AIPLA QUARTERLY JOURNAL

Volume 53, Number 4 Page 659 Fall 2025

NOTE

JUST A GENERATIVE AI USER, STANDING IN FRONT OF A COPYRIGHT TROLL, ASKING IT TO LOVE HER:
GENERATIVE AI, COPYRIGHT TROLLS, AND THE STATUTORY DAMAGES REGIME

*Louise Strange**

* © 2025 Louise Strange, J.D. Candidate, 2026, The George Washington University Law School; B.A., 2021, St Andrews William & Mary Joint Degree Programme. It takes a village, but I would like to thank in particular Professor David Ludwig, Professor Robert Brauneis, and Chase O'Brien for their invaluable guidance in conceptualizing and executing this Note through its (many) drafts. I extend my gratitude to the editorial board and the lovely publication staff for their hard work in bringing this to its current form. And finally, no vanity note would be complete without a thank you to my parents, siblings, and fabulous network of friends.

I. INTRODUCTION

Imagine you are a student: diligently engaged in research and writing morning, noon, and night, but struggling to find time to read all the material relevant to the paper you are currently working on. Whether that is a journal note for law school, a PhD dissertation, or an assignment to fulfill a general education requirement, you are short on time and turn to generative artificial intelligence ("Generative AI") (with the blessing of your professor, of course) to summarize the significant developments in the area of study you are writing on. As a student of the world, you will, of course, return to do further and more comprehensive research, but right now you have an interim deadline to meet. ChatGPT spits out a neat few paragraphs explaining the state of the art, and you end up using some of this information in your writing with a citation to the Generative AI.

You turn in your paper and, sometime later, decide to publish.

Several weeks later, you receive a letter threatening to sue for up to $150,000 for the infringement of copyrighted material of a website you have never even heard of. The letter offers you the option to settle for a paltry $10,000, rather than go to court. You have some inkling of how expensive litigation can be, and you know exactly how much you owe in student loans, so you prostrate yourself before your income-earning loved ones and borrow the money to make this go away, vowing to never use Generative AI again.

The case of the student is one that could soon plague users of Generative AI due to the unique litigation tactics of copyright trolls. These are legal entities that pursue sometimes legitimate copyright claims with an interest in monetizing infringement.[1] Attracted to the possibility of quick and potentially massive out-of-court settlements precipitated by copyright law's statutory damages regime,[2] which allows monetary recovery at rates often far above the injury suffered,[3] copyright trolls flock to areas where actual and perceived infringement runs rampant.[4]

---

1 Brad A. Greenberg, *Copyright Trolls and Presumptively Fair Uses*, 85 UNIV. COLO. L. REV. 53, 59 (2014).

2 *See* 17 U.S.C. § 504.

3 *See* Pamela Samuelson & Tara Wheatland, *Statutory Damages in Copyright Law: A Remedy in Need of Reform*, 51 WM. & MARY L. REV. 439, 456 (2009).

4 *See, e.g.*, *USCG v. The People*, ELEC. FRONTIER FOUND., https://www.eff.org/cases/uscg-v-people [https://perma.cc/9UZR-MBMV].

Due to the often legally questionable development process for Generative AI,[5] copyright trolls could next swarm to this area.[6] Generative AI developers have all but admitted to using copyrighted material without a license.[7] What this means on a practical level is that the input and output of Generative AI often infringe existing copyrights on some level.[8] Copyright trolls may not only take on the developers of Generative AI themselves, but also go after end-users as well.[9]

This Note will investigate first how copyright law relates to Generative AI input and output. Next, it will explore the history of copyright trolls, what they are, and where these entities have often concentrated their efforts. Finally, this Note will suggest potential solutions for the problem posed by this interaction. Keeping in mind the purposes of copyright law and the necessity of access to legal recourse for actual infringement, this Note argues that the solution is to apply a presumption of innocent infringement to Generative AI cases, thereby limiting the damages amount to $200 per work infringed unless willful infringement is shown. This presumption means copyright trolls would have a higher burden to satisfy to achieve a massive windfall, discouraging them from encroaching on end-users of Generative AI. However, when it comes to the developers of Generative AI themselves, copyright trolls would be free to pursue litigation, as this presumption is likely to be overcome.

II. BACKGROUND

To better situate the problem posed by copyright trolls in the developing field of Generative AI, this Part will first dive into the legal underpinnings of copyright law and the statutory damages scheme. This Part will then explain how Generative AI works and why that process is likely to implicate unlicensed copyrighted works. Finally, this Part will define what a copyright troll is for the purposes of this analysis and explore how copyright trolls may enter the field of Generative AI, both from the developer and end-user perspectives.

---

5 *See infra* Section II.B.1.b.

6 *See infra* Section II.B.2.b.

7 *See, e.g.*, Andreessen Horowitz, Comment Letter on Notice of Inquiry on Artificial Intelligence & Copyright (Oct. 30, 2023), https://www.regulations.gov/comment/COLC-2023-0006-9057 [https://perma.cc/232M-2RYA].

8 *See infra* Section II.B.1.b.

9 *See infra* Section II.B.2.b.

A. SOURCES & IMPLEMENTATION OF COPYRIGHT LAW

Copyright law draws its source directly from the so-called Intellectual Property ("IP") Clause of the Constitution, which grants Congress the power to "promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries."[10] In other words, the Constitution provides for a limited monopoly on certain artistic expressions and creations in order to foster innovation and creativity.[11] By giving authors the set of exclusive rights granted by copyright law, legislators hope to "enrich[] the general public through access to creative work."[12]

The most recent statute that governs copyright law was initially passed in 1976 (the "Copyright Act") and extends to "original works of authorship fixed in any tangible medium of expression."[13] Copyrightable works are protected from unauthorized use and reproduction.[14] These exclusive rights exist from the moment a work is fixed in a tangible medium, meaning unpublished works are protected almost identically to published works under copyright law.[15] Greater protections, such as access to an infringement action and statutory damages, are only offered to works registered with the Copyright Office, but the rights

---

[10] U.S. CONST. art. I, § 8, cl. 8; *Intellectual Property Clause,* CORNELL L. SCH.: LEGAL INFO. INST. (June 2023), https://www.law.cornell.edu/wex/intellectual_property_clause [https://perma.cc/R6XP-7VXH] [hereinafter *Intellectual Property Clause*].

[11] *See* U.S. CONST. art. I, § 8, cl. 8. In addition to providing the legal underpinnings for copyright law, this clause also empowers Congress to enact legislation with regard to patents. *See Intellectual Property Clause*, *supra* note 10.

[12] *What is The Purpose of Copyright Law*, COPYRIGHT ALL.: COPYRIGHT L. EXPLAINED, https://copyrightalliance.org/education/copyright-law-explained/copyright-basics/purpose-of-copyright/ [https://perma.cc/X7G9-G555].

[13] 17 U.S.C. § 102; U.S. COPYRIGHT OFF., COPYRIGHT LAW OF THE UNITED STATES (TITLE 17) 8 (2024), https://www.copyright.gov/title17/ [https://perma.cc/E89G-JUK7] [hereinafter COPYRIGHT LAW OF THE UNITED STATES].

[14] *See* 17 U.S.C. §§ 106, 501(a).

[15] *What is Copyright?*, COPYRIGHT.GOV, https://www.copyright.gov/what-is-copyright/ [https://perma.cc/9EBP-DUD8]; *What is the Purpose of Copyright Law*, *supra* note 12.

themselves exist regardless of registration.[16] Copyright protection extends to, most notably for the purposes of this paper, literary, musical, pictorial, and graphic works.[17] The Copyright Act otherwise details how to obtain copyright and how to recover for infringement.[18] This Section will first discuss how copyright infringement is analyzed and then how the Copyright Act structures its statutory damages scheme.

1. *How Direct Copyright Infringement is Analyzed*

"[C]opyright infringement occurs when a copyrighted work is reproduced, distributed, performed, publicly displayed, or made into a derivative work without the permission of the copyright owner."[19] In order to file a copyright infringement action, the potential plaintiff must have a valid copyright registration.[20] In a legal action, copyright infringement is established when a plaintiff demonstrates "(1) ownership of a valid copyright, and (2) copying of constituent elements of the work that are original."[21] All federal courts of appeals have interpreted the second element to include an analysis of whether the works are "substantially similar" to one another, though circuits are split as to how that

---

16 *See* 17 U.S.C. § 411(a); *What Are Statutory Damages and Why Do They Matter?*, COPYRIGHT ALL.: FAQS, https://copyrightalliance.org/faqs/statutory-damages-why-do-they-matter/ [https://perma.cc/A9B5-UGH2] [hereinafter *What Are Statutory Damages*].

17 17 U.S.C. § 102 ("Works of authorship include the following categories: (1) literary works; (2) musical works, including any accompanying words; (3) dramatic works, including any accompanying music; (4) pantomimes and choreographic works; (5) pictorial, graphic, and sculptural works; (6) motion pictures and other audiovisual works; (7) sound recordings; and (8) architectural works.").

18 *See generally* Copyright Act of 1976, 17 U.S.C. §§ 101–1401.

19 *Definitions*, COPYRIGHT.GOV: FAQS, https://www.copyright.gov/help/faq/faq-definitions.html [https://perma.cc/4MKQ-3ZF3]; *see* 17 U.S.C. § 501(a) ("Anyone who violates any of the exclusive rights of the copyright owner . . . is an infringer of the copyright.").

20 17 U.S.C. § 411(a).

21 Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 361 (1991).

analysis is conducted.[22] The Second Circuit[23] approaches the analysis "by comparing the contested design's 'total concept and overall feel.'"[24] Meanwhile, the Ninth Circuit evaluates substantial similarity through a "two-part analysis: an objective extrinsic test and a subjective intrinsic test."[25] The extrinsic test involves "filter[ing] out unprotectable elements and then compar[ing the] remaining" similarities.[26] In short, the crux of any infringement analysis turns on whether the copyrighted and infringing works are substantially similar.[27]

2. *Other Forms of Copyright Infringement Liability*

Beyond direct copyright infringement, a potential plaintiff can also sue for secondary liability under a theory of vicarious or contributory liability. Secondary liability refers to a legal situation where a party can be held responsible for the actions of another due to the nature of their relationship.[28] The Copyright Act itself

---

[22] *See, e.g.*, Tanksley v. Daniels, 902 F.3d 165, 171 (3d Cir. 2018); Airframe Sys., Inc. v. L-3 Commc'ns Corp., 658 F.3d 100, 105 (1st Cir. 2011); Peter F. Gaito Architecture, L.L.C. v. Simone Dev. Corp., 602 F.3d 57, 63 (2d Cir. 2010); Bldg. Graphics, Inc. v. Lennar Corp., 708 F.3d 573, 577 (4th Cir. 2013); Folkens v. Wyland Worldwide, L.L.C., 882 F.3d 768, 774 (9th Cir. 2018); *see also Substantial Similarity in Copyright: It Matters Where You Sue*, DLA PIPER: INTELL. PROP. AND TECH. NEWS (Dec. 22, 2020), https://www.dlapiper.com/en/insights/publications/intellectual-property-news/2022/ipt-news-q4-2020/substantial-similarity-in-copyright [https://perma.cc/9EQP-Z3CQ].

[23] This Section only describes the difference in approach between the Second and Ninth Circuits as these two circuits see the majority of copyright cases. *See Substantial Similarity in Copyright: It Matters Where You Sue*, *supra* note 22.

[24] *Peter F. Gaito Architecture*, 602 F.3d at 66 (quoting Tufenkian Imp./Exp. Ventures, Inc. v. Einstein Moomjy, Inc., 338 F.3d 127, 133 (2d Cir. 2003)).

[25] Gilbert-Daniels v. Lions Gate Ent., Inc., No. 24-153, 2024 WL 5116299, at *1 (9th Cir. Dec. 16, 2024) (quoting Swirsky v. Carey, 376 F.3d 841, 845 (9th Cir. 2004)).

[26] *Id.*

[27] *See Substantial Similarity in Copyright: It Matters Where You Sue*, *supra* note 22.

[28] *Secondary Liability*, CORNELL L. SCH.: LEGAL INFO. INST. (June 2024), https://www.law.cornell.edu/wex/secondary_liability [https://perma.cc/RS8B-P9WL].

is silent on the issue of secondary liability,[29] but a robust body of case law recognizes this theory of recovery.[30]

Vicarious liability is found when a party "profit[s] from direct infringement while declining to exercise a right to stop or limit it," while contributory liability is found when a party knowingly induces or encourages infringement.[31] The Supreme Court has held that "one who distributes a device with the object of promoting its use to infringe copyright, as shown by clear expression or other affirmative steps taken to foster infringement, is liable for the resulting acts of infringement by third parties."[32] However, if the defendant can show that the device is capable of substantial non-infringing uses, it may be able to escape contributory liability.[33]

Once either direct or secondary infringement has been established, the court can continue to evaluate damages.

3. *Statutory Recovery for Copyright Infringement*

17 U.S.C. § 504(a) entitles copyright owners to recover statutory damages, as opposed to actual damages, in a copyright infringement suit.[34] The amount of statutory damages available is calculated based on three tiers of infringers: willful,

---

[29] *Contributory Infringement*, CORNELL L. SCH.: LEGAL INFO. INST. (August 2022), https://www.law.cornell.edu/wex/contributory_infringement [https://perma.cc/VZ9L-3HFF]; s*ee generally* COPYRIGHT LAW OF THE UNITED STATES, *supra* note 13.

[30] *See, e.g.*, Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd., 545 U.S. 913, 930 (2005) ("Although [t]he Copyright Act does not expressly render anyone liable for infringement committed by another, these doctrines of secondary liability emerged from common law principles and are well established in the law." (alterations in original) (internal quotations omitted) (internal citations omitted) (quoting Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 434 (1984))).

[31] *Grokster*, 545 U.S. at 927, 930.

[32] *Id.* at 919.

[33] *Sony Corp. of Am.*, 464 U.S. at 442 ("Accordingly, the sale of copying equipment, like the sale of other articles of commerce, does not constitute contributory infringement if the product is widely used for legitimate, unobjectionable purposes. Indeed, it need merely be capable of substantial noninfringing uses.").

[34] 17 U.S.C. § 504(a).

ordinary, and innocent.[35] Ordinary infringers are liable for damages between $750–$30,000 per work infringed.[36] If the copyright owner makes a sufficient showing of willful infringement, the owner may recover for up to $150,000 per work infringed.[37] Conversely, if the court finds that the infringer was not aware of the infringement, or in other words was an innocent infringer, statutory damages may be lowered to $200 per work infringed.[38] Willful infringement is found when the defendant knew or showed reckless disregard for the fact that their conduct was likely infringing.[39] Innocent infringement is found when the infringer proves that they were "not aware and had no reason to believe that his or her acts constituted an infringement of copyright."[40]

The court is free to determine what exact number is appropriate within those statutory ranges, meaning the actual amount for an ordinary infringer who has infringed one work can be anywhere between $750–$30,000.[41] Some factors that juries consider in deciding this number include the amount of actual damages, perceived good faith, and the applicability of a fair use defense.[42] Juries play a significant role in awarding statutory damages, allowing "the potential for excessive and arbitrary awards when skillful lawyers are able to persuade juries to become outraged about infringing conduct."[43] What this means, practically speaking, is that these awards are inconsistent.[44] This inconsistency is best illustrated by the case of *Capitol Records v. Thomas*, in which defendant Jammie Thomas illegally downloaded and distributed 24 songs owned by the recording

---

35 *Id.* § 504(c).

36 *Id.* § 504(c)(1).

37 *Id.* § 504(c)(2).

38 *Id.* § 504(c)(2).

39 *See, e.g.*, Island Software and Comput. Serv., Inc. v. Microsoft Corp., 413 F.3d 257, 263 (2d Cir. 2005); UMG Recordings, Inc. v. Disco Azteca Distribs., Inc., 446 F. Supp. 2d 1164, 1173 (E.D. Cal. 2006). Note that the analysis differs slightly between the circuits but generally includes the same knowledge or recklessness components. *See id.*

40 17 U.S.C. § 504(c)(2).

41 *Id.* § 504(c)(1).

42 *See* Samuelson & Wheatland, *supra* note 3, at 474–75.

43 *Id.* at 456.

44 *See id.*

company plaintiffs, where actual damages amounted to about $50.[45] Because there was no evidence of innocent infringement, the plaintiffs were entitled to at least $750 per work infringed, amounting to $18,000.[46] However, the jury stepped in at this point and awarded the plaintiffs $1.92 million for the infringement, amounting to 38,400 times the actual damages.[47]

Despite the sometimes-arbitrary nature of statutory damages awards, this recovery mechanism rests on the theory that actual damages can be challenging to prove. For copyright lawsuits in particular, actual damages, representing "the losses suffered by the copyright owner . . . [including] lost sales, lost profits, lost licensing revenue, or any other demonstrable monetary loss," often prove difficult to ascertain.[48] This difficulty can be for various reasons,[49] though the main consideration is that a copyright owner cannot know how many works it would have sold or how much revenue would have been earned through licensing

---

[45] *Id.*; *see also* Capitol Recs., Inc. v. Thomas, 579 F. Supp. 2d 1210, 1227 (D. Minn. 2008) ("[T]he damages awarded in this case are wholly disproportionate to the damages suffered by Plaintiffs. Thomas allegedly infringed on the copyrights of 24 songs—the equivalent of approximately three CDs, costing less than $54 . . . [Damages are] more than ***five hundred*** times the cost of buying 24 separate CDs and more than ***four thousand*** times the cost of three CDs.").

[46] Samuelson & Wheatland, *supra* note 3, at 456; *see also* 17 U.S.C. § 504(c)(1).

[47] Samuelson & Wheatland, *supra* note 3, at 456; *see also* Nate Anderson, *Thomas Verdict: Willful Infringement, $1.92 Million Penalty*, ARS TECHNICA: POLICY (June 18, 2009, at 4:32 PM), https://arstechnica.com/tech-policy/2009/06/jammie-thomas-retrial-verdict/ [https://perma.cc/L4LN-G2MP].

[48] Daniel H. Park, *Copyright Infringement: What Are the Potential Damages?*, BERMAN, FINK, VAN HORN: BLOG (Nov. 6, 2019), https://www.bfvlaw.com/copyright-infringement-what-are-the-potential-damages/ [https://perma.cc/WD7G-MKHN]; *see What Are Statutory Damages*, *supra* note 16 ("Statutory damages are important because the alternative type of damage award is 'actual damages,' which must be proven in court and can be very difficult to establish."); Daniel Berenger-Russell, Note, *The Water Under the Bridge is Darkening—An Analysis of Copyright Law and the Prevalence of Copyright Trolls*, 54 N.M. L. REV. 297, 310 (2024); Samuelson & Wheatland, *supra* note 3, at 499.

[49] Other reasons actual damages may be difficult to prove include that the copyright owner may work in a different market than the infringer or there may be insufficient records to establish actual sales of the infringed work. Samuelson & Wheatland, *supra* note 3, at 446 n.22.

revenue, but for the infringing work's presence in the market.[50] "Actual damages must be based on measurable numbers; an estimate of loss is simply not enough."[51] When designing the Copyright Act of 1909, which predated the current Copyright Act, legislators expressed concern about this very issue, opting to adopt statutory damages as an alternative.[52]

Statutory damages serve an essential function in copyright law, ensuring that authors are able to recover for infringement of their works. However, the fickle nature of their administration, as shown by the *Thomas* case, allows for aggressive litigation tactics to fall through the cracks, especially in areas involving newly developed technologies where the legal ramifications of the work involved are not fully fleshed out.

B. GENERATIVE AI & COPYRIGHT TROLLS

This Section will first explore how Generative AI works, describing the training process and the data used, and then proceed to how outputs are produced and what that means for copyright law. The Section will then turn to copyright trolls and dive into the history of these entities, then investigate how copyright trolls have reacted and may still react to the advent of Generative AI.

1. *What is Generative AI?*

"Generative AI [is] a machine-learning model that is trained to create new data, rather than making a prediction about a specific dataset."[53] This technology took the world by storm in November of 2022, when AI research organization

---

[50] *See* Richard Stim, *Copyright Infringement: How Are Damage Amounts Determined?*, NOLO, https://www.nolo.com/legal-encyclopedia/copyright-infringement-how-damages-determined.html [https://perma.cc/3XWH-6EYT].

[51] Park, *supra* note 47.

[52] Samuelson & Wheatland, *supra* note 3, at 446 n.22 ("The legislative history of the 1909 Act contains numerous expressions of concern about the difficulties of proving actual damages or defendant's profits as a rationale for the adoption of a statutory damage provision.").

[53] Adam Zewe, *Explained: Generative AI*, MASS. INST. OF TECH.: MIT NEWS (Nov. 9, 2023), https://news.mit.edu/2023/explained-generative-ai-1109 [https://perma.cc/6MTE-FW82].

OpenAI[54] released a demo version of ChatGPT for the public to use.[55] Attracting over a million users in the first five days of its release,[56] this software opened the door for many other companies to develop and release their own iterations, including Google's Gemini, Microsoft's Copilot, and more.[57] As of February 2025, about 45% of the United States population uses some form of Generative AI, and about 65% of all users are Millennials or Gen Z.[58] About 70% of Gen Z use Generative AI.[59] At the end of 2023, ChatGPT saw about 100 million users employ

---

[54] *OpenAI*, WIKIPEDIA, https://en.wikipedia.org/wiki/OpenAI [https://perma.cc/Q7G8-T9E6].

[55] Bernard Marr, *A Short History of ChatGPT: How We Got to Where We Are Today*, FORBES (May 19, 2023), https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/ [https://perma.cc/GEA3-VEVC].

[56] *Id.*

[57] Sundar Pichai & Demis Hassabis, *Introducing Gemini: Our Largest and Most Capable AI Model*, GOOGLE BLOG (Dec. 6, 2023), https://blog.google/technology/ai/google-gemini-ai/#sundar-note [https://perma.cc/E8E9-YCCZ]; *Meta Connect 2023: Quest 3, AI Advances, Next-Gen Smart Glasses, & the Road to the Metaverse*, META: BLOG (Sep. 27, 2023), https://www.meta.com/blog/quest/connect-2023-quest-3-ai-ray-ban-smart-glasses-metaverse/ [https://perma.cc/YP5E-GLBG]; Yusuf Mehdi, *Announcing Microsoft Copilot, Your Everyday AI Companion*, THE OFFICIAL MICROSOFT BLOG (Sep. 21, 2023), https://blogs.microsoft.com/blog/2023/09/21/announcing-microsoft-copilot-your-everyday-ai-companion/ [https://perma.cc/4Q3H-MHLE].

[58] *Top Generative AI Statistics for 2025*, SALESFORCE: NEWS (Feb. 2025), https://www.salesforce.com/news/stories/generative-ai-statistics/ [https://perma.cc/976W-5GB9]; *see also* Bergur Thormundsson, *ChatGPT Usage in the U.S. 2023, by Age Group*, STATISTA (Feb. 6, 2024), https://www-statista-com.us1.proxy.openathens.net/statistics/1368577/chatgpt-usage-age-groupr-us/ [https://perma.cc/CT7H-GDGB] (finding 15% of 18–29-year-olds and 17% of 30–44-year-olds had personally used ChatGPT to generate text, compared to 9% of 45–64-year-olds and 5% of 65+-year-olds).

[59] *Top Generative AI Statistics for 2025*, *supra* note 58; Bernard Marr, *10 Mind-Blowing Generative AI Stats Everyone Should Know About*, FORBES (Jan. 29, 2024), https://www.forbes.com/sites/bernardmarr/2024/01/29/10-mind-blowing-generative-ai-stats-everyone-should-know-about/ [https://perma.cc/C5EU-UGUS] [hereinafter *Mind-Blowing Generative AI Stats*].

the tool on a weekly basis.[60] Among the non-users, 70% said they would use Generative AI if they knew more about it.[61] Of those who use Generative AI, 22% use it for content creation and editing, 13% use it for creativity and recreation, and 10% use it for research, analysis, and decision making.[62] The worldwide market size for Generative AI was $138.36 billion in 2023[63] and was expected to generate $128 billion in revenue in 2024.[64] This number is expected to increase by $180 billion by 2032.[65] OpenAI, the creator of ChatGPT, was valued at $157 billion in October 2024.[66]

Generative AI is evidently an enormously profitable industry that has captured the minds of the younger generations.[67] Not only this, but a good portion of its use is directed at creative pursuits that involve generating writing and artwork.[68] Rather than being a temporary fad, all signs point toward the use of Generative AI only expanding as the systems become more complex, assuming that international regulation does not curtail growth. [69]

---

[60] *Mind-Blowing Generative AI Stats*, *supra* note 59.

[61] *Top Generative AI Statistics for 2025*, *supra* note 58.

[62] Marc Zao-Sanders, *How People Are Really Using GenAI*, HARV. BUS. REV. (Mar. 19, 2024), https://hbr.org/2024/03/how-people-are-really-using-genai [https://perma.cc/HTR6-7JDB].

[63] Bergur Thormundsson, *AI Market Size Worldwide From 2020-2031*, STATISTA (June 23, 2025), https://www.statista.com/forecasts/1474143/global-ai-market-size [https://perma.cc/A4GD-7P5Y].

[64] Bergur Thormundsson, *Worldwide Generative AI Revenue 2020-2032*, STATISTA (July 1, 2025), https://www-statista-com.us1.proxy.openathens.net/statistics/1417151/generative-ai-revenue-worldwide/ [https://perma.cc/RF82-H32R].

[65] *Mind-Blowing Generative AI Stats*, *supra* note 59.

[66] Antonio Pequeño IV, *OpenAI Valued at $157 Billion After Closing $6.6 Billion Funding Round*, FORBES (Dec. 23, 2024), https://www.forbes.com/sites/antoniopequenoiv/2024/10/02/openai-valued-at-157-billion-after-closing-66-billion-funding-round/ [https://perma.cc/W3VZ-HX2Q].

[67] *See Mind-Blowing Generative AI Stats*, *supra* note 59; *Top Generative AI Statistics for 2025*, *supra* note 58.

[68] *See* Zao-Sanders, *supra* note 62, at 3.

[69] *See, e.g.*, Pequeño IV, *supra* note 66. There may be rumblings of international regulation on the way soon. The European Union has acted quickly in this area, passing the EU AI Act in June 2024. *EU AI Act: First Regulation on*

a. How Generative AI works

The release of Generative AI represented a major leap forward in the technology available, owing in part to the unique aspects of the software itself as well as the sheer amount of data that is fed into these systems.[70] Generative AI differs from previously extant technology because earlier, non-generative AI models merely predict an outcome, rather than create something new.[71] All Generative AI models rely on the process of converting data into numerical representations, called tokens,[72] and with this data, use machine learning algorithms called neural networks to "identify the patterns and structures within existing data."[73] These neural networks convert inputs to outputs through the identification of patterns[74] and are unique in that they analyze the whole sequence of input at once, rather than word by word.[75]

The process through which the Generative AI learns to produce usable output is called "training," and encompasses doing one of several processes, depending on the model, over and over again until the output begins to resemble the input.[76] For example, in the case of written output, Generative AI will essentially guess what letter will come next when given a certain string of words,

Artificial Intelligence, EUR. PARLIAMENT: TOPICS (Feb. 19, 2025), https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence [https://perma.cc/HQ6P-6VPX]. This Act, set to go fully into effect 24 months after entry into force, would ban AI systems categorized as "high risk" and implements transparency requirements regarding the data used by general purpose AI systems. *Id.*

70 Zewe, *supra* note 53.

71 John T. Kivus, *Generative AI and Copyright Law: A Misalignment That Could Lead to the Privatization of Copyright Enforcement*, 25 N.C. J.L. & TECH. 447, 461 (2024).

72 Zewe, *supra* note 53.

73 *What is Generative AI and How Does it Work?*, NVIDIA, https://www.nvidia.com/en-us/glossary/generative-ai/ [https://perma.cc/8XKH-SDNB].

74 Aatish Bhatia, *Watch an A.I. Learn to Write by Reading Nothing but Jane Austen*, N.Y. TIMES (Apr. 27, 2023), https://www.nytimes.com/interactive/2023/04/26/upshot/gpt-from-scratch.html [https://perma.cc/6JZQ-VTJ4].

75 Kivus, *supra* note 71, at 464.

76 Bhatia, *supra* note 74.

and then compare its output to the input.[77] The Generative AI model is then further refined through a prompting process that ensures the outputs are constantly being evaluated and improved.[78] Notable here is the fact that generating output is dependent on the input—that is, if the input is infringing, then the output, which is trained to resemble the input, will be infringing in some capacity as well.[79]

There are three main Generative AI training models that simulate this process by pairing neural networks.[80] The diffusion model works by using one neural network to add random "noise," or additional extraneous data, to a sample, while another neural network reverses that process.[81] This model offers the highest-quality output and, likewise, takes the longest time to train.[82] A variational autoencoder ("VAE") uses one neural network to condense data samples into smaller packets, preserving the essential data for reconstruction, while the other neural network conducts that reconstruction.[83] This model, though similar to a diffusion model, produces less detailed output.[84] Finally, a generative adversarial network ("GAN") employs one neural network to generate new samples, while another distinguishes between "real" data and data generated by the other neural network.[85] This model works quickly and provides high-quality data, but it is less diverse in the types of data it can produce.[86]

The main innovation that allowed Generative AI to progress so quickly in the past few years is the sheer amount of data these models are trained on.[87] The

---

77 *Id.*

78 Kivus, *supra* note 71, at 466–67.

79 *See* Bhatia, *supra* note 74.

80 *What Is Generative AI*, *supra* note 73.

81 *Id.*

82 *Id.*

83 *Id.*

84 *Id.*

85 *Id.*

86 *What Is Generative AI*, *supra* note 73.

87 *See* Bhatia, *supra* note 74; Jared Kaplan et al., *Scaling Laws for Neural Language Models*, CORNELL UNIV.: arXiv (Jan. 23, 2020), https://arxiv.org/abs/2001.08361v1 [https://perma.cc/W5FZ-M6QW] ("Model performance depends most strongly on scale, which consists of three factors: the number of model parameters . . ., the size of the dataset . . ., and the amount of compute . . . used for training."); Kivus, *supra* note 71, at 465; Zewe, *supra* note 53.

more high-quality data a Generative AI is trained on, the better the output and the more spontaneous abilities, or functionalities that were not programmed into the Generative AI by people, the model can develop.[88] High-quality data is defined based on five characteristics: volume, velocity, variety, veracity, and privacy.[89] In the Generative AI context, each of these factors matter because they correlate directly to performance and, therefore, business value.[90] If the training data is not accurate, usable, or diverse enough, the output of the Generative AI will suffer from inaccuracies or a failure to represent the whole picture.[91] This deficient output is because, by its nature, Generative AI spits out only different combinations of the data it has been trained on, so if it has been trained on faulty data, it will render faulty results.[92]

b. Where Generative AI Interacts with Copyrighted Materials

Often, high-quality data comes in the form of works that have been through the editorial process, such as published books and articles.[93] Where the

---

88 Bhatia, *supra* note 74.

89 *See* Alex Watson, *Solving the Data Quality Problem in Generative AI*, INFOWORLD (June 11, 2024), https://www.infoworld.com/article/2337627/solving-the-data-quality-problem-in-generative-ai.html [https://perma.cc/H3RS-GV9K]. Volume refers to the amount of data; velocity refers to the "speed at which the data is created and how fast it moves"; variety is the diversity of the data; veracity is the accuracy and utility of the data; and privacy refers to the assurance that "the original data is not compromised." *See id.*; Scott Robinson, *5V's of Big Data*, TECH TARGET (Nov. 17, 2023), https://www.techtarget.com/searchdatamanagement/definition/5-Vs-of-big-data [https://perma.cc/Q8RS-PPDX].

90 *See* Watson, *supra* note 89.

91 *See id.*; *see also* Morgan Fluhler, *Why Data Quality Matters in the Age of Generative AI*, DATAIKU: BLOG (Mar. 19, 2024), https://blog.dataiku.com/why-data-quality-matters-in-the-age-of-generative-ai [https://perma.cc/TU5M-BRBD] ("Biases or errors present in your real data will inevitably be reflected in the synthetic data generated by Generative AI models.").

92 *See id.*

93 *See* Cade Metz et al., *How Tech Giants Cut Corners to Harvest Data for A.I.*, N.Y. TIMES (Apr. 6, 2024), https://www.nytimes.com/2024/04/06/technology/tech-giants-harvest-data-artificial-intelligence.html [https://perma.cc/C3W5-EF2S] [hereinafter *Tech Giants*] ("The most prized data, A.I. researchers said, is high-

copyright issues arise is the source of this training data. *The Washington Post* conducted an analysis on Google's C4 dataset, which was used to instruct several neural networks.[94] The newspaper found that most of the content was "from industries including journalism, entertainment, software development, medicine and content creation" and found the copyright symbol more than 200 million times.[95] While some of these uses may be legal, multiple sources have attested to the training process involving unlicensed copyrighted material.[96] It is no wonder that 89.2% of artists believe "copyright laws are inadequate in the age of generative AI."[97]

Several developers have gone on record asserting the necessity of copyrighted works for Generative AI development.[98] Meta, for example, went so

---

quality information, such as published books and articles, which have been carefully written and edited by professionals.").

94 Kevin Schaul et al., *Inside the Secret List of Websites That Make AI Like ChatGPT Sound Smart*, WASH. POST (Apr. 19, 2023), https://www.washingtonpost.com/technology/interactive/2023/ai-chatbot-learning/ [https://perma.cc/AQ8A-F7CH].

95 *Id.*; *see Tech Giants*, *supra* note 93 ("[O]nline information – news stories, fictional works, message board posts, Wikipedia articles, computer programs, photos, podcasts and movie clips – has increasingly become the lifeblood of the booming A.I. industry.").

96 *See, e.g.*, Cade Metz, *Former OpenAI Researcher Says the Company Broke Copyright Law*, N.Y. TIMES (Oct. 23, 2024), https://www.nytimes.com/2024/10/23/technology/openai-copyright-law.html [https://perma.cc/53N8-S8VP] [hereinafter *Company Broke Copyright Law*]; Suchir Balaji, *When Does Generative AI Qualify for Fair Use?*, SUCHIR.NET (Oct. 23, 2024), https://suchir.net/fair_use.html [https://perma.cc/L5GA-A8CD]; *see also* Kadrey v. Meta Platforms, Inc., No. 23-cv-03417-VC, 2025 WL 1752484, at 4 (N.D. Cal. June 25, 2025) ("The upshot is that in many circumstances it will be illegal to copy copyright-protected works to train generative AI models without permission. Which means that companies, to avoid liability for copyright infringement, will generally need to pay copyright holders for the right to use their materials.").

97 *Mind-Blowing Generative AI Stats*, *supra* note 59.

98 Benji Edwards, *OpenAI Says It's "Impossible" to Create Useful AI Models Without Copyrighted Material*, ARS TECHNICA (Jan. 9, 2024), https://arstechnica.com/information-technology/2024/01/openai-says-its-impossible-to-create-useful-ai-models-without-copyrighted-material/ [https://perma.cc/8FSB-H7M9] (explaining that it is impossible to develop Generative AI tools without copyrighted material); Horowitz, *supra* note 7

far as to discuss purchasing Simon & Schuster, one of the biggest literature publishers in the United States, to source more long-form data for training its Generative AI, while Google recently expanded its terms of service to allegedly allow the use of Google Docs information, among other written materials, in its Generative AI training set.[99] One former OpenAI researcher[100] explained that the developers believed they could train on any data, including unlicensed copyrighted material, in the early stages because it was a research project, as opposed to a commercial venture.[101] In one Comment to the Copyright Office, a venture capital firm affirmed the use of unlicensed copyrighted material, asserting that "imposing the cost of actual or potential copyright liability on the creators of AI models will either kill or significantly hamper their development."[102]

("[T]he only practical way generative AI models can exist is if they can be trained on an almost unimaginably massive amount of content, much of which (because of the ease with which copyright protection can be obtained) will be subject to copyright.").

99 *See Tech Giants*, *supra* note 93.

100 This researcher, Suchir Balaji, was found dead in an apartment due to an apparent suicide, mere months after his assertions about OpenAI's practices. *See* Alys Davies, *OpenAI Whistleblower Found Dead in San Francisco Apartment*, BBC (Dec. 14, 2024), https://www.bbc.com/news/articles/cd0el3r2nlko [https://perma.cc/B82Y-JECC].

101 *Company Broke Copyright Law*, *supra* note 96.

102 Horowitz, *supra* note 7. The comment focuses primarily on the use of unlicensed copyrighted material as a fair use, or deemed a non-infringing use based on an analysis of several factors, under the standard articulated in *Authors Guild v. Google*. *See id.* In that case, the Second Circuit held Google's unauthorized digitization of more than 20 million copyright-protected works for the purposes of bolstering their search tool and enabling more sophisticated text research to be a fair use. *See Authors Guild*, 804 F.3d at 229. This holding was based on the finding that Google's elevated textual search function was a transformative purpose because it allowed for greater information about the books themselves as well as the frequency of usage of certain words. *Id.* at 216–19. The court found it persuasive that Google did not make an unauthorized digital copy available to the public and that it did not affect the potential market for the books too drastically. *See id.* at 224–25. While this note will not explore the *Authors Guild* case, nor will it explore fair use, this case and doctrine are likely to play a prominent role in any upcoming Generative AI litigation and are therefore worth keeping an eye on. *See, e.g.*, Horowitz, *supra* note 7. For more about recent developments in this area, see Thomson Reuters Enter. Ctr. GMBH v. Ross Intel. Inc., 765 F.3d 382, 401 (D. Del. 2025) (rejecting the fair use defense in a closely related non-generative AI

Even assuming all data sets used to create Generative AI currently avoid unlicensed copyrighted material, researchers predict data sources will be depleted as soon as 2026, providing further motivation for the developers of Generative AI to infringe on copyrighted material.[103] This motivation presents a clear problem for authors of copyrighted works, but because of unique entities within the copyright system called copyright trolls, end-users of Generative AI may also end up on the metaphorical chopping block.

2. *Copyright Trolls*

The term "copyright troll" has been used to encompass several practices within copyright litigation, most often when plaintiffs are "more interested in gaining income through litigation, or rather the threat of litigation, than selling or licensing [their] work."[104] Trolls see a potential windfall and pursue litigation solely for this monetary gain, often without a strong claim.[105] For the purposes of this background, this paper will recognize a copyright troll when it: "(1) acquires a copyright—either through purchase or act of authorship—for the primary purpose of pursuing past, present, or future infringement actions; (2) compensates authors or creates works with an eye to the litigation value of the work, not the commercial value; […] and [(3)] uses the prospect of statutory damages and litigation expenses to extract quick settlements of often weak claims."[106] However,

---

case); Bartz v. Anthropic PBC, No. C 24-05417 WHA, 2025 WL 1741691, at 12 (N.D. Cal. June 23, 2025) (holding the use of legally obtained books to train Generative AI system to be fair because it was an exceedingly transformative use); *Kadrey*, 2025 WL 1752484, at 5 (finding the use of unlicensed copyrighted material to be fair based on evidentiary issues in the record, but noting that "this ruling does not stand for the proposition that Meta's use of copyrighted materials to train its language models is lawful. It stands only for the proposition that these plaintiffs made the wrong arguments and failed to develop a record in support of the right one"); *see generally* U.S. COPYRIGHT OFFICE, COPYRIGHT AND ARTIFICIAL INTELLIGENCE PART 3: GENERATIVE AI TRAINING (May 2025).

103 *See Tech Giants*, *supra* note 93.

104 Michael P. Goodyear, Note, *A Shield or a Solution: Confronting the New Copyright Troll Problem*, 21 TEX. REV. ENT. & SPORTS L. 77, 83 (2020).

105 *Id.* at 83–84.

106 Greenberg, *supra* note 1, at 59. The original quote includes a fourth condition that the troll "lacks a good faith licensing program," but this is not included in the definition used for this Note. *See id.*

this should not be confused with legitimate copyright enforcement, which simply upholds copyrights as defined by law.[107]

a. Distinguishing Copyright Enforcement from Trolling in the Early Days[108]

While copyright trolls have supposedly existed since the early 19th century,[109] the current iteration of copyright troll began to emerge in the early 2000s, as technology began to change rapidly and copyright law was not able to keep up.[110] Copyright trolls can be found in various areas of copyright, including music sampling and online news, but one of the most common areas is peer-to-peer ("P2P") file sharing.[111]

P2P file sharing refers to a form of internet piracy where users share bits of copyrighted material between their personal devices, making each consumer of this material an infringer.[112] Because each individual user infringes on the material, this opens up the door for copyright owners to sue individual infringers in

---

107 *Id.*

108 A closely related, but separate, phenomenon plagues the patent industry. Patent trolls acquire patents and use them to threaten or pursue litigation against alleged infringers, rather than foster innovation. *See, e.g.*, Patent Trolls, ELEC. FRONTIER FOUND., https://www.eff.org/issues/resources-patent-troll-victims [https://perma.cc/QF3D-33FV]. Often, the cases settle with the alleged infringers agreeing to pay a licensing fee, even if the patent was invalid or not actually infringed, to avoid the costs of litigation. *See id.* Patent trolls have been found to suck funds out of research and development and inhibit innovation by preventing investments in areas known for trolls. *See* Greenberg, supra note 1, at 76. However, patent trolls have significantly more freedom to litigate than copyright trolls. *See id.* at 80. This is because copyright recognizes the infringement defense of independent creation, *see id.*, and because the originality standard in copyright is significantly less stringent than the non-obviousness standard in patents. ROBERT BRAUNEIS & ROGER E. SCHECHTER, COPYRIGHT: A CONTEMPORARY APPROACH 172 (2d ed. 2017).

109 Connie Boutsikaris, *The Rise of Copyright Trolls*, DUNNER LAW (Mar. 28, 2017), https://dunnerlaw.com/the-rise-of-copyright-trolls/ [https://perma.cc/BQ6P-9MGH].

110 *See* Greenberg, *supra* note 1, at 64.

111 James DeBriyn, Note, *Shedding Light on Copyright Trolls: An Analysis of Mass Copyright Litigation in the Age of Statutory Damages*, 19 UCLA ENT. L. REV. 79, 86–92 (2012).

112 *See id.* at 92.

addition to the larger software applications that make this infringement possible on the basis of secondary liability.[113] While some early enforcement actions that, upon first blush, resembled copyright trolls were legitimate exercises of copyright, such as the Recording Industry Association of America ("RIAA")[114] cases,[115] eventually, P2P file sharing became a hotbed for copyright trolls.[116]

What differentiated the copyright enforcement conducted by the RIAA from the behavior of later copyright trolls is the motive of the RIAA. In the RIAA cases, which took place in the 1990s and 2000s, the RIAA first sent cease-and-desist letters to the operators of P2P file-sharing websites.[117] When this failed to enact meaningful change, the RIAA went after the operators of those websites, settling out of court for the operators to cease the activity, rather than accept the massive statutory damage awards on the table.[118] By going after the big players such as the website operators, rather than individual end-users, the RIAA demonstrated that it was interested in the actual problem at hand, namely the illegal downloading of music, rather than the possible monetary windfall.[119] Granted that the big players

---

113 *See id.*; *Grokster*, 545 U.S. at 930.

114 RIAA is a trade organization whose membership includes several hundred companies that work with sound recordings. *About RIAA*, RIAA, https://www.riaa.com/about-riaa/ [https://perma.cc/9C6W-BEJ6]. It works "to protect the intellectual property and First Amendment rights of artists and music labels." *Id.*

115 *See* DeBriyn, *supra* note 111, at 84–85; *see also Trade Group Efforts Against File Sharing*, WIKIPEDIA, https://en.wikipedia.org/wiki/Trade_group_efforts_against_file_sharing [https://perma.cc/QYQ9-JCQP].

116 *See, e.g.*, DeBriyn, *supra* note 111, at 91.

117 *Trade Group Efforts Against File Sharing*, *supra* note 115 (citing Don Jeffrey, *Downloading Songs Subject of RIAA Suit*, BILLBOARD, June 21, 1997, at 83).

118 *Trade Group Efforts Against File Sharing*, *supra* note 115; *see also* Peter Tschmuck, *The Music Industry's Fight Against Napster – Part 4: Napster's Slow Death*, MUSIC BUS. RSCH. (Feb. 17, 2015), https://musicbusinessresearch.wordpress.com/2015/02/17/the-music-industrys-fight-against-napster-part-4-napsters-slow-death/ [https://perma.cc/CH59-63AV] ("Napster . . . announced at press conference to pay US $1bn over five years from a future subscription service to the rights holders to get rid of the lawsuit. . . . RIAA demanded of Napster to stop copyright infringement, to stop delaying court decisions and to strengthen the efforts to build an effective filtering system.").

119 *See id.*

have deeper pockets than the individuals, targeting the hub of the activity is a more effective way of halting the activity than pursuing each individual infringer. Additionally, accepting a permanent injunction in place of a monetary award in some of these cases demonstrates a commitment to combating the infringement, rather than using copyrights to make easy money.[120] In the *Napster* case in particular, when Napster became bankrupt during the case, it began offering a legal download service that the RIAA later sanctioned, again demonstrating the RIAA's interest in the actual infringement, rather than maliciously taking down web applications.[121]

Only when the lawsuits against the website owners did not appear to minimize the rampant copyright infringement did the RIAA begin to target end-users.[122] Early in the process, the RIAA offered an amnesty program for past infringements, so long as past infringers refrained from the practice in the future, but individuals lost eligibility when a suit had been filed against them.[123] The RIAA offered settlements prior to identifying individuals at a rate of $750 per song infringed.[124] When the cases actually made it to the litigation stage, the RIAA sued

---

120 *See Trade Group Efforts Against File Sharing*, *supra* note 115; *see also* Jeffrey, *supra* note 117, at 83 ("'But this is not about collecting damages,' [then-associate director of anti-piracy at the RIAA] says. 'It is about sending a message that these sites are in violation of the law.'").

121 *See Trade Group Efforts Against File Sharing*, *supra* note 115; *see also* Rebecca Viksnins, *Napster 2.0*, CNET (Oct. 8, 2003), https://www.cnet.com/reviews/napster-2-0-preview/ [https://perma.cc/WXF4-5LS3].

122 *Trade Group Efforts Against File Sharing*, *supra* note 115; *see also RIAA v. The People: Five Years Later*, ELEC. FRONTIER FOUND. (Sept. 30, 2008), https://www.eff.org/wp/riaa-v-people-five-years-later [https://perma.cc/EE6N-RA9J].

123 *Clean Slate Program*, RECORDING INDUS. ASS'N OF AM. (2003), https://web.archive.org/web/20070516072653/http://www.riaa.com:80/issues/cleanSlate.asp [https://perma.cc/EN5D-SXEP]; *Trade Group Efforts Against File Sharing*, *supra* note 115.

124 *Trade Group Efforts Against File Sharing*, *supra* note 115; *see also* Elizabeth Lauten, *Recording Industry Battles Piracy*, THE E. CAROLINIAN (Apr. 4, 2007), http://web.archive.org/web/20070927030702/www.theeastcarolinian.com/home/index.cfm?event=displayArticle&ustory_id=4e74bd26-4c1b-4c88-8d84-d915653c9459 [https://perma.cc/Q2PZ-J5AH] ("Instead of threatening a lawsuit right away, the RIAA is now giving students 20 days to settle at a 'discount,' instead of going to court where the fees generally range from $3,000-$5,000.").

for $750 per song infringed, at the minimum of the ordinary infringer statutory damage amount.[125] After several years of this and seeing little progress in the issue they were attempting to solve, the RIAA switched tactics and ultimately worked towards cooperating with internet service providers instead to curb the infringement.[126]

In each instance, the RIAA demonstrated restraint in settlement offers and damages requests, keeping it to the minimum ordinary infringer amount, and changed tactics when the strategy proved ineffective at curbing the infringement.[127] Each action demonstrates a commitment to pursuing strategies that would enforce the copyrights, rather than prey on individual infringers to create a new stream of income.

This is compared to the infamous copyright troll Righthaven LLC ("Righthaven"). Rather than obtain a copyright and then pursue infringement, Righthaven would seek out instances of infringement online and, when it uncovered potentially infringing content, would ostensibly purchase the copyright from the relevant source.[128] This practice alone demonstrates a significant difference from the RIAA cases, pointing towards a motive for monetizing infringement as opposed to enforcing valid copyrights.[129] However, this purchasing was done in such a manner that it was not assured that Righthaven obtained the exclusive rights necessary to sue.[130] To add insult to injury, Righthaven would demand exorbitant sums under the statutory damages regime and pursue thin copyright claims against individual infringers as a regular

---

125 *Trade Group Efforts Against File Sharing*, *supra* note 115 (citing Eric Bangeman, *Judge: RIAA Damages Too High in Innocent Infringement Case*, ARS TECHNICA (Aug. 10, 2008), https://arstechnica.com/tech-policy/2008/08/judge-riaa-damages-too-high-in-innocent-infringement-case/ [https://perma.cc/F5SE-UU3S]); 17 U.S.C. § 504(c).

126 *See Trade Group Efforts Against File Sharing*, *supra* note 115 (citing Sarah McBride & Ethan Smith, *Music Industry to Abandon Mass Suits*, WALL ST. J. (Dec. 19, 2008), https://web.archive.org/web/20170301180415/https://www.wsj.com/articles/SB122966038836021137 [https://perma.cc/HY3Y-EZ45]).

127 *See Trade Group Efforts Against File Sharing*, *supra* note 115.

128 Marcella Ballard, *Copyright Troll Righthaven Kicked Out of Court Again*, VENABLE (Aug. 5, 2011), https://www.venable.com/insights/publications/2011/08/copyright-troll-righthaven-kicked-out-of-court-aga [https://perma.cc/5LHT-DFTJ].

129 *See id.*

130 *See id.*

practice.[131] Righthaven prioritized speed and aggression in obtaining favorable settlements, at the expense of users who were left to pay out often debilitating settlement demands.[132] Interestingly, the eventual downfall of Righthaven's notorious tactics was due to several judges finding a lack of standing.[133] What the Righthaven cases demonstrate is the heart of troll behavior: using aggressive litigation tactics to monetize infringement.

While both copyright enforcers and trolls ultimately aim to settle, the RIAA, as an enforcer, demonstrated a legitimate belief that those being sued had actually infringed on their copyrights and sought to end the rampant illegal downloading.[134] Trolls like Righthaven, on the other hand, "target massive groups of potential defendants, some backed by valid claims, some not, sending threatening letters to scare the alleged copyright infringers, who, overwhelmed by the legal jargon and potential damages, often settle for an arbitrary sum of money."[135] These entities do not necessarily have a reasonable belief that there is infringement, nor is the litigation designed to halt it, but rather to produce a new source of revenue.[136]

In the specific arena of internet infringement, trolls will often file suit without any idea of who specifically has infringed to compel exact identities and contact information via the court system.[137] Obtaining this information allows trolls to then send threatening demand letters directly to the alleged infringers.[138] Here, the statutory damages scheme comes into play, allowing trolls to threaten the statutory maximum and scare potential defendants into settling quickly for a smaller sum.[139]

This process can be better explained through the U.S. Copyright Group ("USCG") cases, which involved troll-like behavior against individual users of P2P file-sharing programs who were obtaining unauthorized copies of several films.

---

131 *See id.*

132 *See id.*

133 *See Righthaven*, WIKIPEDIA, https://en.wikipedia.org/wiki/Righthaven [https://perma.cc/RU79-7A76].

134 *See* DeBriyn, *supra* note 111, at 84–85.

135 Goodyear, *supra* note 104, at 79.

136 *See* Ballard, *supra* note 128.

137 *See, e.g.*, DeBriyn, *supra* note 111, at 94. These suits are usually filed identifying IP addresses instead of named defendants. *See id.*

138 *See id.* at 98.

139 *See id.* at 99.

USCG was a law firm that, in early 2010, filed several copyright infringement lawsuits against thousands of users of the P2P file-sharing program BitTorrent.[140] In the complaints, USCG did not identify any specific plaintiffs, but sued thousands of John Does.[141] Following the filing of the complaints, USCG sought to subpoena the internet service providers for the identities of these individuals.[142] After this, USCG would theoretically be able to sue the actual named individuals; however, USCG instead would send these named individuals demand letters.[143] These demand letters would leverage the willful infringement statutory maximum, $150,000 per work infringed, to "pressure the alleged infringers to settle quickly for $1,500 to $2,500."[144] Sometimes, USCG would even offer an initial settlement amount of $1,500 and then bump that sum up to $2,500 if the money was not paid within three weeks, before threatening litigation and the statutory damage maximum of $150,000.[145]

The RIAA cases were often in similar situations as the USCG cases, and the RIAA often employed identical tactics to identify infringers, but the salient difference, as discussed above, is the indication of a motive to actually enforce copyright.[146] Here, while $2,500 may seem like a small sum compared to $150,000, when that number is shared across thousands of defendants, it adds up to a massive windfall. In fact, at the time the suits were transpiring, USCG was projected to make $25 million in settlements for copyright infringement of *The Hurt Locker*.[147] This number represents almost $10 million more than the box office revenue, meaning the trolling generated more money than the creative work itself.[148] That is exactly the quantity over quality strategy employed by copyright

---

140 *USCG v. The People*, *supra* note 4.

141 *Id.*

142 *Id.*

143 *See id.*

144 *Id.*

145 Ernesto Van der Sar, *Piracy Will Earn Hurt Locker More Than the Box Office*, TORRENTFREAK (May 30, 2010), https://torrentfreak.com/piracy-will-earn-hurt-locker-more-than-the-box-office-100530/ [https://perma.cc/A9TT-EHCZ].

146 *See supra* notes 134–136 and accompanying text.

147 Van der Sar, *supra* note 145.

148 *See id.*

trolls and, again, demonstrates the interest in monetizing infringement, as opposed to actually working to stamp out that infringement.

b. The Interaction of Copyright Trolls and Generative AI

Within the theatre of Generative AI-created work, the legal landscape is still developing. A seminal difference between the old era of cases discussed above versus the new era of cases that may arise with the advent of Generative AI lies in the fact that the most popular software applications that facilitated P2P file sharing, such as Napster or Grokster, are bankrupt or have converted to other models now.[149] Generative AI, however, is massively profitable and is, at this point, so ingrained in the fabric of everyday life that bankrupting the companies would be disruptive to society.[150] The P2P file-sharing issue, at least with regard to music, has, for the most part, disappeared with time, as legitimate services have become the main forum for listening to music.[151] The Generative AI problem, however, is unlikely to do the same.[152]

Additionally, Napster was found to know that its conduct was infringing when the court found it liable for contributory infringement.[153] End-users of Generative AI, however, are not necessarily informed about how the training process works.[154] There is a greater imbalance in the relationship between end-

[149] *See Napster*, WIKIPEDIA, https://en.wikipedia.org/wiki/Napster [https://perma.cc/2WM5-LUAZ]; *Grokster*, WIKIPEDIA, https://en.wikipedia.org/wiki/Grokster [https://perma.cc/Y9GR-CEB7].

[150] *See supra* Section II.B.1.

[151] Gebre Waddell, *From Napster to the Cloud: The Evolution of Music File Sharing in the Music Industry*, SOUND CREDIT (June 9, 2025), https://www.soundcredit.com/blog/the-music-industrys-file-sharing-odyssey-from-piracy-to-ethical-collaboration [https://perma.cc/S85S-778U].

[152] *See supra* Section II.B.1.

[153] Michael Brick, *Court Rules Napster Users Infringe on Copyrights*, N. Y. TIMES (Feb. 12, 2001), https://www.nytimes.com/2001/02/12/technology/court-rules-napster-users-infringe-on-copyrights.html [https://perma.cc/QN6V-5PMG].

[154] *See Proactive Risk Management in Generative AI*, DELOITTE, https://www.deloitte.com/content/dam/assets-zone3/us/en/docs/services/consulting/2024/us-ai-institute-responsible-use-of-generative-ai.pdf [https://perma.cc/Z5ZZ-CMMB] ("End users can include people who have limited understanding of AI generally, much less the complicated workings of large language models.").

users of Generative AI and potential copyright trolls created by this lack of knowledge.

Early copyright cases involving Generative AI-related parties seem to be targeting the developers of this technology, rather than individual end-users.[155] In these cases, the plaintiffs are individual artists, ranging from authors to reporters, whose work has allegedly been infringed in the Generative AI training process without authorization.[156] While the parties targeted are not the traditional targets of trolls, this does not exclude these parties from troll-like behavior due to the nature of the claims and the likelihood that a fair use defense may apply.[157]

Even with the statutory minimum for an ordinary infringer set at $750 per work infringed,[158] a Generative AI developer could be liable for millions due to the sheer amount of data copied and reproduced in the training process.[159] Given that Generative AI improves the more data it acquires,[160] and many creators and developers are likely not engaging in fully legal practices when it comes to

---

155 *See, e.g.*, Michael M. Grynbaum & Ryan Mac, *The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work*, N. Y. TIMES (Dec. 27, 2023), https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html [https://perma.cc/DR6Y-HTFW]; *Lawsuit Seeks to Hold OpenAI and Microsoft Liable for "Rampant Theft" of Authors' Works*, SUSMAN GODFREY (Nov. 21, 2023), https://www.susmangodfrey.com/news/lawsuit-seeks-to-hold-openai-and-microsoft-liable-for-rampant-theft-of-authors-works/ [https://perma.cc/R5LY-VCRZ] [hereinafter *Lawsuit Seeks to Hold OpenAI Liable*].

156 *See id.*

157 *See Authors Guild*, 804 F.3d at 229 (finding Google's copying of the entirety of thousands of books a fair use because Google did not make the copy available to the public and used the data for the transformative use of textual analysis). This paper will not explore the fair use defense.

158 17 U.S.C. § 504(c)(1).

159 *See, e.g.*, Bhatia, *supra* note 74; *but see Kadrey*, 2025 WL 1752484, at *3–4 ("Another argument offered in support of the companies is more rhetorical than legal: Don't rule against them, or you'll stop the development of this groundbreaking technology. The technology is certainly groundbreaking. But the suggestion that adverse copyright rulings would stop this technology in its tracks is ridiculous. These products are expected to generate billions, even trillions, of dollars for the companies that are developing them. If using copyrighted works to train the models is as necessary as the companies say, they will figure out a way to compensate copyright holders for it.").

160 *See id.*

licensing copyrights,[161] many artists are experiencing textbook copyright infringement.[162]

Lawyers could flock to these lawsuits in the hopes of monetizing the infringement. So far, most of the Generative AI copyright infringement cases have been filed by a handful of lawyers.[163] Though each has artists attached to the suit who allege genuine harm, the fact that few lawyers file these suits suggests these lawyers are soliciting plaintiffs, rather than the other way around.[164] This practice appears to mimic the behavior exhibited by Righthaven, which looked proactively for online sites that were infringing copyrighted works.[165] This suggests that the lawyers may be offensively looking for nominal plaintiffs who will allow them to file suit and partake in the enormous windfall offered by statutory damages, rather than seeking to genuinely curb an infringing practice. These actions align the Generative AI lawyers closer with copyright trolls than copyright enforcers.[166]

This practice, however, is not as concerning, considering the Generative AI companies have exhibited some knowledge of infringing practices.[167] This is not to mention the fact that the technology giants behind these infringing practices have the financial backing to defend their interests in court[168]—a notable difference from the usual victims of copyright trolls.

---

161 *See supra* Section II.B.1.b.

162 *See* 17 U.S.C. § 501.

163 *See* Grynbaum & Mac, *supra* note 155 (describing a suit filed by Susman Godfrey attorneys); *Lawsuit Seeks to Hold OpenAI Liable*, *supra* note 157 (describing a separate lawsuit filed by Susman Godfrey); Joseph Saveri & Matthew Butterick, *We've Filed Lawsuits Challenging ChatGPT, LLaMA, and Other Language Models for Violating the Legal Rights of Authors. Because AI Needs to Be Fair & Ethical for Everyone*, LLM LITIGATION, https://llmlitigation.com/ [https://perma.cc/4ANN-7ZTP] (summarizing Joseph Saveri and Matthew Butterick's seven ongoing AI class action lawsuits).

164 *See id.*

165 *See* Ballard, *supra* note 128.

166 *See* Goodyear, *supra* note 104, at 83.

167 *See supra* Section II.B.1.b.

168 *See supra* Section II.B.1.

Where this paper is primarily concerned is with the individual end-users[169] who could be liable for infringement as well.[170] In fact, some Generative AI work may already be caught by trolls known for using reverse image searching.[171] Because most output of Generative AI is likely inherently infringing,[172] copyright trolls have a colorable case against essentially any user who generates an output. If the output renders an image that is substantially similar to a copyrighted work used to train the Generative AI model,[173] then by prompting Generative AI to produce images, the end-user presses the figurative volitional "button" in

[169] Trolls generally prefer to go after deeper pockets than individual infringers, where the creators of Generative AI are the most likely targets; however, the P2P file sharing infringement actions suggest that end-users may be the next target for trolls on a quantity over quality basis. *See* Kivus, *supra* note 71, at 486 n.207.

[170] *Id.* ("There are other legal issues related to Generative AI that are tangentially related to copyright law. For example, if Generative AI memorization creates an image for an end-user that is substantially similar to a copyrighted work, how is liability between the model and the end-user determined under copyright's strict liability regime? Though it is traditionally understood that plaintiffs like to go after parties with deep pockets, the Author submits that some copyright plaintiffs prefer a 'quantity over quality' strategy regarding infringement.").

[171] *See* Jeffrey M. Allen & Ashley Hallene, *Are Copyright Trolls Hunting Your Website?*, AM. BAR ASSOC. (May 25, 2021), https://www.americanbar.org/groups/gpsolo/resources/ereport/archive/are-copyright-trolls-hunting-your-website/ [https://perma.cc/J6R5-S6BV] (explaining that copyright trolls can use reverse image search engines to identify copyright infringements).

[172] *See supra* Section II.B.1.b.

[173] Where going after the individual infringers may be complicated is with their potential liability for the training process. While users can likely be held liable for their conduct in potentially prompting a Generative AI model to create an infringing work, they likely cannot be held liable for the input used to train that Generative AI model if their prompting is non-specific, due to the substantial similarity test. *See, e.g.*, *Tanksley*, 902 F.3d at 171. However, because of the reality that trolls often go after thin or non-existent copyright claims, this does not necessarily prevent the threatened practice from happening. *See* Goodyear, *supra* note 104, at 84. Even a meritorious defense to a copyright troll may not help in this situation because settling is so much less expensive than defending a claim in court. *See Dealing with a Copyright Troll*, THE ROTHMAN L. FIRM: BLOG (Apr. 27), https://rothman.law/blog/dealing-with-a-copyright-troll [https://perma.cc/5FW3-3XVX].

copyright infringement.[174] This volitional element becomes especially apparent when users engage in practices where they explicitly prompt for infringing content—for example, asking for art in the style of a particular artist,[175] asking for a particular article on a particular subject,[176] or prompting AI to generate content including a certain character.[177] The output need not necessarily be substantially similar to the copyrighted work to invite the attention of a copyright troll; due to the preference for quantity over quality in copyright troll litigation,[178] many pursue thin or nonexistent claims against individual infringers.[179] Creators of Generative AI have foreseen this practice, attesting to the potential problems it may present in this arena, with some expressly offering to indemnify any user of their model in the case they are sued for copyright infringement by a third party.[180]

---

174 *See* Cartoon Network, LP v. CSC Holdings, Inc., 536 F.3d 121, 131 (2d Cir. 2008) ("In the case of a VCR, it seems clear—and we know of no case holding otherwise—that the operator of the VCR, the person who actually presses the button to make the recording, supplies the necessary element of volition, not the person who manufactures, maintains, or, if distinct from the operator, owns the machine.").

175 *See, e.g.*, Kashmir Hill, *This Tool Could Protect Artists From A.I.-Generated Art That Steals Their Style*, N.Y. TIMES (Feb. 17, 2023), https://www.nytimes.com/2023/02/13/technology/ai-art-generator-lensa-stable-diffusion.html [https://perma.cc/5ZS3-TCGG] (describing how artist Greg Rutkowski's name was used to generate art within his style).

176 *See, e.g.*, Grynbaum & Mac, *supra* note 155 (explaining that users can prompt about current events and generate answers that not only rely on journalism by *The New York Times*, but sometimes provide near-verbatim excerpts from articles).

177 *See, e.g.*, Lakshmi Varanasi, *Generative AI Models Can Now Create Replicas of Trademarked Characters with Prompts as Simple as 'Videogame Italian,'* BUSINESS INSIDER (Jan. 7, 2024), https://www.businessinsider.com/generative-ai-models-trademarked-characters-images-study-text-phrase-prompts-2024-1 [https://perma.cc/K4DZ-83FH] (describing how prompts such as "black armor with light sword" result in generative AI producing characters that share likeness to those from Star Wars).

178 *See supra* Section II.B.2.a.

179 Goodyear, *supra* note 104, at 84 ("The dubious nature of a claim does not deter an opportunistic copyright troll from taking advantage of the high potential damages in copyright litigation to achieve a lucrative, quick settlement.").

180 *See* Brad Smith & Hossein Nowbar, *Microsoft Announces New Copilot Copyright Commitment for Customers*, MICROSOFT: BLOGS (Jan. 5, 2024), https://blogs.microsoft.com/on-the-issues/2023/09/07/copilot-copyright-

The potential copyright trolling activity could be disastrous in a community where Generative AI is becoming more ingrained in the everyday fabric of life. The analysis will explore how the statutory damages scheme can be modified to mitigate or potentially end the practice of copyright trolling with regard to Generative AI.

III. ANALYSIS

This Part will first look at the overall motivations of copyright trolls and why those motivations point towards a solution in the statutory damages scheme. Next, this Part will address two potential, but ultimately ineffective, solutions: total elimination of the statutory damages scheme and further articulated guidelines for this scheme. Finally, this Part will argue that the solution to the copyright troll problem is to apply a presumption of innocent infringement to Generative AI cases, setting the default statutory damages number at $200 per work infringed.

A. REWARD STRUCTURE FOR COPYRIGHT TROLLS

When looking at the problem presented by copyright trolls, the motivation behind their conduct appears to be the low-risk, high-reward nature of copyright infringement lawsuits against individual infringers.[181] The natural solution, therefore, is to eliminate that low-risk, high-reward nature by either increasing the risk associated with copyright lawsuits or minimizing the reward for frivolous actions.[182] The low risk is offered by the relative freedom exhibited by copyright trolls in past cases to track down and/or sue hundreds of alleged infringers and then settle quickly, without protracted litigation.[183] However, increasing the risk is unlikely to serve the purposes of copyright law, which is namely to foster

commitment-ai-legal-concerns/ [https://perma.cc/E3MY-LS3F] ("[I]f a third party sues a commercial customer for copyright infringement for using Microsoft's Copilots or the output they generate, we will defend the customer and pay the amount of any adverse judgment or settlements that result from the lawsuit."); *see also* Grynbaum & Mac, *supra* note 155 (explaining that Microsoft would indemnify and cover legal costs of customers of its AI products).

181 Goodyear, *supra* note 104, at 81 ("[The underlying causes of trolling are] the lack of risk for copyright trolls and the high potential cost for alleged copyright infringers.").

182 *See id.*

183 *See id.* at 84.

creativity and innovation, so the solution must come from lowering the reward.[184] The main reason trolling has such a high reward is because of the statutory damages scheme, which allows a plaintiff to recover at minimum $200 per work infringed[185] or, owing to the rarity of innocent infringement being recognized,[186] at least $750 per work infringed.[187] In other words, copyright trolls have access to a high reward because, should the case continue to suit, they are almost guaranteed at least $750 per work infringed, and likely more, because very few courts have chosen to recognize the innocent infringement defense.[188]

In law review articles addressing the broader problem of copyright trolls in general, statutory damages are the predominant arena in which solutions[189] are found.[190] Many who have weighed in suggest the elimination of statutory damages,[191] while fewer recommend some further guidelines for how to afford statutory damages.[192] For the reasons stated below, this Note will argue that the best way to combat copyright trolls is to apply a rebuttable presumption of innocent infringement in cases involving Generative AI.

---

184 *See What is the Purpose of Copyright Law*, *supra* note 12.

185 17 U.S.C. § 504(c)(2).

186 *See, e.g.*, Samuelson & Wheatland, *supra* note 3, at 456 (explaining that in *Capital Records v. Thomas-Rasset*, the "jury had no choice, given the implausibility of an innocent infringement defense, but to award [] at least $750").

187 17 U.S.C. § 504(c)(1).

188 Samuelson & Wheatland, *supra* note 3, at 456.

189 Some articles suggest the fair use defense, but as that is a facts-based defense it would require the defendant to go through litigation, which targets of copyright trolls are generally trying to avo*id. See, e.g.*, *U.S. Copyright Office Fair Use Index*, COPYRIGHT.GOV (Aug. 2025), https://www.copyright.gov/fair-use/ [https://perma.cc/C9AE-GM6P] ("Courts evaluate fair use claims on a case-by-case basis, and the outcome of any given case depends on a fact-specific inquiry.").

190 *See, e.g.*, DeBriyn, *supra* note 111, at 106–10.

191 *See id.*

192 *See, e.g.*, Samuelson & Wheatland, *supra* note 3, at 500–05.

B. THE PITFALLS OF THE ELIMINATION OF STATUTORY DAMAGES

Eliminating statutory damages would effectively fix the problem of copyright trolls.[193] Through total elimination of the potential windfall represented by the statutory damages maximum of $150,000 per work infringed, there would be little reason to pursue litigation for the sole purpose of damages.[194] Instead, copyright owners would have to sue for actual damages and make a genuine showing of the monetary loss experienced because of the infringement. Copyright trolls, who often pursue thin or nonexistent claims, would likely be unable to show actual damages anywhere near the statutory maximum.[195] However, therein lies the problem with the complete elimination of statutory damages.

As explored above, actual damages often prove hard to ascertain in copyright cases.[196] Relying solely on actual damages would therefore leave most copyright owners without a feasible chance to recover, and owing to the expensive nature of litigation, effectively no legal recourse for copyright infringement. Additionally, copyright is a field in which "actual damages suffered by any one individual may be so small that the law would be persistently underenforced in the absence of a statutory damage remedy."[197] Practically speaking, statutory damages are a way to ensure that copyright owners can justify the cost of litigation when their work is being infringed upon, even if the infringement itself may not be commercially viable.[198] Removing one of the main avenues through which authors can recover from infringement of their copyrights would disincentivize authors from sharing their work by preventing them from protecting their work. In doing so, the removal of statutory damages would work directly against the purposes of copyright law by preventing the dissemination of art.[199]

In addition, the Copyright Office has shown a reluctance to do away with statutory damages entirely, despite the pitfalls of the system.[200] Enacting any

---

193 DeBriyn, *supra* note 111, at 108.

194 *See id.*

195 *See* Goodyear, *supra* note 104, at 84.

196 *See supra* Section II.a.629.

197 Samuelson & Wheatland, *supra* note 3, at 492–93.

198 *See id.* (implying that actual damages may be so small such that it would not justify the cost of litigation).

199 *See What is the Purpose of Copyright Law*, *supra* note 12.

200 Matthew Sag & Jake Haskell, *Defense Against the Dark Arts of Copyright Trolling*, 103 IOWA L. REV. 571, 637 (2018) ("The Copyright Office has

legislative change is a process that can take years, but enacting one facing direct opposition would be untenable.

Proponents of eliminating statutory damages may argue that actual damages are easier to prove today than they were in the days the statutory damages regime was developed,[201] and therefore, the argument underlying statutory damages no longer applies. However, the problem with proving actual damages has not disappeared simply because some of the evidentiary issues are mitigated or solved.[202] The central issue of proving what could have been sold in the absence of a particular infringing product is still as relevant as it was in 1909, and eliminating statutory damages is therefore an ineffective solution to the copyright troll problem.[203] The change must come from within the statutory damages regime, so this paper turns next to the merits of keeping the statutory damages regime and, instead of wholesale elimination, creating further guidelines on how to apportion statutory damages.

C. FURTHER GUIDELINES FOR HOW TO ADMINISTER STATUTORY DAMAGES

The next solution, besides the complete elimination of statutory damages, is to draw up further judicial guidelines for how to administer those statutory damages.[204] As proposed, these edits would further delineate when it is appropriate for courts to award damages on the lower and upper ends of the spectrum, ideally by applying a series of factors to determine willfulness and the extent to which the damages are compensatory or purely punitive.[205]

While, in conception, this idea would help mitigate the arbitrary damages that are often found in cases involving statutory damages, it would ultimately

---

traditionally been an enthusiastic proponent of statutory damages and has resisted calls for their reform.").

201 *See* Greenberg, *supra* note 1, at 122 ("[I]t likely has become easier to prove actual damages following the data revolution."); Samuelson & Wheatland, *supra* note 3, at 496 ("In the overwhelming majority of copyright cases, the harm to the rights holder (such as lost license fees) and any unjust enrichment to the defendant attributable to infringement are reasonably discernible.").

202 *See* Greenberg, *supra* note 1, at 122.

203 *See* Stim, *supra* note 50.

204 *See* Samuelson & Wheatland, *supra* note 3, at 500–05.

205 *See, e.g.*, *id.* at 501–08.

suffer for the same reasons the current regime is ineffective.[206] While guidelines may potentially help juries have a better measuring stick by which to determine what is appropriate, ultimately, they are not expected or even able to measure those damages against previous cases, nor are they experts in what is or is not appropriate in the grand scheme of statutory damages.[207]

However, even if juries were able to seamlessly implement a series of guidelines with regard to how statutory damages are to be apportioned, the maximum penalty of $150,000 per work infringed is still on the table for an alleged infringer in this solution.[208] There is no effective bar on that large sum, meaning copyright trolls could still, with a semblance of legal footing, threaten alleged infringers with that amount.[209] Because the statutory maximum is still introduced at the beginning of the process when the initial demand letters are sent to potential defendants, the pressure to settle is just as great for end-users under the guidelines regime as it is for end-users of the current regime. In short, to fully benefit from the judicial guidelines, potential infringers would have to roll the dice and pursue litigation, hoping that a jury would adhere to those guidelines in a fair manner. This solution does not seem best poised to stop copyright trolls from the beginning, but rather provides a check on actions in the event that a potential defendant can muster up the funds to stage a defense.

The solution, therefore, lies somewhere between the wholesale elimination and better regulation of the apportionment of statutory damages. It must ostensibly take the maximum statutory damages amount off the table at the early stages, such that copyright trolls cannot threaten the number and pressure end-users into settling, but it cannot wholly eliminate that pathway for genuinely willful infringers.

D. Rebuttable Presumption of Innocent Infringement

Courts should apply a common law rebuttable presumption of innocent infringement to any case involving Generative AI. What this would mean on a practical level is that, should a plaintiff in a case involving Generative AI work elect to recover statutory damages, these damages would be set at the $200

---

206 *See id.* at 509.

207 *See id.* at 510 (explaining that lack of principles in guiding juries have often resulted in arbitrary, inconsistent, and even excessive awards for statutory damages).

208 *See* 17 U.S.C. § 504(c).

209 *See id.*

minimum per work infringed[210], until ordinary or willful infringement is proven. Because plaintiffs may not always know that the work is sourced from Generative AI, defendants would have to offer proof of the source of the work to benefit from the presumption.

Statutory damages are set up in a tripartite structure that classifies an infringer as either innocent, ordinary, or willful, with corresponding tiers of damages.[211] However, the innocent infringement tier, which ranges from $200-$30,000 per work infringed, is almost never used in practice.[212] The statute itself includes innocent infringement as a tier that requires a showing of proof of a lack of awareness of infringement, but in practice, this showing is so burdensome as to render the tier effectively useless.[213] Rather, the court begins damages discussions at the ordinary infringement tier, and therefore, the minimum statutory damage amount is, in practice, the $750 minimum for ordinary infringers.[214]

By applying a rebuttable presumption of innocent infringement, any potential plaintiff would have to evaluate whether they could put forth sufficient evidence of willfulness to rebut the presumption and, if not, whether filing suit would be worth the small sum of $200 per work infringed. In the case of copyright trolls, the answer to the first question would likely be "no" because of the thin nature of the claims that copyright trolls pursue.[215] The answer to the second would be no as well because at that low a sum, trolls would lose money through the costs of litigation.[216] This is not to mention the presumption effectively hamstrings the willful maximum by requiring a rebuttal first of innocent infringement and a further showing of willfulness. Without any monetary

---

210 *See id.*

211 *See* 17 U.S.C. § 504; *see also* Samuelson & Wheatland, *supra* note 3, at 444–45.

212 Samuelson & Wheatland, *supra* note 3, at 453–54 ("However, th[e innocent infringement] part of the statutory damage framework has virtually no significance in litigation, not even in the fair use context.").

213 *See* 17 U.S.C. § 504(c)(2); Samuelson & Wheatland, *supra* note 3, at 474–75 (stating that the authors could only find two cases where statutory damages were lower than the ordinary infringer standard).

214 *See* 17 U.S.C. § 504(c); Samuelson & Wheatland, *supra* note 3, at 453–54.

215 *See* Goodyear, *supra* note 104, at 84.

216 *See Copyright Litigation 101*, THOMSON REUTERS LEGAL: BLOG (Dec. 16, 2022), https://legal.thomsonreuters.com/blog/copyright-litigation-101/ [https://perma.cc/D7S8-X9VT] ("[T]he average cost of litigating a copyright infringement case in federal court from pre-trial through the appeals process is $278,000.").

windfall at the end of the litigation tunnel, copyright trolls would be discouraged from pursuing frivolous suits against end-users of Generative AI.[217]

This presumption would be rebuttable with a showing of the opposite of what is currently required to show innocent infringement: an awareness of infringement or a lack of reasonable belief that the conduct was non-infringing.[218] Upon a showing of awareness or lack of reasonableness, the damages would be increased to the ordinary infringement level with a minimum of $750 per work infringed or to the willful infringement level if appropriate.[219] If innocent infringement is truly as rare as the case law would suggest,[220] then the presumption should be rejected in virtually every case and there is no harm done. This is particularly true in the case of the Generative AI creators, who, as described above, have likely knowingly engaged in some form of copyright infringement during the training process—they are not shielded by the innocent infringement presumption.[221]

While statutory change is preferable as a permanent solution, lobbying for any legislative change can take years, if not decades. In this area in particular, where the statutory damages regime is favored by the Copyright Office, it may be even more difficult to lobby for change.[222] However, because Generative AI is a

---

217 Copyright trolls do not usually proceed to a protracted trial, however, and rely on scary demand letters and quick settlements. *See* DeBriyn, *supra* note 111, at 98–99. These demand letters usually include the statutory maximum as a scare tacti and offer a lower (but still egregious) sum as a settlement, which recipients often take. *See id.* Without total elimination of statutory damages, copyright trolls could still feasibly use this statutory maximum. *See* 17 U.S.C. § 504. However, with a presumption of innocent infringement, any potential recipient of such a letter would be able to dispel that fear with a quick Google search and evaluate the situation with a more level head. A presumption of innocent infringement would mean cases likely would not reach the often egregious and arbitrary levels of statutory damages that likely induced recipients to settle in the first place.

218 *See* 17 U.S.C. § 504(c)(2).

219 *See id.* § 504(c)(1).

220 *See* Samuelson & Wheatland, *supra* note 3, at 453–54 (explaining that innocent infringement "has virtually no significance in litigation").

221 *See supra* Section II.B.1.b.

222 *See* Sag & Haskell, *supra* note 200, at 637 (stating that the Copyright Office is "an enthusiastic proponent of statutory damages and has resisted calls for their reform").

field that is growing at an exponential rate,[223] there needs to be a solution before irreparable harm occurs. The quickest way to achieve this is through courts using the common law to enact this presumption, though this solution is not without its challenges. Enactment presents a challenge in that the application of a common law presumption may differ across jurisdictions until the Supreme Court can intervene. Additionally, judges who establish the presumption may be accused of judicial activism and, therefore, usurping the role of legislators.[224] However, common law will bring relief faster than any statutory revision, making it the most viable option.

A rebuttable presumption of innocent infringement would simultaneously protect end-users while ensuring that creators of Generative AI can still be liable for their efforts to encroach on the property of authors. Take, for example, the illustration from the Introduction, where the student used output from a Generative AI that, unbeknownst to the student, generated text that was substantially similar to the work of another website. In that case, were the court to apply a presumption of innocent infringement in Generative AI cases, when the student is hit with the demand letter, the outcome can go one of three ways.

In the first scenario, the copyright troll would be wholly discouraged from pursuing the student in the first place, knowing that the infringement case was shaky at best because the student had no knowledge of the training process. As the presumption of innocent infringement is unlikely to be rebutted, there is little monetary gain for the copyright troll, and they are therefore less likely to even engage in troll-like behavior in the first place. In this outcome, the presumption of innocent infringement has halted the copyright troll before it even sent a demand letter.

The second option is where the copyright troll sends a demand letter for a much smaller sum, knowing that it cannot threaten $150,000 but can still threaten the innocent infringer with a sum in addition to the cost of litigation. But here, the student can make a fair evaluation of whether it is worth going to court, or rather, worth calling the copyright troll's bluff, or whether the settlement number is low enough to be worth paying. In this situation, the balance between the student and the copyright troll is somewhat more even because the number demanded is not as exorbitant and therefore does not immediately scare the student into settling.

---

223 *See supra* Section II.B.1.

224 *See Judicial Activism*, CORNELL L. SCH.: LEGAL INFO. INST. (June 2023), https://www.law.cornell.edu/wex/judicial_activism [https://perma.cc/6L5R-8LFP] ("Judicial activism refers to the practice of judges making rulings based on their policy views rather than their honest interpretation of the current law.").

The final option is where the copyright troll embellishes the potential damages available and immediately includes the statutory maximum in the demand letter, glossing over the presumption of innocent infringement. In this case, the potential defendant could do a quick Google search and find some legal website summarizing the presumption. From here, the potential defendant is left with the same choice as the second option, where they can weigh whether it is worth going to court or paying the settlement, but knowing that the power between the two parties is somewhat more evenly distributed.

In any event, innocent infringers are allowed the opportunity to either be completely free from predatory demand letters as in the first instance or have a more level playing field as in the last two because they are not exposed to as much risk when it comes to statutory damages. The student can evaluate the demand letter and choose to fully ignore it, resting comfortably in the knowledge that the cost of litigation is unlikely to be worth it for a $200 settlement for the troll.

IV. CONCLUSION

As it becomes more necessary for Generative AI to be trained on unlicensed copyrighted material, the likelihood of end-users generating infringing output, or output that could be loosely construed as infringing, grows. This infringement creates the ever-growing possibility that end-users are caught in the crosshairs of copyright trolls. The solution to this problem lies in the statutory damages regime. By applying a presumption of innocent infringement to cases involving Generative AI, copyright trolls would have a much higher evidentiary burden to reach the statutory maximum and would likewise find less potential reward in the pursuit of innocent infringers. With this burden heightened, copyright trolls would be discouraged from going after end-users and would be forced to engage only in legitimate copyright enforcement efforts.