundamental purposes behind the [IP Clause] was to promote national uniformity in the realm of intellectual property."); The Federalist No. 43 (James Madison); Goldstein v. California, 412 U.S. 546, 556 (1973).

[446] *See supra* note 443.

[447] Copyright Act of 1790, ch. 15, 1 Stat. 124 (copyrights).

[448] *Are Tweets Protected by Copyright?*, Copyright All., https://copyrightalliance.org/faqs/tweets-protected-copyright/ [https://perma.cc/6MXX-8KEL] (last visited Oct. 15, 2024).

[449] Star Athletica, L.L.C. v. Varsity Brands, Inc., 580 U.S. 405, 417–18 (2017); Staci Zaretsky, *Supreme Court Says Decorative Fashion Design Elements Protected by Copyright Law*, Above the L. (Mar. 22, 2017), https://abovethelaw.com/2017/03/supreme-court-says-decorative-fashion-design-elements-protected-by-copyright-law/ [https://perma.cc/8UMB-DYXV].

[450] Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340 (1991).

AI companies and their Big Tech backers, as seen in the class action complaints, brazenly take works of authorship, from literary to nonfiction, photographs, and images to music works, to develop their dataset. They steal to build their models. They disregard the constitutionally mandated copyright protection for human creativity as expressed in eight categories of works of authorship.[451] Their AI foundation models generate output infringing on copyrighted works and populating the online content with synthetic commons where fakes are quickly replacing human-created text, images, and sounds. The theft of human creativity should be addressed by compensating the authors for their losses. Before we discuss "how," a detour on copyright law on damages is illuminating.

Federal protection of works of authorship encompasses a robust enforcement system against infringers. Indeed, copyright law accords the copyright owner with the option to elect statutory damages in the amount of no less than $750 and no more than $30,000 per work being infringed.[452] If the infringer committed the act(s) willfully, the statutory damages per work can increase up to $150,000.[453] Statutory damages are only available in copyrights, not other types of intellectual property infringements, because Congress recognized that actual damages and profits are difficult to establish by copyright owners, and desired to compensate the copyright owners for the losses to their creativity as expressed in works of authorship. Statutory damages also serve to prevent copyright infringements which are easy to accomplish by copyists. To obtain statutory damages, copyright owners must register their copyrights within three months of publication or before the infringement begins.

Returning to "the how," we suggest fair licensing to compensate the creative working class who are the authors of copyrighted works of authorship being taken with neither permission nor compensation from AI companies and their Big Tech backers. What is licensing? What is fair?

Licensing is an old business model in the intellectual property field. Bill Gates sold licenses or the right to use Microsoft's software which is protected

[451] 17 U.S.C. § 102.
[452] 17 U.S.C. § 504(c)(1).
[453] 17 U.S.C. § 504(c)(2).

under copyright law to end-users who are individuals and business entities.[454] In demanding that people pay for his copyrighted software instead of using it without permission, Gates asked, "Who can afford to do professional work for nothing?" and sparked the licensing of software.[455] Microsoft dictates the prices of different licenses, and the individuals and business entities pay. The prices are fair market value because Microsoft, the seller, and the individuals and business entities, the buyers, are unrelated, and they agree to engage in the transactions without pressure.[456] In other words, Microsoft and other Big Tech companies know the licensing of copyrights well and their lawyers understand the clients' licensing-centered business.

Moreover, the music industry is known for its savviness in licensing copyrights; their experience can shape licensing deals. Further, under the compulsory licenses for music, everyone can use others' songs for programming, digital, and physical sound recording releases, without permission from the composition rightsholders but must pay the statutory mechanical rate for each copy sold on the song covered. The rate is 9.1 cents per song or 1.75 cents per minute of playing time, whichever is greater.[457] Spotify, Apple Music, and other digital streaming services pay blanket licenses for mechanical royalties, and parties on both sides of blanket licenses can furnish the licensing terms as samples to AI copyright class action cases.[458]

---

[454] Irfan Mirza, *Bill Gates and Paul Allen's Greatest Invention*, LinkedIn (Apr. 8, 2015), https://www.linkedin.com/pulse/bill-gates-paul-allens-greatest-invention-irfan-mirza/ [https://perma.cc/BS3X-QYNA] ("[O]n February 3, 1976, Bill wrote an open letter to computer hobbyists explaining that unless there is an equitable licensing model for computer software, it will 'prevent good software from being written.' Therein lies the magic of Bill and Paul's greatest invention—the software licensing model.").

[455] William Henry Gates III, *An Open Letter to Hobbyists*, N.Y. Times (Feb. 3, 1976), https://archive.nytimes.com/www.nytimes.com/library/cyber/surf/072397mind-letter.html [https://perma.cc/58V7-XMEZ].

[456] James Chen, *Fair Market Value (FMV): Definition and How to Calculate It?*, Investopedia (June 15, 2024), https://www.investopedia.com/terms/f/fairmarketvalue.asp [https://perma.cc/5BQZ-FDES] (Fair market value (FMV) is the price of an asset when "buyer and seller are reasonably knowledgeable about the asset" and "are free of undue pressure.").

[457] *What Is a Compulsory License?*, Song Trust, https://help.songtrust.com/knowledge/what-is-a-compulsory-license [https://perma.cc/KP3W-DYUB] (last visited Oct. 15, 2024).

[458] Virginie Chelles, *How Music Licensing Works for Music Streaming Services?*, TunedGlobal, https://blog.tunedglobal.com/how-music-licensing-works-for-music-streaming-services [https://perma.cc/BP64-LNVS] (last visited Oct. 15, 2024); Jason Koransky, *Digital Dilemmas: The Music Industry Confronts Licensing for On-Demand Streaming Services*, Am. Bar Ass'n (Jan. 1,

Likewise, the number of licensing deals AI companies have secured with various social media platforms, media entities, and publishing companies intimates the current pricing lists of different types of content for AI companies to train their models.[459] These deals demonstrate that AI companies know how to value and pay for the content types that they need. What constitutes fair market value is what the parties decide in the deals that they both have reasonable knowledge of the copyrighted works, the use of the copyrighted works, and their value, and they act under no undue pressure.

Regarding representation in negotiating for fair terms in licensing for the creative working class, the lawyers for the plaintiffs in the AI class action cases are sophisticated law firms. They are capable of negotiating with AI companies and their Big Tech backers for a fair market value of different licenses of copyrights. Using the statutory damages provisions as guideposts, $750 to $150,000 per work of authorship, the parties through their lawyers can reach sensible licensing terms to achieve the widely recognized fair market value concept.

Moreover, for the licensing to be fair to the creative class who typically possess neither resources nor sophistication for negotiation, we suggest three key terms for consideration. First, the scope of the license grant must satisfy four components: (i) the royalty setting must be fairly commensurate with the individual creator's works, (ii) the license grant must preserve the author's control over the use of their copyrights in subsequent technologies, unknown at the time of the license but later developed, (iii) the authors are entitled to receive notice of any proposed changes to the scope of the license grant, and (iv) and licensees' unilateral expansion of the scope of the license grant constitutes intentional copyright infringement. These components will help tilt the balance from AI defendants/licensees to the creators.

Second, the creators must also retain control over their copyrights in addition to the control relating to the scope of the license grant. Here, the authorship of the

2016), https://www.americanbar.org/groups/intellectual_property_law/publications/landslide/2015-16/january-february/digital-dilemmas-music-industry-confronts-licensing-on-demand-streaming-services/ [https://perma.cc/68DV-WDSD].

[459] Spangler, *supra* note 17; Robertson, *supra* note 17; David, *supra* note 17; Christy et al., *supra* note 17; Ohlheiser, *supra* note 17; Morris, *supra* note 17.

creators' works must be acknowledged. This requires appropriate attribution to the original human creators who created the works before machines.

Lastly, creators' autonomy must be respected. That means creators have the right to opt-in or opt-out of the licensing scheme at any time during the duration of the licensing scheme.

Licensing, though not a perfect model, is the only viable option to righting the wrongs, the stealing by AI companies of works of authorship produced by the creative class. The massive stealing must end. With the three key terms outlined above, we believe that the lawyers of both sides can negotiate towards an agreement to conclude the AI input claims, the first in the two steps in the preservation of humanity.

Conclusion

Upholding copyright law in the AI age will not stymie innovation; this is not about fiction authors barring technology from accurately diagnosing cancer. Nor is upholding copyright law against AI companies administratively impossible. In a world where AI can instantaneously accomplish so many tasks, it can also accomplish the task central to the constitutional concept of copyright: preserving and respecting the small, individual, and fundamentally human contributions of creators to national and international culture. The pending AI suits allow these individual creators to aggregate to take on AI defendants—the ultimate technological aggregators. We believe that courts can, and should, use long-existing copyright law to recognize and vindicate the copyrights of the creative working class.

****

Appendix A

Conducted on October 1, 2024, 12:45 p.m.

Sent by you:
generate multiple choice questions for sorting Hogwarts houses

Sent by Copilot:
Sure! Here are some multiple-choice questions to help sort someone into a Hogwarts house:

1. **Which quality do you value most?**
    - A) Bravery
    - B) Loyalty
    - C) Intelligence
    - D) Ambition
2. **What would you do if you found a lost wallet?**
    - A) Try to find the owner yourself
    - B) Hand it over to the authorities
    - C) Look for clues to find the owner
    - D) Keep it for yourself
3. **Which magical creature do you find most fascinating?**
    - A) Phoenix
    - B) Hippogriff
    - C) Thestral
    - D) Basilisk
4. **How do you prefer to spend your free time?**
    - A) Going on adventures
    - B) Helping friends
    - C) Reading books
    - D) Planning your future

5. **What is your greatest fear?**
    - A) Failure
    - B) Betrayal
    - C) Ignorance
    - D) Weakness
6. **Which of these subjects would you excel in?**
    - A) Defense Against the Dark Arts
    - B) Herbology
    - C) Charms
    - D) Potions
7. **How do you handle conflict?**
    - A) Face it head-on
    - B) Seek a peaceful resolution
    - C) Analyze the situation
    - D) Use it to your advantage
8. **What is your ideal pet?**
    - A) Lion
    - B) Badger
    - C) Eagle
    - D) Snake
9. **Which of these quotes resonates with you the most?**
    - A) “Courage is not the absence of fear.”
    - B) “Loyalty is the strongest bond.”
    - C) “Knowledge is power.”
    - D) “Ambition drives success.”

10. **What is your approach to teamwork?**
    - A) Lead the team
    - B) Support the team
    - C) Strategize for the team
    - D) Ensure the team wins

These questions can help identify traits associated with Gryffindor, Hufflepuff, Ravenclaw, and Slytherin. Have fun sorting! 🧙‍♀️🧙

****

Sent by you:
generate crypto-related questions sorting Hogwarts houses

Sent by Copilot:
Sure! Here are some crypto-related multiple-choice questions to help sort someone into a Hogwarts house:

1. **Which aspect of cryptocurrency excites you the most?**
    - A) The potential for revolutionary change
    - B) The sense of community and collaboration
    - C) The innovative technology behind it
    - D) The opportunities for financial gain
2. **How would you handle a sudden drop in the value of your crypto investments?**
    - A) Stay calm and hold on
    - B) Discuss strategies with fellow investors
    - C) Analyze the market trends and data
    - D) Look for ways to capitalize on the situation

3. **What type of blockchain project would you most likely support?**
    - A) One that aims to disrupt traditional systems
    - B) One that focuses on social good and inclusivity
    - C) One that introduces groundbreaking technology
    - D) One that promises high returns on investment
4. **How do you prefer to learn about new cryptocurrencies?**
    - A) Through hands-on experience and experimentation
    - B) By joining online forums and communities
    - C) By reading whitepapers and technical documents
    - D) By following market trends and expert advice
5. **What is your approach to securing your crypto assets?**
    - A) Use a combination of methods to ensure safety
    - B) Share tips and best practices with others
    - C) Stay updated with the latest security technologies
    - D) Invest in the most secure and profitable options
6. **Which of these crypto-related activities appeals to you the most?**
    - A) Mining or staking
    - B) Participating in airdrops and giveaways
    - C) Developing or coding blockchain solutions
    - D) Trading and investing
7. **How do you view the future of cryptocurrency?**
    - A) As a tool for global change and empowerment
    - B) As a way to build stronger communities
    - C) As a field ripe for innovation and discovery
    - D) As a lucrative investment opportunity

8. **What would you do if you discovered a new, promising cryptocurrency?**
    - A) Invest immediately and take the risk
    - B) Share the information with your network
    - C) Research thoroughly before making a decision
    - D) Evaluate its potential for high returns
9. **Which of these statements resonates with you the most?**
    - A) "Fortune favors the bold."
    - B) "Together, we are stronger."
    - C) "Knowledge is the key to success."
    - D) "Ambition knows no bounds."
10. **What is your strategy for staying informed about the crypto market?**
    - A) Follow the latest news and trends
    - B) Engage with the crypto community
    - C) Study market analysis and reports
    - D) Monitor investment opportunities closely

These questions can help identify traits associated with Gryffindor, Hufflepuff, Ravenclaw, and Slytherin in the context of cryptocurrency. Enjoy sorting!

****

Sent by you:
Generate law school-related questions sorting Hogwarts houses

Sent by Copilot:
Sure! Here are some law school-related multiple-choice questions to help sort someone into a Hogwarts house:

1. **Which area of law interests you the most?**
    - A) Criminal Law
    - B) Family Law
    - C) Intellectual Property Law

- D) Corporate Law

2. **How do you handle a challenging case?**
    - A) Tackle it head-on with determination
    - B) Collaborate with colleagues for the best outcome
    - C) Research extensively to find the best solution
    - D) Strategize to turn the case in your favor
3. **What motivates you to become a lawyer?**
    - A) Fighting for justice
    - B) Helping those in need
    - C) The intellectual challenge
    - D) Achieving success and recognition
4. **How do you prepare for a big exam?**
    - A) Study diligently and practice under pressure
    - B) Form study groups and support each other
    - C) Create detailed outlines and review notes
    - D) Focus on key areas to maximize your score
5. **What is your approach to legal ethics?**
    - A) Uphold justice at all costs
    - B) Ensure fairness and compassion
    - C) Follow the rules and guidelines meticulously
    - D) Navigate the system to benefit your client
6. **Which legal figure do you admire the most?**
    - A) Ruth Bader Ginsburg
    - B) Thurgood Marshall
    - C) Oliver Wendell Holmes Jr.
    - D) Clarence Darrow

7. **How do you handle courtroom pressure?**
    - A) Stay calm and assertive
    - B) Rely on teamwork and support
    - C) Use logic and evidence to persuade
    - D) Use your wit and charm to influence
8. **What is your ideal legal career path?**
    - A) Public defender or prosecutor
    - B) Legal aid or non-profit work
    - C) Academic or legal scholar
    - D) Corporate lawyer or partner at a firm
9. **Which of these quotes resonates with you the most?**
    - A) "Justice delayed is justice denied."
    - B) "Injustice anywhere is a threat to justice everywhere."
    - C) "The law is reason, free from passion."
    - D) "The ends justify the means."
10. **How do you approach legal writing?**
    - A) Write with passion and conviction
    - B) Ensure clarity and empathy
    - C) Focus on precision and detail
    - D) Make persuasive and compelling arguments

These questions can help identify traits associated with Gryffindor, Hufflepuff, Ravenclaw, and Slytherin in the context of law school. Good luck with your sorting!