Copyright and Artificial Intelligence:
Human Authorship, Copyrightability, and the Emerging Boundaries of Protection

Benjamin D. Burge, Amedeo Zmarandoiu – Rupp Pfalzgraf LLC

Introduction

Generative artificial intelligence ("AI") systems now produce text, images, audio, and video that resemble the creative works traditionally protected by copyright. The resulting policy question is not whether machines should be "authors," but how existing doctrines of human authorship and originality apply when AI systems participate in the creative process. The U.S. Copyright Office's ongoing multi-part report series crystallizes this inquiry. Part 1 (July 2024) addresses digital replicas; Part 2 (January 2025) examines the copyrightability of AI-generated outputs; and Part 3 (pre-publication version May 2025) considers the legality of training on copyrighted works and the contours of fair use.[1] The Office's Part 2 Report confirms a baseline: copyright protects human-created expression, and the assessment of AI-assisted works turns on whether a human contribution qualifies as authorship of expressive elements.[2]

This Article proceeds in four parts. Part I synthesizes the human authorship requirement and the Office's treatment of authorship in AI-assisted works, including the role of prompts, expressive inputs, and human modifications. Part II addresses training, infringement, and fair use as outlined in the Office's Part 3 Report and recent analyses. Part III turns to data acquisition and scraping constraints as they intersect with AI development. Part IV situates these U.S. debates alongside the EU's regulatory posture. Throughout, the aim is practical clarity: to identify where current doctrine is already sufficient and where line-drawing remains fact-intensive.

I. Human Authorship & Copyrightability

A. The baseline: human authorship remains essential

U.S. copyright law protects original expression created by human authors. The Office's Part 2 Report—drawing on both longstanding doctrine and recent litigation—states that outputs generated wholly by AI are not copyrightable, while works incorporating AI material may be

---

[1] U.S. COPYRIGHT OFFICE, *Copyright and Artificial Intelligence, Part 1: Digital Replicas* (July 2024); U.S. COPYRIGHT OFFICE, *Copyright and Artificial Intelligence, Part 2: Copyrightability* (Jan. 2025); U.S. COPYRIGHT OFFICE, *Copyright and Artificial Intelligence, Part 3: Generative AI Training* (pre-publication version, May 2025).

[2] *See* U.S. COPYRIGHT OFFICE, *Part 2: Copyrightability*, Executive Summary ("outputs generated by AI systems"), focusing on the human contribution sufficient to bring outputs within protection.

protected to the extent of human-authored contributions.[3] The Office situates this principle in the broader arc of copyright's responsiveness to technology (from photography to computer programs) without abandoning the human authorship premise.[4]

Consistent with that premise, the D.C. Circuit's decision in *Thaler v. Perlmutter* affirmed that an AI system cannot be deemed the legal "author" of a work, and that human authorship is a bedrock requirement for registration.[5] In *Thaler*, the applicant identified a computer system as the sole author; the court agreed with the Office that the Copyright Act, taken as a whole, presupposes a human author capable of owning property, signing transfers, living and dying (for term computation), and otherwise exercising legal capacities a machine lacks.[6] These holdings do not resolve every question about the degree of human contribution required in AI-assisted works; they do foreclose the threshold proposition that a machine alone can author protectable expression.[7]

B. Prompts, control, and authorship

A common misconception in generative-AI practice is that sufficiently intricate prompting or "prompt engineering" can itself supply the creativity necessary for copyright protection in the resulting output. The Office's Part 2 analysis rejects a bright-line rule but explains why prompts, as such, generally function as unprotectable instructions or ideas rather than human determination of expressive elements.[8] Because current systems can vary outputs from identical prompts and "fill in the gaps" beyond the user's specifications, prompts typically do not give the user the requisite control over how an idea is expressed.[9]

The Office illustrates this dynamic with a prompt that directs a "bespectacled cat in a robe reading the Sunday newspaper and smoking a pipe" in a particular style and lighting. The resulting image reflects some instructions and not others; where instructions are absent, the

---

[3] *Id*. (noting consensus that material generated wholly by AI is not copyrightable); see also U.S. COPYRIGHT OFFICE, *Part 2: Copyrightability* (analysis of human contribution as authorship).

[4] U.S. COPYRIGHT OFFICE, *Part 2: Copyrightability*, Introduction (technology and copyright's adaptability).

[5] *Thaler v. Perlmutter*, D.C. Cir. (Mar. 18, 2025), summarized in Stuart D. Levi, Jordan Feirman & Mana Ghaemmaghami, *Appellate Court Affirms Human Authorship Requirement for Copyrighting AI-Generated Works*, Skadden, Arps, Slate, Meagher & Flom LLP (Mar. 21, 2025).

[6] *See id*. (Act presupposes a human author—ownership capacity, inheritance/term-of-life, signature for transfers).

[7] Skadden, *Appellate Court Affirms Human Authorship Requirement*, supra note 5 (noting remaining "line-drawing" questions about the degree of permissible AI contribution are distinct from whether a machine can be an "author").

[8] U.S. COPYRIGHT OFFICE, *Part 2: Copyrightability* (Prompts) (rejecting bright-line rules; prompts generally convey unprotectable ideas and usually do not supply control over expressive elements).

[9] Id. (example and analysis that identical prompts can produce varying outputs; systems "fill gaps" beyond user specifications).

system selects the expressive details.[10] The Office's point is not that unpredictability defeats authorship per se, but that authorship requires the human to be "principally responsible" for the appearance of expressive elements—something that mere prompting of today's systems usually does not achieve.[11]

C. Assistive use, expressive inputs, and post-generation modifications

The Office draws careful distinctions among (1) AI used as an assistive tool to enhance human expression, (2) human expressive inputs that remain perceptible in the output, and (3) human selection, coordination, arrangement, or material modification of AI outputs.

Assistive use. Using AI as a tool to implement or enhance a human creative vision (for example, where a human author determines expressive elements and employs AI to realize them) does not undermine copyright protection for the human-authored work as a whole.[12]

Expressive inputs. When a human's preexisting, copyrightable work is used as an input and remains perceptible in the output, the human generally retains protection at least for the original expression embedded in the input (and potentially for compilation authorship).[13]

Selection, arrangement, and modification. Human authors may obtain protection in the creative selection or arrangement of AI-generated material, and in material, human-driven modifications of AI outputs that rise to *Feist*'s originality threshold.[14] By contrast, time and effort alone ("sweat of the brow") are not a path to protection.[15]

This framework aligns with the Office's technology-as-tool view dating to *Burrow-Giles* and modern joint-authorship doctrine: copyright protects the human party who translates ideas into fixed expression.[16] Commissioning or directing another human who actually fixes the work may support authorship or joint authorship depending on control over expression; directing an AI tool by supplying prompts generally does not.[17]

---

[10] *Id.* (cat-image prompt example).

[11] *Id.* (humans must be "principally responsible" for expressive elements; unpredictability alone is not dispositive).

[12] U.S. COPYRIGHT OFFICE, *Part 2: Copyrightability*, Introduction (assistive AI does not affect availability of protection for the work as a whole).

[13] *Id.* (expressive inputs perceptible in the output remain protected; potential compilation rights).

[14] Perkins Coie LLP, *Copyright Office Solidifies Stance on the Copyrightability of AI-Generated Works* (discussing material modifications and original selection/arrangement meeting *Feist*).

[15] U.S. COPYRIGHT OFFICE, *Part 2: Copyrightability* (rejecting "sweat of the brow" as a path to protection).

[16] Id. (technology-as-tool; *Burrow-Giles* on authorship; *Feist* on originality).

[17] Id. (joint-authorship analogies and *CCNV v. Reid*; commissioning and suggestions are unprotectable ideas).

D. Case-by-case, evidence-driven determinations

The Office emphasizes that no single factor is dispositive and that each registration inquiry must be assessed on its facts. There is no categorical rule tied to model type or interface. Instead, the focus remains on whether and where the applicant's human contributions determined expressive elements in the final work, including the extent of any perceptible expressive inputs and the nature of post-generation modifications.[18]

Secondary commentary from the legal market underscores these points without expanding the doctrine. Analyses surveying the Office's Part 2 Report observe that the Office declined to recommend new sui generis rights for AI-generated outputs, reiterated the centrality of human control over expression, and highlighted documentation of human involvement as increasingly important in registration practice.[19] For comparative context, practical guides also note that other jurisdictions take divergent approaches to AI-generated output ownership, though the U.S. default is no authorship absent human expression.[20]

II. Training Generative AI Models and the Boundaries of Fair Use

A. The Copyright Office's Part 3 Report

The Copyright Office's pre-publication Part 3 Report focuses on the central question animating current litigation: whether training generative AI systems on copyrighted works constitutes infringement and, if so, whether fair use may shield such conduct.[21] The Report identifies multiple acts in the development pipeline that may implicate exclusive rights, including the collection and curation of training data, the creation of intermediate copies during training, and the generation of outputs substantially similar to training inputs.[22] It also raises the controversial question of whether the model "weights" themselves—numeric parameters encoding linguistic relationships—may embody infringing expression where outputs replicate training material.[23]

---

[18] Id. (no model-type bright line; case-by-case registration assessment).

[19] Hogan Lovells LLP, *U.S. Copyright Office Issues Report on Copyrightability of AI Assisted and Generated Works* (Feb. 4, 2025); Perkins Coie LLP, *Copyright Office Solidifies Stance on the Copyrightability of AI-Generated Works* (documentation of human involvement).

[20] Cooley LLP, *Copyright Ownership of Generative AI Outputs Varies Around the World* (Jan. 29, 2024) (U.S. default: no ownership absent human authorship; registration protects only human-authored aspects).

[21] U.S. COPYRIGHT OFFICE, *Copyright and Artificial Intelligence, Part 3: Generative AI Training* (pre-publication version, May 2025).

[22] *Id*. at 21–27.

[23] *Id*. at 35–38.

The Office acknowledges the unsettled state of doctrine but stresses that prima facie infringement occurs when copyrighted works are reproduced for training purposes without authorization.[24] While recognizing the availability of the fair-use defense in certain circumstances, the Office squarely rejects arguments that training is inherently transformative merely because it is for machine learning or because it can be analogized to human learning.[25] In the Office's view, human learners retain imperfect impressions shaped by their own worldviews, while AI systems create perfect, scalable copies capable of generating outputs at industrial speed.[26]

B. The Four Fair-Use Factors in the AI Context

The Office applies the traditional four-factor test to training activity. Under the **first factor** (purpose and character of the use), the analysis depends on whether the model is trained to generate diverse outputs across contexts (more likely transformative) or to replicate works appealing to the same audience as the originals (less likely transformative).[27] The Office is especially skeptical of claims that training is "quintessentially transformative" if the system's purpose overlaps with that of the underlying works.[28]

The **second factor** (nature of the copyrighted work) disfavors fair use where training data consists of highly expressive or unpublished works, though this factor is rarely dispositive.[29] The **third factor** (amount and substantiality) acknowledges that entire works may be copied for training, which ordinarily weighs against fair use, but this may be mitigated by the transformativeness of the use or by guardrails that limit infringing outputs.[30] Finally, the **fourth factor** (effect on the potential market) is analyzed broadly, encompassing not only substitutional harms (lost sales of identical works) but also market dilution, where AI systems generate works in the same style or genre, potentially undermining demand for the originals.[31]

The Report thus articulates a cautious approach: fair use remains available, but the scope is narrow when training data is used for commercial purposes and when outputs threaten existing or derivative markets.[32]

---

[24] *Id.* at 19–20.

[25] *Id.* at 45–52.

[26] *Id.* at 50–51.

[27] *Id.* at 60–62.

[28] *See id.*; see also *Bartz v. Anthropic PBC,* No. 24-cv-05417 (N.D. Cal. June 23, 2025), summarized in Baker Botts LLP, *AI Legal Watch: July 2* (July 2, 2025).

[29] U.S. COPYRIGHT OFFICE, *Part 3 Report* at 64.

[30] *Id.* at 67–70.

[31] *Id.* at 71–74.

[32] *Id.* at 75–79.

C. Early Litigation Signals

The first merits-level decisions in the United States confirm the unpredictability of these questions. In *Thomson Reuters Enterprise Centre GmbH v. ROSS Intelligence Inc.*, the District of Delaware rejected the defendant's fair-use defense, holding that training a legal research engine on Westlaw headnotes constituted direct infringement.[33] The court emphasized the non-transformative nature of the use, which replicated the purpose of Westlaw's own service, and the market harm from creating a competing product.[34] While the tool at issue was not generative AI, the case established that training-based copying—even at intermediate steps—may constitute actionable infringement.[35]

Subsequent cases involving generative systems have cut the other way. In *Bartz v. Anthropic PBC*, the Northern District of California granted summary judgment for the defendant, finding that training its language model on copyrighted books was "quintessentially transformative" because the outputs did not reproduce the works verbatim but generated new text.[36] Similarly, in *Kadrey v. Meta Platforms, Inc.*, the same court concluded that training Meta's LLaMA model constituted fair use, though the court emphasized the fourth factor—market harm—more heavily than the *Anthropic* court and cautioned that market dilution theories might be more promising in future litigation.[37]

These divergent outcomes illustrate the fact-intensive nature of the fair-use analysis. Courts have analogized training to human reading and writing on the one hand, and to direct reproduction on the other, producing inconsistent results. The Copyright Office's Part 3 Report attempts to impose analytical coherence by focusing on outputs, market substitution, and dilution, but until appellate guidance emerges, fair use will remain contested.[38]

D. Licensing and Market Solutions

The Office also considers possible licensing mechanisms to address the training question, ranging from voluntary private licensing to collective or statutory regimes and opt-out models.[39] For now, the Office concludes that additional legislation is not required and that voluntary

---

[33] *Thomson Reuters Enter. Ctr. GmbH v. ROSS Intelligence Inc.*, No. 1:20-cv-00613 (D. Del. Feb. 11, 2025), discussed in Ropes & Gray LLP, *Does Training an AI Model Using Copyrighted Works Infringe the Owners' Copyright? An Early Decision Says, "Yes."* (Mar. 6, 2025).

[34] *Id.*

[35] *Id.* (rejecting intermediate copying as fair use).

[36] *Bartz v. Anthropic PBC*, No. 24-cv-05417 (N.D. Cal. June 23, 2025), summarized in Baker Botts LLP, *AI Legal Watch: July 2* (July 2, 2025).

[37] *Kadrey v. Meta Platforms, Inc.*, No. 23-cv-03417 (N.D. Cal. June 25, 2025), summarized in Baker Botts LLP, *AI Legal Watch: July 2* (July 2, 2025).

[38] Compare *Thomson Reuters v. ROSS*, supra note 33, with *Bartz* and *Kadrey*, supra notes 36–37.

[39] U.S. COPYRIGHT OFFICE, *Copyright and Artificial Intelligence, Part 3: Generative AI Training* (pre-publication version, May 2025), at 83–88.

markets, reinforced by targeted guardrails against infringing outputs, may suffice.[40] This position reflects a balance between acknowledging infringement risks and preserving innovation incentives, though litigation outcomes may shift that balance.

III. Web Scraping and Data Acquisition in the AI Training Pipeline

A. Scraping as a Threshold Question of Access

Before training generative models, developers must obtain the data. Much of that data comes from scraping publicly accessible websites, raising immediate legal and contractual questions. Courts and commentators have long wrestled with the boundary between data that is lawfully "public" and data that is off-limits because of contractual terms of service, technical access controls, or statutory prohibitions such as the Computer Fraud and Abuse Act ("CFAA").[41]

The Ninth Circuit's decisions in *hiQ Labs, Inc. v. LinkedIn Corp.* illustrate these tensions. In its most recent iterations, the court held that scraping publicly accessible LinkedIn profiles likely did not violate the CFAA, because that statute is directed at "hacking" protected computers, not at gathering information from pages visible to any internet user.[42] Yet these holdings were cabined to the CFAA and left open other claims, including breach of contract and trespass to chattels.[43] For AI developers, the lesson is that while scraping may avoid criminal liability under federal anti-hacking law, it may still trigger civil liability depending on the website's terms and state-law theories.

B. The Role of Website Terms and Post-*Van Buren* CFAA Interpretation

The Supreme Court's decision in *Van Buren v. United States* reframed CFAA's "exceeding authorized access" prong, limiting liability to breaches of technological barriers rather than mere violations of use restrictions.[44] Applying this reasoning, courts reviewing scraping disputes have distinguished between "public" sites and gated content protected by passwords or paywalls.[45] The White & Case analysis of *hiQ* following *Van Buren* emphasizes that violations of website terms alone, without circumventing technical restrictions, are

---

[40] *Id.* at 90–93; see also Stuart D. Levi, Mana Ghaemmaghami & MacKinzie M. Neal, *Copyright Office Weighs In on AI Training and Fair Use*, Skadden, Arps, Slate, Meagher & Flom LLP (May 15, 2025).

[41] Quinn Emanuel Urquhart & Sullivan, LLP, *The Legal Landscape of Web Scraping* (Apr. 28, 2023).

[42] *See* Gibson, Dunn & Crutcher LLP, *Ninth Circuit Issues Decision in Closely Watched Data Scraping Case* (Sept. 16, 2019).

[43] See id.

[44] *Van Buren v. United States*, 593 U.S. 374 (2021).

[45] *See* Quinn Emanuel, *The Legal Landscape of Web Scraping*, supra note 41.

insufficient to create CFAA liability.[46] Nevertheless, companies hosting valuable data continue to enforce contractual prohibitions aggressively, and litigation strategies frequently combine breach of contract and state-law claims with CFAA allegations.

C. Practical Considerations for AI Developers

The Quinn Emanuel survey of the "legal landscape of web scraping" frames scraping disputes as multifaceted: CFAA and state computer access statutes, contract law, copyright law, and potential misappropriation claims.[47] For AI developers, this means that compliance requires more than an analysis of CFAA exposure. Scraping practices must be reviewed for consistency with terms of service, and contractual licensing models are emerging as risk-mitigating alternatives.

Moreover, even when CFAA liability is avoided, large-scale scraping may draw claims of trespass to chattels (where scraping burdens servers), unjust enrichment, or unfair competition.[848] The Gibson Dunn commentary on a "closely watched data scraping case" underscores that courts have not settled on bright-line rules for the permissibility of scraping, leaving AI companies navigating a patchwork of overlapping doctrines.[49]

D. Implications for AI Training

Because generative AI models require massive corpora of text, images, and other content, the legality of web scraping is not a peripheral issue but a foundational one. Companies that rely on scraping risk exposure under multiple legal theories, and the resolution of scraping disputes will shape the scope of permissible training data acquisition. The Copyright Office's Part 3 Report explicitly notes that much training data is obtained through scraping, often without authorization, and that this fact has spurred many of the pending infringement suits.[50]

Taken together, the current landscape signals that AI developers should treat scraping as a compliance-sensitive activity requiring contractual diligence, technological guardrails, and risk allocation in vendor and licensing agreements. The absence of CFAA liability does not equate to a green light for unrestricted scraping; the broader web of private ordering and state-law claims ensures continuing litigation risk.

---

[46] S. Burr Eckstut & Erin Hanson, *Web Scraping, Website Terms and the CFAA: hiQ's Preliminary Injunction Affirmed Again Under Van Buren*, White & Case LLP (Apr. 22, 2022).

[47] Quinn Emanuel, *The Legal Landscape of Web Scraping*, supra note 41.

[48] See id.

[49] Gibson Dunn, *Ninth Circuit Issues Decision in Closely Watched Data Scraping Case*, supra note 42.

[50] U.S. COPYRIGHT OFFICE, *Copyright and Artificial Intelligence, Part 3: Generative AI Training* (pre-publication version, May 2025).

IV. The EU AI Act and the Global Regulatory Overlay

A. A Risk-Based Framework for Artificial Intelligence

While U.S. copyright doctrine grapples with authorship, infringement, and fair use, the European Union has advanced a comprehensive regulatory framework through the Artificial Intelligence Act ("AI Act"). The Act adopts a risk-based classification of AI systems: *unacceptable risk* (prohibited outright), *high risk* (subject to stringent obligations), *limited risk* (subject to transparency obligations), and *minimal risk* (largely unregulated).[51]

Providers of high-risk AI systems must implement quality management systems, undergo third-party conformity assessments, maintain detailed technical documentation, and ensure human oversight.[52] The Baker Botts July 2025 litigation roundup contextualizes this framework, noting that the European Commission has already launched consultations to refine which systems should be classified as high risk and how regulatory sandboxes may facilitate compliant experimentation.[53]

B. Initial Prohibitions and Early Implementation

Initial prohibitions under the AI Act have already taken effect. As noted in the July 2025 Quinn Emanuel commentary, certain AI practices deemed to present unacceptable risks—such as social scoring and manipulative subliminal techniques—are now expressly barred within the Union.[54] The Act's phased implementation means that additional obligations will come into force over the next several years, but providers must already comply with these prohibitions.

For companies developing or deploying generative models, this regulatory overlay is significant. Even if U.S. copyright doctrine permits or limits certain uses, an EU-based or globally deployed system must also comply with the EU's classification, documentation, and oversight requirements.

C. "Build Once, Comply Twice": Strategic Compliance

Commentary on the AI Act urges companies to "build once, comply twice": designing governance and compliance systems that simultaneously address EU obligations and anticipated

---

[51] Vinson & Elkins LLP, *Build Once, Comply Twice: The EU AI Act's Next Phase is Around the Corner* (June 13, 2025).

[52] *Id.*

[53] Baker Botts LLP, *AI Legal Watch: July 2* (July 2, 2025).

[54] Quinn Emanuel Urquhart & Sullivan, LLP, *Initial Prohibitions Under EU AI Act Take Effect* (July 3, 2025).

U.S. regulatory frameworks.[55] Although the United States has not yet enacted AI-specific legislation, the Copyright Office's reports and federal agency inquiries indicate that legislative and regulatory initiatives are under active consideration. Developing compliance programs with EU standards in mind may therefore position companies for cross-jurisdictional resilience.

D. Implications for Copyright and Training Practices

The AI Act does not directly regulate copyright issues, but its obligations intersect with data governance and transparency in ways that bear on copyright compliance. High-risk system obligations include documenting training data provenance, implementing safeguards against harmful outputs, and disclosing system limitations.[56] These requirements dovetail with the U.S. Copyright Office's emphasis on provenance, human involvement, and guardrails in assessing copyrightability and fair use.

As a result, copyright doctrine and AI regulation are converging: one defines what may be protected and infringed; the other prescribes how AI systems must be designed, tested, and documented. Together, they form the twin axes of compliance for companies operating across jurisdictions.

Conclusion

The debates surrounding artificial intelligence and copyright are not abstract: they directly implicate how creative industries, technology companies, and regulators conceptualize authorship, ownership, and compliance. The U.S. Copyright Office's multipart reports provide the most authoritative framework to date. Part 1 identifies the urgent need for a federal right against unauthorized digital replicas, warning that state publicity and privacy laws leave significant gaps.[57] Part 2 reaffirms that copyright remains tethered to human authorship, rejecting claims that AI systems themselves may author protectable works and clarifying that prompts, without more, rarely suffice to establish originality.[58] Part 3 addresses training and fair use, cautioning that unauthorized use of copyrighted data to train models may constitute infringement and that fair use is fact-intensive, with transformativeness and market effects as the most contested factors.[59]

---

[55] Vinson & Elkins, *Build Once, Comply Twice*, supra note 51.

[56] *See id.*; see also Baker Botts, *AI Legal Watch: July 2*, supra note 53.

[57] U.S. COPYRIGHT OFFICE, *Copyright and Artificial Intelligence, Part 1: Digital Replicas* (July 2024).

[58] U.S. COPYRIGHT OFFICE, *Copyright and Artificial Intelligence, Part 2: Copyrightability* (Jan. 2025).

[59] U.S. COPYRIGHT OFFICE, *Copyright and Artificial Intelligence, Part 3: Generative AI Training* (pre-publication version, May 2025).

The broader legal ecosystem echoes these themes. Judicial developments such as *Thaler v. Perlmutter* and *Thomson Reuters v. ROSS Intelligence* reinforce the primacy of human authorship and the vulnerability of fair-use defenses in training litigation.[60] Law firm analyses emphasize the evidentiary importance of documenting human contributions, the unsettled status of scraping as a method of acquiring training data, and the practical risks of unlicensed ingestion at scale.[61] At the same time, the EU AI Act signals that compliance obligations are not limited to copyright. Providers must build governance structures that satisfy risk-based classifications, transparency duties, and prohibitions already in force in Europe.[62]

Taken together, these sources delineate the emerging boundary lines. **First**, copyright protection remains available for works in which humans exercise creative control, but absent such involvement, AI outputs fall outside protection. **Second**, training practices that rely on scraped or unlicensed copyrighted material carry significant litigation risk; while fair use may be a viable defense in narrow circumstances, it will not reliably shield commercial AI systems. **Third**, compliance cannot be siloed: copyright obligations and regulatory duties intersect, particularly in provenance documentation and guardrail design.

For practitioners, the imperative is clear. Companies developing or deploying generative AI should (1) preserve robust evidence of human contributions in creative workflows, (2) assess training pipelines for copyright exposure and contractual compliance, (3) consider voluntary licensing or opt-out regimes to mitigate litigation risk, and (4) align governance systems with the EU AI Act's requirements to ensure global compliance. In this hybrid environment of doctrinal uncertainty and regulatory acceleration, anticipating convergence between copyright law and AI regulation is not only prudent—it is essential.

---

[60] *Thaler v. Perlmutter*, No. 20-613 (D.D.C. Mar. 18, 2025), discussed in Skadden, Arps, Slate, Meagher & Flom LLP, *Appellate Court Affirms Human Authorship Requirement for Copyrighting AI-Generated Works* (Mar. 21, 2025); *Thomson Reuters Enter. Ctr. GmbH v. ROSS Intelligence Inc.*, No. 1:20-cv-00613 (D. Del. Feb. 11, 2025), discussed in Ropes & Gray LLP, *Does Training an AI Model Using Copyrighted Works Infringe the Owners' Copyright? An Early Decision Says, "Yes."* (Mar. 6, 2025).

[61] *See, e.g.*, Hogan Lovells LLP, *U.S. Copyright Office Issues Report on Copyrightability of AI Assisted and Generated Works* (Feb. 4, 2025); Perkins Coie LLP, *Copyright Office Solidifies Stance on the Copyrightability of AI-Generated Works*; White & Case LLP, *Web Scraping, Website Terms and the CFAA: hiQ's Preliminary Injunction Affirmed Again Under Van Buren* (Apr. 22, 2022); Quinn Emanuel Urquhart & Sullivan, LLP, *The Legal Landscape of Web Scraping* (Apr. 28, 2023); Gibson, Dunn & Crutcher LLP, *Ninth Circuit Issues Decision in Closely Watched Data Scraping Case* (Sept. 16, 2019).

[62] Vinson & Elkins LLP, *Build Once, Comply Twice: The EU AI Act's Next Phase Is Around the Corner* (June 13, 2025); Quinn Emanuel Urquhart & Sullivan, LLP, *Initial Prohibitions Under EU AI Act Take Effect* (July 3, 2025).