U.C.L.A. Law Review

Fair Use Defenses in Disruptive Technology Cases

Pamela Samuelson

ABSTRACT

The fair use limitation on the scope of authorial exclusive rights is expected to be the main defense in lawsuits charging generative artificial intelligence (AI) developers with copyright infringement. These lawsuits—brought by and on behalf of some authors, artists, and copyright industry rights holders—challenge the legality of the unauthorized use of in-copyright works for purposes of developing generative AI systems, which they allege has greatly harmed the plaintiffs. The outcome of these copyright infringement suits will likely turn on courts' assessments of the fair use doctrine's market effects factor: whether and to what extent generative AI is harming or is likely to harm plaintiffs' existing and expected markets.

Controversies about new and unexpected uses of copyrighted works enabled by technological advances are far from new; they continue to play an integral part in the development of modern copyright law. This Article revisits numerous disputes over uses of in-copyright works enabled by disruptive technologies since the 1970s. These disputes involved copying of research and educational materials, home video and audio taping, and storage and processing of digital copies for the purpose of developing search engines and other research tools. Tracing the evolution of the concept of market harm and the role it has played in copyright fair use cases involving new technologies, this Article lays out the arguments about market effects—harmful and beneficial effects, as well as the requirements of proof—that have been presented in the courts and debated in Congress. The Article then assesses the current landscape of generative AI copyright litigation. Insights gained from the past can be instrumental in thinking about conflicts involving generative AI, as well as any future technologies that are new, innovative, and disruptive.

AUTHOR

Richard M. Sherman Distinguished Professor of Law, Berkeley Law School.

I am grateful for the opportunity to deliver the annual Melville Nimmer Lecture at UCLA, given in November 2023, for which this Article was prepared.

TABLE OF CONTENTS

INTRODUCTION 1486
I. FAIR USE AND THE CONCEPT OF MARKET HARM 1488
II. FAIR USE DEFENSES AND CLAIMS OF COPYRIGHT HARMS IN DISRUPTIVE TECHNOLOGY CASES 1493
A. Copying of Research and Educational Materials 1494
1. Library Photocopying of Scientific Articles 1494
2. Photocopying Controversies in Congress 1499
3. An Interlude on First Generation Photocopying Controversies 1502
4. For-Profit Firm Photocopying of Scientific & Technical Journal Articles 1504
5. Educational Coursepack Cases 1507
6. Online Course Materials 1510
7. An Interlude on the Role of Licensing in Research and Educational Use Copying Cases 1513
B. Home Video and Audio Taping 1516
1. Home Video Taping 1517
2. An Interlude on Sony and the Concept of Copyright Harms 1522
3. Home Video Recording Revisited 1527
4. Home Audio Taping 1529
5. Peer-to-Peer Filesharing 1534
6. An Interlude on Audio Taping and P2P File-Sharing 1537
C. Computational Search and Research Tools 1538
1. Search Engine Webcrawling and Storage of Copyrighted Works 1539
2. Thumbnail Images and In-Line Linking for Web Search 1541
3. Mass Digitization to Enable Search Indexes and Analytic Research Tools 1542
III. GENERATIVE AI AS A COPYRIGHT-DISRUPTIVE TECHNOLOGY 1547
A. Copyright Lawsuits Challenging Generative AI 1548
1. Overview of the Generative AI Copyright Claims 1549
2. Outputs-as-Derivative-Works Claims 1551
3. Training Data Ingestion Claims 1556
B. Questioning Proposals for a Generative AI Collective License Regime 1564
CONCLUDING THOUGHTS 1569

INTRODUCTION

Generative artificial intelligence (AI) is the latest disruptive technology to enable new and unexpected uses of in-copyright works.[1] Some authors and artists strenuously object to the use of their works as training data for generative AI models and to their outputs, such as texts or pictures, which they say unfairly competes with their human creations.[2] Several complaints have charged makers of generative AI systems with copyright infringement.[3] They allege, among other things, that the ingestion of in-copyright works as training data infringes on authors' exclusive right to reproduce their works in copies and that AI outputs infringe on their exclusive rights to prepare derivative works.[4] The complaints typically claim that generative AI has greatly harmed the plaintiffs, for which they are owed many millions or even

1. *See, e.g.*, David De Cremer, Nicola Morini Bianzino & Ben Falk, *How Generative AI Could Disrupt Creative Work*, HARV. BUS. REV. (Apr. 13, 2023), https://hbr.org/2023/04/how-generative-ai-could-disrupt-creative-work [https://perma.cc/6ERC-8M56]. For an excellent primer about generative artificial intelligence (AI) systems and how models are trained, as well as the copyright issues they raise, see Katherine Lee, A. Feder Cooper & James Grimmelmann, *Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain*, J. COPYRIGHT SOC'Y (forthcoming 2024).
2. *See, e.g.*, Naomi Klein, *AI Machines Aren't 'Hallucinating.' But Their Makers Are*, THE GUARDIAN (May 8, 2023, 4:02 AM), https://www.theguardian.com/commentisfree/2023/may/08/ai-machines-hallucinating-naomi-klein [https://perma.cc/GE5K-NEVM]; Michael Cuenco, *We Must Declare Jihad Against AI*, COMPACT (Apr. 28, 2023), https://compactmag.com/article/we-must-declare-jihad-against-a-i [https://perma.cc/5LGG-TPNE].
3. *See, e.g.*, Complaint Class Action at 1, Andersen v. Stability AI Ltd., No. 3:23-cv-00201 (N.D. Cal. Jan. 13, 2023) [hereinafter Andersen Complaint]; Class Action Complaint at 2, Authors Guild v. OpenAI, Inc., No. 1:23-cv-8292 (S.D.N.Y. Sept. 19, 2023) [hereinafter Guild Complaint]; Complaint at 4, Getty Images (US), Inc. v. Stability AI, Inc., No. 1:23-cv-00135 (D. Del. Feb. 3, 2023) [hereinafter Getty Complaint]; Complaint ¶ 67, New York Times Co. v. Microsoft Corp., No. 1:23-cv-11195 (S.D.N.Y. Dec. 27, 2023) [hereinafter NYT Complaint]; Complaint Class Action at 1, Silverman v. OpenAI, Inc., No. 3:23-cv-03416 (N.D. Cal. July 7, 2023) [hereinafter Silverman Complaint]. Most of the legal challenges are class actions, but the *Getty Images* and *New York Times* cases are not. Copies of complaints against generative AI companies can be found at Master List of Lawsuits v. AI, ChatGPT, OpenAI, Microsoft, Meta, Midjourney & Other AI Cos., CHAT GPT IS EATING THE WORLD (Mar. 11, 2024), https://chatgptiseatingtheworld.com/2023/12/27/master-list-of-lawsuits-v-ai-chatgpt-openai-microsoft-meta-midjourney-other-ai-cos [https://perma.cc/4ZD4-M7VA] [hereinafter Master List of Lawsuits].
4. *See, e.g.*, Guild Complaint, *supra* note 3, ¶¶ 111–28. Section 106 of the 1976 Copyright Act (1976 Act) sets forth the copyright owner's exclusive rights. *See* 17 U.S.C. § 106(1) (covering reproduction rights), § 106(2) (covering derivative work rights). For an overview of the copyright issues in the generative AI litigations, see, for example, Pamela Samuelson, *Generative AI Meets Copyright*, 381 SCIENCE 158 (2023) (describing copyright lawsuits and likely defenses).

billions of dollars in damages.[5] Some assert they have lost licensing fees for use of their works as training data,[6] while other critics of generative AI urge the U.S. Congress to establish a collective licensing regime so that authors and artists can get paid for the value their works have contributed to generative AI systems.[7] Fair use is likely to be the main defense generative AI companies will raise in the lawsuits.[8]

Copyright controversies arising from new uses of copyrighted works enabled by technological advances are far from new. Virtually all of the copyright challenges brought against developers of disruptive technologies that enable new uses of copyrighted works have called for an interpretation of the fair use limitation on copyright's exclusive rights. This Article traces the evolution of judicial fair use analyses in the disruptive technology cases and shows that fair uses are likely to be found unless there is evidence that copyright owners' markets have been harmed or are likely to be harmed as a result of the copying enabled by the challenged technologies. Part I provides an overview of the law of fair use and the significance of harm to copyright markets in fair use cases. Part II discusses three types of disruptive technologies that have triggered copyright controversies addressing perceived harms that the technologies or their use allegedly caused. Part III considers lessons that might be gleaned from the resolution of past copyright controversies involving disruptive technologies when courts address claims

5. *See, e.g.*, Andersen Complaint, *supra* note 3, ¶¶ 153–244.
6. *See, e.g.*, Getty Complaint, *supra* note 3, ¶ 56.
7. *See, e.g.*, Emell Adolphus, *Authors Guild Meets With Legislators Over Generative AI Concerns*, PUBLISHERS WKLY. (Apr. 5, 2023), https://www.publishersweekly.com/pw/by-topic/industry-news/licensing/article/91920-authors-guild-meets-with-legislators-over-generative-ai-concerns.html [https://perma.cc/JVH6-M48C].
8. As of this writing, all but one of the generative AI lawsuits are in very early stages, so fair use defenses have not yet been raised, but several legal articles anticipate that fair use defenses will be raised. *See, e.g.*, Peter Henderson, Xuechen Li, Dan Jurafsky, Tatsunori Hashimoto, Mark A. Lemley & Percy Liang, *Foundation Models and Fair Use*, 24 J. MACH. LEARNING RSCH., no. 400, 2023, at 1, 4; Matthew Sag, *Copyright Safety for Generative AI*, 61 HOUS. L. REV. 295, 301 (2023); Mark A. Lemley & Bryan Casey, *Fair Learning*, 99 TEX. L. REV. 743, 745–46 (2021); Benjamin L. W. Sobel, *Artificial Intelligence's Fair Use Crisis*, 41 COLUM. J.L. & ARTS 45, 49–50 (2017); Andrew W. Torrance & Bill Tomlinson, *Training Is Everything: Artificial Intelligence, Copyright, and "Fair Training"*, 128 DICKINSON L. REV. 233, 236 (2023); *see also* CHRISTOPHER T. ZIRPOLI, CONG. RSCH. SERV., LSB10922, GENERATIVE ARTIFICIAL INTELLIGENCE AND COPYRIGHT LAW 3–4 (2023), https://crsreports.congress.gov/product/pdf/LSB/LSB10922 [https://perma.cc/AR9C-PRAR]; Oren Bracha, *The Work of Copyright in the Age of Machine Production*, 38 HARV. J.L. & TECH (forthcoming 2024) (explaining why non-expressive use defenses should prevail over fair use); Andres Guadamuz, *A Scanner Darkly: Copyright Liability and Exceptions in Artificial Intelligence Inputs and Outputs*, 73 GRUR INT'L 111, 118 (2024) (discussing input and output copyright issues under United Kingdom and European law).

of harms to copyright markets resulting from uses of in-copyright works for generative AI.

I. FAIR USE AND THE CONCEPT OF MARKET HARM

Fair use is a limitation on all of the exclusive rights that Congress has granted to authors in their original works of authorship.[9] Congress codified the fair use limitation in § 107 of the Copyright Act of 1976 (1976 Act).[10] It states that fair uses are not infringements and identifies several favored purposes—criticism, comment, news reporting, teaching, scholarship, and research—that tip in favor of fair use, but not dispositively so.[11] Noncommercial uses also tend to be favored over commercial uses, but courts have recognized that many common fair uses are "indisputably commercial."[12] They have interpreted the statutorily favored uses as illustrative, not exhaustive; consequently, unanticipated purposes (including those in some of the new technology cases discussed in this Article) may nonetheless qualify as fair uses.[13]

Section 107 identifies four factors to consider in fair use cases: (1) the purpose and character of the challenged use; (2) the nature of the copyrighted work; (3) the amount and substantiality of the taking; and (4) the effect of the challenged use on the potential market for and value of the copyrighted work.[14] This fourth factor is sometimes said to be the most

9. 17 U.S.C. § 107. All authorial exclusive rights—as set forth in §§ 106, 106A, and 602(a)—are "[s]ubject to" the limitations and exceptions set forth in §§ 107–122. Violations of exclusive rights are infringements under § 501(a).
10. Before the 1976 Act, fair use was a common law limitation on copyright. *See* ALAN LATMAN, STUDY NO. 14: FAIR USE OF COPYRIGHTED WORKS, *in* STAFF OF S. COMM. ON THE JUDICIARY, 86TH CONG., COPYRIGHT LAW REVISION (Comm. Print 1960) (discussing fair use case law). Latman's was one of several studies prepared during the two-decades-long copyright revision process leading to the enactment of the 1976 Act.
11. 17 U.S.C. § 107; 4 WILLIAM F. PATRY, PATRY ON COPYRIGHT § 10:13 (2024).
12. Google LLC v. Oracle Am., Inc., 593 U.S. 1, 32 (2021) (discussing commercial and noncommercial uses listed in 17 U.S.C. § 107); *see also id.* at 40 (concluding that the use of parts of the Java application program interface (API) was a fair use despite its commercial purpose). *But see* Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith, 598 U.S. 508, 537 (2023) ("The undisputed commercial character of AWF's use, though not dispositive, 'tends to weigh against a finding of fair use.'") (quoting Harper & Row, Publishers v. Nation Enters., 471 U.S. 539, 562 (1985)).
13. *See, e.g.*, Pamela Samuelson, *Unbundling Fair Uses*, 77 FORDHAM L. REV. 2537, 2602–15 (2009) (giving examples of unforeseen fair uses).
14. 17 U.S.C. § 107(1)–(4). Courts sometimes consider other factors. *See, e.g.*, Harper & Row Publishers, Inc. v. Nation Enters., 471 U.S. 539, 563 (1985) (noting, in addition to the four

important,[15] although all factors must be considered and balanced together to resolve the fair use issue. Courts often take account of the public interests affected by fair use rulings and construe fair use in light of copyright's constitutional purposes.[16]

The plain text of § 107 refers to market effects as the fourth factor to consider in fair use cases. This language implies that challenged uses might have positive effects on plaintiffs' markets,[17] and a few cases have considered positive effects.[18] But fair use cases generally focus on whether the allegedly infringing uses have harmed or are likely to harm markets for plaintiffs' works. Courts commonly consider whether the challenged use "supplants . . . the normal market" or "supersed[es] . . . demand for the original" work, as well as the demand for "potentially licensed derivatives."[19] Although courts and commentators in the 1980s sometimes opined that evidence of harm should be inferred from copying,[20] more recent rulings have required a showing of actual harm or a meaningful likelihood of future harm to undermine fair use defenses, sometimes adding that the harm must be substantial enough to affect authorial incentives.[21]

---

fair use factors, that the manuscript used to prepare defendant's article was "purloined"). Most fair use cases are decided on summary judgment. *See* Barton Beebe, *An Empirical Study of U.S. Copyright Fair Use Opinions, 1978–2005*, 156 U. PA. L. REV. 549, 570 (2008). But juries sometimes decide fair use cases, as in *Google*, 593 U.S. at 15–16.

15. *See Harper & Row*, 471 U.S. at 566.

16. *See, e.g.*, *Google*, 593 U.S. at 30, 35–36. See *infra* Part II for further examples, including *Williams & Wilkins Co. v. United States*, 487 F.2d 1345, 1349 (Ct. Cl. 1973), *aff'd by an equally divided Court*, 420 U.S. 376 (1975), *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417, 451 (1984), *Authors Guild v. HathiTrust*, 755 F.3d 87, 96 (2d Cir. 2014), *Authors Guild v. Google, Inc.*, 804 F.3d 202, 207 (2d Cir. 2015), and the search engine cases. This Article refers to the *Authors Guild, Inc. v. HathiTrust* and *Authors Guild v. Google* cases collectively as "*Google Books*."

17. *See* David Fagundes, *Market Harm, Market Help, and Fair Use*, 17 STAN. TECH. L. REV. 359, 377–85 (2014) (discussing the possible positive effects of certain uses, such as creator recognition and affirmation).

18. *See* Jeanne C. Fromer, *Market Effects Bearing on Fair Use*, 90 WASH. L. REV. 615, 638–39 (2015).

19. *Harper & Row*, 471 U.S. at 568 (quoting S. REP. NO. 94-473, at 65 (1975)); Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 587–88 (1994).

20. *See, e.g.*, Apple Comput., Inc. v. Franklin Comput. Corp., 714 F.2d 1240, 1254 (3d Cir. 1983) (recognizing "the prevailing view that a showing of a prima facie case of copyright infringement . . . raises a presumption of irreparable harm" in the preliminary injunction context); PAUL GOLDSTEIN, GOLDSTEIN ON COPYRIGHT § 1.14.2.3(c), at 1:57–1:58 (3d ed. Supp. 2023-2); *see also infra* Part II.B.2.

21. Incentivizing authors to create and disseminate ideas has often been stated as an important purpose of copyright. *See, e.g.*, Harper & Row, 471 U.S. at 558; *see also infra* Parts II.A.5, II.B.2 (discussing incentives).

Various theories about market effects have been tested in new technology cases discussed in Part II. Owing to the influence of U.S. Supreme Court fair use precedents, courts have given considerable weight to evidence of harm or its absence. When there is evidence defendants' uses caused market harm, as the Court found in *Harper & Row Publishers, Inc. v. Nation Enterprises*, fair use defenses tend to fail.[22] When courts decide that plaintiffs' claims of market harm are speculative, as the Court concluded in *Sony Corporation of America v. Universal City Studios, Inc.*,[23] fair use defenses tend to succeed. When evidence of market effects is more equivocal, as in *Google LLC v. Oracle America, Inc.*,[24] courts tend to weigh evidence of harms and benefits in relation with other factors, including public benefits.

The new technology fair use cases have often grappled with the market harm circularity problem. When copyright owners sue for infringement, they often claim that unauthorized uses of their works caused harm because the defendants should have, but did not, license rights to use their works.[25] By that logic, plaintiffs could always defeat fair use defenses by asserting that the failure to license harmed their markets. But this argument is circular since it would, in effect, preclude defendants from raising a successful fair use defense as long as plaintiffs expressed a desire to grant a license for the use. Yet, courts generally recognize that acceding to such claims would defeat Congress's purpose in codifying fair use.[26] Of course, if licensing markets already exist and defendants bypass them, that is likely to tip against a fair use defense. Yet, failure to license may not be dispositive.[27] When licensing

22. 471 U.S. at 567 (detailing a magazine's cancellation of an agreement to publish excerpts of a memoir). The Court did not directly address the market effects factor in its recent decision in *Warhol v. Goldsmith* because the only issue it addressed was whether a particular use was transformative. Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith, 598 U.S. 508, 525 (2023). Yet, the Court did emphasize that whether a second work supplants demand for the original is relevant to whether it is transformative. *Id.* at 528.
23. 464 U.S. 417, 451 (1984). *Sony* is discussed *infra* Part II.B.1, II.B.2.
24. 593 U.S. 1, 35–40 (2021).
25. *See, e.g.*, Am. Geophysical Union v. Texaco, Inc., 60 F.3d 913, 929–32 (2d Cir. 1994) (holding that it was not fair use to make copies of journal articles for which a licensing market existed).
26. *See, e.g.*, *infra* text accompanying notes 143, 201–203 (describing the courts' recognition of a circularity problem); *see also Google*, 593 U.S. at 38 (quoting 4 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 13.05[A][4]) (describing the "danger of circularity" in a plaintiff's licensing argument).
27. *See, e.g.*, Bill Graham Archives v. Dorling Kindersley Ltd., 448 F.3d 605, 615 (2d Cir. 2006) (holding transformative fair use where publishers reproduced small images of concert posters to illustrate the history of the Grateful Dead band despite their failure to license

markets do not already exist, courts may consider whether they are reasonably likely to develop.[28] Courts may also consider possible market failures, such as when transaction costs are very high relative to the usage value, such that a market is unlikely to form.[29]

The Court's 1994 decision in *Campbell v. Acuff-Rose Music, Inc*. refined fair use analysis, including the concept of market effects, in several respects. The ruling is best known for holding that "transformative" uses of copyrighted works weigh in favor of fair use defenses.[30] The Court defined that term as uses that "add[] something new, with a further purpose or different character, altering the first with new expression, meaning, or message."[31] When follow-on creators make transformative uses of preexisting works, such as Luther R. Campbell's parody of Roy Orbison's popular song, the Court thought there was less reason to infer that the challenged work would serve as a market substitute.[32] No one, for example, would purchase 2 Live Crew's rap parody if they wanted to listen to Roy Orbison's version of "Oh, Pretty Woman." After *Campbell*, numerous new technology cases have viewed defendants' uses as transformative because the defendants had very different purposes compared to the purpose of the original, even if they copied the whole or substantial parts of the plaintiff's work.[33]

---

the images); Sega Enters. Ltd. v. Accolade, Inc., 977 F.2d 1510, 1514 (9th Cir. 1992) (determining the fair use defense was not undermined by the existence of a licensing market in which Accolade chose not to participate); *see also Google*, 593 U.S. at 37–38 (observing that the existence of a licensing market for uses of the Java API did not undermine Google's fair use defense).

28. See *infra* Part II.A.4 for a discussion of *Texaco*.
29. The significance of market failure in fair use contexts is discussed *infra* text accompanying notes 261–265.
30. Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 579 (1994).
31. *Id. But see* Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith, 598 U.S. 508, 525, 541–42 (2023) (recognizing a new meaning or message may be relevant to whether a second comer's use will be considered transformative, but dropping that phrase from its definition of transformation).
32. The Court identified several reasons why transformative uses pose a lesser risk of market substitution, such as "the large extent of transformation of the earlier work, the new work's minimal distribution in the market, the small extent to which it borrows from an original, or other factors." *Campbell*, 510 U.S. at 580 n.14.
33. *See infra* Part II.C (discussing *Field v. Google, Inc.*, 412 F. Supp. 2d 1106, 1109 (D. Nev. 2006), *A.V.* ex rel. *Vanderhye v. iParadigms, LLC*, 562 F.3d 630, 633–34 (4th Cir. 2009), *Kelly v. Arriba Soft Corp.*, 336 F.3d 811 (9th Cir. 2003), *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146 (9th Cir. 2007), and two *Google Books* cases); *see also Warhol*, 598 U.S. at 525 (stating a second work is more likely to be considered transformative if it has a different purpose than the original).

Campbell also rejected Acuff-Rose's argument that the defendant's use was unfair because he had sought, but not obtained, a license to make the rap parody of the Orbison song.[34] The Court suggested that seeking a license might, in fact, be a sign of good faith in an effort to avoid litigation.[35] The Court further recognized that copyright owners are unlikely to license lampoons or other critical commentaries of their works, which then "removes such uses from the very notion of a potential licensing market."[36] The only markets courts should consider are those in which copyright owners would generally license or develop a licensing market. Market substitution, the Court noted, is "the only harm to derivatives that need concern us."[37] Moreover, the Court emphasized the need to ensure that the challenged use was causing copyright-relevant harms. For example, if "a lethal parody, like a scathing theater review, kills demand for the original, it does not produce a harm cognizable under the Copyright Act."[38] These refinements to the market effects factor in fair use cases have been influential in subsequent cases.[39]

Yet, the Court remanded *Campbell* for an evidentiary inquiry into the possible effects of the use on a market for rap music derivatives: "Evidence of substantial harm to [such a market] would weigh against a finding of fair use, because the licensing of derivatives is an important economic incentive to the creation of originals."[40] Thus, the Court insisted that the market effects factor be decided based on evidence of actual market harm or harm to foreseeable markets or the lack thereof, further linking market effects with incentives to create.[41] It repudiated the presumption of market harm on which the Sixth Circuit relied in ruling against Campbell's fair use defense.[42]

The Supreme Court further refined the market effects factor in *Google LLC v. Oracle America, Inc.* by directing courts to consider "the public benefits the [defendant's] copying will likely produce,"[43] not just potential harms, and

34. *See Campbell*, 510 U.S. at 593.
35. *See id.* at 585 n.18.
36. *Id.* at 592.
37. *Id.* at 593.
38. *Id.* at 591–92.
39. After *Campbell*, numerous courts have concluded that copyright owners are not entitled to exercise control over transformative fair use markets. *See, e.g.*, Bill Graham Archives v. Dorling Kindersley Ltd., 448 F.3d 605, 614–15 (2d Cir. 2006) (holding fair use to include unlicensed images of concert posters in a book about the Grateful Dead).
40. *Campbell*, 510 U.S. at 593 (citation omitted).
41. *See id.*
42. *See id.* at 590–91. The Sixth Circuit had cited *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417, 451 (1984), for this presumption.
43. Google LLC v. Oracle Am., Inc., 593 U.S. 1, 35 (2021).

what effect a court's fair use ruling would have for ongoing creativity as part of the market effects inquiry.[44] The Court's most recent fair use decision, *Andy Warhol Foundation for the Visual Arts v. Goldsmith*, reiterated that public benefits should be given due consideration in fair use cases, although giving more weight than in *Google* to second comers' commercial purposes.[45]

II. FAIR USE DEFENSES AND CLAIMS OF COPYRIGHT HARMS IN DISRUPTIVE TECHNOLOGY CASES

In the last quarter of the twentieth century, numerous lawsuits and legislative proposals targeted unauthorized copying of in-copyright works that were made possible and more efficient because of technological advances. As discussed in this Part, some lawsuits also challenged the development of infringement-enabling technologies. Several appellate court decisions resolving these disputes were closely divided.[46] In virtually all cases, evidence of market harms (or lack thereof) resulting from unauthorized copying of plaintiffs' works was a very significant consideration.[47] Challenges to technologies rarely succeeded in litigation, but copyright industry groups sometimes persuaded Congress to regulate technologies to make them less efficient at making unauthorized copies of copyrighted works and to impose compulsory licensing aimed at compensating copyright owners for unauthorized copying of their works.[48] This Part focuses first on photocopying disputes, second on home taping controversies, and third on computational use lawsuits.

44. *See id.* at 39–40.
45. *See* Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith, 598 U.S. 508, 530–31 (2023). *Warhol* addressed only whether the Foundation made a transformative use of a photograph on which a Warhol print was based when it licensed use of the print as a magazine cover. *See id.* at 525–26. Because Goldsmith's and the Foundation's works had the same purpose and competed in the same market, the Court concluded the use was nontransformative. *Id.* at 525–26.
46. *See, e.g.*, *infra* Parts II.A.1 (discussing *Williams & Wilkins*); II.A.4 (discussing *Texaco*); II.B.1 (discussing *Sony*).
47. *See, e.g.*, *infra* text accompanying notes 67–72 (discussing *Williams & Wilkins*); *infra* Part II.B.2 (discussing *Sony* and copyright harms); *infra* Part II.C.3 (discussing market effects (harm) in the *Google Books* cases).
48. *See infra* text accompanying 309–315 (discussing the Audio Home Recording Act of 1992); *infra* text accompanying notes 500–503(describing existing copyright compulsory license regimes).

A. Copying of Research and Educational Materials

Xerox machines, which produce high quality photocopies, were the bane of publishers' existence during much of the second half of the twentieth century. Several copyright lawsuits and legislative proposals challenged the widespread use of these machines by librarians and educators to make unauthorized copies of in-copyright works.[49] While publishers claimed all copies of in-copyright works were infringements, librarians and educators defended them as fair uses.[50] Publishers expressed concerns that photocopying threatened their core businesses, but their main goal in challenging photocopying was their desire to develop new licensing markets, which could happen only if they could persuade courts or Congress to characterize the copying as infringement.[51] Clashes over photocopying (and later, online course reserves[52]) have yielded sharp disagreements among the judges deciding them. Publishers won some cases and lost others, depending on whether courts were persuaded by their market harm arguments.[53]

1. Library Photocopying of Scientific Articles

To test its theory that library photocopying of research articles was copyright infringement,[54] Williams & Wilkins Co. (W&W), a publisher of

49. *See, e.g.*, Stephen Breyer, *The Uneasy Case for Copyright: A Study of Copyright in Books, Photocopies, and Computer Programs*, 84 HARV. L. REV. 281, 329–30 (1970). In the late 1960s, there were more than 800,000 photocopy machines in businesses, libraries, and other organizations. *See* Project, *New Technology and the Law of Copyright: Reprography and Computers*, 15 UCLA L. REV. 939, 943 (1968) [hereinafter UCLA Project]. The New York Public Library made more than one million photocopies per year for patrons. *Id.* at 942.
50. *See* Breyer, *supra* note 49, at 329–30.
51. *See* UCLA Project, *supra* note 49, at 941–44, 961–75.
52. Professors teaching courses at universities sometimes ask librarians to post digitized copies of course readings on library websites so that students can access them. Online course reserve copying of in-copyright works was the subject of a publisher lawsuit against a university, discussed *infra* Part II.A.6 (reviewing *Cambridge Univ. Press* (*CUP*) *v. Patton*, 769 F.3d 1232 (11th Cir. 2014)).
53. *See infra* Part II.A. In the cases discussed in this section, parties won or lost at different stages of litigation, which, together with the split decisions in these cases, reflect the range of disagreement on the market effects (harm) factor.
54. The UCLA Project may have emboldened Williams & Wilkins Co. (W&W) to sue because its report opined that photocopying of scientific and educational materials constituted infringement. *See* UCLA Project, *supra* note 49, at 952–57. The author-editors claimed that the project was an "empirical study" because it was based on interviews with various stakeholders, including with W&W's president. *Id.* at 939, 1029. The report,

thirty-seven medical research journals, sued the United States because librarians at the National Institutes of Health (NIH) and National Library of Medicine (NLM) made unauthorized copies of eight articles from four of its journals.[55] Discovery disclosed that the NIH and NLM routinely made copies of scientific articles for affiliated researchers from thousands of journals to which they subscribed.[56] In 1970, for example, the NIH satisfied 85,744 researcher requests for copies of scientific articles, resulting in more than 930,000 copied pages.[57] The eight copies from W&W journals obviously represented a very small part of these libraries' photocopying activities.

Initially, a trial judge found this photocopying to be an infringement.[58] But by a 4–3 majority, an appellate panel of the Court of Claims decided that the NIH and NLM had made fair use copies of those eight articles, all but one of which were more than two years old.[59] The majority noted that it was only deciding this case on its facts and not making broader statements about photocopying of journal articles or other materials.[60] The court first considered the constitutional purpose of copyright, in service of which courts "must occasionally subordinate the copyright holder's interest in a maximum financial return to the greater public interest in the development of art, science and industry."[61] It emphasized that the NIH and NLM were both nonprofit institutions "devoted solely to the advancement and dissemination of medical knowledge," and that the copies were made for scientific

---

however, unequivocally endorsed publishers' positions and dismissed librarian and educator perspectives. *See id.* at 954 ("[T]he doctrine of fair use, at least as presently construed, ought not survive the reprographic revolution."); *see also* Melville B. Nimmer, *Foreword: Two Copyright Crises*, 15 UCLA L. REV. 931, 931 (1968) (predicting that "[t]he day may come when conventional book and periodical publishing will be almost completely supplanted, individually or in combination, by machine copying devices and sophisticated computer operations[,]" causing "a copyright crisis of profound dimensions"). In the modern era, publishers have been able to use digital rights management (DRM) technology and paywalls to protect their works from unbridled copying.

55. *See* Williams & Wilkins Co. v. United States, 487 F.2d 1345, 1349 (Ct. Cl. 1973), *aff'd by an equally divided Court*, 420 U.S. 376 (1975).
56. *Id.* at 1347–48.
57. *Id.* at 1348. To satisfy researcher requests, the National Institute of Health employed four staff and had an annual budget of $86,000 for library photocopying in 1970. *Id.*
58. *Id.* at 1347.
59. *See id.* at 1349, 1363. The eighth article copied was twenty-one or twenty-two months old. *Id.* at 1349. Although W&W emphasized the large number of copies made by these libraries, it lacked standing to sue on behalf of other publishers and their copyrights. Only the eight articles in which it held copyright were before the Court of Claims.
60. *See id.* at 1362–63.
61. *Id.* at 1352 (quoting Berlin v. E.C. Publ'ns, Inc., 329 F.2d 541, 544 (2d Cir. 1964)).

research purposes.[62] It then asserted that "the law gives copying for scientific purposes a wide scope."[63] The majority pointed to long-accepted historical library practices favoring research copying,[64] a 1935 "gentlemen's agreement" between publishers and libraries about the permissibility of copying for scholarly research purposes,[65] and implicit support from certain previous congressional enactments.[66] Copying entire articles did not undermine fair use, given that the libraries had imposed strict limits on the number and types of requests researchers could make.[67] The libraries also considered page lengths and the publication's age in deciding whether to copy articles for patrons.[68]

Concerning the market effects of library photocopying, the majority regarded as "very important that [the] plaintiff has failed to prove its assumption of economic detriment, in the past or potentially for the future."[69] Although the trial judge inferred that publishers had lost revenues from the large volume of NIH and NLM photocopying, the majority thought the "record made in this case does not sustain that assumption."[70] It observed that W&W had experienced a rise in subscriptions and taxable income; in addition, the four journals at issue constituted a small part of its growing business.[71] W&W's willingness to grant a license to the NIH and NLM was not evidence that it was suffering harm from the libraries' photocopying.[72] The majority was very aware that Congress was considering what to do about library and educational use copying as part of the copyright revision process, and opined that it was for Congress, not the courts, to decide whether to create a compulsory licensing regime for copying of the sort at stake in *W&W*.[73] An additional consideration tipping toward fair use was harm to the progress of

62. *Id.* at 1354.
63. *Id.*
64. *See id.* at 1351 (discussing Library of Congress policies dating back to 1909).
65. *Id.* at 1355–56.
66. *Id.* at 1356.
67. *Id.* at 1348–49.
68. *Id.* at 1354–55.
69. *Id.* at 1359.
70. *Id.* at 1357.
71. *See id.*
72. *Id.* at 1358–60.
73. *See id.* at 1360.

science, which the majority foresaw if it ruled that library photocopying was unfair.[74]

The dissenting judges regarded the scale of the libraries' copying to be "so vast that it dwarfs the output of many small publishing companies."[75] They thought the copying of entire articles, each of which was separately copyrighted, was unfair because it supplanted demand for the original.[76] They characterized the libraries' claim of a fair use privilege for scientific research purposes as "far-fetched."[77] They were sympathetic to difficulties copyright plaintiffs have in proving harm and asserted that proof of actual damages was not required as long as challenged uses tended to diminish or prejudice the copyright owners' interests.[78] The dissenters also quoted from a UCLA School of Law study that unequivocally denounced fair use defenses for library copying of scientific and technical articles.[79] They pointed out that W&W was not trying to stop the libraries from making copies; it just wanted a reasonable royalty, the setting of which was within the court's power.[80]

Given the deep split in the appellate panel's ruling and the strong public interest in resolution of this important issue, the Supreme Court granted W&W's petition to review the fair use ruling. During oral argument, noted copyright scholar Alan Latman,[81] representing W&W, argued that library

74. *See id.* at 1356–57. The court cited a study expressing that scientific researchers had an urgent need to access literature relevant to their work. *See id*. at 1356 (citing BORGE VARMER, STUDY NO. 15: PHOTODUPLICATION OF COPYRIGHTED MATERIALS BY LIBRARIES, *in* STAFF OF S. COMM. ON THE JUDICIARY, 86TH CONG., COPYRIGHT LAW REVISION 49 (Comm. Print 1960)).

75. *Id.* at 1364 (Cowen, C.J., dissenting). Chief Judge Cowen's dissent included large portions of the trial judge's opinion determining that the libraries had infringed the journal articles' copyrights. *See id.* at 1372–86.

76. *Id.* at 1366. The dissenters noted that exceptions to limits on photocopying were routinely granted. *Id.* at 1365–66.

77. *Id.* at 1367.

78. *Id.* at 1368 (quoting MELVILLE B. NIMMER, NIMMER ON COPYRIGHT § 145, at 646 (1973)).

79. *Id.* (quoting UCLA Project, *supra* note 49, at 951).

80. *Id.* at 1371–72. One dissenter decried the majority's opinion as the "*Dred Scott* decision in copyright law." *Id.* at 1387 (Nichols, J., dissenting). Professor Nimmer picked up on this inappropriate metaphor in a subsequent article. *See* Melville B. Nimmer, *Photocopying and Record Piracy: Of* Dred Scott *and Alice in Wonderland*, 22 UCLA L. REV. 1052, 1052–54 (1975) (writing critically of the majority's decision, especially its failure to address "whether wholesale copying of plaintiff's journals by any and all libraries and similar institutions would decimate the plaintiff's potential market"); *see also* Paul Goldstein, *The Private Consumption of Public Goods: A Comment on* Williams & Wilkins Co. v. United States, 21 BULL. COPYRIGHT SOC'Y U.S.A. 204, 205–07 (1974) (criticizing *W&W* for its narrow analysis of harm); WILLIAM F. PATRY, THE FAIR USE PRIVILEGE IN COPYRIGHT LAW 179–80 (1985) (reproaching *W&W* for failing to recognize the potential harm to W&W's licensing market).

81. Latman authored the Copyright Law Revision Study on fair use. *See supra* note 10.

photocopying of articles posed a serious threat to the future existence of scientific and technical journals.[82] Latman asserted that fair use was simply unavailable when copies were effective substitutes for the originals.[83] Like the *W&W* dissenters, he emphasized that W&W only wanted a reasonable royalty from the libraries, and thought the Court of Claims could set a reasonable royalty as a pilot for a licensing regime that could have broader applications.[84] Solicitor General Bork, arguing on behalf of the United States, emphasized that W&W "asks for a sweeping rule that is going to upset the status quo in medical research."[85] Bork mainly focused on the practical difficulties of achieving a judicial resolution of the photocopying issue, which would affect thousands of libraries as well as 500 to 600 publishers of medical journals, suggesting—as had the Court of Claims majority—that Congress was the more appropriate venue to address the issue.[86]

The Justices were ultimately unable to resolve the fair use question, affirming the lower court without setting a precedent.[87] The Court thus kicked the photocopying can back to Congress.

82. *See* Transcript of Oral Argument at 3–4, Williams & Wilkins Co. v. United States, 420 U.S. 376 (1975) (No. 73-1279) [hereinafter W&W Transcript]. In response to one Justice's question about whether he would be an infringer if he made a photocopy of a book chapter, Latman indicated that this presented a harder question, but thought it would be impractical to sue individuals for such copying. *See id.* at 21–22.

83. *See id*. at 5–6.

84. *See id.* at 10, 16–17. Courts cannot enjoin the U.S. government from engaging in copyright infringement under 28 U.S.C. § 1498.

85. *Id.* at 28. Bork could have (but did not) emphasize that only eight articles were actually at issue in *W&W*, which would have yielded at most a miniscule damage award if the copies were found to infringe. *W&W* was not a class action lawsuit, so copies of other publishers' works were not properly before the Court.

86. *See id.* at 34–39, 42. Justice Breyer had previously expressed doubts that a compulsory license regime for library photocopying could be developed. *See* Breyer, *supra* note 49, at 332–35.

87. Williams & Wilkins Co. v. United States, 420 U.S. 376 (1975) (per curiam). Justice Blackmun participated in the *W&W* oral argument, but recused himself from the Court's ruling. Given his dissent in *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417, 457–500 (1984) (Blackmun, J., dissenting), which regarded ordinary copying of the whole of copyrighted works as infringements, Justice Blackmun would likely have been the fifth vote to overturn the Court of Claims ruling. *See also* Wihtol v. Crow, 309 F.2d 777 (8th Cir. 1962) (holding an educational use copying to infringe, with Blackmun on panel). Given that Justices Marshall, Powell, and Rehnquist joined Blackmun's *Sony* dissent, they probably voted for W&W, while Brennan, White, and Burger were likely votes for the U.S., given their support for the Stevens majority opinion in *Sony*. I suspect Douglas was the fourth vote for the U.S. and Stewart was the fourth vote for W&W. *See also* Nimmer, *supra* note 80, at 1056–59 (speculating on how Justice Blackmun might have voted in *W&W*).

2. Photocopying Controversies in Congress

In the early stages of Congress' two decades-long process of revising the Copyright Act of 1909, which culminated in enactment of the 1976 Act (hereinafter the copyright revision process)—and well before W&W initiated its lawsuit against NIH and NLM—librarians and educators pressed for exemptions from liability so that they could make copies of in-copyright materials for nonprofit research and educational purposes.[88] The Register of Copyrights was initially sympathetic to the library concerns and proposed in 1963 a special library privilege for making single-use copies.[89] Publisher opposition to such an exemption resulted in its being dropped from the next version of the draft bill.[90] Instead, the 1964 draft fair use provision responded to libraries' and educators' concerns by adding "teaching, scholarship, or research" to the draft fair use provision as favored uses.[91] Educators objected to qualifying language "to the extent reasonably necessary or incidental to a legitimate purpose," which seemed both ambiguous and restrictive.[92]

As talks continued, educators proposed an exemption to allow teachers to make "no more than one copy or phonorecord of the work" in the course of an educational use and to make "a reasonable number of copies or phonorecords of excerpts or quotations from the work," as long as they were "not substantial in length in proportion to their source."[93] The Register opposed this exemption, saying that:

> [A] statutory recognition of fair use would be sufficient to serve the reasonable needs of education with respect to the copying of short extracts from copyrighted works, and that the problem of obtaining clearances for copying larger portions or entire works could best be solved through a clearinghouse arrangement worked

88. *See* PATRY, *supra* note 80, at 218–19 (describing efforts by library associations in 1950s and 1960s to gain an exemption with regard to library photocopying for purposes of research).
89. PATRY, *supra* note 80, at viii.
90. *Id.* at 235. Publishers thought the library privilege went too far, and librarians thought it did not go far enough.
91. *Id.* at 229.
92. *Id.* at 229–31.
93. *Id.* at 245 (detailing a draft exemption). Teachers had not prevailed with fair use defenses in some cases decided before and during the copyright revision process. *See, e.g.*, *Wihtol* v. Crow, 309 F.2d 777, 783 (8th Cir. 1962) (regarding a new musical arrangement); Macmillan Co. v. King, 223 F. 862 (D. Mass. 1914) (regarding a textbook outline).

> out between the educational groups and the author-publisher interests.[94]

One Congressman quipped that "the publishers and authors on the one side and the educational groups on the other have now dug in and are engaged in positional or trench warfare."[95]

U.S. House and Senate committees held numerous hearings on the educational and library photocopying issues in the hope of reaching some consensus, but this proved elusive.[96] Outside of Congress, representatives of publishers, authors, educators, and librarians continued to explore some options.[97] But whatever modest progress these efforts achieved suffered a setback when the trial judge ruled against fair use in *W&W*.[98] Despite this development, a committee representing educators and libraries put forth a new proposed copyright exception for educational and scholarly uses of copyrighted works,[99] which the author and publisher representatives once again rejected.[100] The fair use pendulum swung back briefly toward the libraries (and by analogy, educators whose nonprofit copying was similar to the libraries') when the *W&W* appellate majority ruled in favor of fair use.[101] But the Supreme Court's inability to resolve the *W&W* fair use controversy meant that neither publishers nor libraries had attained a clear victory.

Frustrated that intense disagreements about photocopying and other new technology issues were holding up enactment of the revised copyright law, Congress decided to create a National Commission on New Technological Uses of Copyrighted Works (CONTU) to address some photocopying and computer-related issues.[102]

CONTU commissioned a study of library photocopying.[103] The study estimated that in 1976, library staff had made 53.9 million copies of in-

94. PATRY, *supra* note 80, at 236–37.
95. *Id.* at 246 (quoting the remarks of Congressman Celler).
96. *See id.* at 249–64.
97. *See id.* at 265–71.
98. *Id.* at 272–73.
99. *Id.* at 273–74.
100. *Id.* at 275–76.
101. *See supra* Part II.A.1 (discussing *W&W*).
102. *See* NAT'L COMM'N ON NEW TECH. USES OF COPYRIGHTED WORKS, FINAL REPORT (1978) [hereinafter CONTU REPORT], https://digital-law-online.info/CONTU/PDF/index.html [https://perma.cc/VPG6-HZ9Hhttps].
103. Under 17 U.S.C. § 108(g)(2), libraries and archives could make copies for interlibrary exchanges if such arrangements "do not have, as their purpose or effect, that the library or archives receiving such copies or phonorecords for distribution does so in such aggregate quantities as to substitute for a subscription to or purchase of such work." 17

copyright materials (which did not include patrons' uses of photocopy machines in libraries to make their own copies).[104] None of this copying was licensed, and yet CONTU reported that publishers were not in crisis or asking for legislative changes to address machine copying.[105]

Even after creating CONTU, members of Congress and various stakeholder representatives continued to seek compromises about educational and scholarly uses through fair use.[106] In March 1976, two significant changes were made to the text of the fair use provision, which arguably bolstered nonprofit educational fair use claims.[107] One was the insertion of "including multiple copies for classroom use" after "teaching" in the preamble.[108] A second was the insertion of "including whether such use is of a commercial nature or is for a nonprofit educational use" in the first factor which considers the purpose and character of the challenged use.[109] Librarians succeeded in getting a series of exemptions for single-use copying and preservation purposes.[110] Under pressure from key members of Congress, representatives of author-publisher and educator organizations negotiated a set of minimum guidelines for asserting fair use when making

U.S.C. § 108(g)(2). CONTU published guidelines detailing how libraries could satisfy this proviso. CONTU REPORT, *supra* note 102, at 54–55 (setting forth the CONTU Guidelines on Photocopying Under Interlibrary Loan Arrangements).

104. *See* CONTU REPORT, *supra* note 102, at 56; *see also id.* app. H, at 131–33. CONTU also commissioned an economic study that concluded the publishing industry was earning adequate returns and that students, teachers, researchers, or libraries should not have to pay fees to publishers as long as they did not sell their photocopies. *See id.* at 129.

105. CONTU recommended an amendment to 17 U.S.C. § 107 that would have required commercial copy shops to post a prominent warning to the public that "in most cases it is copyright infringement, even for purposes of private study, to reproduce more than one article or other contribution to a copyrighted collection or periodical, or more than a small part of any other copyrighted work." CONTU REPORT, *supra* note 102, at 50. The U.S. Congress did not act on this recommendation. Copy shops became defendants in later litigation. *See infra* Part II.A.5.

106. Earlier in the copyright revision process, nonprofit education institutions got exemptions for public performances and displays of in-copyright works for face-to-face teaching. *See* 17 U.S.C. § 110(1). In addition, the 1976 Act authorized courts to remit statutory damage awards against nonprofit educators or librarians if they had reasonable grounds for believing their use of in-copyright works was fair. *See* 17 U.S.C. § 504(c)(2).

107. *See* PATRY, *supra* note 80, at 304.

108. *Id.*

109. *Id.*

110. *See* 17 U.S.C. § 108. One of these exemptions, § 108(a), codified the appellate holding in *W&W*. Subsection (f)(4) preserved libraries' rights to raise fair use defenses for copies outside of the specified exemptions.

multiple copies for classroom use.[111] These guidelines have had significance in subsequent cases.

3. An Interlude on First Generation Photocopying Controversies

In retrospect, the *W&W* appellate decision upholding the libraries' fair use defenses seems sound, notwithstanding prominent critics' denunciations.[112] The only photocopies before the court were of eight articles identified in W&W's complaint. W&W produced no evidence that these photocopies harmed its business, and it was implausible that eight copies could have meaningfully done so.[113] The majority noted that W&W did not stock back issues to sell to individuals and referred requests for article reprints to their authors.[114] It based its business on sales of journal subscriptions.[115] Occasional photocopies of individual articles for scientific research purposes were unlikely to supplant demand for subscriptions.

The *W&W* majority did not discuss, but was surely aware, that notwithstanding W&W's emphasis on the large scale of NIH and NLM photocopying activities, its lawsuit was not a class action for which W&W had been certified as an adequate representative of a class of publishers having the same interests and claims. Nor did W&W (or the Association of American Publishers (AAP)) during the seven years of this litigation commission a study to test the theory that photocopying was harming or about to harm journal publishing. What W&W and publisher allies actually wanted was a new income stream to increase their profits and to scale back the consumer surplus from which the libraries were benefiting.[116]

111. *See* PATRY, *supra* note 80, at 307–09 (setting forth the Agreement on Guidelines for Classroom Copying in Not-for-Profit Educational Institutions With Respect to Books and Periodicals). Multiple copies could be made if brief and spontaneous, and the cumulative effect of copying was limited. *Id.* at 308; *see also id.* at 296–317 (discussing the negotiations leading up to guidelines).
112. *See supra* note 80 (citing *W&W* critics).
113. Because the libraries were nonprofit entities, there would be no profits from photocopying for disgorgement, even if infringement had been found.
114. Williams & Wilkins Co. v. United States, 487 F.2d 1345, 1357 n.17 (Ct. Cl. 1973), *aff'd by an equally divided Court*, 420 U.S. 376 (1975).
115. *Id.* at 1347.
116. *See* Brief of Amici Curiae—Ass'n of Research Libraries Medical Library Ass'n, American Ass'n of Law Libraries, Ass'n of American Universities, American Medical Ass'n, American Dental Ass'n, American Ass'n of Dental Schools, Robert H. Ebert, M.D. (In His Capacity as Dean of the Faculty of Medicine, Harvard Medical School), the University of Michigan Medical School, American Sociological Ass'n, Modern Language Ass'n of

Because the Court of Claims had before it only one publisher, two libraries, and at most eight infringements, it could not discern what would be a reasonable royalty to set for those copies, let alone a future royalty rate that NIH and NLM should pay to make copies of articles from W&W journals. The court had no power to issue an order directly affecting other publishers. Yet, whatever royalty the court might have set for W&W would almost certainly be considered a benchmark affecting other publishers and libraries.[117]

Moreover, Congress was actively considering what to do about machine copying of in-copyright works as part of the copyright revision process. Hence, it was prudent to leave the photocopying controversies to be resolved by the legislature, which at least could hear from more than the two parties in the *W&W* litigation. In the absence of a clearinghouse authorized to grant blanket licenses to allow photocopying of articles from all or virtually all scientific and technical journals to which the NIH and NLM subscribed, there was a market failure; a ruling in one litigated case involving one publisher could not cure it.[118]

Numerous participants in the legislative debates and commentators on the copyright photocopy controversies were hoping that W&W would win its lawsuit and that this might lead to the development of a collective licensing regime to enable one-stop shopping to get licenses to make machine copies of in-copyright works.[119] Professor Melville Nimmer and the UCLA Law Review Project on machine reproductions were early and enthusiastic endorsers of the formation of a copyright collecting society, modeled on organizations such as the American Society of Composers, Authors and Publishers (ASCAP), which would act as an agent for copyright owners and collect photocopying fees.[120] The UCLA Project also considered whether a flat tax on the sales of machines would be a viable way to provide

America, History of Science Society at 60–66, Williams & Wilkins Co. v. United States, 420 U.S. 376 (1975) (No. 73-1279) (arguing that publishers were seeking, through litigation, additional payments for their publications); Brief Amici Curiae of Ass'n of Am. Publishers, Inc., the Ass'n of Am. Univ. Presses, Inc., Educ. Media Producers Council et al. 14–15, Williams & Wilkins Co. v. United States, 420 U.S. 376 (1975) (No. 73-1279).

117. Several publishers and library associations filed amicus curiae briefs to the appellate panel and the Court on the fair use issue, not about how to set a reasonable royalty.

118. *See* Wendy J. Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of the* Betamax *Case and its Predecessors*, 82 COLUM. L. REV. 1600 (1982).

119. *See* UCLA Project, *supra* note 49, at 967, 1027.

120. *See id.* at 964–67, 973–75. Paul Goldstein also endorsed this approach. *See* Goldstein, *supra* note 80, at 211.

compensation to publishers for copies of in-copyright materials that machine purchasers might make.[121]

Other commentators were skeptical about an ASCAP-like collective licensing solution to the photocopying controversies.[122] At the very least, such a clearinghouse would have to persuade a large number of publishers to participate and then "would have to sell licenses, monitor unlicensed copiers to detect violations, bring infringement actions where piracy is found, make random samples to determine the relative frequency with which works are copied, and distribute collected income to numerous writers and publishers."[123] Such a regime would impose substantial transaction costs on all stakeholders, especially if the regime required documentation of every copy.[124] There was also reason to doubt that a photocopy clearinghouse would be able to collect as much revenue as ASCAP insofar as its client base might largely be budget-strapped nonprofit libraries and educational institutions.[125]

In any event, the *W&W* majority was wise to leave the resolution of these complicated issues to Congress, even though Congress too was unable to fully resolve them. Yet, in time, the Copyright Clearance Center (CCC)—formed as a voluntary collecting society through which publishers could collect revenues for granting permission to photocopy in-copyright works—became a significant actor in the ongoing photocopying wars.[126]

4. For-Profit Firm Photocopying of Scientific & Technical Journal Articles

Two decades after *W&W*, a class action lawsuit on behalf of eighty-three journal publishers challenged the legality of photocopying scientific and technical journal articles—this time, by employees of a large corporation. As in *W&W*, the judicial resolution of this issue in *American Geophysical Union v. Texaco, Inc.*[127] resulted in deeply split opinions, although this time the publishers prevailed. The harm they claimed was lost permission fees.

121. *See* UCLA Project, *supra* note 49, at 971–75.
122. *See, e.g.*, Jon Vander Ploeg, Note, *Copyright Law and Library Photocopying: Striking a Balance Between Profit Incentive and the Free Dissemination of Research Information*, 48 IND. L.J. 503, 512–14 (1973).
123. Breyer, *supra* note 49, at 332.
124. Vander Ploeg, *supra* note 122, at 513.
125. Breyer, *supra* note 49, at 332.
126. *See infra* Parts II.A.4 (discussing *Texaco*), II.A.6 (discussing *CUP*).
127. 60 F.3d 913 (2d Cir. 1994).

During the lawsuit, Texaco employed four to five hundred research scientists and paid subscription fees for print copies of various scientific and technical journals, one copy of each issue circulated to its researchers,[128] who could make photocopies of articles relevant to their research. The litigants agreed to focus their fair use battle on one Texaco researcher's photocopies. The files of that scientist contained eight photocopied articles from the *Journal of Catalysis*.[129] As in *W&W*, that publisher sold subscriptions to entire journals and did not sell reprints of individual articles.

The litigants and their amici tried to frame the *Texaco* case around "the broad issue of whether photocopying of scientific articles is fair use, or at least the only slightly more limited issue of whether photocopying of such articles is fair use when undertaken by a research scientist engaged in his own research."[130] The Second Circuit majority rejected these broad frames, saying that the issue was whether Texaco researchers' photocopying of journal articles was fair use,[131] adding that this issue included "the question whether such institutional, systematic copying increases the number of copies available to scientists while avoiding the necessity of paying for license fees or for additional subscriptions."[132]

Texaco emphasized the scientific research purposes of the copying and argued that this practice had become customary. But the panel majority characterized the copying as nontransformative, archival in nature, and "part of a systematic process of encouraging employee researchers to copy articles so as to multiply available copies while avoiding payment," which superseded the objects of the originals.[133] These considerations disfavored Texaco's fair use defense. The fact-intensive nature of the articles may have favored Texaco, but the copying of entire works cut against fair use.

As in so many fair use cases, the market effects factor was the most determinative in *Texaco*. The majority first considered whether this photocopying had impacted journal subscriptions, on the theory that if its researchers could not photocopy articles, Texaco might have increased the number of its subscriptions.[134] It regarded this prospect as "of somewhat

128. *Id.* at 915.
129. *Id.* Each of this journal's monthly issues was about two hundred pages long and contained twenty to twenty-five articles. *Id.*
130. *Id.* at 916.
131. *See id.*
132. *Id.*
133. *Id.* at 919–20; *see also id.* at 921–25. The court also mentioned the negotiated classroom guidelines a few times. *Id.* at 917 n.3, 919 n.5.
134. *See id.* at 927–28.

limited significance."[135] While acknowledging there was no market to enable purchases of individual journal articles, the majority noted that § 107 of the 1976 Act directs courts to consider potential market effects. When assessing such effects, courts should consider "traditional, reasonable, or *likely to be developed* markets."[136] CCC was developing such a market for journal photocopy permissions, and the legislative history of the 1976 Act indicated that Congress had wanted a clearinghouse such as CCC to be created.[137] The majority concluded that Texaco's failure to pay CCC an annual license fee to allow its researchers to make photocopies of scientific articles from CCC-licensed journals had harmed the publishers' emerging photocopy permissions market.[138] It mattered that the *Catalysis* was one of the journals for which CCC had photocopying license rights. The majority observed that it was not deciding "how the fair use balance would be resolved if a photocopying license for *Catalysis* articles were not currently available."[139]

Judge Jacobs's dissent disagreed with the majority's analysis of the purpose and market effects factors. He viewed favorably photocopies of journal articles in service of scientific research and questioned the majority's conclusions that the copying was "systematic" and "archival."[140] Texaco librarians' circulation of journals to its researchers may have been systematic, but each scientist made an "insubstantial" number of copies, which was "simply a technologically assisted form of note-taking."[141] Judge Jacobs pointed to the lack of evidence that widespread photocopying was impairing the marketability of journals; nor was there evidence that increased publisher revenues from CCC would "fuel scientific creativity."[142] He was bothered by the circularity of the publishers' arguments: "[T]he market [for licensing] will not crystalize unless courts reject the fair use

135. *Id.* at 928.
136. *Id.* at 930 (emphasis added).
137. *See id.* at 931 (citations omitted) ("Congress apparently prompted the development of CCC by suggesting that an efficient mechanism be established to license photocopying. It is difficult to understand why Congress would recommend establishing such a mechanism if it did not believe that fees for photocopying should be legally recognized as part of the potential market for journal articles.").
138. *See id.* at 930 (recognizing CCC as a workable market for photocopying). The trial court noted that eleven other oil companies, as well as several hundred other corporations, had taken CCC photocopy licenses. Am. Geophysical Union v. Texaco, Inc., 802 F. Supp. 1, 8 (S.D.N.Y. 1992).
139. *Texaco*, 60 F.3d at 931.
140. *Id.*
141. *Id.* at 934 (Jacobs, J., dissenting).
142. *Id.* at 936, 940.

argument," but there was no infringement without evidence of harm.[143] More troubling was CCC's inability to reassure Texaco that a photocopy license would cover all photocopying its researchers might do, as only 30 percent of the journals to which Texaco subscribed had licensing deals with CCC.[144] Hence, the dissent questioned whether CCC had developed a workable market for licensed photocopying.

5. Educational Coursepack Cases

In the early 1990s, several book publishers brought test cases challenging commercial copy shop sales of compilations of excerpts of in-copyright works (known as "coursepacks") to university students for their classes. Professors prepared these compilations as reading assignments and gave them to copy shops, such as Kinko's, which then made and bound multiple copies and sold them to students. Since § 107 plainly states that teaching is a favored purpose, "including multiple copies for classroom use,"[145] many thought that this kind of nonprofit educational use was fair use.

Basic Books, Inc. v. Kinko's Graphics Corp. was the first lawsuit to challenge copy shop coursepacks.[146] Eight major New York book publishers identified twelve books whose chapters Kinko's copied for specific classes taught at New York City universities.[147] Kinko's defended this service as fair use because it was serving an educational purpose by providing course materials requested by faculty to their students.[148]

The trial judge, however, regarded the copying as nontransformative and commercial, so the purpose factor weighed against Kinko's fair use defense.[149] Although the factual nature of the books favored Kinko's, the court found the amount of copying to be quantitatively and qualitatively substantial.[150] A significant consideration, however, was that coursepack copying went far beyond the modest scale of multiple copies for classroom

143. *Id.* at 937.
144. *Id.*
145. 17 U.S.C. § 107.
146. 758 F. Supp. 1522 (S.D.N.Y. 1991).
147. The number of pages assigned from these books ranged from 14 to 110 pages, constituting between 5 percent to 28 percent of the books' contents. *Id.* at 1527–28. Some of the books were in-print and some out-of-print. *Id.*
148. *Id.* at 1531.
149. *See id.* at 1530–31.
150. *Id.* at 1532–34. The court did not consider that individual book chapters constituted small parts of the books, which students would be unlikely to buy because they only needed to read those small parts.

use that publishers and educators had agreed to during the legislative history of the 1976 Act.[151] Publishers and educators had agreed on those guidelines in the mid-1970s as a floor-like safe harbor, but in *Kinko's*, the court treated them as a ceiling. It concluded that the coursepack copies harmed the sales of the plaintiffs' books and deprived the plaintiffs of permission fees they should have received for licensing book chapters for coursepacks.[152]

Because Kinko's operated hundreds of copy shops across the nation, the publishers' victory had a substantial impact. The judgment against Kinko's also persuaded many other copy shops to produce coursepacks only if permissions had been obtained or if the contents were public domain materials.[153] The *Kinko's* judgment provided an opportunity for CCC to expand its service from granting licenses for corporate photocopying of journal articles to establishing permissions fees for coursepacks.[154]

Some copy shop operators, however, were unpersuaded by the *Kinko's* ruling and continued to prepare coursepacks for university classes. Michigan Document Service (MDS) was one of the holdouts. A few years after *Kinko's*, Princeton University Press (PUP) and two other publishers sued MDS for copyright infringement of six works.[155] As in *Kinko's*, a trial judge ruled against the copy shop's fair use defense.[156] Although MDS's appeal was initially successful, a split en banc panel of Sixth Circuit judges ultimately ruled in PUP's favor.[157]

151. *Id.* at 1534–37. The court found Kinko's infringement to be willful and awarded substantial statutory damages and attorney fees and issued an injunction against further copying of substantial parts of plaintiffs' books. *See id.* at 1542–47. The court was "not attempting to measure the financial loss to the plaintiff but to deter the defendant." *Id.* at 1545.

152. *See id.* at 1534.

153. *See* Princeton Univ. Press v. Mich. Document Servs., Inc., 99 F.3d 1381, 1384 (6th Cir. 1996).

154. *See Campus Guide to Copyright*, *Chapter 3: Using Content*, COPYRIGHT CLEARANCE CTR., https://www.copyright.com/education-campus-guide-to-copyright/using-content [https://perma.cc/77PP-M55A] [hereinafter CCC Campus Guide] (discussing coursepack policies). The *Kinko's* decision did not mention CCC.

155. Macmillan and St. Martin's Press were fellow plaintiffs. *Mich. Document Servs.*, 99 F.3d at 1384–85. The pages copied ranged from 17 to 95 pages, constituting between 5 percent and 30 percent of the books at issue. *Id.*

156. Princeton Univ. Press v. Mich. Document Servs., Inc., 855 F. Supp. 905 (E.D. Mich. 1994), *aff'd in part, vacated in part, and remanded*, 99 F.3d 1381 (6th Cir. 1996) (en banc).

157. The first panel reversed the district court, but upon rehearing the lower court's ruling was affirmed in an 8–5 decision. *Mich. Document Servs.*, 99 F.3d at 1383. Three judges wrote dissenting opinions, which were joined by two others. *Id.* at 1393–412 (Martin, Jr., C.J., Merritt, J., and Ryan, J., separately dissenting).

As in *Kinko's*, the *MDS* majority concluded that the photocopying was nontransformative and commercial, the amounts copied were substantial, and the unauthorized copying was harming the publishers' market for permissions fees.[158] As in *Kinko's*, the court also took note of the negotiated guidelines for making multiple copies for classroom use, characterizing MDS's copying as "light years away from the safe harbor" established by the guidelines.[159] As in *Kinko's*, the Sixth Circuit treated the classroom guidelines as if they set a ceiling, not a floor, as a safe harbor for educators.

Although MDS received support from book authors, the majority was unmoved because the publishers owned the copyrights, and their incentives mattered. The majority observed that most of the works at issue were marginally profitable and suggested that publishers might discontinue publishing such works unless they could receive permission fees for coursepack uses.[160]

One of the dissenters criticized the majority's analysis of the purpose, amount, and market effect factors.[161] The dissenting opinion regarded MDS's purpose as nonprofit and educational because professors, not MDS, selected and arranged the in-copyright materials for coursepacks and chose those materials "solely in order to enrich the educational experiences of their students."[162] Nor did the dissent consider MDS's use as commercial because its "profit is attributable entirely to its provision of a mechanical service."[163] Evidence was lacking that "any of the six excerpts in the coursepacks are so substantial as to supersede the original works."[164] Having been persuaded MDS's use was noncommercial, this dissenter thought the court should presume the use was fair and require the publishers to show actual harm to their markets, which they had not done.[165] Under the majority's circular reasoning, the dissenter thought, publishers would prevail in every fair use case in which they asserted the existence of a licensing market for the

158. *See id.* at 1386–91 (majority opinion).
159. *Id.* at 1391.
160. *See id.* at 1390–91. The majority was unpersuaded that MDS's infringement was willful and directed the lower court to reconsider the statutory damage award. *See id.* at 1392. It also directed the trial court to issue a more precise injunction. *See id.* It did not mention whether CCC licenses were available for the excerpts at issue.
161. *See id.* at 1397–412 (Ryan, J., dissenting).
162. *Id.* at 1404.
163. *Id.* at 1401–02.
164. *Id.* at 1405.
165. *See id.* at 1407 ("To prove entitlement to permission fees, the publishers must show market harm and the market harm they claim is the loss of permission fees.").

challenged use,[166] which was inconsistent with Congress's intent when including teaching and research as favored purposes in § 107.

6. Online Course Materials

A dozen years after MDS lost its coursepack fair use defense, Cambridge University Press (CUP) and two other publishers of academic books sued the President of Georgia State University (GSU) and various GSU officials for copyright infringement.[167] The publishers claimed that GSU's copyright policy too liberally allowed professors to post in-copyright materials, such as book chapters, on library electronic course reserves or course websites. GSU officials defended the online postings as fair uses.[168]

CUP's initial complaint charged that hundreds of GSU professors had put thousands of in-copyright works on GSU websites and that university officials had facilitated, encouraged, and induced infringements.[169] They noted that when GSU professors prepared coursepacks, permissions fees were paid for uses of in-copyright works.[170] They analogized online postings of in-copyright works on GSU online reserve and course websites as the virtual equivalent of coursepacks and argued, hence, that permission fees should also be paid for these materials either through CCC or licenses with the books' publishers.[171]

The trial court directed the publishers to produce a list of each allegedly infringing book chapter that GSU professors or librarians had posted online for three school terms in 2009.[172] A bench trial focused on fair use defenses

166. *See id.* The dissent questioned whether permission fees "affects in any meaningful way the publishers' decisions about whether to publish any given work." *Id.* at 1410. It also took issue with the majority's reliance on the legislative history pertinent to the negotiated classroom guidelines. Their inclusion in U.S. House and U.S. Senate reports did not make them authoritative because those reports "do not accurately indicate congressional intent because they do not 'necessarily say anything about what Congress as a whole thought,' *even if* all the members of the Committee . . . agree[d]." *Id.* at 1411 (quoting Wis. Pub. Intervenor v. Mortier, 501 U.S. 597, 620 (1991) (Scalia, J., concurring)).

167. Cambridge Univ. Press v. Becker, 863 F. Supp. 2d 1190 (N.D. Ga. 2012), *rev'd sub nom.* Cambridge Univ. Press v. Patton, 769 F.3d 1237 (11th Cir. 2014).

168. *See Patton*, 769 F.3d at 1237–41 (recounting district court proceedings).

169. *Id.* at 1241.

170. *Id.* at 1239–41.

171. *Id.* CUP noted that in 2009, only fifteen professors prepared photocopied coursepacks, but hundreds were using the online services instead. *Id.* at 1240.

172. The court ordered the publishers to identify the title of the work at issue, the name of the course, the identity of the instructor, a description of the work, the number of pages

as to seventy-four online postings.[173] The trial court ruled that the publishers had failed to make a prima facie case of infringement as to twenty-six postings,[174] that forty-three uses were fair, and that five were infringements.[175] It made detailed findings as to each of the challenged uses.

On appeal, the publishers argued that the trial judge erred by making excerpt-by-excerpt findings when the focus should have been on the extensive up- and downloading of in-copyright works by GSU faculty, students, and librarians under GSU's copyright policy.[176] A majority of the panel found no fault in the lower court's ruling on that issue, although one judge agreed with the publishers that the focus should be on the "university-wide practice to substitute 'paper coursepacks' (the functional equivalent of textbooks) that contained *licensed* copyrighted works with 'digital coursepacks' that contained *unlicensed* copyrighted works" in the "vast majority of courses offered at GSU."[177] If publishers got permissions fees for GSU paper coursepacks, the judge thought they should also get them for digital coursepacks.[178]

On the purpose factor, the appellate majority agreed with the trial court's analysis. Although the publishers insisted that online copies of their book chapters were nontransformative and for the same intrinsic purpose as the originals, the majority thought that the more salient consideration was that GSU faculty and librarians were making copies for nonprofit educational purposes, which favored fair use.[179]

As for market effects, the majority was unpersuaded that digital excerpts had undermined sales of the publishers' books, and more significantly, that GSU's failure to pay permissions fees necessarily

copied, the retail price of the entire work, and the potential cost of a license. *Id.* at 1243–44. The publishers initially listed 126 works. *Id.* at 1244.

173. *Id.* at 1237.

174. *Id.* The publishers failed to show that they owned some of the copyrights or there was evidence that students had not accessed some copies or had been required to buy the books. *Id.* at 1251.

175. *Id.* at 1251–52. The trial court determined that five uses were unfair because professors posted more than one chapter of particular books; it found small but actual damage to existing licensing markets for those works. *Id.* at 1251, 1278.

176. *See id.* at 1259–60.

177. *Id.* at 1285 (Vinson, J., concurring).

178. *See id.* at 1287.

179. *See id.* at 1262–67 (majority opinion). While the publishers tried to argue that GSU's uses were commercial because it was not paying the customary price, the panel majority disagreed, noting also that "the use provides a broader public benefit—furthering the education of students at a public university." *Id.* at 1264–67.

constituted harm to publishers' markets.[180] It agreed with the trial judge that when publishers failed to make digital licenses available for certain excerpts, that was evidence that they were disinterested in digital licensing markets and hence not harmed by the online posting.[181] It regarded the crux of the market effect factor to be whether economic losses from copying were substantial enough to impair the publishers' incentives, not whether challenged uses might cause some loss of revenues.[182] When licenses were available, the trial court had taken them into account, and in close cases, it "examined the amount of permissions income a work had generated in order to determine how much this particular revenue source contributed to the value of the copyright in the work," concluding that "where there is no significant demand for excerpts, the likelihood of repetitive unpaid use is diminished."[183] The majority thought that the lower court had engaged in a nuanced analysis of the market effects factor.[184]

The publishers emphasized that the excerpt copying at issue was incompatible with the negotiated classroom guidelines. But the majority noted that the guidelines were minima, not maxima, that they did not carry the force of law, and that the Supreme Court in *Campbell* had rejected hardline presumptions and called for a sensitive balancing of all fair use factors.[185] It agreed with the trial court's decision to weigh the reasonableness of the amount of each excerpt in light of the professor's pedagogical purpose.[186] The panel disagreed, however, with the trial court's one-chapter-or-no-more-than-10-percent of the book criterion for assessing the substantiality of the taking.[187] The court remanded the case for further proceedings consistent with its opinion.[188]

180. *See id.* at 1276.
181. *See id.* at 1277.
182. *See id.* at 1276.
183. *Id.* at 1279.
184. *See id.* at 1278.
185. *See id.* at 1273–74 (citing Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 584 (1994)).
186. *See id.* at 1275.
187. *See id.* at 1271–72.
188. *See* Cambridge Univ. Press v. Becker, 446 F. Supp. 3d 1145 (N.D. Ga. 2020) (relating the subsequent history of this lawsuit and the trial court's revised fair use analysis on an excerpt-by-excerpt basis, ultimately holding that the publishers should prevail as to eleven of the forty-eight claimed infringements). The publishers chose not to appeal from this ruling. Porter Anderson, *AAP and CCC End Georgia State 'E-Reserves' Copyright Litigation*, PUBL'G PERSPS. (Nov. 12, 2020), https://publishingperspectives.com/2020/11/aap-and-copyright-clearance-center-end-georgia-state-e-reserves-litigation [https://perma.cc/HEB8-SHMZ].

7. An Interlude on the Role of Licensing in Research and Educational Use Copying Cases

Unlike *W&W*, *Texaco* was a class action lawsuit on behalf of eighty-three publishers of scientific and technical journals, all of whom presumably had signed up with CCC to receive permissions fees for corporate photocopying of journal articles.[189] Even if a substantial number of publishers of journals to which Texaco subscribed had not authorized CCC to include its articles in the photocopy license program, CCC could attract more publishers to join once the court found Texaco's use unfair. Although the Second Circuit did not say so, its ruling, in effect, regarded the development of CCC's photocopy service as having at least partially cured a market failure that may have justified the fair use ruling in *W&W*.[190]

Nonprofit library and educational groups filed amicus briefs in *Texaco*.[191] Perhaps to allay their concerns about the implications of a broad ruling against fair use, the *Texaco* majority amended its decision to clarify that it was not addressing "copying by an individual, for personal use in research or otherwise (not for resale)" because under the fair use or the *de minimis* doctrines, "such a practice by an individual might well not constitute an infringement."[192] It was, in other words, "not decid[ing] the case that would arise if [the copyist] were a professor or an independent scientist engaged in copying and creating files for independent research, as opposed to being employed by an institution in the pursuit of his research on the institution's behalf."[193] Notwithstanding this cautionary note, some worried that this precedent would embolden publishers to bring a comparable lawsuit to test the legality of university and nonprofit library photocopying of in-copyright materials.[194] Yet, none have materialized.

Unlike *Texaco*, the coursepack cases were not class action lawsuits. The eight *Kinko's* plaintiffs were well-known major publishers of nonfiction

189. Am. Geophysical Union v. Texaco, Inc., 60 F.3d 913, 916 (2d Cir. 1994).
190. *See* Gordon, *supra* note 118, at 1614–15, 1656 (explaining market failures and market cures).
191. *See Texaco*, 60 F.3d at 914 (listing amici briefs submitted by the American Library Association and by various university library and research organizations).
192. *Id.* at 916. The opinion was amended twice. *Id.* at 913.
193. *Id.* at 916.
194. *See, e.g.*, Maureen Ryan, *Fair Use and Academic Expression: Rhetoric, Reality, and Restriction on Academic Freedom*, 8 CORNELL J.L. & PUB. POL'Y 541, 565–57 (1999) (surmising that the legal reasoning in *Texaco* could apply in a university context).

books on topics likely to be relevant to university courses.[195] PUP was the lead plaintiff in *MDS*, rather than its large publisher co-plaintiff Macmillan. This strategic decision was perhaps to make the lawsuit seem more sympathetic for publishers, for PUP's argument that it relied on permissions fees to justify publishing marginally profitable books resonated with the court. In both cases, the courts agreed with the publishers that copying beyond the negotiated guidelines for making multiple copies for classroom use cut against the commercial copy shop's fair use argument.[196] In neither case was CCC mentioned as a possible source of licensing of in-copyright materials in coursepacks, although in the aftermath of these cases, CCC served this function for many copy shops. Copy shops, not universities, typically did rights clearances,[197] which meant that the permissions fees were invisible to the students and faculty who used coursepacks instead of textbooks for assigned readings.

CUP was the first educational use copying lawsuit brought directly against a university. It teed up the issue long latent in *Texaco* and the commercial coursepack cases: whether nonprofit educational use copying was fair use or infringement. The publishers' victories in the coursepack cases made them confident they would prevail in *CUP*; after all, online course websites and reserves were functional equivalents of coursepacks.

Yet, after a dozen years of litigation, publishers prevailed as to only eleven (21 percent) of the forty-eight claimed infringements that went to trial.[198] The *CUP* majority was more impressed with the nonprofit educational purpose of the excerpts than with the publishers' nontransformative or same intrinsic purpose assertions.[199] The majority

195. The eight plaintiffs in *Kinko's* were "all major publishing houses in New York City." Basic Books, Inc. v. Kinko's Graphics Corp., 758 F. Supp. 1522, 1526 (S.D.N.Y. 1991). College professors had selected the works that were the subject of this litigation for their particular courses. *See id.* at 1546; *see also id.* at 1527–28 (titles of twelve works at issue).

196. *See supra* text accompanying notes 151 & 159.

197. *See, e.g.*, Princeton Univ. Press v. Mich. Document Servs., Inc., 99 F.3d 1381, 1394 (6th Cir. 1996) (detailing how copy shops could obtain permissions); *Basic Books*, 758 F. Supp. at 1532 (describing how Kinko's services included assuming responsibility for permission process). More recently, other parties may be responsible for coursepack permissions, including the course instructor or librarian. *See* CCC Campus Guide, *supra* note 154.

198. *See supra* note 188 and accompanying text.

199. The court mentioned that GSU required every instructor to fill out a fair use checklist before posting book chapters online. *See* Cambridge Univ. Press v. Patton, 769 F.3d 1232, 1242–43 (11th Cir. 2014), *rev'g* Cambridge Univ. Press v. Becker, 863 F. Supp. 2d 1190 (N.D. Ga. 2012). Most book chapters at issue were assigned for small seminars on specialized topics, and professors had often assigned books to be purchased for the course or seminar. *See Becker*, 863 F. Supp. 2d at 1218.

was satisfied with the trial court's consideration of the reasonableness of the amount used in light of the professor's pedagogical purpose. It was notably more skeptical of the publishers' lost licensing fee arguments than the courts in *Texaco* and the coursepack cases. It focused on whether the university's failure to pay permission fees would cause "*substantial* economic harm such that allowing it would frustrate the purposes of copyright by materially impairing [the plaintiffs'] incentive to publish the work."[200] It expressed concern that "licensing poses a particular threat that the fair use analysis will become circular."[201] It rejected the publishers' argument that a failure to pay for permissions constituted market harm,[202] observing that "it is not determinative that programs exist through which universities may license excerpts of Plaintiffs' works. In other words, the fact that Plaintiffs have made paying easier does not automatically dictate a right to payment."[203]

Although the *CUP* majority mentioned CCC only in passing, it may have been aware that CUP and its fellow plaintiffs were not funding that lawsuit. The litigation costs were borne by the AAP and CCC,[204] both of which were probably hoping for a test case victory that would enable a very large new market to develop for obtaining permissions fees from universities across the nation. The far narrower victory than the publishers anticipated has not resulted in the emergence of that market.

Since the *CUP* litigation ended, educators have had less reason to worry about making multiple copies of parts of in-copyright works for classroom use because AAP and CCC are unlikely to bring another test case for online postings of course materials. It thus appears that fair use has accommodated the interests of educators even without the special exemption for which they lobbied in the 1960s and 1970s. The educational materials ecosystem has, moreover, evolved considerably since the *CUP* lawsuit, for open access textbooks, articles, and other materials posted on the internet have become plentiful.

Yet, through digital licensing, publishers have been able to exercise greater control over certain uses of their works than they could with analog copies. E-books are typically licensed with restrictions on many kinds of

200. *Patton*, 769 F.3d at 1276.
201. *Id.* at 1278.
202. *Id.* at 1276.
203. *Id.*
204. *See Becker*, 863 F. Supp. 2d at 1213.

uses.[205] E-books are also typically protected with technical measures to inhibit users' abilities to make copies.[206] Moreover, journal publishers now have greater leverage with libraries because libraries that decide not to renew subscriptions will lose access to back issues and volumes.[207] Digitizing print books to enable a library to make book contents available through controlled digital lending has been found an unfair use of the books.[208]

B. Home Video and Audio Taping

Xerox machines were not the only disruptive reprographic technology that deeply upset certain copyright industry sectors. Tape recording machines were another such technology. Starting in the 1960s, these machines became cheap enough that millions of people were able to purchase and easily use them to make copies of their favorite songs and, later, movies, which they could then listen to or watch in the privacy of their own homes or other nonpublic spaces.

Copyright owners of musical sound recordings and movies regarded home copying as infringement and sometimes sought to ban or strictly regulate the machines that enabled this copying.[209] They typically claimed that home tapers were harming their markets, but because the copying was difficult to detect, let alone to enforce, lawsuits against individual copiers were rare. Some proposed the creation of compulsory licenses so that the copying could continue while enabling some compensation to copyright owners, but the only legislative experiment with compulsory licensing for home audio recording in the United States failed to live up to its promise.[210]

205. *See Authors Alliance Supports Model Ebook Legislation*, AUTHORS ALL. (July 7, 2022), https://www.authorsalliance.org/2022/07/07/authors-alliance-supports-model-ebook-legislation [https://perma.cc/A4N9-6TM7].

206. *See* Mike Harman, *Ebook Protection: Important Things You Should Know to Prevent Anyone From Copying and Selling Your eBook*, KITABOO (Dec. 7, 2022), https://kitaboo.com/protect-ebooks-from-copying-and-selling [https://perma.cc/624B-XBLM].

207. *See* COMM. ON INTELL. PROP. RTS. & THE EMERGING INFO. INFRASTRUCTURE ET. AL., THE DIGITAL DILEMMA: INTELLECTUAL PROPERTY IN THE INFORMATION AGE 100–04 (2000).

208. *See* Hachette Book Grp., Inc. v. Internet Archive, 664 F. Supp. 3d 370, 391 (S.D.N.Y. 2023), *argued*, No. 23-1260 (2d Cir. June 28, 2024). For an explanation of controlled digital lending, see *FAQ*, CONTROLLED DIGIT. LENDING, https://controlleddigitallending.org/faq [https://perma.cc/A8TM-9CN5].

209. *See supra* Parts II.B.1, II.B.3, II.B.4.

210. *See infra* text accompanying notes 309–322 for a discussion of the Audio Home Recording Act (AHRA). The EU imposes a tax on tape recording machines to generate

Regulating home taping was especially difficult because it largely happened in people's homes or other nonpublic spaces, not in institutional settings such as libraries or large corporations where copying could be monitored and licensing was feasible. As with photocopying, a serious problem for the copyright industries was that ordinary people who made these copies typically thought they were doing nothing wrong.

Although each individual's home taping did not harm copyright markets, there was some logic behind the industry's assertions that the aggregate effect of millions of people making such copies constituted market harm. There were, however, reasons to question industry estimates about lost revenues from home copies, as it was unclear that home tapers would purchase authorized copies if unable to tape music or programs they liked, and much of the copying may have made listening or viewing more convenient (for example, taping music to listen in one's car).

1. Home Video Taping

Sony began selling video tape recording (VTR, also known as Betamax) machines in the U.S. market in 1975.[211] Soon thereafter, Universal City Studios and Walt Disney sued Sony for contributory infringement, claiming that it sold these VTRs knowing that it was materially contributing to the infringing acts of consumers who purchased VTRs to make unauthorized (and allegedly infringing) copies of Universal and Disney movies from broadcast television.[212] They also claimed Sony had vicariously infringed their copyrights, because Sony had the right and ability to control the design of the machines and, hence, how consumers were able to use these machines to make copies of programs, and because Sony financially benefited from consumers' infringements.[213] They further asserted that they would suffer "great monetary damage" if Sony was allowed to continue to sell VTRs.[214] They also sought an injunction to prevent further sales and an order requiring Sony to redesign Betamax machines so they could no longer be

revenues for copyright owners. *See* WORLD INTELL. PROP. ORG., INTERNATIONAL SURVEY ON PRIVATE COPYING 1, 29 (2012), https://www.wipo.int/edocs/pubdocs/en/copyright/1037/wipo_pub_1037_2013.pdf [https://perma.cc/T86X-UR7N].

211. *See, e.g.*, Jessica Litman, *The* Sony *Paradox*, 55 CASE W. RSRV. L. REV. 917, 921 (2005).

212. Universal City Studios, Inc. v. Sony Corp. of Am., 480 F. Supp. 429 (C.D. Cal. 1979), *aff'd in part, rev'd in part*, 659 F.2d 963 (9th Cir. 1981), *rev'd*, 464 U.S. 417 (1984). For a history of this controversy, see, for example, JAMES LARDNER, FAST FORWARD: HOLLYWOOD, THE JAPANESE, AND THE VCR WARS (1988).

213. *Sony*, 480 F. Supp. at 432, 461–62.

214. *Id.* at 432.

used to infringe copyrights.[215] More grotesquely, Jack Valenti, then head of the Motion Picture Association of America, told members of Congress that VTRs are "to the American film producer and the American public as the Boston strangler is to the woman home alone."[216]

Sony was the first lawsuit to charge a manufacturer of a disruptive reprographic technology with copyright infringement. Following a five week trial, a trial court ruled against Universal and Disney.[217] It agreed with Sony that in 1971, Congress had created an implied exemption from liability for home taping,[218] relying in part on a House report stating that "it is not the intention of the Committee to restrain the home recording, from broadcasts or from tapes or records, of recorded performances, where the home recording is for private use and with no purpose of reproducing or otherwise capitalizing commercially on it."[219] The trial court believed that Congress had carried over this implied exemption into the 1976 Act.[220]

Yet, the *Sony* trial court also engaged in a lengthy fair use analysis. Because of its importance in fair use cases, the court addressed the market effects factor first, noting that Universal had conceded that it had suffered no harm to date.[221] Universal's experts testified about several types of potential future harms, including that people might go out to see movies less often if they could watch home-taped movies instead. Although the experts claimed that future harms were "probable" and "imminent," they were unable to say when the alleged harms would manifest.[222] Because the experts relied heavily on numerous assumptions, some "based on neither fact nor experience," the court found the plaintiffs' harm evidence unpersuasive.[223]

215. *See id.* at 468–69.
216. *See Home Recording of Copyrighted Works: Hearings on H.R. 4783, H.R. 4794, H.R. 4808, H.R. 5250, H.R. 5488 and H.R. 5707 Before the Subcomm. on Cts., C.L., and the Admin. of Just. of the H. Comm. on the Judiciary*, 97th Cong. 8 (1982) (statement of Jack Valenti, President, MPAA).
217. *Sony*, 480 F. Supp. at 432–33.
218. *Id.* at 444–46.
219. H.R. REP. NO. 92-487, at 7 (1971), *quoted in Sony*, 480 F. Supp. at 444; *see also id.* at 445–47 (discussing testimony). Professor Nimmer strongly criticized this aspect of the *Sony* opinion. *See* Melville B. Nimmer, *Copyright Liability for Audio Home Recording: Dispelling the* Betamax *Myth*, 68 VA. L. REV. 1505 (1982). He pointed out that there was no counterpart to the House Report statement in the Senate Report and discounted the court's reliance on other legislative-related statements. *Id.* at 1509–14.
220. *See Sony*, 480 F. Supp. at 445. Nimmer also criticized the trial court in *Sony* on this point, noting that there was nothing in the statute or legislative history of the 1976 Act to support this assertion. Nimmer, *supra* note 219, at 1514–19.
221. *See Sony*, 480 F. Supp. at 451.
222. *Id.* at 451.
223. *Id.* at 451–52.

Yet, even if some future harm was probable, the court said home copying of television programs would not necessarily be unfair.[224]

Like the eventual Supreme Court majority in *Sony*, the trial court was impressed by the private noncommercial nature of the copying and by Universal's decision to license its movies for television broadcasts. It also concluded that copying of entire programs did not negate fair use because home copies were not harming Universal's markets.[225] It regarded "library" copies (that is, programs permanently recorded for repeat viewing) as equally fair as time-shift copies (that is, programs recorded to watch at a later time, which were temporary in character, as home tapers reused videotapes the next time they wanted to time-shift other programs).[226] Given this fair use ruling, neither Universal's contributory nor vicarious infringement claims could survive.[227]

The Ninth Circuit reversed. It first rejected the implied exemption theory for two principal reasons: First, the 1976 Act was "unambiguous; the grant of exclusive rights is only limited by the statutory exemptions," and second, the legislative history of the 1971 Act concerned home audio taping, not videotaping.[228] It disagreed sharply with the trial court's fair use analysis. Home copying, it noted, was not among the kinds of productive uses of preexisting works to make new works (for example, criticism and comment) that tend to favor fair use; nor were home copies for nonprofit educational purposes.[229] It regarded the second and third factors as disfavoring fair use because Universal's movies were creative works and the copying was of entire programs.[230] Because these three factors disfavored Sony's fair use

224. *See id.* at 452. Professor Nimmer recognized that fair use defenses were more plausible as to home video taping than home audio taping. Nimmer, *supra* note 219, at 1524 n.68. There was in the 1970s and early 1980s no existing market for movie videotapes, although there were active markets for audio tapes of sound recordings. Besides, the most common use of VTRs was for temporary time-shift purposes (to watch programs at a later time), whereas home audio tapers typically made permanent copies of sound recordings.
225. *See Sony*, 480 F. Supp. at 452–56.
226. *Id.* at 466–68.
227. *See id.* at 459–62. The court invoked patent law's staple article of commerce limit on contributory infringement liability, 35 U.S.C. § 271(c), which allows developers to make and sell technologies having substantial noninfringing uses. *Id.* at 461.
228. Universal City Studios, Inc. v. Sony Corp. of Am., 659 F.2d 963, 966–69 (9th Cir. 1981), *rev'd*, 464 U.S. 417 (1984). The Ninth Circuit cited Nimmer's work twenty times.
229. *See id.* at 970–72. The Ninth Circuit said the lower court "stretches fair use beyond recognition and undermines our traditional reliance on the economic incentives provided to authors by the copyright scheme." *Id.* at 970.
230. *See id.* at 972–73.

defense, the Ninth Circuit thought it unnecessary to address the market effects factor. It criticized the lower court, however, for being "much too strict in requiring [Universal] to establish its degree of harm," saying a tendency to diminish or prejudice copyright markets should suffice, and for failing to appreciate the difficulties copyright plaintiffs have in proving potential harms to markets.[231] Given its unfair use ruling, the Ninth Circuit held that Betamax machines lacked substantial noninfringing uses and ruled that Sony was a contributory infringer because it knew its customers would use their VTRs to make infringing copies of programs.[232]

Justice Stevens' *Sony* decision for a 5–4 majority reversed the Ninth Circuit, holding that private noncommercial home copying of Universal movies for time-shifting purposes was fair use and that Betamax machines were staple articles of commerce which consumers should be able to buy for their noninfringing uses.[233] This established a safe harbor from copyright liability for technologies capable of substantial noninfringing uses.[234]

The main difference between Justice Stevens's and the trial court's analysis of fair use was that Stevens focused the Court's fair use analysis only on copying for time-shifting purposes and said nothing about "library" copying.[235] His opinion did not engage in a factor-by-factor fair use analysis,[236] although it addressed the market effects factor at some length. It recounted the trial court's reasons for concluding that Universal's potential market harm arguments were unpersuasive, noting that there was ample support in the record for its findings.[237] Justice Stevens articulated a standard for assessing market effects in noncommercial copying cases:

> A challenge to a noncommercial use of a copyrighted work requires proof either that the particular use is harmful, or that if it should become widespread, it would adversely affect the potential market

231. *Id.* at 973–74.
232. *See id.* at 975–76.
233. *See* Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 456 (1984).
234. *See id.* at 442. The majority did not address Sony's implied exemption for home copying defense. *See id.* at 430 n.11.
235. Given how deeply split the Court was on the time-shift copying issue, the implication was that "library" home taping was probably not fair use.
236. The majority concluded that the purpose of the time-shift copying was private and noncommercial, that Universal had chosen to license its programs for broadcasts over the air waves, and that copying of entire programs did not defeat fair use given the lack of market harm. *See id.* at 442, 449–51. The dissent criticized the majority for virtually ignoring the nature and amount factors. *See id.* at 496–97 (Blackmun, J., dissenting). It also disagreed with the majority on the purpose and harm factors. *See id.* at 472–85.
237. *See id.* at 452–54 (majority opinion).

> for the copyrighted work. Actual present harm need not be shown; such a requirement would leave the copyright holder with no defense against predictable damage. Nor is it necessary to show with certainty that future harm will result. What is necessary is a showing by a preponderance of the evidence that *some* meaningful likelihood of future harm exists.[238]

The Court thought Universal should bear the burden of proving that noncommercial time-shift copying would cause potential harm to the studios' markets and agreed with the trial court that the plaintiffs had failed to proffer more than speculation to support their claim.[239] In discussing market effects, Justice Stevens pointed to a social benefit resulting from the *Sony* ruling: the public would have greater access to broadcast television programs.[240]

Justice Blackmun dissented, saying that fair use should generally not be available to ordinary consumptive uses such as time-shift copying and fair use should be reserved for productive uses which resulted in the creation of new works that build on preexisting works.[241] When assessing market effects in ordinary use cases, Blackmun thought a showing of a potential for harm should suffice.[242] Universal had, after all, identified some ways in which VTR recording could cause harm:

> VTR recording could reduce their ability to market their works in movie theaters and through the rental or sale of prerecorded videotapes or videodiscs; it also could reduce their rerun audience, and consequently the license fees available to them for repeated showings. Moreover, advertisers may be willing to pay for only "live" viewing audiences, if they believe VTR viewers will delete commercials or if rating services are unable to measure VTR use; if this is the case, VTR recording could reduce the license fees the Studios are able to charge even for first-run showings. Library-building may raise the potential for each of the types of harm identified by the Studios, and time-shifting may raise the potential for substantial harm as well.[243]

238. *Id.* at 451.
239. *See id.* at 449–51 (asserting a presumption that noncommercial copying is fair); *see also id.* at 456 (concluding that Universal failed to show time-shifting caused it harm).
240. *See id.* at 454.
241. *See id.* at 472–79 (Blackmun, J., dissenting).
242. *See id.* at 482.
243. *Id.* at 483 (citation omitted).

The dissent also observed that people who wanted to watch programs at different times or at their convenience might be willing to pay for the right to watch programs at a different time, and Universal should be able to share in the value of that market.[244]

2. An Interlude on *Sony* and the Concept of Copyright Harms

During the 1960s through 1980s, when many in the copyright industries perceived machine copying—such as library photocopying for patrons and home taping of entertainments—as serious problems, some commentators were scornful of the idea that copyright owners should have to prove some actual economic harm or provide evidence of likely harm to potential markets to defeat fair use defenses. In the early 1980s, David Ladd, then-Register of Copyrights, regarded the view that economic harm was a necessary precondition to imposing liability for copyright infringement or for remedies for infringement as "a dangerous concept" that "may alter and could destroy the role of copyright as a principal mechanism for fostering intellectual activity."[245] He was critical of the Court of Claims in *W&W* and the trial court in *Sony* for their assertions that the absence of evidence of harm supported fair use defenses.[246] In his view, "potential market" should extend as broadly as the exclusive rights granted in the 1976 Act.[247]

Along similar lines, a law firm memorandum prepared for a legislative hearing stated that the economic harm concept "conflict[ed] with the fundamental concept underlying copyright law—the absolute right of the

244. *See id.* at 485.

245. David Ladd, *Economic Harm: A Trojan Horse in Copyright*, *Appendix C to the* 1982 A.B.A. SECTION PAT. TRADEMARK & COPYRIGHT L. PROC. 157, 157 (1982) (hereinafter Ladd, *Trojan Horse*); *see also* David Ladd, Thirteenth Donald C. Brace Memorial Lecture, *The Harm of the Concept of Harm in Copyright*, 30 J. COPYRIGHT SOC'Y U.S.A. 421 (1983). Ladd thought the concept of harm was itself harmful, "amorphous and practically impossible to define or to apply consistently," adding that "it threatens to warp the universe of both authors and publishers." *Id.* at 422–23. For a contemporary expression of similar views, see PAUL GOLDSTEIN, GOLDSTEIN ON COPYRIGHT § 1.14.2.3(c), at 1:57–1:58 (3d ed. Supp. 2023-2) (positing that copyright owners do not need to show harm to recover for infringement). But a strict liability rule without regard to a showing of harm is inconsistent with recent U.S. Supreme Court rulings stating that "concrete" harms must exist for plaintiffs to have standing to bring a lawsuit in federal court. *See* TransUnion LLC v. Ramirez, 594 U.S. 413, 427 (2021) ("Only those plaintiffs who have been *concretely harmed* by a defendant's statutory violation may sue that private defendant over that violation in federal court."); Spokeo, Inc. v. Robins, 578 U.S. 330, 340–41 (2016) (discussing requirements to establish "injury in fact").

246. *See* Ladd, *Trojan Horse*, *supra* note 245, at 157.

247. *Id.* at 159.

copyright owner to control and therefore prohibit all unauthorized uses of his property."[248] The absence of evidence of harm had, in this view, no bearing on liability or an entitlement to damages.[249] It pointed to the availability of statutory damages, which required no proof of harm, to support this assertion.[250] It made dire predictions about eroded incentives to create and disseminate works as well as the prospect of bankruptcy or massive losses for copyright industries if plaintiffs had to prove harm to justify monetary awards.[251]

Economic harm became, however, a much more salient consideration in fair use cases after *Sony*. That decision changed copyright law discourse about market effects in several respects. First, the Court paid much more attention to this factor than to the others, signaling its considerable importance in fair use cases.[252] Second, it reviewed at length the evidentiary basis for the trial court's conclusion that Universal had tried, but failed, to show that time-shifting copying would cause market harm.[253] Third, it

248. AKIN, GUMP, STRAUSS, HAUER & FELD, ECONOMIC HARM: AN INAPPROPRIATE AND DAMAGING PREREQUISITE FOR COPYRIGHT PROTECTION, *prepared for Video and Audio Home Taping: Hearing on S. 31 and S. 175 Before the Subcomm. on Pats., Copyrights, and Trademarks of the S. Comm. on the Judiciary*, 98th Cong. 483, 483 (1983) (appended memorandum) [hereinafter Akin Gump Memorandum]. Some commentators have analogized infringement of copyright to trespass to real property, in which the act of trespass is itself deemed a harm, but this analogy has been debunked in subsequent literature. *See, e.g.*, Wendy J. Gordon, *The Concept of "Harm" in Copyright*, *in* INTELLECTUAL PROPERTY AND THE COMMON LAW (Shyamkrishna Balganesh ed., 2013); *see also* eBay, Inc. v. MercExchange, LLC., 547 U.S. 388 (2006) (rejecting the theory that violating exclusive rights entitles owner to injunctive relief and putting the burden on plaintiffs to show entitlement to injunctive relief, including by proving they will suffer irreparable harm unless an injunction issues).

249. *See* Akin Gump Memorandum, *supra* note 248, at 485. But this view is inconsistent with the text of § 107(4), which directs courts to consider market effects as part of the fair use analysis. Nor, in this view, would the absence of harm affect the copyright owners' entitlement to injunctive relief, otherwise infringers would enjoy a "free mandatory judicial license" for their misconduct, which would "virtually negate" the exclusive nature of copyright rights. *Id.* at 486–87; *see also* Daniel M. Jochnowitz, Note, *Proof of Harm: A Dangerous Prerequisite for Copyright Protection*, 10 COLUM.-VLA J.L & ARTS 153, 157–62 (1985). *But see* eBay Inc. v. MercExchange, L.L.C., 547 U.S. 388, 391 (2006) (requiring plaintiffs to prove irreparable harm and inadequacy of damages to qualify for injunctive relief).

250. *See* Akin Gump Memorandum, *supra* note 248, at 487–88. *But see* RESTATEMENT OF COPYRIGHT § 9.03 (AM. L. INST., Tentative Draft No. 3, 2022) (stating that awards of statutory damages should generally approximate actual damages).

251. Akin Gump Memorandum, *supra* note 248, at 489–91.

252. *Sony* discussed the market effects factor at length. *See* Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 450–54 (1984); *cf. id.* at 448–50 (discussing other three factors).

253. *See id.* at 451–54.

recognized that "[a]ctual present harm need not be shown" because "such a requirement would leave the copyright holder with no defense against predictable damage"; nor was it "necessary to show with certainty that future harm will result."[254] It thereby allayed concerns raised by Ladd and others about putting too heavy a burden on plaintiffs to show economic harm. Fourth, it directed courts to assess whether market harm would result if the challenged use became widespread, recognizing that even if one particular use by one user did not harm copyright markets, the cumulative effect of many such uses might cause harm.[255] Fifth, the Court distinguished commercial and noncommercial uses, indicating that the former were more likely than the latter to harm copyright markets.[256] Sixth, it endorsed presumptions that commercial uses are unfair and that noncommercial uses are fair, putting the burden on the copyright owner to overcome the latter presumption.[257] Seventh, it linked the assessment of market effects to copyright owner incentives, observing that "a use that has no demonstrable effect upon the potential market for, or the value of, the copyrighted work need not be prohibited in order to protect the author's incentives to create."[258] Eighth, it focused attention on standing issues. It pointed out that Universal and Disney had not brought a class action lawsuit against Sony, although the relief they were requesting would affect other copyright owners, including many who did not object to time-shifting or actively supported it.[259] Ninth, it repudiated the analogy of unauthorized copying of in-copyright works to theft.[260]

While the Court in *Sony* did not invoke the economic concept of market failure, *Sony* has become associated with the market failure rationale for fair use under the influence of Wendy Gordon's article, which offered some

254. *Id.* at 451. The Court did not endorse the Ninth Circuit's "tend[ency] to diminish or prejudice … potential sale" standard, which derived from the Nimmer treatise. *See* Universal City Studios, Inc. v. Sony Corp. Am., 659 F.2d 963, 974 (9th Cir. 1981), *rev'd*, 464 U.S. 417 (1984).
255. *Sony*, 464 U.S. at 451.
256. *Id*.
257. *Campbell* repudiated the commercial purpose presumption. *See supra* text accompanying note 42.
258. *Sony*, 464 U.S. at 450–51.
259. *See id.* at 434. Recall that *Texaco* was a class action. Had the Justices paid attention to this issue in *W&W*, there should have been a unanimous Court affirmance of the Court of Claims appellate panel.
260. *See Sony*, 464 U.S. at 450 n.33.

support for fair use in cases such as *Sony*.[261] Gordon set forth a three-part market failure inquiry to assess fair uses: (1) whether the user could purchase the right to engage in the challenged use, or whether a market failure impeded voluntary transactions; (2) whether some societal benefit would occur if the user was permitted to engage in the challenged use; and (3) whether the copyright owner's incentives to invest in creation and dissemination of protected works would be substantially impaired if the use was held fair.[262]

While Gordon suggested that *Sony* should have been remanded for a detailed inquiry about a market failure basis for fair use, she noted that home users of Betamax machines could not negotiate with copyright owners to obtain licenses to tape programs in a timely fashion, and enforcing a ban on home taping would be virtually impossible.[263] Increased public access to television programs was a public benefit of home taping, and Universal's incentives to invest in making motion pictures were unlikely to be substantially impaired by home taping. Gordon recognized that a market cure might occur through imposing a tax on the sale of machines and tapes, but raised concerns about the practicability of this solution.[264] Under the market failure conception of fair use, economic harm (or its absence) is a very pertinent consideration.[265]

During oral argument, Sony's counsel emphasized that American households already owned an estimated 9.5 million VTRs and questioned whether their owners should be deemed infringers for watching broadcast programs at a later time.[266] Sony's lawyer pointed out that Universal and Disney had said they would not sue individual home tapers for infringement and questioned whether it was equitable to hold Sony liable as a contributory

261. *See* Gordon, *supra* note 118. The article reacted to the Ninth Circuit ruling in *Sony*. Gordon also applied this framework to *W&W*. *Id.* at 1647–52. For recognition of the influence of Gordon's article interpreting *Sony*, see, for example, Glynn S. Lunney, Jr., *Fair Use and Market Failure:* Sony *Revisited*, 82 B.U. L. REV. 975, 985–88 (2002).
262. *See* Gordon, *supra* note 118, at 1601.
263. *See id.* at 1655, 1657.
264. *See id.* at 1655–56.
265. In subsequent fair use cases, courts have taken market failure and the potential for market cure into account. *See, e.g.*, Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 592 (1994) (recognizing that copyright owners are unlikely to want to license lampoons or other critical uses of their works, which "removes such uses from the very notion of a potential licensing market"); Am. Geophysical Union v. Texaco, Inc., 60 F.3d 913, 930 (2nd Cir. 1994) (acknowledging CCC as workable solution for obtaining photocopying permissions).
266. *See* Transcript of Oral Argument at 4–5, Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417 (1984) (No. 81-1687) [hereinafter Sony Oral Argument].

infringer if individual tapers were off the hook.[267] Universal's counsel emphasized that Sony had built a billion dollar business that Sony knew was built on infringement, pointing to an estimate that forty to fifty million VTRs would be in American homes by the end of the 1980s and fifty-five million blank tapes would be sold that very year.[268] Given the huge volume of tape sales and the tens of millions of copies being made of copyrighted programs, the economic loss to the movie industry "will be enormous," argued Universal, which was why copyright owners should get some compensation for this copying.[269] Universal's lawyer explained that they had not put forward evidence of harm because they were relying on the presumption of harm from the fact of copying and seeking only statutory damages.[270] On rebuttal, Sony's lawyer pointed out that not only had Universal's counsel asked for an injunction against further manufacture of VTRs, but it had also asked the trial court to order a recall of all of the VTRs that Sony had sold so they could be "disemboweled"—that is, disabled, so they could no longer be used to record programs off the air—and to award Universal and Disney the statutory damage minimum of $250 for each unauthorized copy made by Betamax owners.[271] Sony's lawyer ended by reminding the Court that broadcasters had handsomely paid Universal for making its films available to millions of American households, an audience made wider because of Betamax time-shifting.[272] Even so, the Court split 5–4 on the holding that time-shifting was fair use.[273]

In retrospect, Universal was arguably lucky to have lost this lawsuit because VTRs opened a new market for video cassette tapes of movies, followed by new markets for DVD and then Blu-Ray versions of movies, all of which became huge revenue sources for motion picture studios.[274] The installed base of VTR machines made home viewing of movies a much greater

267. *See id.* at 18. Universal actually did sue William Griffiths—who used his Betamax VTR to make copies of broadcast programs—as a nominal defendant in the *Sony* case, but sought no remedy against him. *Sony*, 464 U.S. at 422 n.2.
268. *See* Sony Oral Argument, *supra* note 266, at 30, 36.
269. *Id.* at 29–30.
270. *Id.* at 30–31.
271. *Id.* at 47.
272. *Id.* at 51–52; *see also id.* at 33–35.
273. Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 419, 457 (1984).
274. *See* Derek Khanna, *A Look Back at How the Content Industry Almost Killed Blockbuster and Netflix (and the VCR)*, TECHCRUNCH (Dec. 27, 2013, 9:00 AM), https://techcrunch.com/2013/12/27/how-the-content-industry-almost-killed-blockbuster-and-netflix [https://perma.cc/YPM2-Q8NB] (describing how, post-*Sony*, revenues from video tape sales and rentals eclipsed box office revenues and rose to $18 billion in 2012).

success than if the studios had prevailed. Over time, on-demand streaming services have provided motion picture studios with new opportunities to recoup investments by making valuable audiovisual works available through subscriptions to very large audiences.[275]

3. Home Video Recording Revisited

Home recording of video programs gave rise to a second copyright lawsuit in *Cartoon Network LP v. CSC Holdings, Inc.*[276] Cartoon Network, among others, sued for a declaratory judgment that the remote-server digital video recording (RS-DVR) feature that Cablevision was planning to offer its customers would infringe their copyrights. Cablevision had a license to transmit video content from these plaintiffs,[277] but the plaintiffs brought suit because the RS-DVR service would allow Cablevision customers to make time-shift copies of programs to be stored on and later played back from Cablevision's servers, which the plaintiffs claimed was infringement without additional licensing.[278]

Cartoon Network claimed that the RS-DVR system infringed the reproduction right in two ways, both involving Cablevision's computer buffers. The first focused on alleged "copies" made in RS-DVR's primary ingest buffer (a form of Random Access Memory, or RAM).[279] The second focused on copies stored in user-designated hard drives for later playback.[280] The Second Circuit was unpersuaded by both claims.

Addressing the primary ingest buffer issue, the court observed that a reproduction of a copyrighted work must be "fixed" in a tangible medium "for

275. Yet, copyright owners still challenged home taping in subsequent litigation. *See infra* Part II.B.3 (discussing litigation concerning copies of programs stored on company's servers).

276. 536 F.3d 121 (2d Cir. 2008) (hereinafter Cablevision). Although CSC Holdings (Cablevision) did not raise a fair use defense, the case merits attention because it demonstrates the continuing efforts by program producers to demand licensing revenues when technology advances make it possible for consumers to make time-shift copies of television programs. *See also* Fox Broad. Co. v. Dish Network LLC, 747 F.3d 1060, 1067–69 (9th Cir. 2014) (affirming denial of a preliminary injunction that would have prohibited a cable company from providing customers with tools to make time-shift copies of programs and to skip commercials, holding that Fox was unlikely to prevail with direct or indirect infringement claims, invoking *Sony*).

277. *See Cablevision*, 536 F.3d at 123.

278. *See id.* at 124 (describing RS-DVR feature, which would allow customers to record programs and receive playback at their choosing). The district court awarded summary judgment to the plaintiffs and enjoined operation of the RS-DVR system unless it obtained the necessary licenses. *See id.*

279. *See id.* at 128.

280. *See id.* at 125.

a period of more than [a] transitory duration" to infringe that right.[281] Because the buffer copies were stored for no more than 1.2 seconds, the court decided they were too transitory to satisfy the duration requirement and hence did not infringe.[282]

As for the playback copies, Cartoon Network argued that Cablevision made those copies on the company's servers, likening this to the coursepack copy shops in *MDS*.[283] The Second Circuit decided that Cablevision's customers' volitional conduct caused those copies, similar to shops that supply photocopy machines so customers could make their own copies.[284]

The court found Cartoon Network's public performance claim more challenging. Cablevision argued that customers who received stored copies of programs were performing them, but the Second Circuit ruled that Cablevision was the performer because the 1976 Act focuses on whoever "transmit[s] or otherwise communicate[s] a performance."[285] Customers could not "perform" stored programs because they were incapable of transmitting the programs to themselves.

The trickier question was whether the performance was *to the public* or was *private*. If the former, Cartoon Network would have established a prima facie case of infringement.[286] Although the statute contemplated that a performance could be "public" even if members of the public did not receive it at the same time or in the same place, the Second Circuit focused on the statutory reference to the person who was "capable of receiving the performance."[287] It concluded that "[g]iven that each RS-DVR transmission is made to a given subscriber using a copy made by that subscriber, we conclude that such a transmission is not 'to the public.'"[288] Under Cartoon Network's interpretation, no digital transmission would ever be a private performance; the Second Circuit doubted that Congress intended that

281. *Id.* at 127–29 (citing to the definitions of "copy" and "fixed" in 17 U.S.C. § 101 and to § 106(1)).
282. *See id.* at 129–30. Cablevision also raised a de minimis defense to this claim, which the court did not reach. *See id.* at 130.
283. *See id.* at 131.
284. *See id.* at 132–33.
285. *Id.* at 134–35 (referring to the 17 U.S.C. § 101 definition of "publicly" with respect to performances).
286. *See* 17 U.S.C. § 106(4) (granting the copyright owner the exclusive right to publicly perform the copyrighted work).
287. *Cablevision*, 536 F.3d at 134–35.
288. *Id.* at 138.

result.[289] Thus, the development of a technology enabling customers to time-shift television programs for private noncommercial purposes was once again held to be noninfringing.

The Second Circuit's *Cablevision* decision quoted at length from *Sony*, especially about the importance of maintaining a distinction between direct and indirect infringement claims in lawsuits against technology developers.[290] The court's ruling also, in effect, reaffirmed the fair use holding about time-shifting in *Sony*. The court was unconvinced that the RS-DVR features and consumer time-shifting should give rise to infringement liability,[291] even if the plaintiffs may have perceived an opportunity to receive a new revenue stream to share in (or capture) the value Cablevision and its customers would receive by offering this new feature.

4. Home Audio Taping

Even after passage of a 1971 law extending copyright protection to sound recordings,[292] the record industry continued to be plagued not only by counterfeiting, but also, it thought, by massive volumes of home audio taping by members of the public.[293] One widely cited study by Warner Communications reported that in 1980, 39 million people in the U.S. had home-taped audio recordings, claiming these copies had an aggregate value of $2.85 billion.[294]

289. *See id.* at 136. *But see* Am. Broad. Cos. v. Aereo, Inc., 573 U.S. 431 (2014). Aereo developed a service to enable members of the public to watch broadcast television on computing devices. *Id.* at 432. It designed the service so that customers made copies of the programs that were transmitted only to them. *Id.* The Court decided that Aereo was publicly performing the programs. *Id.* at 433–34. Although the Court's interpretation of the transmit clause of the public performance right is difficult to square with the Second Circuit's interpretation in *Cablevision*, Westlaw does not flag it as having been overturned by *Aereo*. *Aereo* may be distinguishable from *Cablevision* in that Cablevision had licensed content from television program providers and its customers had lawful access to the programs through its service and were therefore entitled to have programs transmitted to them.

290. *Cablevision*, 536 F.3d at 132–33. Cartoon Network did not make an indirect infringement claim against Cablevision, almost certainly because the cable company would have qualified for the *Sony* safe harbor.

291. *Id.* at 140.

292. Sound Recording Act of 1971, Pub. L. No. 92-140, 85 Stat. 391.

293. *See* David H. Horowitz, *The Record Rental Amendment of 1984: A Case Study in the Effort to Adapt Copyright Law to New Technology*, 12 COLUM.-VLA J.L. & ARTS 31, 32 (1987).

294. *Id.* at 46; *see also* Nimmer, *supra* note 219, at 1523 (reporting Warner's and other studies of home taping).

During the pendency of the *Sony* case, numerous bills about home taping were introduced; some aimed at exempting home taping from copyright liability and some at creating a compulsory license for home taping of sound recordings.[295] Once the Supreme Court decided *Sony*, the momentum for these bills subsided.[296]

A few years later, members of Congress asked the Office of Technology Assessment (OTA) to study home taping and identify options for legislative action. In 1989, OTA published a report which concluded that the legal status of home taping was "ambiguous," pointing out that Congress and the courts had provided only "meager guidance" about it.[297] OTA reported on a survey which estimated that 4 in 10 Americans over the age of ten were making over a billion tapes of recorded music annually, mostly to play music in their cars or on portable devices.[298] OTA discounted previous industry-sponsored studies of the economic impacts of home taping for using flawed methodologies and commissioned three studies of the economic impacts of home taping, one of which reported:

> Although home taping may reduce the recording industry's revenues, [the study's] analysis suggests that in the short term a ban on audiotaping would reduce blank-tape revenues, be more harmful to consumers than beneficial to the recording industry, and result in a loss of benefits to society in the billions of dollars.[299]

The OTA report set forth numerous options for possible congressional action on home taping, including exempting it, banning it, and creating a compulsory license to allow home taping with some compensation to right holders.[300]

295. Horowitz, *supra* note 293, at 43–44. The record industry might have preferred a ban on home taping, but compulsory licensing, with taxes on machines and tapes, was a more realistic and enforceable option. *Id.*

296. The record industry had one legislative success in 1984 when Congress passed legislation to ban commercial rentals of sound recordings. Record Rental Amendment of 1984, Pub. L. No. 98-450, § 2, 98 Stat. 1727, 1727 (codified at 17 U.S.C. § 109(b)).

297. OFF. TECH. ASSESSMENT, OTA-CIT-422, COPYRIGHT AND HOME COPYING 5 (1989).

298. *See id.* at 3, 11.

299. *Id.* at 14. The industry's studies typically assumed a one-to-one ratio of unauthorized home taping as lost sales, which is an unrealistic assumption. *See* Lewis Kurlantzick & Jacqueline E. Pennino, *The Audio Home Recording Act of 1992 and the Formation of Copyright Policy*, 45 J. COPYRIGHT SOC'Y U.S.A. 497, 505–07 (1998). Such studies also ignore possible positive effects of copying on sales. *Id.* at 508–09. Much home taping aimed to make music portable. *Id.* at 514.

300. *See* OFF. TECH. ASSESSMENT, *supra* note 297, at 20–21.

Although OTA's survey showed that people widely believed home taping was lawful,[301] a view arguably reinforced by the *Sony* decision, Professor Nimmer was among those who strongly believed that home audio taping infringed both musical work and sound recording copyrights. Audio home taping was, in his view, unfair because it was an unproductive use of creative works, which were typically copied in their entirety, thereby supplanting demand for the original.[302] He thought it was "empirically obvious" that "audio home recording does have a devastating impact on the market for music and sound recordings."[303] He regarded home audio taping to be more commercially harmful than home video taping because virtually all home taping of sound recordings made permanent copies, not time-shifting ones.[304] Even if each individual noncommercial audio tape did not harm markets, he thought the cumulative effect of billions of home copies must have had a significant effect on record industry markets.[305]

The emergence of digital audio tape (DAT) recording technologies in the late 1980s brought the long-simmering home taping debate to a boiling point. The record industry perceived DAT devices as substantial threats to its business because these devices could make multiple perfect digital copies of commercially valuable sound recordings; in effect, each DAT player could become a factory for making digital copies.[306] These machines would obviously have substantial noninfringing uses, such as backup copying, so it was questionable whether a lawsuit challenging their legality could succeed after *Sony*. But because the primary use of DAT machines was likely to make permanent or "library" copies of sound recordings—the legality of which was left hanging in *Sony*—there was a nontrivial chance that courts might uphold a contributory infringement claim against their makers.

The songwriter Sammy Cahn and some music publishers decided to test this by suing Sony to stop it from selling DAT recorders.[307] The *Cahn v. Sony* litigation settled with an agreement between the music and consumer

301. *Id.* at 3. This was true both of home-tapers and non-home-tapers. *Id.*
302. *See* Nimmer, *supra* note 219, at 1520–25.
303. *Id.* at 1524.
304. *See id.* at 1524 n.68.
305. *See id.* at 1524.
306. *See* OFF. TECH. ASSESSMENT, *supra* note 297, at 3 (summarizing "Reasons for Concern"). Analog audio tapes, by contrast, would degrade in quality if users tried to make multiple copies from existing taped copies. *Id.*
307. *See* Complaint Class Action, Cahn v. Sony Corp., No. 90 Civ. 4537 (S.D.N.Y. filed July 9, 1990); *see also* Kurlantzick & Pennino, *supra* note 299, at 500–01.

electronics industries, under which it would be lawful to make and sell DAT machines and tapes capable of making first-generation copies of digital sound recordings, as long as the devices included serial copy management system (SCMS) chips to prevent DAT machines from making multiple copies of music from first-generation copies.[308] The settling parties agreed to support legislation to implement this compromise, which became the Audio Home Recording Act of 1992 (AHRA).[309] The computer industry agreed to support this legislation as long as computers and hard drives were excluded from AHRA's strictures.[310]

Several features of the AHRA, apart from the SCMS chip requirement,[311] are noteworthy. First, AHRA was designed to generate revenues for copyright owners by requiring a levy of 2 percent of the price of DAT machines and 3 percent of the price of DAT tapes be deposited into funds for distribution to owners of recorded music and music copyrights.[312] Second, AHRA provided that no actions could be brought that charged infringement based on consumers' noncommercial use of audio recording devices or media, whether against makers of such devices or media or against consumers themselves.[313] This provision resolved the longstanding controversy about the legality of noncommercial consumer home taping.[314] Third, another provision prohibited development of technologies or

308. *See* Gary S. Lutzker, Note, *Dat's All Folks:* Cahn v. Sony *and the Audio Home Recording Act of 1991—Merrie Melodies or Looney Tunes?*, 11 CARDOZO ARTS & ENT. L.J. 145, 170, 174–75 (1992). In this respect, the serial copy management system (SCMS) chip made DAT machines more like analog cassette recorders whose users rarely made tapes of tapes because of quality degradation.

309. Audio Home Recording Act of 1992, Pub. L. No. 102-563, 106 Stat. 4237 (codified as amended at 17 U.S.C. §§ 1001–1010). Commentators have heavily criticized AHRA for failing to ground its levy regime mandating royalty payments on digital recorders and blank tapes in an appropriate economic analysis and for various economic distortions and inefficiencies. *See* Kurlantzick & Pennino, *supra* note 299, at 515–44. In 1994, only $450,000 was collected for distribution to copyright owners. *Id.* at 497 n.3. *But see* Xiyin Tang, *The Class Action as Licensing and Reform Device*, 122 COLUM. L. REV. 1627, 1647–49 (2022) (discussing the *Cahn* lawsuit and the settlement as copyright reform).

310. *See* S. Rep. No. 102–294, at 122 (1992).

311. *See* 17 U.S.C. § 1002(a).

312. *See* 17 U.S.C. §§ 1004(a)(1), 1004(b); *see also* 17 U.S.C. §§ 1003–07 (setting forth rules and procedures for making royalty payments and for distributing the proceeds).

313. 17 U.S.C. § 1008.

314. *See* 4 PATRY, *supra* note 11, § 11:47.

services, the primary purpose or effect of which was to circumvent SCMS chips.[315]

Recording Industry Association of America v. Diamond Multimedia System, Inc.[316] was the first lawsuit to test the scope of AHRA. Diamond sold cassette-sized portable MP3 Rio players that allowed consumers to load copies of digital music from their hard drives onto Rio players, from which they could listen to music through headphones.[317] RIAA claimed that Diamond violated the AHRA because it had not installed SCMS chips in its Rio players and was not paying royalties into funds for copyright owners.[318]

The court decided that the Rio player was not a "digital audio recording device" subject to AHRA's strictures because Rio players could only make copies of digital audio files from computer hard drives, and "computers (and their hard drives) are not digital audio recording devices because their 'primary purpose' is not to make digital audio copied recordings."[319] The court observed that "the exclusion of computers from the Act's scope was part of a carefully negotiated compromise between the various industries with interests at stake, and without which, the computer industry would have vigorously opposed passage of the Act."[320]

Technology had, to RIAA's chagrin, bypassed the DAT regulatory regime established in 1992.[321] Although Congress has not repealed the AHRA, that law failed to live up to the promise of providing substantial revenues to music and sound recording copyright owners, for the most recent payment into the DAT fund was a whopping $56.43 in 2020.[322]

315. *See* 17 U.S.C. § 1002(c). This provision was the precursor for a more general set of anti-circumvention rules now codified in 17 U.S.C. § 1201. Circumvention of the SCMS chips would enable DAT machines to make multiple perfect copies and thereby defeat AHRA's main purpose.
316. 180 F.3d 1072 (9th Cir. 1999).
317. *Id.* at 1073.
318. *Id.* at 1075.
319. *Id.* at 1078. Nor could a Rio make copies of music from transmissions; it could only make copies from computer hard drives, which Congress had excluded from the scope of the AHRA. *Id.* at 1081.
320. *Id.* at 1078 n.6.
321. *See* All. of Artists & Recording Cos. v. DENSO Int'l Am., Inc., 947 F.3d 849 (D.C. Cir. 2020) (affirming summary judgment against AARC, a collecting society for AHRA funds, which challenged car companies for installing in-car CD-copying systems and holding that these systems were not subject to AHRA requirements because the copying was done on computer hard drives).
322. *See* U.S. COPYRIGHT OFF., LICENSING SECTION REPORT OF RECEIPTS 09/30/2023 UNAUDITED (2023), https://www.copyright.gov/licensing/receipts.pdf [https://perma.cc/5AUZ-GKAJ].

5. Peer-to-Peer Filesharing

Peer-to-peer (P2P) software-based technologies have been available on the internet since its inception.[323] P2P architecture enables efficient sharing of computing resources and data without the need for centralized controls. P2P did not become a disruptive technology for copyright law until the 1999 launch of Napster, an online platform through which users could share MP3 files of recorded music.

To become a Napster user, a person would open an account by registering as a user of the platform, agreeing to its terms of service, and then downloading a client-side app that would enable that user to access Napster's centralized directory of MP3 files available on other Napster users' computers, which would allow the user to search for songs in which the user was interested.[324] What made Napster a P2P system was that digital files were shared by direct transfers from one user's computer to another's without the digital copies passing through Napster's servers; that is, User A could obtain copies of desired MP3 files by downloading them directly from User B's computer.[325] At its peak, Napster had 80 million registered users[326] who shared their music with other Napster users and enjoyed getting other users' MP3 files in turn.

A&M Records was the lead plaintiff in the record industry's lawsuit against Napster for contributory and vicarious infringement.[327] Napster's main defense was that its software had and was capable of substantial noninfringing uses and hence qualified for the *Sony* safe harbor.[328] Napster claimed three uses of its service were noninfringing: sampling music to decide whether to purchase it, space-shifting copies of music users already owned, and permissive uses by artists.[329]

The Ninth Circuit affirmed an indirect infringement ruling against Napster. It (incorrectly) regarded the purpose of users' sharing as commercial because they were getting something for free that they would

323. *See, e.g.*, Genan Zilkha, Note, *The RIAA's Troubling Solution to File-Sharing*, 20 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 667, 669–70 (2010) ("The Internet began as a series of peer-to-peer, or 'P2P' networks").
324. *See* A&M Recs., Inc. v. Napster, Inc., 239 F.3d 1004, 1011–12 (9th Cir. 2001) (describing Napster's system).
325. *See, e.g.*, Michael A. Carrier, *Copyright and Innovation: The Untold Story*, 2012 WIS. L. REV. 891, 901–02 (describing the Napster service).
326. *Id.*
327. Napster, 239 F.3d at 1011.
328. *Id.* at 1025.
329. *Id.* at 1014.

ordinarily have had to purchase.[330] The nature and amount factors disfavored fair use because the recorded music was creative and entire works were copied.[331] Citing three commissioned studies of the economic impacts of Napster users' file-sharing, the Ninth Circuit agreed that Napster was causing harm to the record industry: First, because of evidence of reduced sales of audio CDs, and second, because Napster was interfering with the record companies' efforts to build their own online music services.[332] The court concluded that Napster had caused actual damage to record markets and there was a meaningful likelihood of harm to authorized online marketplaces.[333]

After Napster's demise, several new ad-supported P2P file-sharing systems launched, which unlike Napster relied on decentralized directories of available songs; they actively sought to attract former Napster users to their platforms.[334] Among them were Grokster and Streamcast, whose success drew not only many former Napster users, but also a new round of copyright lawsuits. When sued for contributory and vicarious infringement, these services raised *Sony* safe harbor defenses, claiming their technologies were capable of substantial noninfringing uses. Owing to their more decentralized architectures as compared with Napster, the safe harbor defenses prevailed at the trial court and Ninth Circuit.[335] However, the Supreme Court in *Metro-Goldwyn-Mayer Studios, Inc. v. Grokster, Inc*. ruled that Grokster and Streamcast were liable for actively inducing copyright

330. *Id.* at 1016 (citing Harper & Row Publishers, Inc. v. Nation Enters., 471 U.S. 539, 566 (1985)). Under this rationale, virtually all copying could be considered commercial. The users' downloads of in-copyright songs were, however, nontransformative, which tends to weigh against fair use.

331. *Id.* at 1016.

332. *See id.* at 1016–17.

333. On remand, the trial court issued a preliminary injunction requiring Napster to install filtering technology to block infringing uses of its system, which the Ninth Circuit affirmed. A&M Recs., Inc. v. Napster, Inc., 284 F.3d 1091 (9th Cir. 2002); *see also In re* Aimster Copyright Litig., 334 F.3d 643 (7th Cir. 2003) (holding a P2P filesharing service liable as an indirect infringer).

334. *See* Tim Wu, *When Code Isn't Law*, 89 VA. L. REV. 679, 730–34, 749 (2003) (describing various companies and technologies that attempted to improve on the Napster model and companies' reliance on the advertising model for revenue); *see also* Metro-Goldwyn-Mayer Studios, Inc. v. Grokster, Ltd., 545 U.S. 913, 939–40 (recognizing StreamCast and Grokster's advertisement-based revenue model).

335. Metro-Goldwyn-Mayer Studios, Inc. v. Grokster, Ltd., 380 F.3d 1154, 1157 (9th Cir. 2004), *rev'd*, 545 U.S. 913 (2005). The Ninth Circuit was persuaded by evidence of substantial noninfringing uses and the decentralized architecture of their systems, which meant the services could not have actual knowledge of specific infringements or supervise or control their users.

infringements,[336] opining that there was ample evidence in the record of their efforts to support this finding, such as distributing an electronic newsletter containing links to articles promoting its software's ability to access popular copyrighted music.[337]

The most significant question raised in *Grokster* was whether the Court would interpret the *Sony* safe harbor broadly, narrowly, or abandon it altogether. If the Court did not disavow the Ninth Circuit's assessment of Metro–Goldwyn–Mayer (MGM)'s contributory infringement claim—despite evidence that roughly 90 percent of the traffic on these services was of infringing content and users were transferring literally billions of files[338]—the broad interpretation of the safe harbor would remain intact. The Solicitor General recommended a narrow interpretation of *Sony*, under which the safe harbor would be available to technology developers only if less than half of the systems' uses were infringing.[339] MGM and its fellow plaintiffs were hoping the Court would decide that technology developers could be sued based on such theories as having intentionally designed their systems to enable user infringements or having built their businesses on infringing contents.[340] In the end, the Court reaffirmed its broad interpretation of the *Sony* safe harbor,[341] in keeping with recommendations of many technology developer amici.[342] To MGM's disappointment, neither the high volume of infringements Grokster enabled nor its profits from infringement persuaded the Court to abandon the *Sony* safe harbor. Yet, it articulated another legal

336. *See Grokster*, 545 U.S. at 913, 914.
337. *See id.* at 913, 919, 923–24 (describing the defendants' activities such as responding directly to user email inquiries and promoting their services to former Napster users).
338. *Id.* at 922 (estimated by a study of traffic on one of the defendants' networks).
339. *See* Brief for the United States as Amicus Curiae Supporting Petitioners at 17, Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd., 545 U.S. 913 (2005) (No. 04-480).
340. *See* Pamela Samuelson, *Three Reactions to* MGM v. Grokster, 13 MICH. TELECOMMS. & TECH. L. REV. 177, 186 (2006) (describing several alternative liability standards for which MGM and its amici argued).
341. Justice Ginsburg, however, disagreed with the Ninth Circuit's analysis of the contributory infringement claim and would have sent it back for trial. *Grokster*, 545 U.S. at 942–49 (Ginsburg, J., concurring).
342. Intel and several technology industry associations supported affirmance of the *Sony* safe harbor. *See id.* at 929 (majority opinion) (citing Grokster-side amici concerns about the negative impacts of narrowing *Sony* and imposing secondary liability on technology developers).

theory—namely, active inducement of user infringements, through which MGM could obtain compensatory remedies and injunctive relief.[343]

6. An Interlude on Audio Taping and P2P File-Sharing

Until enactment of the AHRA in 1992, home audio taping had become de facto lawful. Much of this taping, as OTA reported, was done to make consumers' music more portable. After AHRA became law, home audio taping was now statutorily exempt from liability and devices that enabled home taping were also lawful unless they fell within the narrow definitions set forth in the Act. The Ninth Circuit in *Diamond* regarded Rio players as "entirely consistent with the [AHRA's] main purpose—the facilitation of personal use," such as music space-shifting that Rio players enabled.[344]

Although Napster users may have been sharing music files in the privacy of their homes, this sharing of MP3 files of music from users across the global internet was a far cry from the personal use copies in *Sony*. The Ninth Circuit found that the AHRA exemption for home copying did not apply to Napsters' users' file-sharing activities.[345]

One significant difference between home copying and file-sharing was that the former was undetectable, whereas the latter could be detected by agents for the recording industry. Not only did the industry sue Napster as an indirect infringer, but the RIAA also took legal action against roughly 35,000 individual file-sharers, although it finally gave up this tactic and began urging online service providers to do more to stop file-sharing.[346]

The staggeringly large volume of file-sharing of recorded music signaled the existence of a substantial demand for legitimate online services that could offer all-you-can-eat repertoires of popular music for relatively modest monthly charges. Scholars made numerous proposals for compulsory

343. Another disruptive technology blocked by copyright litigation was a service designed to allow consumers to resell used sound recordings. *See* Capitol Records, LLC v. ReDigi, Inc., 910 F.3d 649 (2d Cir. 2018) (holding that ReDigi made unauthorized copies of sound recordings in the course of transferring files from users to its servers, which did not qualify as fair uses).

344. Recording Indus. Ass'n Am. v. Diamond Multimedia Sys., 180 F.3d 1072, 1079 (9th Cir. 1999).

345. *See* A&M Recs. v. Napster, Inc., 239 F.3d 1004, 1024 (9th Cir. 2001).

346. Andre Paine, *RIAA Stops Suing Individuals Over File-Sharing*, HOLLYWOOD REP. (Dec. 19, 2008, 10:06 AM), https://www.hollywoodreporter.com/business/business-news/riaa-stops-suing-individuals-file-124702 [https://perma.cc/N3GJ-3NE3]. *See generally* Annemarie Bridy, *Why Pirates (Still) Won't Behave: Regulating P2P in the Decade After Napster*, 40 RUTGERS L.J. 565 (2009).

licensing of P2P file-sharing,[347] but none got traction with Congress or the record industry. Apple was the first company to persuade the major labels to license their recordings to enable a download service.[348] Over time, numerous music services have enabled music lovers to enjoy their favorite artists' works, first through downloading, and now through online streaming, both of which have greatly benefited the music industry.[349] While file-sharing of music has not disappeared, the shutdown of services such as Grokster has at least lessened the leakage, and other services have provided users with alternative and more lawful ways to get access to music without incurring RIAA's wrath.

C. Computational Search and Research Tools

Significant advancements in the development of computer software and the availability of vast quantities of information on the internet has made it possible to develop information-locating and -processing tools that were unimaginable in the pre-digital analog world.[350] Because building such tools may require making copies of in-copyright works, several copyright infringement lawsuits have challenged these disruptive technologies, whose defendants routinely relied on fair use defenses.

This section discusses seven cases in which courts have been called upon to decide whether technology developers' copying was infringement or

347. *See, e.g.*, Neil Weinstock Netanel, *Impose a Noncommercial Use Levy To Allow Free Peer-to-Peer File Sharing*, 17 HARV. J.L. & TECH. 1 (2003).

348. *See* Brad Hill, *The iTunes Influence, Part One: How Apple Changed the Face of the Music Marketplace*, ENGADGET (Apr. 29, 2013), https://www.engadget.com/2013-04-29-the-itunes-influence-part-one.html [https://perma.cc/F74W-J7T4].

349. MICHAEL MASNICK & LEIGH BEADON, COMPUT. & COMMC'NS INDUS. ASS'N, THE SKY IS RISING: A DETAILED LOOK AT THE STATE OF THE ENTERTAINMENT INDUSTRY 5–12 (Apr. 8, 2019), https://skyisrising.com/report.html [https://perma.cc/JPX8-FBRV].

350. Only rarely were copyright lawsuits brought to challenge the unauthorized development of analog indexes. One such case was *New York Times Co. v. Roxbury Data Interface, Inc.*, 434 F. Supp. 217 (D.N.J. 1977), in which the Times sued Roxbury for copyright infringement because it published a personal name index covering entries from many years copied from the annual indexes that the Times published of content appearing in each year's index. *Id.* at 218–19. Roxbury successfully raised a fair use defense to the Times' claim. *See id.* at 226–27. Although Roxbury had a commercial purpose in using entries from the Times' indexes, the court concluded that its publication had the social benefit of helping people find information of use to them. *See id.* at 221. The court was persuaded that Roxbury's publication would not harm the market for the Times' indexes because the Roxbury index was only useful to people who had access to the Times' indexes and did not supplant demand for the original. *Id.* at 226–27; *see also* Kipling v. G.P. Putnam's Sons, 120 F. 631, 635 (2d Cir. 1903) (holding index did not infringe plaintiff's copyright).

fair use. Subsection C.1 reviews two cases challenging the unauthorized storage of digital copies of the plaintiffs' works. Subsection C.2 considers two other cases focused on unauthorized copying of the plaintiffs' works to enable online users to locate relevant images on the internet. Subsection C.3 examines another three lawsuits that claimed infringement based on the creation of digital databases of in-copyright materials to enable computational analyses. Copyright plaintiffs claimed these uses caused market harm because the defendants failed to obtain licenses to authorize the challenged uses. Yet courts in all but one case characterized as speculative claims of actual harm or potential harm to copyright owner markets when there were reasons to doubt that viable markets could be created given large transaction costs in relation to the relatively low value of the potentially licensed uses.[351] In only one case was a court convinced that the defendant could and should have licensed the challenged uses.[352]

Although *Campbell* did not involve digital copying for information-locating or -processing purposes,[353] the Court's analysis of fair use significantly affected the fair use analyses in these computational use cases. Without *Campbell*'s refinements of fair use analysis, the information-locating and -processing purpose cases might have been decided differently. Six of the seven defendants were commercial entities, and their uses of preexisting works were arguably for commercial purposes; the takings were generally of the entire works; and if commercial uses were presumed unfair and lack of harm was irrelevant—as the Sixth Circuit in *Campbell* had opined—the fair use defenses in these cases might have failed.

1. Search Engine Webcrawling and Storage of Copyrighted Works

Field v. Google, Inc.[354] challenged Google's unauthorized reproduction and storage of the plaintiff's online content. Field posted fifty-one textual works on his personal website.[355] He then sued Google because its webcrawling software made copies of his works and stored copies of them in caches so that users could access this content in response to queries more rapidly than if the users had to follow links to the content on his local

351. See the discussions on *Field v.* Google, iParadigms, Kelly v. Arriba, Perfect 10, and the two Google Books cases discussed *infra* Part II.C.
352. See *TVEyes*, discussed *infra* Part II.C.3.
353. *See* Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569 (1994).
354. 412 F. Supp. 2d 1106, 1109 (D. Nev. 2006).
355. *Id.* at 1110. Field created the fifty-one works over three days. *Id.* at 1114.

website.[356] Field demanded $50,000 of statutory damages per infringed work and an injunction.[357]

The court granted Google's motion for summary judgment, finding that Google's copying and storage of Field's content was fair use. Google's purpose in making these copies was, it found, transformative because it made and stored the copies for a different purpose than Field's.[358] Field had voluntarily posted his works on an open website and knew that Google's webcrawling software would make copies of the site's contents unless he took active steps to prevent such copying.[359] The court likened cache copies to time-shift copies in *Sony*.[360] There was no evidence that Google's copying had harmed Field's markets for or the value of his works; moreover, no market existed or was likely to develop for licensing the kind of cache copying for which Field sued Google.[361]

Equally unavailing were several students' complaints against iParadigms, the maker of plagiarism detection software. They claimed that iParadigms' archival storage of their papers and consequent use of the contents of their papers for its computational analysis tool for detecting plagiarism constituted infringement.[362] As in *Field*, the court found iParadigms' purpose to be transformative because of its very different purpose than the students' in writing the papers.[363] The students' lawyer emphasized the unpublished nature of the papers, but the court gave this little weight because iParadigms was not trying to exploit the plaintiffs' creative expression, nor trying to supplant their right of first publication.[364] Although student papers were copied in their entirety, the court regarded their claims of market harm as speculative.[365]

356. *Id.* at 1110–12.
357. *Id.* at 1110.
358. *See id.* at 1118–19.
359. *Id.* at 1116, 1120. Field knew that if he wanted to stop Google from copying the contents of his website, he could tag the site with a code (known as the robots.txt metatag) that would signal that he did not want the site to be crawled. An alternative rationale for the ruling in Google's favor was that by posting his works online and not using the robots.txt tag, Field had impliedly licensed Google to make this use. *Id.* at 1109, 1114.
360. *See id.* at 1120–21.
361. *Id.* at 1121–22.
362. A.V. *ex rel.* Vanderhye v. iParadigms, LLC, 562 F.3d 630, 633–34 (4th Cir. 2009). The students' schools had required them to submit their papers to iParadigms for processing and they agreed to iParadigms' terms of service which granted it the right to archive submitted papers. *Id.* at 634–35.
363. *See id.* at 638–40.
364. *See id.* at 640–42.
365. *See id.* at 642–45.

2. Thumbnail Images and In-Line Linking for Web Search

Kelly v. Arriba Soft Corp.[366] was the first appellate court decision to address the legality of search engine creation and public display of low-resolution, thumbnail-sized images of photographs posted on the internet. Kelly, a professional photographer, sued Arriba for infringement because its webcrawling software copied thirty-five of his photographs of the American West and created and displayed thumbnails of them in response to user search queries.[367] Arriba claimed that making and displaying the thumbnails was fair use.[368]

Because Arriba had a very different purpose than Kelly in making and displaying the copies, the court found Arriba's purpose was transformative.[369] Kelly's photographs were creative, but the court decided Arriba's making and display of the thumbnails was reasonable in light of its transformative purpose of enabling displays of images responsive to user search queries.[370] Moreover, the thumbnails were sufficiently small and low enough in resolution that they were unlikely to supplant demand for the original.[371] The court implied that Arriba's thumbnails might actually help Kelly's market; after all, Kelly was trying to attract customers to his website and Arriba's image search engine enabled internet users to find Kelly's photographs.[372] The court praised Arriba for facilitating public access to knowledge.[373]

Perfect 10, Inc. v. Amazon.com, Inc.[374] was a second search engine thumbnail case. It differed from *Kelly* in three respects. First, *Perfect 10* addressed whether Google's providing of in-line links to full-sized images on

366. 336 F.3d 811 (9th Cir. 2003). Kelly also claimed infringement of his public display right because Arriba provided inline links to Kelly's full-sized images on his site, which Arriba users might think were displayed on Arriba's site. Kelly v. Arriba Soft Corp., 77 F. Supp. 2d 1116, 1118 (C.D. Cal. 1996). The appellate court reversed and remanded summary judgment as to full-size images because the parties had not moved for summary judgment on this claim. *Kelly*, 336 F.3d at 822. The Ninth Circuit addressed this issue in a later case. *See infra* note 374 and accompanying text (discussing the *Perfect 10* case).
367. *Kelly*, 336 F.3d at 815–16.
368. The court did not address the claim of infringement based on Arriba's initial webcrawling copies, perhaps because Arriba deleted them after making thumbnails of the images. *See id.* at 815–16.
369. *See id.* at 817–20.
370. *See id.* at 820–21.
371. *Id.* at 821–22.
372. *See id.*
373. *See id.* at 818–19.
374. 508 F.3d 1146 (9th Cir. 2007).

internet sites violated Perfect 10's public display right.[375] Second, the copyright owner (here, Perfect 10) had not voluntarily posted its photographs on the internet, and Google provided links to images that infringed Perfect 10's copyrights.[376] Third, Perfect 10 tried to distinguish *Kelly* by asserting that Google was interfering with its licensing market for downloading small images on smartphones.[377]

As in *Kelly*, Perfect 10 claimed that Google's thumbnails of Perfect 10's photographs were infringing, to which Google raised a fair use defense. Although characterizing Google's thumbnails as more commercial than Arriba's (owing to Google's AdSense program), the court decided that Google's overall purpose in serving up thumbnails was transformative and "provided a significant benefit to the public" by incorporating the original work into an online reference tool.[378] The court also determined that any potential harm to Perfect 10's market for cell phone downloads was merely hypothetical, and concluded that Google's use of thumbnails was fair.[379]

3. Mass Digitization to Enable Search Indexes and Analytic Research Tools

Authors Guild, Inc. v. HathiTrust[380] was the first of two fair use cases involving Google Book Search (GBS) databases to reach the Second Circuit. The Guild sued HathiTrust for creating in 2008 a digital library containing digital copies of several million in-copyright books that Google provided to its university library partners which Google scanned from the libraries' collections.[381] The Guild regarded the very existence of HathiTrust as an illegal database.[382]

375. This is the issue the Ninth Circuit did not address in *Kelly*. In *Perfect 10*, the court rejected the plaintiff's in-line-linking public display claims because Google did not store copies of any full-sized images on its servers and instead supplied HTML instructions to users' computers. *Id.* at 1155–56, 1159–61. For the court, what the technology actually did was pertinent, not how it might look to ordinary internet users. *Id.* at 1161.
376. *Id.* at 1157. The court did not give any weight to the infringing character of the photographs, perhaps because Google's webcrawling program cannot discern the legal status of images it finds on the open internet.
377. *Id.* at 1164.
378. *Id.* at 1164–68 (following Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 579 (1994)).
379. *See id.* at 1168 (noting that the trial court had not found actual downloading of Perfect 10 images for cell phone use).
380. 755 F.3d 87 (2d Cir. 2014).
381. *Id.* at 90–91.
382. *See id.* at 97.

HathiTrust defended its digital library as fair use, asserting that three types of its uses were fair: First, full-text searching to identify books containing entries that satisfied researcher's queries to enable them to locate relevant books in physical stacks; second, making full texts of the books accessible to print-disabled persons; and third, preserving books in library collections so they might be replaced if badly damaged or destroyed, for example, by natural disasters.[383]

The Second Circuit characterized HathiTrust's creation of a full-text searchable database as "a quintessentially transformative use."[384] To create such a database, the court thought it was necessary to copy entire books.[385] Because HathiTrust's search function did not provide patrons access to the books' contents, only to information about their contents, use of the database would not supplant demand for the books or otherwise harm the books' markets.[386] The court dismissed as speculative the Guild's arguments that HathiTrust should have licensed this use and that risks of security breaches might cause future harm, as there was unrebutted evidence that HathiTrust had adopted strong security measures.[387]

As for print-disabled access, the court found this use of the HathiTrust corpus to be fair.[388] "By making copyrighted works available in formats accessible to [people who are] disabled, [HathiTrust] enables a larger audience to read those works, but the underlying purpose of [HathiTrust's] use is the same as the author's original purpose."[389] Despite its nontransformative purpose, that purpose factor nonetheless favored HathiTrust's fair use defense because it was socially beneficial to enable print-disabled persons to have the same access to book contents as sighted persons.[390] The court thought that the amount was not excessive in light of its beneficial purpose.[391] The court credited evidence that publishers had made only a minute percentage of books available to individuals who are print-disabled, so the market effects factor also favored fair use.[392]

383. *Id.* at 91–92.
384. *Id.* at 97.
385. *See id.* at 98–99.
386. *Id.* at 99–100.
387. *See id.* at 100–01.
388. *See id.* at 101.
389. *Id.* HathiTrust required documentation of print-disabilities before providing access. *Id.*
390. *Id.* at 101–02.
391. *See id.* at 102–03.
392. *See id.* at 103.

The Second Circuit vacated the lower court's ruling on HathiTrust's preservation defense, indicating that the record was insufficient to resolve this issue on summary judgment.[393] But it also directed the lower court to consider whether the Guild had standing to litigate the preservation issue.[394] The Guild owned only 116 copyrights to books in the HathiTrust corpus, so it could not seek relief for potential infringements as to books whose copyrights it did not own.[395]

The most closely watched of the *GBS* fair use cases was *Authors Guild v. Google, Inc*.[396] The Guild commenced this lawsuit in 2005, three years before creation of the HathiTrust library.[397] Joining the Guild as plaintiffs were three named authors who claimed to represent the interests of all authors whose books Google had copied for GBS.[398] The complaint challenged Google's scanning of the contents of an estimated twenty million books from university library collections and serving up snippets from the books in response to search queries.[399] Google defended this mass digitization project, claiming that it had made fair use of the books by indexing their contents so it could serve up a few snippets of relevant texts in response to user search queries.[400]

Authors Guild appealed the fair use ruling in Google's favor, arguing that Google's scans of entire books and displays of expressive snippets from the books were not transformative under *Campbell*; that Google's commercial motivation and the revenues it derived from the books' contributions to its dominance in the search market undermined its fair use defense; that class members were entitled to license fees for derivative uses that Google made of the books, which constituted harm to this potential market; that there was

393. *See id.* at 93–94, 105.
394. *Id.* at 103–04. The case settled after it was remanded to the trial court. Krista Cox, *Authors Guild v. HathiTrust Litigation Ends in Victory for Fair Use*, ASS'N. OF RSCH. LIBRS. (Jan. 8, 2015), https://www.arl.org/news/authors-guild-v-hathitrust-litigation-ends-in-victory-for-fair-use [https://perma.cc/K7JS-2Y6G].
395. Authors Guild, Inc. v. HathiTrust, 902 F. Supp. 2d 445, 458 (S.D.N.Y. 2012) (acknowledging that plaintiffs had asserted copyright ownership as to 116 works). *See* 17 U.S.C. § 501(b) (stating that only copyright owners can sue for infringement).
396. 804 F.3d 202 (2d Cir. 2015). The opinion briefly recounted the long history of the case, which included a proposed class settlement that the district court rejected and a vacated class certification before the court's summary judgment ruling. *Id.* at 211–12.
397. *Id.* at 211.
398. *See id.* at 208. The case caption lists the three plaintiffs "individually and on behalf of all others similarly situated." *Id.* at 202.
399. *See* Authors Guild, Inc. v. Google Inc., 954 F. Supp. 2d 282, 285–89 (recounting the facts and procedural history of the case).
400. *See id.* at 289.

a risk of hacking into the storage of digital copies; and that providing Google's library partners with copies of the scanned books was unfair.[401]

The Second Circuit rejected all five arguments.[402] It concluded that Google's "making of a digital copy to provide a search function is a transformative use, which augments public knowledge by making available information *about* Plaintiffs' books without providing the public with a substantial substitute for matter protected by the Plaintiffs' copyright interests in the original works or derivatives of them," and the same was true "of Google's provision of the snippet function."[403] Nor did Google's commercial motivation for scanning the books undermine its fair use defense.[404] The court was unpersuaded that "Google has usurped their opportunity to access paid and unpaid licensing markets for substantially the same functions that Google provides," saying that the argument failed "in part because the licensing markets in fact involve very different functions than those that Google provides, and in part because an author's derivative rights do not include an exclusive right to supply information (of the sort provided by Google) about [their] works."[405] Nor was it convinced by the plaintiffs' arguments about possible future infringement because of alleged security risks or the prospect that Google's library partners might infringe the digital copies Google provided to them.[406]

The court offered a detailed response to the plaintiffs' claim that snippets had substitutive effects and thereby harmed book markets. While acknowledging that snippets do display some expressive parts of the books in the corpus, it observed that "[s]nippet view, at best and after a large commitment of manpower, produces discontinuous, tiny fragments, amounting in the aggregate to no more than 16% of a book."[407] Because of this, snippet view "does not threaten the rights holders with any significant harm to the value of their copyrights or diminish their harvest of copyright

401. *Id.* at 207.
402. *See id.* at 207–08.
403. *Id.* at 207. The court took note of numerous ways that Google limited snippets, such as not providing snippets from some books (for example, dictionaries) for which snippets might have substitutive effects and limiting the number of snippets per book. *See id.* at 210, 222.
404. *Id.* at 207. Google did not serve up ads next to snippets; instead, it provided links to sites where the books could be purchased or lent from libraries. *Id.* at 209.
405. *Id.* at 207.
406. *Id.* at 207–08.
407. *Id.* at 224.

revenue."[408] Even if Google's supply of snippets does cause some lost sales of books, the court stated:

> [T]he possibility, or even the probability or certainty, of some loss of sales does not suffice to make the copy an effectively competing substitute that would tilt the weighty fourth factor in favor of the rights holder in the original. There must be a *meaningful or significant effect* "upon the potential market for or value of the copyrighted work."[409]

The Second Circuit was also impressed that GBS enabled researchers to find those books among the twenty million that were most relevant to their search queries, identifying information about which Google "instantaneously supplied [that] would otherwise not be obtainable in lifetimes of searching."[410] GBS also made possible "new forms of research, known as 'text mining' and 'data mining,'" including Google's "ngram" research tool that "draws on the [GBS] corpus to furnish statistical information to internet users about the frequency of word and phrase usage over centuries."[411] The court thought these new capabilities supported Google's fair use defense.

Fox News Network LLC v. TVEyes, Inc.[412] challenged not the mass digitization of television news programs, but rather several features that allowed TVEyes subscribers to make and distribute copies of clips of these programs.[413] TVEyes was a commercial service that constantly recorded programs on more than 1400 television and radio stations, copies of which it stored in a database, and allowed its subscribers to view, archive, download, and email up to ten minutes of clips of programs in response to their search queries.[414] This helped government officials and companies learn about what was being said and by which commentators on any particular topic of interest to them.[415] Because of the *Google Books* decisions, Fox chose not to challenge the creation of the database of news programs or TVEyes' search function, but it argued that some features of

408. *Id.*
409. *Id.* (emphasis added) (quoting 17 U.S.C. § 107(4)).
410. *Id.* at 209.
411. *Id.*
412. 883 F.3d 169 (2d Cir. 2018).
413. *See also* Associated Press v. Meltwater U.S. Holdings, Inc., 931 F. Supp. 2d 537, 560–61 (S.D.N.Y. 2013) (finding a proprietary news media monitoring service's excerpting of news articles unfair in part because the plaintiff had established a licensing market for such uses).
414. *TVEyes*, 883 F.3d at 175.
415. *Id.* at 174–75.

TVEyes' system that permitted viewing and distribution of Fox programs went beyond what fair use allows.[416]

The Second Circuit agreed. The court characterized TVEyes' search functionality as transformative because "it enables users to isolate, from an ocean of programming, material that is responsive to their interests and needs, and to access that material with targeted precision."[417] But the features that allowed clients to watch, archive, download, and email clips had "only a modest transformative character because, notwithstanding the transformative manner in which it delivers content, it essentially republishes that content unaltered from its original form, with no 'new expression, meaning or message.'"[418] The amount and market effects factors favored Fox because TVEyes enabled clients to watch virtually all of Fox's programs and it sold access to those programs, thereby "depriv[ing] Fox of revenues to which Fox is entitled as the copyright holder."[419]

This Part's exploration of three types of disruptive technologies in fair use cases—those enabling the copying of research and educational materials, home video and audio taping, and search and research tools—has revealed some insights about the way courts have evolved, pre- and post-*Sony*, in their approach to the perceived harms that these technologies allegedly caused. While courts before *Sony* often assumed market harm due to unauthorized copying, courts after *Sony* examined more closely the market effects of a secondary use, seeking evidence of actual or likely harm or scrutinizing the significance of licensing claims, for example. The next Part will consider how these insights might apply in the context of in-copyright works and generative AI.

III. GENERATIVE AI AS A COPYRIGHT-DISRUPTIVE TECHNOLOGY

That generative AI is a disruptive technology for copyright law is evident from the spate of copyright lawsuits filed since late 2022. These technologies are very similar to the computational search and research tool fair use cases because the generative AI developers' purposes in making copies of in-copyright works are generally quite different from the purposes

416. *Id.* at 176.
417. *Id.* at 177.
418. *Id.* at 178 (citing Authors Guild v. HathiTrust, 755 F.3d 87, 96 (2d Cir. 2014) (quoting Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 579 (1994))).
419. *Id.* at 180. TVEyes had once approached Fox about a license, but Fox refused because it wanted TVEyes to stop recording its programs. *Id.* at 175. Nonetheless, the court characterized Fox as having a "plausibly exploitable market." *Id.* at 180.

authors had for their works. Moreover, generative AI developers aim to make nonexpressive uses of the works' contents, as in the *iParadigms* and *HathiTrust* cases. The main difference is that generative AI developers make copies of works to train computational models to produce new content, not just to display parts of existing contents, as in the search engine and *Google Books* cases. Section A considers the merits of the copyright claims in the generative AI lawsuits and defenses that developers will likely raise. Section B explains why some critics of generative AI systems favor legislation to establish a collective licensing regime for use of in-copyright works as training data, and why it is doubtful such a regime would be workable.

A. Copyright Lawsuits Challenging Generative AI

Virtually all of the generative AI lawsuits charge their developers with infringement of the plaintiffs' copyrights for reproducing in-copyright works for purposes of training generative AI models.[420] Once these models are trained, generative AI systems are able to produce outputs in response to user prompts, that is, a set of directions to the system to produce outputs having specified characteristics (for example, a proposed itinerary for a three-day trip to Milan). Depending on the type of generative AI model, these outputs may consist of lines of computer code, documents, images, or video.[421] Because these outputs are, as a factual matter, based to some extent on the contents of the data on which the system was trained, and recast or

420. *See, e.g.*, Getty Complaint, *supra* note 3, ¶¶ 36, 62–73 (alleging infringement of 17 U.S.C. § 106(1)). The exception is Complaint, Doe 1 v. GitHub, Inc., No. 22-cv-06823 (N.D. Cal. June 24, 2024) (proposed class action) [hereinafter GitHub Complaint], filed in November 2022, which charged GitHub with violations of 17 U.S.C. § 1202 for unlawful removal of copyright management information (CMI), but not with copyright infringement, although the complaint characterizes GitHub as engaged in "software piracy." *See* GitHub Complaint, *supra* note 420, ¶¶ 8, 138–67. Because of the § 1202 and piracy claims, the *GitHub* case has a copyright dimension. As this Article was preparing to go to press, a trial court granted defendants' motion to dismiss this claim. *See* Doe 1 v. GitHub, Inc., No. 22-cv-06823, slip op. at 15 (N.D. Cal. June 24, 2024). As it had previously dismissed plaintiffs' § 1202 claim on the same ground, the court now dismissed the claim with prejudice. *Id*. at 6. Several other generative AI complaints, including those in *Andersen* and *Getty Images*, include § 1202 claims as well as copyright infringement claims. Because removal or alteration of CMI must be knowing and must facilitate copyright infringement, it is unlikely that the § 1202 claims will prevail, but a full assessment of them is beyond the scope of this Article.

421. *See* Lee et al., *supra* note 1 (describing what the authors call the generative AI supply chain).

transform those contents, the generative AI lawsuits generally claim that outputs of these systems are infringing derivative works.[422]

Subsection A.1 briefly reviews some details about the generative AI copyright lawsuits. Most common are claims that uses of works as training data infringe copyright's reproduction right, but several lawsuits also claim that outputs generated by AI models in response to user prompts are infringing derivative works. Subsection A.2 analyzes the AI output derivative work claims first, because although these charges are serious, the major defenses are unlikely to argue fair use, on which this Article is focused. Subsection A.3 considers the merits of the claims of copying of in-copyright works as training data. Fair use will almost certainly be the main defense to the training data charges. Given how courts have dealt with fair use defenses in other disruptive technology cases, especially the computational use precedents, the generative AI developers would generally seem to have a reasonable chance to prevail with their fair use defenses. Yet, every fair use case must be decided on its facts. At this early stage, it is difficult to predict how courts might distinguish the fair use claims in generative AI cases and the other disruptive technology cases. With many authors and artists railing against generative AI for "stealing" their works,[423] generative AI developers may encounter more resistance in the courts than they may expect.

1. Overview of the Generative AI Copyright Claims

Numerous lawsuits have charged major developers of generative AI technologies with copyright infringement in U.S. courts.[424] The defendants so far are OpenAI (eight cases), Meta (three cases), Stability AI (two cases),

422. *See, e.g.*, Getty Complaint, *supra* note 3, ¶¶ 51, 67 (alleging infringement of 17 U.S.C. § 106(2)).

423. *See, e.g.*, Kyle Chayka, *Is A.I. Art Stealing From Artists?*, NEW YORKER (Feb. 10, 2023), https://www.newyorker.com/culture/infinite-scroll/is-ai-art-stealing-from-artists [https://perma.cc/2ELD-T2LZ].

424. Because the copyright claims in these lawsuits are very similar, this Article focuses on the *Andersen*, *Authors Guild*, *Getty*, *New York Times*, and *Silverman* claims, which are typical of the rest, although they differ in some respects likely to affect the analysis of fair use defenses. As noted above, numerous complaints also allege violations of 17 U.S.C. § 1202. *See supra* note 420. The complaints typically include several other claims such as unfair competition, unjust enrichment, and negligence, which are likely to be preempted under 17 U.S.C. § 301(a) because they rely on facts pleaded in the copyright claims. The main claim in the generative AI lawsuits against OpenAI, Microsoft, and GitHub is that these defendants breached open source licenses when developing a code-generating AI. *See, e.g.*, GitHub Complaint, *supra* note 420, ¶¶ 168–83.

Anthropic (one case), and Alphabet (one case).[425] All but two of these cases are class actions brought on behalf of authors, artists, or software developers whose works were used as training data for building models for generative AI systems.[426] Most are pending in the most influential circuits that have decided several of the disruptive technology fair use cases.[427]

The complaints typically allege, first, that the defendants made copies of the plaintiffs' in-copyright works when ingesting them for purposes of training generative AI models capable of producing outputs—such as texts, images, or software code, in response to user prompts—and second, that generated outputs are infringing derivative works because they are based on the works on which the models were trained.[428] Some of the generative AI

425. Midjourney and Deviant Art are co-defendants with Stability in *Andersen v. Stability AI Ltd.*, No. 3:23-cv-00201, 2023 WL 7132064 (N.D. Cal. Oct. 30, 2023). Microsoft, Bloomberg LP, and EleutherAI Institute are co-defendants in *Huckabee v. Meta Platforms, Inc.*, No. 23-cv-09152 (S.D.N.Y. filed Oct. 17, 2023). Microsoft is also a co-defendant in five other generative AI copyright cases. A much lesser known generative AI developer that has been charged with copyright infringement is Ross Intelligence. Thomson-Reuters, owner of the Westlaw database, alleges Ross made copies of Westlaw headnotes when training its legal analytics model. *See* Thomson Reuters Enter. Ctr. GmbH v. Ross Intel. Inc., No. 20-cv-613, 2023 WL 6210901 (D. Del. Sept. 25, 2023). As this Article is going to press, the number of generative AI cases has expanded to twenty-eight cases. *See* Master List of Lawsuits, *supra* note 3.

426. Andersen and Silverman claim to represent the interests of all authors of U.S. works used as training data for the defendants' models. *See* Andersen Complaint, *supra* note 3, ¶ 41; Silverman Complaint, *supra* note 3, ¶ 43. The *Authors Guild* initial complaint limits its class definition to professional authors whose U.S. works of fiction sold at least five thousand copies that were used to train OpenAI's models. *See* Guild Complaint, *supra* note 3, ¶¶ 310–13. It is far from clear that it will be possible to certify any of the classes, as some authors and artists favor generative AI. *See, e.g.*, *Open Letter: Artists Using Generative AI Demand Seat at Table From US Congress*, CREATIVE COMMONS (Sept. 29, 2023), https://creativecommons.org/about/program-areas/policy-advocacy-copyright-reform/open-letter-artists-using-generative-ai-demand-seat-at-table-from-us-congress [https://perma.cc/ZE7S-UHCM]; *Copyright and Generative AI: Our Views Today*, AUTHORS ALL. (Aug. 30, 2023), https://www.authorsalliance.org/2023/08/30/copyright-and-generative-ai-our-views-today [https://perma.cc/22Y4-CNFJ]. Moreover, publication contracts differ about which rights authors retain, which may also affect class certifications. As of yet, there are no book publisher lawsuits against developers of generative AI.

427. Through the end of 2023, nine cases were pending in the Northern District of California, four in the Southern District of New York, two in Delaware, and one in the Middle District of Tennessee. The Joseph Saveri law firm filed five of the nine California lawsuits, including the *Andersen*, *Silverman*, and *GitHub* cases.

428. *See, e.g.*, Andersen Complaint, *supra* note 3, ¶¶ 153–77. The *Andersen* and *Silverman* complaints assert that the models embodied in Stable Diffusion and OpenAI are derivative works. *See id.* ¶ 18; Silverman Complaint, *supra* note 3, ¶ 57. *See infra* notes 443–448 and accompanying text on dismissal of *Andersen*'s derivative work claims.

complaints raise vicarious infringement claims as well as direct infringement claims.[429]

The complaints also typically claim that the plaintiffs have been harmed by generative AI systems and ask for awards of actual damages, defendants' profits, statutory damages, and injunctive relief.[430] Some also ask the court to order impoundment and destruction of models trained on the plaintiffs' data.[431]

2. Outputs-as-Derivative-Works Claims

Fair use is likely to play a lesser role in the output infringement claims than in the training data infringement claims in the generative AI cases. A more significant defense may be that outputs of the generative AI systems—whose developers are defendants in these cases—may focus on the lack of substantial similarity in expressive elements of particular outputs and particular works used as training data. Insofar as generative AI systems occasionally produce outputs that are substantially similar or close copies of training data, those outputs may be fair uses. Or if they are not, the users whose prompts caused the outputs to be substantially similar may be the direct infringers. Insofar as generative AI developers made significant efforts to adopt technical measures—such as input or output filters—to prevent infringements, the developers may be able to persuade courts not to treat them as indirect infringers of copyrights.

One respect in which generative AI technologies disrupt copyright markets differently than with past disruptive technology cases, such as *Sony*, *Napster*, and *Grokster*, is that generative AI systems generally do not produce exact copies of expression in preexisting works. In fact, outputs of the generative AI systems in litigation are overwhelmingly not substantially similar in expression to the works on which the models were trained,[432]

429. *See, e.g.*, Andersen Complaint, *supra* note 3, ¶¶ 153–77.

430. *See* Andersen Complaint, *supra* note 3, at 42. One complaint asks the court to order destruction of models built on infringing content. *See* Getty Complaint, *supra* note 3, at 35.

431. *See, e.g.*, Getty Complaint, *supra* note 3, at 35; *see also* Pamela Samuelson, *How to Think About Remedies in the Generative AI Copyright Cases*, LAWFARE (Feb. 15, 2024, 1:00 PM), https://www.lawfaremedia.org/article/how-to-think-about-remedies-in-the-generative-ai-copyright-cases [https://perma.cc/39FE-QHG6] (discussing model destruction remedy claims).

432. *See* Nicholas Carlini et al., *Extracting Training Data from Diffusion Models*, ARXIV 4–7 (2023), https://arxiv.org/pdf/2301.13188.pdf [https://perma.cc/RVN4-XD83]. The

although the extent to which models are likely to produce similar outputs depends on the diversity (or lack thereof) and volume of data used to train the models and the extent to which developers have used technical safety measures to prevent infringing outputs.[433] These systems can produce outputs faster and more cheaply than human authors, and the AI outputs are generally of high enough quality to be perceived as competitive in the marketplace with works of human authors. This means that there is a nontrivial potential effect on human-authored work markets, even though the kind of market harms copyright law has typically addressed have focused on the defendant's appropriation of the plaintiff's expression, not the production of different and non-infringing works.

Fair use defenses have succeeded in the past when a defendant's intermediate copying enabled them ultimately to produce noninfringing products that competed with the plaintiff's products. The most pertinent example is the competition enabled by a videogame developer's decompilation copying of another firm's platform software for purposes of reverse engineering it to discern interface specifications so it could create a version of its games capable of running on a popular platform. In *Sega Enterprises Ltd. v. Accolade, Inc.*, the Ninth Circuit held that this intermediate copying of Sega's code for a legitimate purpose of extracting unprotectable elements was fair use.[434] It noted, but did not regard as significant, that Sega had licensed interface information to other videogame developers. Accolade did not like Sega's license terms, so it obtained interface information by reverse engineering.[435] In response to Sega's argument that Accolade's commercial and competitive purpose undercut its fair use defense, the court said:

> Accolade's identification of the functional requirements for Genesis compatibility has led to an increase in the number of independently designed video game programs offered for use with the Genesis console. It is precisely this growth in creative expression, based on the dissemination of other creative works

risk of potentially infringing outputs is greater in respect of visual images. *See* Gary Marcus & Reid Southern, *Generative AI Has a Visual Plagiarism Problem*, IEEE SPECTRUM (Jan. 6, 2024), https://spectrum.ieee.org/midjourney-copyright [https://perma.cc/5S2C-BEHS].

433. Sag, *supra* note 8, at 295–96.

434. *See* Sega Enters. Ltd. v. Accolade, Inc., 977 F.2d 1510, 1522–24 (9th Cir. 1992); *see also* Sony Comput. Ent., Inc. v. Connectix Corp., 203 F.3d 596 (9th Cir. 2000) (discussing reverse engineering to develop competing platform held fair use).

435. *Sega*, 977 F.2d at 1514–15.

> and the unprotected ideas contained in those works, that the Copyright Act was intended to promote.[436]

As long as Accolade's videogames did not infringe any Sega copyrights, Accolade was free to compete in the marketplace with Sega's and Sega licensees' products, even though that competition was enabled, in part, by the intermediate copying Accolade did of Sega's software.[437]

The Authors Guild complains that OpenAI's ChatGPT has already caused some professional authors to lose income because generated outputs are displacing opportunities for their services and otherwise undermining their ability to earn a living from their work.[438] Yet the Guild's complaint emphasizes that these plaintiffs are concerned about threats to their livelihood because of competition from generative AI systems, rather than pointing to specific examples of outputs they believe are infringing derivative works.[439] In the absence of evidence of substantial similarity in expression between these inputs and those outputs,[440] the derivative work claims in the generative AI cases are novel, and given many precedents interpreting the derivative work right,[441] they seem quite weak. Two courts have rejected generative AI output-derivative-work claims.[442] The initial Andersen complaint, for example, conceded that "[i]n general, none of the Stable Diffusion output images provided in response to a particular Text Prompt is

436. *Id.* at 1523.
437. *Id.*; *see also* Brief of Amicus Curiae American Committee for Interoperable Systems at 12–14, Sega Enter. Ltd. v. Accolade, Inc., 977 F.2d 1510 (9th Cir. 1992) (No. 92-15655) (criticizing an apparent "fruit of the poisonous tree" rationale for a district court's preliminary injunction against Accolade's noninfringing products, citing congressional reports tying copyright remedies only to infringing products). For more on this doctrine, see *infra* note 459 and accompanying text.
438. *See* Guild Complaint, *supra* note 3, ¶¶ 113–22.
439. *See id.* ¶¶ 117–22.
440. For a discussion of the origins and proper scope of copyright's derivative work right, see Pamela Samuelson, *The Quest for a Sound Conception of Copyright's Derivative Work Right*, 101 GEO. L.J. 1505 (2013).
441. *See, e.g.*, Litchfield v. Spielberg, 736 F.2d 1352, 1357 (9th Cir. 1984); *see also* 2 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 8.09 (Matthew Bender rev. ed. 2023) ("Unless enough of the pre-existing work is contained in the later work to constitute the latter an infringement of the former, the latter, by definition, is not a derivative work."); 1 *id.* § 3.01 ("A work is not derivative unless it has substantially copied from a prior work.").
442. *See* Kadrey v. Meta Platforms, Inc., No. 23-cv-03417, 2023 WL 8039640, at *1 (N.D. Cal. Nov. 20, 2023); Andersen v. Stability AI Ltd., No. 3:23-cv-00201, 2023 WL 7132064, at *7–8 (N.D. Cal. Oct. 30, 2023) ("[T]he alleged infringer's derivative work must still bear some similarity to the original work or contain the protected elements of the original work.").

likely to be a close match for any specific image in the training data."[443] Insofar as this is true, derivative work claims cannot be sustained. Andersen may argue that AI outputs are "based upon one or more pre-existing works" and "recast, transform[], or adapt" training data works and fall within the statutory definition of "derivative work."[444] While it is true that outputs are, in a lay sense, "based upon" the works on which generative AI models were trained, this has never sufficed to support derivative work claims; courts have required a finding of substantial similarity in expression to infringe that right.[445] Nor does it suffice to claim that the AI outputs have recast, transformed, or adapted training works because that clause does not do away with the requirement that the defendants' works must be substantially similar to expressive aspects of the plaintiffs' works to infringe the derivative work right.[446] Andersen's complaint fails to show any examples of Stable Diffusion outputs that are substantially similar to expressions in Andersen's works. Her theory has been that *all* outputs are infringing derivatives.[447] This is why a trial judge dismissed Andersen's derivative work claims for lack of substantial similarity.[448]

The Silverman complaint seemingly focuses its derivative work claim on summaries of the plaintiffs' works that OpenAI produced in response to user prompts.[449] Summaries of book contents, as in Wikipedia or magazine reviews, are generally noninfringing. While summaries of in-copyright

443. Andersen Complaint, *supra* note 3, ¶ 93. Andersen's amended complaint proffers exhibits aimed at showing some substantial similarities. *See* First Amended Complaint, Andersen v. Stability A.I., Ltd., No. 3:23-cv-00201 (N.D. Cal. Nov. 29, 2023). Getty claims that some outputs of Stable Diffusion closely resemble particular Getty images but does not provide specific examples of what it considers infringing derivatives. *See* Getty Complaint, *supra* note 3, ¶ 51.

444. 17 U.S.C. § 101.

445. *See, e.g.*, Skidmore v. Led Zeppelin, 952 F.3d 1051, 1064 (9th Cir. 2020) (explaining that substantial similarity in the expression that the defendant allegedly copied from the plaintiff's work is necessary to establish infringement); Litchfield v. Spielberg, 736 F.2d at 1357 (characterizing a claim of infringement of the derivative work right in the absence of substantial similarity as "frivolous"); *see also* Authors Guild v. Google, Inc., 804 F.3d 202, 207 (2d Cir. 2015) (rejecting Guild's derivative work claims).

446. *See* 4 PATRY, *supra* note 11, § 12:13 (explaining that "there must be substantial similarity in protectible expression between the parties' works" to infringe the derivative work right); 2 NIMMER & NIMMER, *supra* note 441, § 8.09 (explaining the same); *see also* Samuelson, *supra* note 440, at 1525–27 (discussing the last clause of § 101's definition).

447. *See* Andersen Complaint, *supra* note 3, ¶ 4.

448. *See* Andersen v. Stability A.I. Ltd., No. 23-cv -00201, 2023 WL 7132064 (N.D. Cal. Oct. 30, 2023); *see also* Kadrey v. Meta Platforms, Inc., No. 23-cv-03417, 2023 WL 8039640 (N.D. Cal. Nov. 20, 2023) (dismissing claims on the same grounds).

449. *See* Silverman Complaint, *supra* note 3, ¶¶ 39–42, exhibit B. Silverman relies on the summaries as evidence that OpenAI used her book as training data. *See id.* ¶ 5.

works have sometimes been held to infringe derivative work rights,[450] users of ChatGPT who prompted the model to produce book summaries would seem to be responsible for those outputs, not OpenAI. For almost two decades, courts have withheld copyright infringement liability from developers of technology unless there was evidence that the copying complained of was the result of a volitional act of a human.[451] The Silverman and most other complaints do not articulate what volitional conduct would justify holding OpenAI liable for potentially infringing acts of its 180 million users. If a user asked for summaries for private noncommercial use or litigation purposes, the outputs, even if substantially similar to book expressions, may well be fair uses.[452] Moreover, OpenAI may have a *Sony* safe harbor defense because its technology has and is capable of substantial noninfringing uses.

It is, of course, possible for outputs of generative AI to be prima facie infringements of copyright owners' reproduction and derivative work rights.[453] This is particularly likely if users prompt AI systems to produce images of popular characters, such as Snoopy, Mickey Mouse, or other content the model may have "memorized" because of multiple duplicates in

450. *See* Twin Peaks Prods., Inc. v. Publ'ns Int'l, Ltd., 996 F.2d 1366, 1375 (2d Cir. 1993) (holding that summaries of certain plots of the television series Twin Peaks infringed).

451. *See, e.g.*, VHT, Inc. v. Zillow Grp., Inc., 918 F.3d 723, 732 (9th Cir. 2019) ("[T]o demonstrate volitional conduct, a [plaintiff] must provide some 'evidence showing the alleged infringer exercised control (other than by general operation of [its website]), selected any material for upload, download, transmission, or storage, or instigated any copying, storage, or distribution' of [its works]."(second alteration in original)); Fox Broad. Co. v. Dish Network LLC, 747 F.3d 1060, 1067 (9th Cir. 2014) (affirming denial of a preliminary injunction because even though Dish "exercises a degree of discretion over the copying process," it was the user, not Dish, who caused the copies to be made); Religious Tech. Ctr. v. Netcom On-Line Commc'n Servs., Inc., 907 F. Supp. 1361 (N.D. Cal. 1995) (finding Netcom not liable for infringing copies on its site because it engaged in no volitional conduct that contributed to the unlawful copying).

452. One does not get to fair use defenses unless a prima facie case of infringement has been established.

453. The most plausible derivative work right claims against a generative AI developer is Concord Music Group's complaint against Anthropic. Complaint ¶¶ 66–79, Concord Music Grp., Inc. v. Anthropic PBC, No. 3:23-cv-01092 (M.D. Tenn. Oct. 18, 2023) [hereinafter Concord Complaint] (giving examples of substantially similar lyrics generated by Anthropic's AI, Claude). The NYT Complaint, *supra* note 3, includes examples of near-verbatim copies, although that raises more of a reproduction than derivative work claim. The derivative work right in the context of generative AI is explored more fully in Oren Bracha, *Generating Derivatives: AI and Copyright's Most Troublesome Right*, 25 N.C. J.L. & TECH. 345 (2024).

the training dataset.[454] A firm that used memorized outputs in commercial advertising would likely infringe. But generative AI developers can take steps to prevent infringing outputs, for example, by removing duplicate images from its training dataset or using input or output filters to prevent infringing outputs.[455] AI developers' explanations of their mitigation efforts might well support their defenses against direct or indirect claims of infringement as to memorized content.[456] Alternatively, courts might instead view AI developers' mitigation efforts as indications that they had the ability to control infringement and their failure to effectively mitigate could count against indirect infringement defenses.

3. Training Data Ingestion Claims

Before discussing the contours of the fair use defenses likely to be raised to infringement claims based on generative AI systems' uses of in-copyright works as training data, it is important to understand that training datasets and generative AI software that embody trained models are distinct objects,[457] as are generative AI outputs. Claims of infringement of copyright's reproduction right are plausible against those who compile and curate datasets and against those who use such datasets in the process of training models. But once models have been trained, training datasets may be set aside or used only infrequently to retrain or fine-tune the models. The copies made for training purposes may thus be intermediate in character. Moreover, the datasets themselves may not be commercialized products, but even if they are, the datasets will remain distinct objects from models on which they were trained. Assembling and curating datasets for training

454. *See* Sag, *supra* note 8, at 321–28 (providing examples of character outputs); Guadamuz, *supra* note 8, at 25–26 (reporting low memorization rates). For a technical discussion of this phenomenon, see Carlini et al., *supra* note 432.

455. *See* Sag, *supra* note 8, at 337–41 (suggesting ways to mitigate infringement risks); Henderson et al., *supra* note 8, at 20–25; *see also* Lee et al., *supra* note 1, at 79 (noting difficulties in removing duplicates).

456. *See* Authors Guild v. Google, Inc., 804 F.3d 202, 221–23 (2d Cir. 2015) (describing various safeguards Google incorporated in the search and snippet features of GBS to avert potentially infringing uses). Andersen is among the class action plaintiffs who have claimed that generative AI defendants should be held liable for direct and vicarious infringement. Andersen Complaint, *supra* note 3, ¶¶ 169–77. Such claims would depend not only on the existence of infringing outputs, but also on a showing that those outputs were within the developers' ability to control. *See* Perfect 10, Inc. v. Visa Int'l Serv. Ass'n, 494 F.3d 788 (9th Cir. 2007) (holding that a lack of control over the infringing activity precluded the plaintiff's vicarious infringement claim).

457. *See, e.g.*, Lee et al., *supra* note 1, at 67.

models may, moreover, be done by different entities than model developers and deployers.[458] One important implication of these considerations is that even if a court ruled that the use of in-copyright works in a training dataset infringed copyrights, that infringement might not carry over to the generative AI system insofar as the models do not embody the data on which they were trained in recognizable form. Copyright case law has thus far not embraced a "fruit of the poisonous tree" theory of liability,[459] although plaintiffs in the generative AI cases may urge courts to adopt this theory if they prevail with their training data claims.

Fair use will almost certainly be the main defense raised in litigations in response to claims that the use of in-copyright works as data for training models is copyright infringement.[460] Fair use inquiries typically begin with consideration of the purpose and character of the challenged use. Since *Campbell*, an important and relevant consideration is whether the challenged use had a transformative character.

Generative AI defendants will likely assert that training data uses of in-copyright works is transformative, relying on the Second Circuit's *Authors Guild v. Google* decision, which involved a commercial firm's ingestion of in-copyright works for the purpose of creating a large database of works

458. The dataset on which Stable Diffusion's model was built, for instance, was prepared by a nonprofit German research organization known as LAION (Large-Scale Artificial Intelligence Open Network). *See* Andersen Complaint, *supra* note 3, ¶¶ 101–09. LAION makes its datasets available to the public for free for use as training data for AI models. *Id*.

459. In *Liu v. Price Waterhouse LLP*, No. 97 CV 3093, 2000 WL 1644585, at *2 (N.D. Ill. Oct. 30, 2000), the court rejected a proposed jury instruction because it was based on a "fruit of the poisonous tree" theory. The court explained "[t]hat defendants may have viewed or studied plaintiff's program is irrelevant if defendants' resulting work is not substantially similar to plaintiff's." *Id*. *Accord* Real View, LLC. v. 20–20 Techs., Inc., 811 F. Supp. 2d 553, 561 (D. Mass. 2011) (ordering remittitur to disallow the award of profits on noninfringing products despite their illegal download); *see also* Mark A. Lemley, *The Fruit of the Poisonous Tree in IP Law*, 103 IOWA L. REV. 245, 260 (2017).

460. Several commentators have discussed fair use defenses for training generative AI models. *See, e.g.*, Henderson et al., *supra* note 8 (writing as generally supportive of fair use defenses); Lee et al., *supra* note 1 (discussing fair use defenses but raising questions about them); Lemley & Casey, *supra* note 8 (writing as generally supportive of fair use defenses); Sag, *supra* note 8 (writing the same); Sobel, *supra* note 8 (writing as skeptical of fair use defenses, but also of full liability); Torrance & Tomlinson, *supra* note 8 (writing as supportive of fair use defenses); *see also* Bracha, *supra* note 8 (explaining why nonexpressive use defenses should succeed without reaching fair use); Joshua S. Gans, *Copyright Policy Options for Generative Artificial Intelligence* 2–3 (NBER Working Paper No. 32106, 2024), https://www.nber.org/system/files/working_papers/w32106/w32106.pdf [https://perma.cc/L2QX-SUJ5] (showing that economic analysis supports fair use as the socially optimal copyright outcome for the use of works to train large language models because of prohibitive transaction costs).

enabling the developer to produce outputs in response to user queries.[461] Insofar as generative AI systems' uses of in-copyright works are for very different purposes than the originals' purposes—as in *Field, iParadigms*, *Google Books*, and the search engine cases—the AI training uses are likely to be considered transformative. These developers may have had commercial purposes when using works as training data, but courts have generally given commerciality less weight when defendants' uses were transformative, as in *GBS* and the search engine cases.[462]

To properly assess the fair use purpose factor, it is important to understand how models are trained and how information from training data is represented in models. Computer scientist Christopher Callison-Burch explained the training process in his testimony at a congressional hearing.[463] He observed that the goal of training is to enable generative AI models to analyze, among other things, "the structure, syntax, and semantics of language," including "grammar, sentence construction, and how words and phrases are related to each other" in order to facilitate the generation of "coherent and contextually appropriate text."[464] The training process exposes the models to "a vast array of factual information, which they internalize and use to generate relevant responses or content," including "knowledge about geography, history, science, and various other domains."[465] During training, generative AI systems "gain some capacity for common sense reasoning, which allows them to understand basic cause and effect relationships, infer missing information, and make simple deductions."[466] AI models discern "underlying patterns, relationships, and

461. *See Authors Guild*, 804 F.3d at 208–09.

462. Generative AI plaintiffs may argue that the Court's *Warhol v. Goldsmith* decision has re-elevated commerciality as a significant consideration in fair use cases. Andy Warhol Found. for the Visual Arts v. Goldsmith, 598 U.S. 508, 531 (2023) (discussing commerciality). But training datasets may not be commercial products. Defendants can be expected to argue that the Court's *Google* decision, which also involved computer software and a commercial use, is the more apt precedent in generative AI cases than *Warhol*, especially if they can persuade courts that their uses of copyrighted works are for very different purposes than the originals'.

463. *See Understanding Generative Artificial Intelligence and Its Relationship to Copyright: Hearing on Artificial Intelligence and Intellectual Property: Part I – Interoperability of AI and Copyright Law Before H. Subcomm. on Cts., Intell. Prop., & the Internet*, 118th Cong. (2023) (statement of Christopher Callison-Burch, Associate Professor, University of Pennsylvania) [hereinafter Callison-Burch Testimony].

464. *Id.* at 9.

465. *Id.*

466. *Id.* at 10.

structures from the data, which allow them to generate entirely novel sentences, images, and other content."[467]

It is well established that the scope of copyright protection does not extend to "statistical information" such as "word frequencies, syntactic patterns, and thematic markers."[468] Computational processing of in-copyright works to extract "information about the original [work]" has been held noninfringing because this kind of extraction does not "replicat[e] protected expression."[469] Courts sometimes emphasize that not extending copyright protection to information embodied by works is essential to achieving copyright's overall purpose of promoting progress in the creation and dissemination of new knowledge so that "each author [can] build[] on the work of others."[470] The extractive purpose of training would seem to favor generative AI fair use defenses.[471]

Yet, one trial court has decided that a generative AI case, *Thomson Reuters Enterprise Centre GmbH v. Ross Intelligence Inc.*, should go to a jury because there was a triable issue of fact as to whether Ross used Westlaw headnotes to extract information about language patterns in Westlaw headnotes, which would be transformative, or used the headnotes to construct outputs that replicated the headnotes, which would not be transformative.[472] Generative AI plaintiffs can be expected to urge courts to view generative AI systems holistically, asserting that their outputs that have the same purpose as the input works (for example, stock images of soccer players) may, under the Supreme Court's *Warhol v. Goldsmith* decision, be deemed nontransformative.

As in the computational search and research tools cases, courts may give the nature-of-the-work factor little weight in the generative AI cases because those systems' uses of works as training data are nonexpressive—that is,

467. *Id.*
468. Authors Guild v. Google, Inc., 804 F.3d 202, 209 (2d Cir. 2015). *See generally* Bracha, *supra* note 8 (arguing that nonexpressive uses of in-copyright works for the purposes of extracting information and other unprotectable elements do not infringe); Molly Van Houweling, *The Freedom to Extract in Copyright Law*, 103 N.C. L. REV. (forthcoming 2025).
469. *Authors Guild*, 804 F.3d at 220.
470. Nash v. CBS, Inc., 899 F.2d 1537, 1540 (7th Cir. 1990) (holding that a television program based on information and speculations in nonfiction books about historical figure did not constitute infringement).
471. Van Houweling, *supra* note 468 (explaining various copyright doctrines as well as constitutional policies that support the freedom to extract data and other unprotectable elements in protected works, even when some expression must be copied to enable these extractions).
472. *See* Thomson Reuters Enter. Ctr. GmbH v. Ross Intel. Inc., No. 20-cv-613, 2023 WL 6210901, at *8 (D. Del. Sept. 25, 2023).

they use the works as data, not as works.[473] Moreover, insofar as generative AI systems train models on in-copyright works available on the open internet,[474] this consideration would seem to be relevant not only to the nature of the work factor, but also to the market effects factor, as in *Field*.[475] Yet, complaints filed by some plaintiffs, including the New York Times and the Authors Guild, emphasize that the high quality of their works when used for training data purposes is a consideration that supports their infringement claims.[476]

Another counterbalancing consideration in the *Silverman* case is the allegation that OpenAI trained its model on "shadow libraries" of pirated books.[477] Silverman can be expected to argue this undermines OpenAI's fair use defense, especially if OpenAI knew its models were being trained on pirated works.[478] Some judges and juries may be swayed by this argument.

Yet, OpenAI may argue, as in *iParadigms* and *Google Books*, that it was not using ingested works to exploit their expressions, but only to analyze them and to extract facts and other unprotectable elements.[479] OpenAI and other generative AI defendants facing similar claims will likely assert that models do not embody pirated books' expressions because their training datasets are distinct from the models,[480] and that shadow library books

473. *See* Bracha, *supra* note 8 (discussing non-expressive use defenses in generative AI cases); Matthew Sag, *Copyright and Copy-Reliant Technology*, 103 NW. U. L. REV. 1607 (2009) (discussing non-expressive uses of in-copyright works).

474. Search engines such as Google routinely make copies of digital works available on the open internet, as in *Field v. Google Inc.*, 412 F. Supp. 2d 1106 (D. Nev. 2006). A nonprofit organization, Common Crawl, uses webcrawling software to make copies of content on the open internet, which it makes available to anyone who wants a copy. *See* COMMON CRAWL, https://commoncrawl.org [https://perma.cc/S6DR-GDLE].

475. *See supra* text accompanying notes 354–361.

476. *See, e.g.*, Guild Complaint, *supra* note 3, ¶¶ 60–63; NYT Complaint, *supra* note 3, ¶ 90.

477. Silverman Complaint, *supra* note 3, ¶ 35. For a discussion of one such library, see, for example, Kate Knibbs, *The Battle Over Books3 Could Change AI Forever*, WIRED (Sept. 4, 2023, 6:00 AM), https://www.wired.com/story/battle-over-books3 [https://perma.cc/7BZW-XJUH].

478. *See* Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 563 (1985) (pointing out the trial court's finding that the defendant "knowingly exploited a purloined manuscript" of Ford memoir); *see also* Lee et al., *supra* note 1, at 103 (speculating that training on shadow libraries may undermine fair use defenses). Recall that the Ninth Circuit gave no weight to the infringing character of Perfect 10 images when considering Google's fair use defense in *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146 (9th Cir. 2007). *See supra* note 374–379 and accompanying text.

479. *See, e.g.*, Michael W. Carroll, *Copyright and the Progress of Science: Why Text and Data Mining Is Lawful*, 53 U.C. DAVIS L. REV. 893 (2019) (arguing that text and data mining (TDM) is lawful even if done on a corpus of infringing works).

480. *See* Lee et al., *supra* note 1, at 67.

constitute a very small fraction of the works on which generative AI models were trained. To the extent that Silverman raises the "shadow library" issue to inject "bad faith" into the fair use calculus, generative AI defendants may point to the Supreme Court's "justifiable" skepticism about whether such considerations are relevant to fair use determinations.[481]

The amount and substantiality factor will likely depend, as in other computational search and research tools cases, on how the courts weigh those factors in relation to the purpose and market effects factors. Copying entire works generally has not weighed against fair use in transformative use cases if the amount taken is reasonable in light of transformative purposes.[482] Generative AI defendants can be expected to argue that training datasets are intermediate steps in the AI model development process, akin to iParadigms' uses of student papers or Accolade's decompilation of Sega's videogame software, which suggests that training datasets may have no effect on copyright markets.[483]

The market effects factor is likely to weigh differently in the class action than in the individual plaintiff cases. Claims that uses of class members' works as training data have caused harm to markets for their works seem weak insofar as the data on which models were trained comes from copies of works available on the open internet. Moreover, no class yet exists that is capable of granting a license to use class members' works as training data, and none certainly existed when the defendants' datasets were created and used to train AI models. Nor is it possible for AI developers to license the rights to use all in-copyright works available on the internet, given the exceptionally large number of works and copyright owners at stake.[484] Transaction costs would be prohibitive relative to the value of use of works as training data.[485] Hence, market failure considerations would seem to undercut the class action plaintiffs' market harm arguments. The infeasibility of market formation was relevant to fair use defenses in *W&W*, *Field*, *HathiTrust*, and the *Google Books* cases, as it may be in generative AI cases as well.

Getty's complaint, by contrast, alleges that "Stability AI is stealing a service that Getty Images already provides to paying customers in the

481. *See, e.g.*, Google LLC v. Oracle Am., Inc., 593 U.S. 1, 31–32 (2021); Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 585 n.18 (1994).
482. *See, e.g.*, Authors Guild v. Google, Inc., 804 F.3d 202, 221–22 (2d Cir. 2015).
483. Guadamuz, *supra* note 8, at 10–12.
484. Callison-Burch Testimony, *supra* note 463, at 13; Lemley & Casey, *supra* note 8.
485. *See* Gans, *supra* note 460, at 2.

marketplace for that very purpose," adding that it has licensed "millions of suitable digital assets for a variety of purposes related to artificial intelligence and machine learning in a manner that respects personal and intellectual property rights."[486] It complains that "Stability AI has taken that same content from Getty Images without permission, depriving Getty Images and its contributors of fair compensation."[487] As in *Texaco*, the existence of an established market for licensing uses of works will certainly be relevant to Getty's unfairness argument and may weigh heavily with some judges.

Yet, whether Stability's failure to get a license from Getty will defeat its fair use defense remains to be seen. If other factors weigh in Stability's favor and Stability's use does not undermine Getty's market incentives, Stability's defense might still succeed. The market effects factor might also be neutral if Stability trained Stable Diffusion before Getty developed a training data licensing market. Consider also that Stable Diffusion was trained on an open source dataset created by LAION.[488] Under a European Directive exempting uses of in-copyright materials for text and data mining (TDM),[489] LAION's development of that dataset may well be lawful, which may have spillover effects for Stability's fair use defense.

The New York Times complaint does an impressive job linking the harm the company perceives from OpenAI's and Microsoft's uses of the newspaper's contents as training data with competitive harm to its market for news. The Times supported its allegation of harm by showing that it was able to use ChatGPT and Microsoft's Bing search engine (which draws on OpenAI models) to generate exact or near-exact copies of stories published in the newspaper that might divert users away from the Times' website.[490]

486. Getty Complaint, *supra* note 3, ¶ 56; *see also* Concord Complaint, *supra* note 453, ¶¶ 105–10 (claiming harm from the failure to license use of music lyrics). It would not be surprising if the Getty and Concord cases settle.

487. Getty Complaint, *supra* note 3, ¶ 56. The only decision to have addressed the merits of fair use defenses in a generative AI lawsuit is *Thomson Reuters Enter. Ctr. GmbH v. Ross Intel. Inc.*, No. 20-cv-613, 2023 WL 6210901 (D. Del. Sept. 25, 2023) (denying cross-motions for summary judgment on infringement claims for the use of Westlaw headnotes to train generative AI system because of triable issues of fact, including those over market effects).

488. Andersen Complaint, *supra* note 3, ¶¶ 101–09.

489. *See* Directive 2019/790, of the European Parliament and of the Council of 17 April 2019 on Copyright and Related Rights in the Digital Single Market and Amending Directives 96/9/EC and 2001/29/EC, 2019 O.J. (L 130) 92, 113–14.

490. *See* NYT Complaint, *supra* note 3, ¶¶ 99–100 (describing examples of near-identical outputs), 105–07 (describing the same), 110 (alleging that these outputs divert traffic away from the Times' site); *see also* Concord Complaint, *supra* note 453, ¶¶ 66–79 (offering examples of near-exact copies of lyrics generated by Claude).

That complaint also refers to failed negotiations between OpenAI and the newspaper for a license to use Times' content as training data,[491] which the newspaper will certainly raise as a highly relevant consideration in the analysis of the market effects factor in response to OpenAI's and Microsoft's fair use defenses.

All of the generative AI defendants are likely to invoke several defendant-friendly statements from the Supreme Court's *Google v. Oracle* decision.[492] Recall that Google persuaded the Court that the market effect factor weighed in favor of its fair use defense, despite Oracle's argument that Google knew that it needed a license to make use of 11,500 declarations from the Java Standard Edition for the Android platform.[493] In discussing market effects, *Google* emphasized that courts should consider public benefits that may arise from challenged uses and balance them along with any lost revenues the plaintiffs may claim.[494] *Google* also directed consideration of how much creativity the challenged use enabled as relevant to assessment of market effects because millions of programmers were able to write apps for the Android platform, as they could take advantage of their expertise with Java's application program interface.[495] The generative AI defendants will almost certainly draw upon these statements in their fair use arguments, for trained models are, like the Android platform, highly creative and have enabled millions of people to generate desired outputs. Yet, if courts ultimately view generative AI defendants as making parasitic use of the plaintiffs' creative works,[496] they may be much more sympathetic to the plaintiffs' claims of infringement. As in several of the past disruptive technology cases addressing fair use defenses, it would not be surprising if appellate courts rendered split decisions in the generative AI cases or if some fair use defenses prevailed while others did not.

491. *See* NYT Complaint, *supra* note 3, ¶¶ 8–9.
492. *See* Google LLC v. Oracle Am., Inc., 593 U.S. 1 (2021).
493. *See id.* at 35–40; *see also* Authors Guild v. Google, Inc., 804 F.3d 202, 226 (2d Cir. 2015) (rejecting plaintiffs' arguments that the possibility of licensing markets supported their derivative right claims).
494. *Google*, 593 U.S. at 35–36.
495. *Id.* at 39.
496. *See, e.g.*, Guild Complaint, *supra* note 3, ¶¶ 127–30.

B. Questioning Proposals for a Generative AI Collective License Regime

Because courts will almost certainly take several years to definitively resolve the many thorny copyright issues the generative AI lawsuits pose, interest has been growing in the possibility of a collective licensing regime for AI training data. The U.S. Copyright Office has issued a Notice of Inquiry about generative AI copyright issues, asking for comments about collective licensing, among other things.[497]

The Authors Guild is among the organizations representing creators who have met with congressional personnel to express support for legislation to create a collective licensing regime. Under this regime, makers of generative AI systems would be required to negotiate with a yet-to-be-named collective management organization (CMO), such as the CCC, to represent registered right holders and have authority to grant a blanket license to use copyright works as training data.[498] These groups are not alone in endorsing a collective license regime for generative AI.[499]

Congress has adopted compulsory collective licensing regimes to regulate copyright-disruptive technologies in the past.[500] The AHRA is one example and perhaps an object lesson in that it was unsuccessful in achieving the goal of getting meaningful compensation to right holders.[501] The Music Modernization Act (MMA) has more recently created such a regime for

497. *See* Artificial Intelligence and Copyright, 88 Fed. Reg. 59942, 59947 (Aug. 30, 2023).

498. *See* Adolphus, *supra* note 7. For a summary of the Guild's conception of the collective license regime for professional writers, see *FAQs on the Authors Guild's Positions and Advocacy Around Generative AI*, AUTHORS GUILD, https://authorsguild.org/advocacy/artificial-intelligence/faq [https://perma.cc/7BEL-28JS].

499. *See* Christophe Geiger & Vincenzo Iaia, *The Forgotten Creator: Toward a Statutory Remuneration Right for Machine Learning of Generative AI*, 52 COMPUT. L. & SEC. REV., Apr. 2024, at 1, 2; Martin Senftleben, *Generative AI and Author Remuneration*, 54 INT'L REV. INTELL. PROP. & COMPETITION L. 1535 (2023); Sobel, *supra* note 8, at 91–93; *see also* Memorandum Prepared for Copyright Clearance Center, Inc., by Prof. Daniel Gervais (2023) (on file with author). Gervais, Geiger, and Sobel recommend a levy system for training data; Senftleben favors a levy on generative AI services.

500. Copyright law was, of course, adopted to respond to the disruptive technology of printing presses. The disruptive technology that led to the creation of the first compulsory license in U.S. copyright law was player pianos. After the Supreme Court in *White-Smith Music Publishing Co. v. Apollo Co.*, 209 U.S. 1 (1908), ruled against music publishers' claims that Apollo's piano rolls infringed copyrights in their musical compositions, Congress granted music copyright owners a right to control the mechanical reproductions of these compositions, but subjected that right to a compulsory license after the first recording of the compositions, now codified at 17 U.S.C. § 115(a).

501. *See supra* text accompanying notes 321–322.

noninteractive digital audio services,[502] which was carefully targeted to address one narrow type of exploitation of expressions in sound recordings. Perhaps the most successful such license has been the complex regime that regulates cable television systems' retransmission of broadcast signals.[503] Congress tends to adopt compulsory collective licenses when it would be virtually impossible to clear rights with every affected right holder, but thinks some compensation should be provided for second comers' commercial exploitations of copyrighted works.

While the generative AI class action plaintiffs' actual damage claims seem weak, critics of generative AI may emphasize unjust enrichment to a legislative audience. Consider the Authors Guild's open letter to CEOs of generative AI companies:

> Generative AI technologies built on large language models owe their existence to our writings. These technologies mimic and regurgitate our language, stories, style, and ideas. Millions of copyrighted books, articles, essays, and poetry provide the "food" for AI systems, endless meals for which there has been no bill. You're spending billions of dollars to develop AI technology. It is only fair that you compensate us for using our writings, without which AI would be banal and extremely limited.[504]

Generative AI outputs may compete with human-creator outputs, which is why some believe they pose dire threats to the ability of authors and artists to earn a living from their creative work.[505] That Big Tech companies are obtaining revenues for consumers' and commercial customers' uses of generative AI models to produce outputs, when they paid nothing for the inputs, is more than a little grating on the author and artist critics of generative AI.[506]

Yet, the logistical challenges likely to attend efforts to develop a mandatory collective license regime for the right to use in-copyright works as training data may be much more daunting than those attending the cable

502. *See* Orrin G. Hatch–Bob Goodlatte Music Modernization Act, Pub. L. No. 115-264, 132 Stat. 3676 (2018) (compulsory license codified at 17 U.S.C. § 115(d)).
503. *See* 17 U.S.C. § 111.
504. *Open Letter to Generative AI Leaders*, AUTHORS GUILD, https://actionnetwork.org/petitions/authors-guild-open-letter-to-generative-ai-leaders [https://perma.cc/6U2M-UKWM].
505. *See id.*; Senftleben, *supra* note 499, at 1539.
506. *See* Klein, *supra* note 2.

and broadcast, AHRA, and MMA regimes.[507] For one thing, each of those regimes regulate particular types of works in particular industry sectors. Generative AI systems, by contrast, may train on works of all kinds. Second, existing regimes base compensation on ongoing exploitations of the expressions in works covered by the licenses,[508] whereas generative AI systems may train on in-copyright works only once (to build a model) or may reuse training data only infrequently. Third, unlike other industries regulated through compulsory collective licenses, the generative AI industry is very young and evolving very rapidly.[509] The economics of this young industry are likely to evolve in ways that legislators cannot predict and make the alignment of regulatory goals difficult to achieve.

A collective license would, moreover, be very difficult to administer given the staggeringly large number of copyright owners and of works and of types of works used as AI training data—literally in the billions. The Authors Guild proposes a collective license for professional writers,[510] but they are far from the only authors whose works are contributing to the value of AI models. Generative AI systems are, moreover, being trained on many types of digital data, including computer programs, blogs, digital chat, digital audio, digital video, and digital visual art. While some generative AI systems may focus on only one type of digital data, other generative AI systems are multi-modal (that is, they process multiple types of inputs, such as images, videos, and text). Creating a collective license regime for each data type would likely

507. The Copyright Office once considered possible adoption of an extended collective license regime to enable mass digitization of books, but there was no consensus among stakeholders about the desirability of such a regime and how it should be structured. *See* Letter from Karyn A. Temple Claggett, Acting Reg. Copyrights & Dir., U.S. Copyright Off., to Senators Grassley and Feinstein (Sept. 29, 2017) (explaining reservations about the proposal). A collective license regime to allow use of digitized in-copyright works as training data may encounter more resistance than the mass digitization project.

508. *See, e.g.*, 17 U.S.C. § 111(d)(1) (detailing cable retransmission system statements of account on a semiannual basis); 37 C.F.R. pt. 386 (2023) (setting 2024 royalty rates for satellite carrier compulsory license as established under 17 U.S.C. § 119(c)(2)(C). If some form of compensation were to become necessary, then paying royalties to right holders based on some sampling of the exploitations of protected works might make sense, though still problematic. *See also* Jacob Victor, *Reconceptualizing Compulsory Copyright Licenses*, 72 STAN. L. REV. 915, 982–88 (2020) (advocating a policy-driven rate setting schema in the music streaming context, weighting public access).

509. Martin Senftleben recommends the adoption of a levy focused on the outputs of generative AI, even though he recognizes that the outputs will generally not be substantially similar in expression to works ingested as inputs. He argues this would create a new revenue stream to support authors and artists. *See* Senftleben, *supra* note 499, at 1550–57.

510. *See FAQs on the Authors Guild's Positions*, *supra* note 498.

be unworkable, but so would be trying to include all data types in one regime, for the value of inputs may vary significantly.

Administering such a regime would be very costly, even if it could be set up. A significant proportion of revenues collected from generative AI companies would be needed to cover administrative costs. The licensing entity, whether a CMO or a governmental agency, would almost certainly be the major beneficiary of revenue streams from generative AI firms.[511] The amounts paid to individual copyright owners would likely be very modest, and would be unlikely to provide significant financial support to authors and artists.[512]

Also challenging would be determining how to measure value and set prices for licensing. Would all works be valued the same? If not, on what criterion could differential value and payments be assessed? Would there be a one-time payment, a payment each time the model was retrained on those data, or would generative AI systems have to submit payments on a regular basis? Would payments depend on who has registered with a CMO or governmental entity administering the licensing regime? What would the administrator do with orphan works or works as to which ownership rights were in dispute?[513]

The global nature of the internet and much of the data on which generative AI foundation models have been trained would make a single nation's regime radically incomplete. If one country or region developed a mandatory collective license regime, that might precipitate adoption of similar regimes elsewhere.[514] A proliferation of collective license regimes, each of which would aim to collect significant sums to support authors and artists, might make development of generative AI too expensive to be feasible. Moreover, some nations may choose not to impose mandatory collective license regimes to attract AI developers, as Israel has seemingly

511. *See* Jonathan Band & Brandon Butler, *Some Cautionary Tales About Collective Licensing*, 21 MICH. ST. U. COLL. INT'L L. REV. 687 (2013) (discussing corruption, lack of transparency, and inefficiencies of CMOs).

512. *See* Senftleben, *supra* note 499, at 1548–53 (explaining why a levy on training data uses is unlikely to benefit individual authors and why a levy on AI services would generate meaningful remuneration for authors).

513. Similar questions arose in connection with the proposed settlement of the *Authors Guild v. Google* litigation. *See* Pamela Samuelson, *The Google Book Settlement as Copyright Reform*, 2011 WIS. L. REV. 479 (discussing orphan works and ownership disputes).

514. *See, e.g.*, Senftleben, *supra* note 499, at 1542–57 (anticipating the adoption of an EU collective license).

done.[515] This might result in innovation arbitrage with AI developers and development shifting to less regulated environments.

Competition policy considerations may also caution against a compulsory collective license regime for use of works as training data. While Meta, Alphabet, and Microsoft can afford to pay license fees, a licensing mandate would likely create entry barriers for startups and entrench the large incumbents (who did not have to pay when training their models). Nonprofit research organizations and individual researchers might be exempted, as the EU did with its TDM rules.[516] If, however, nonprofits can lawfully make models and then release them on an open-source basis, can a commercial entity use those models for their commercial purposes?[517] And what rules should apply when companies integrate generative AI into other software applications, as Microsoft has done with ChatGPT in Bing? Would the collective license regime apply to firms that use generative AI systems in house without licensing their use to the public?

As challenging as collective license design issues may be, the political reality is that legislation to adopt such a regime is unlikely, at least in the United States. Most major copyright industries have not been galvanized against generative AI, and they, not the Authors Guild, have the political clout to advance copyright legislation.[518] Private licensing deals between large right holders and large generative AI firms are far more likely. This seems to be happening already.[519]

Insofar as the plea for collective licensing is based on concerns about displacement of employment opportunities for individual authors and

515. MINISTRY OF JUST., STATE OF ISRAEL, OPINION: USES OF COPYRIGHTED MATERIALS FOR MACHINE LEARNING (2022) (opining that the use of in-copyright works as training data is fair use).
516. *See* Directive 2019/790, *supra* note 489, at 11.
517. This is apparently what Stability AI did with a training data set created by LAION.
518. For example, in its comment submitted in response to the Copyright Office's Notice of Inquiry on generative AI, *supra* note 497, the Motion Picture Association (MPA) characterized as unhelpful any generalizations about using copyrighted works to train AI models as all fair or unfair, and accepted that some training data uses might qualify as fair use depending on the particular facts. *See* Motion Picture Ass'n of Am., Inc., Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 2, 15–19 (Dec. 5, 2023). The comment added that MPA members do not favor collective or compulsory licensing, preferring instead open market solutions, and generally believe that no changes to the Copyright Act or other legislative changes are presently needed. *See id.* at 28–30.
519. The Associated Press, for example, recently announced a deal with OpenAI for use of its works as training data. *See* Matt O'Brien, *ChatGPT-Maker OpenAI Signs Deal With AP to License News Stories*, AP NEWS (July 13, 2023, 8:41 AM), https://apnews.com/article/openai-chatgpt-associated-press-ap-f86f84c5bcc2f3b98074b38521f5f75a [https://perma.cc/9TE3-7E5F].

artists, it is important to realize that they are far from the only workers whose jobs may be displaced by generative AI.[520] Government policymakers should formulate plans for promoting employment opportunities for human workers displaced by AI.[521] This is, however, a broad policy challenge that copyright law is ill suited to address.

CONCLUDING THOUGHTS

Before *Sony*, courts and commentators often assumed harm from the existence of unauthorized copies. One school of thought posited that harm was irrelevant to copyright liability and entitlement to relief.[522] *Sony* represented a sea change in judicial understanding of the market effects factor in fair use cases, later refined in *Campbell* and *Google v. Oracle*. Since *Sony*, courts have consistently looked for evidence of actual or likely harm before ruling against fair use defenses. In some past disruptive technology cases, courts found that challenged uses caused actual harm to copyright markets, as in *Kinko's* and *Napster*, or likely harm to present or future licensing markets, as in *Napster* and *TVEyes*. Fair use defenses have tended to prevail, however, when claimed harms were speculative or when courts perceived licensing markets as unlikely to form, as in *Field*, *iParadigms*, and the two *Google Books* cases. Courts have sometimes been skeptical of claims that a failure to license caused harm, as in *W&W*, *CUP*, and *Cablevision*, in part because of the circularity problem with market harm claims. Moreover, the mere existence of licensing opportunities that defendants declined did not sink fair use defenses in *CUP*, *Sega v. Accolade*, and *Google v. Oracle*. Some courts, including *Sony*, have opined that the harm must be substantial enough to affect authorial incentives. Moreover, courts may be wary of licensing claims when defendants make transformative uses of earlier works, sometimes saying that copyright owners do not have rights to control transformative markets, as in *Google v. Oracle*. Yet, when copyright owners

520. *See* Ben Wodecki, *Goldman Sachs: Generative AI Could Replace 300 Million Jobs*, AI BUS. (Mar. 29, 2023), https://aibusiness.com/nlp/goldman-sachs-generative-ai-could-replace-300-million-jobs [https://perma.cc/7U9R-44UU]. Most at risk are office and administrative jobs, followed by jobs in law, architecture, and engineering.

521. *See* Daron Acemoglu, David Autor & Simon Johnson, MIT SHAPING FUTURE WORK INITIATIVE, POLICY MEMO: CAN WE HAVE A PRO-WORKER AI? (2023), https://shapingwork.mit.edu/wp-content/uploads/2023/09/Pro-Worker-AI-Policy-Memo.pdf [https://perma.cc/428N-DVBC].

522. *See* Leo J. Raskind, *A Functional Interpretation of Fair Use*, 31 J. COPYRIGHT SOC'Y U.S.A. 601, 607–14 (1984) (discussing cases); *see also supra* Part II.B.2 (discussing Ladd and Akin Gump commentaries).

have successfully developed new licensing markets through CCC, as in *Texaco*, the market effects factor has weighed against fair use defenses.

The generative AI training data fair use defenses will likely turn on how courts assess the market effects factor. The class action market harm allegations against Alphabet, Meta, OpenAI, and Stability seem at this stage too speculative to weigh against fair use. CCC might welcome the opportunity to license uses of in-copyright works as training data for generative AI systems,[523] but CCC does not have an existing portfolio of rights with which to license uses of works as training data to generative AI companies, and it is unlikely to be able to get a substantial quantum of such rights any time soon. For reasons discussed above, collective licensing markets for training data uses may, moreover, be very difficult or impossible to form or administer.

Getty may have a reasonable chance of defeating a generative AI fair use defense if it can persuade a court that Stability is harming an existing market for licensing uses of its collection of images as training data. Yet, to prevail on its claim that Stable Diffusion infringes and its model should be destroyed, courts would have to embrace—as they have never previously done—a fruit of the poisonous tree theory that outputs of generative AI systems infringe copyrights without any showing of substantial similarity in expression in particular works on which the system was trained and in outputs that may compete with the Getty's works.[524] The New York Times complaint does an impressive job tying its training data and output infringement claims together and explaining why OpenAI's system is likely to have detrimental effects on its markets; but of course, the Times will have to prove these allegations to defeat OpenAI's defenses.[525]

It remains to be seen what weight courts in the generative AI cases will give to public benefits enabled by unlicensed copying for training data purposes when assessing market effects, as the Court directed in *Google v. Oracle* and as courts have done in several disruptive technology cases—including *W&W*, *Sony*, the two *Google Books* cases—and search engine cases. Among the many socially beneficial uses of generative AI systems are to plan gardens, trips to exotic places, workouts, and meals, write wedding speeches and emails, sort through archives of photographs, organize research for one's

523. *See* Gervais, *supra* note 499.
524. This Article does not address any of Getty's non-copyright claims.
525. It would not be surprising if these two generative AI cases settle to avoid the risk of staggeringly large damage awards and court orders that the defendants must destroy models trained on the plaintiffs' works, as those complaints request. *See* Samuelson, *supra* note 431.

thesis, fix software bugs, and build Spotify playlists.[526] None of the lawsuits acknowledge that if generative AI developers are infringers, users of these technologies would likely be infringers too, although this is implicit in the lawsuits' vicarious infringement claims.[527] As in *Sony*, courts may be reluctant to extend copyright liability so broadly as to treat tens or hundreds of millions of users of generative AI systems as infringers. Moreover, companies are widely adopting AI systems to improve their operations.[528] They too will be affected by the outcomes of the generative AI lawsuits.

Generative AI plaintiffs will emphasize another public interest affected by generative AI systems: the effects of generative AI on future livelihoods of authors and artists. Are generative AI systems endangering the ability of authors and artists to make a living, or are these claims overblown, as have been similar past predictions of doom when disruptive technologies enabled unexpected uses of in-copyright works?[529] Recall Alan Latman's argument in *W&W* that unlicensed photocopying posed a serious threat to the future of journal publishing. Jack Valenti likened Betamax machines to the Boston Stranger, implying that the machines would kill the motion picture industry. Universal's lawyer told the *Sony* Court that the movie industry was suffering "enormous" losses from home taping.[530] Yet, the motion picture industry greatly benefited from the installed base of VTRs. Professor Nimmer opined that home audio taping was having "devastating" impacts on the music industry,[531] and the RIAA persuaded Congress of such a threat to its industry that it should mandate levies and technological limits on DAT machines. Yet the development of MP3 players and other digital devices have enabled new revenue streams

526. Francesca Paris & Larry Buchanan, *35 Ways Real People Are Using A.I. Right Now*, N.Y. TIMES (Apr. 14, 2023), https://www.nytimes.com/interactive/2023/04/14/upshot/up-ai-uses.html [https://perma.cc/5V3N-T2VT].

527. In recognition of this risk, some generative AI companies have announced their intent to indemnify users against infringement claims. *See, e.g.*, Blake Brittain, *Google to Defend Generative AI Users From Copyright Claims*, THEPRINT (Oct. 12, 2023), https://theprint.in/tech/google-to-defend-generative-ai-users-from-copyright-claims/1801743 [https://perma.cc/A5YF-JTK4].

528. *See, e.g.*, Paul Smith-Goodson, *The Extraordinary Ubiquity of Generative AI and How Major Companies Are Using It*, FORBES (July 21, 2023), https://www.forbes.com/sites/moorinsights/2023/07/21/the-extraordinary-ubiquity-of-generative-ai-and-how-major-companies-are-using-it/?sh=6df79e6e2124 [https://perma.cc/LC7X-UK9L].

529. *See* Nate Anderson, *100 Years of Big Content Fearing Technology—In Its Own Words*, ARS TECHNICA (Oct. 11, 2009), https://arstechnica.com/tech-policy/2009/10/100-years-of-big-content-fearing-technologyin-its-own-words [https://perma.cc/758S-N9GD].

530. *See supra* note 269 and accompanying text.

531. *See supra* note 303 and accompanying text.

that greatly benefit the music industry. Of course, the fact that past predictions of doom have not been overblown is no guarantee that the plaintiffs' predictions about the effects of generative AI are unsound, but they might be.

Courts have no choice but to predict what outcome in the generative AI cases will best promote the constitutional purposes of copyright. On the answer to that question may well hang the future of generative AI.