

RESEARCH ARTICLE

Generative Artificial Intelligence and the Reconstruction of Intellectual Property Law: Copyright, Patent, Trademark, Trade Secret, and Governance in the Age of Machine Production

Dmitry Erokhin

Research Scholar, International Institute for Applied Systems Analysis (IIASA), Austria

Corresponding Author: Dmitry Erokhin (erokhin@iiasa.ac.at)

ABSTRACT

Generative artificial intelligence (Gen AI) has placed intellectual property (IP) law by affecting the entire chain of creative, technical, and commercial production. Copyright law must determine the extent of human control required to protect AI-assisted work; whether extensive ingestion of protected works for model training constitutes infringement, fair use, text and data-mining, or licensed use; and how courts should evaluate outputs that resemble or compete with existing works. Patent law confronts a related question: how to preserve human inventorship as generative technologies increasingly assist with hypothesis formation, molecular design, code generation, engineering optimization, claim drafting, and experimental planning. Trademark law is also changing, although it has received less attention. In AI-mediated markets, consumers may no longer encounter trademarks through ordinary searches and comparisons. Instead, they may receive ranked answers, conversational recommendations, generated product descriptions, synthetic endorsements, or automated purchase suggestions. Trade secret law has become equally central because model weights, datasets, prompts, safety procedures, filtering methods, benchmarks, and fine-tuning practices are often protected as secrets, even when creators, regulators, courts, and downstream users require additional information to evaluate legal compliance and public risk.

This study argues that generative AI should be regulated by reconstructing existing intellectual property principles rather than granting legal personhood to AI systems or creating new, exclusive rights for machine-generated outputs. Human authorship and inventorship should remain the legal baseline. Meanwhile, the law must exercise greater caution regarding the stages at which AI creates legal risks. Training, model development, output generation, and market deployment raise different questions and should not be reduced to a singular debate about AI and IP. This study employs doctrinal, comparative, and policy analyses using materials from the U.S., EU, U.K., and international organizations. It examines recent case law and administrative guidance alongside academic scholarship on authorship, fair use, text and data mining, augmented inventorship, algorithmic trademarks, deepfakes, transparency, and trade secrets. It concludes that the most affective AI era framework is layered and emphasizes human-centered attribution, contextual training approaches, meaningful transparency, sector-sensitive licensing, output-stage accountability, disciplined trade secret protection, and coordination among IP law, competition law, consumer protection, privacy, and AI regulation.

Open Access

Received: May 27, 2026**Revised:** June 19, 2026**Accepted:** September 04, 2026**Published:** September 30, 2026

Funding: The author received manuscript fees for this article from Korea Institute of Intellectual Property.

Conflict of interest: No potential conflict of interest relevant to this article was reported.

© 2026 Korea Institute of Intellectual Property



This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

KEYWORDS

generative artificial intelligence, intellectual property, copyright, fair use, text and data mining, authorship, patent inventorship, trademarks, trade secrets, AI transparency, licensing, AI governance

원저

생성형 인공지능과 지식재산법의 재구성: 기계 생산 시대의 저작권, 특허, 상표, 영업비밀 및 거버넌스

드미트리 예로힌

국제응용시스템분석연구소 연구원

교신저자: 드미트리 예로힌 (erokhin@iiasa.ac.at)

차례

1. Introduction
2. Methodology
3. Human Authorship and the Copyright Status of AI-Assisted Works
4. Training Data, Fair Use, Text and Data Mining, and Licensing
5. Patent Law and AI-Assisted Invention
6. Trademark Law, AI-Mediated Markets, and Synthetic Endorsement
7. Trade Secrets, Confidentiality, and the Transparency Dilemma
8. A Layered Governance Framework for AI and IP
9. Conclusion

국문초록

생성형 인공지능은 창작·기술·상업적 생산의 전 과정에 영향을 미친다는 점에서 지식재산법에 이례적인 압박을 가하고 있다. 이제 저작권법은 인공지능의 도움을 받아 제작된 저작물이 보호받기 위해 어느 정도의 인간의 통제가 필요한지, 모델 학습을 위해 보호받는 저작물을 대규모로 수집·이용하는 행위를 저작권 침해, 공정이용, 텍스트·데이터 마이닝 또는 이용허락에 따른 이용 중 무엇으로 보아야 하는지, 그리고 기존 저작물과 유사하거나 경쟁 관계에 있는 결과물을 법원이 어떻게 평가해야 하는지를 판단해야 한다. 특허법 역시 표현의 영역이 아닌 기술적 영역에서 이와 유사한 문제에 직면하고 있다. 즉, 생성형 시스템이 가설 수립, 분자 설계, 코드 생성, 공학적 최적화, 청구항 작성 및 실험 계획을 점점 더 많이 보조하는 상황에서 인간의 발명자 지위를 어떻게 유지할 것인가 하는 문제이다. 상표법도 변화하고 있지만, 이에 관한 논의는 상대적으로 주목을 덜 받아 왔다. 인공지능이 매개하는 시장에서 소비자는 더 이상 일반적인 검색과 비교 과정을 통해 상표를 접하지 않을 수 있다. 그 대신 순위가 매겨진 답변, 대화형 추천, 자동 생성된 상품 설명, 합성된 추천·보증 표현 또는 자동화된 구매 제안을 제공받을 수 있다. 영업비밀법 역시 그에 못지않게 중요한 영역으로 부상하였다. 창작자, 규제기관, 법원 및 후속 이용자가 법적 준수 여부와 공공의 위험을 평가하기 위해 더 많은 정보를 필요로 하는 경우에도 모델 가중치, 데이터셋, 프롬프트, 안전 절차, 필터링 방법, 벤치마크 및 미세조정 방식은 흔히 비밀유지 방식으로 보호되기 때문이다. 본 논문은 인공지능 시스템에 법인격을 부여하거나 기계 생성 결과물에 광범위한 새로운 배타적 권리를 인정하는 방식이 아니라, 기존 지식재산법 원칙을 재구성하는 방식으로 생성형 인공지능을 규율해야 한다고 주장한다. 인간 저작자성과 인간 발명자성은 계속해서 법적 판단의 기본 원칙으로 유지되어야 한다. 이와 동시에 법은 인공지능이 법적 위험을 초래하는 단계들을 더욱 세밀하게 구분해야 한다. 학습, 모델 개발, 결과물 생성 및 시장 배포는 각각 서로 다른 문제를 제기하므로, 이를 인공지능과 지식재산권에 관한 하나의 포괄적인 논쟁으로 통합해서는 안 된다. 본 논문은 미국, 유럽연합, 영국 및 국제적 자료를 바탕으로 법리적·비교법적·정책적 분석 방법을 활용한다. 또한 최근의 판례와 행정지침을 저작자성, 공정이용, 텍스트·데이터 마이닝, 증강된 발명자성, 알고리즘 상표, 딥페이크, 투명성 및 영업비밀에 관한 학술적 논의와 연계하여 검토한다. 결론적으로 인공지능 시대에 가장 강력한 규율 체계는 인간 중심의 귀속, 학습에 대한 맥락적 접근, 실질적인 투명성, 분야별 특성을 고려한 이용허락, 결과물 생성 단계에서의 책임성, 절제된 영업비밀 보호, 그리고 지식재산법·경쟁법·소비자보호법·개인정보보호법 및 인공지능 규제 간의 조정으로 구성되는 다층적 체계라고 주장한다.

주제어

생성형 인공지능, 지식재산, 저작권, 공정이용, 텍스트·데이터 마이닝, 저작자성, 특허 발명자성, 상표, 영업비밀, 인공지능 투명성, 이용허락, 모델 거버넌스

1. Introduction

Generative artificial intelligence has become the most comprehensive challenge to intellectual property law since the rise of the internet. Earlier technologies placed pressure on particular doctrines. Photography changed the law's understanding of mechanical reproduction and authorship. Sound recording and broadcasting reshaped the economics of performance and distribution. Software strained the boundary between expression, function, and technical invention. Search engines and online platforms forced courts to reconsider copying, linking, indexing, intermediary liability, and digital markets. Generative AI reaches across all of these fields at once. A single model may be trained on books, images, music, software code, scientific publications, product data, trademarks, design materials, and confidential information. The same model may later produce essays, songs, illustrations, legal summaries, technical drawings, patent claims, advertising copy, logos, synthetic voices, or product recommendations. It may do so through systems whose datasets, weights, filters, prompts, and safety layers remain largely inaccessible to the public. Institutional work now reflects that breadth. The OECD treats scraped training data as an issue cutting across several IP regimes¹⁾, WIPO documents generative AI as a rapidly expanding field of patenting and scientific publication²⁾, and the U.S. Copyright Office treats generative-AI training as a major copyright-policy problem³⁾.

The public debate has often moved too quickly toward slogans. One common response describes AI as merely another tool. That comparison is useful up to a point. Law has never denied copyright or patent protection simply because a person used a camera, microscope, word processor, synthesizer, computer-aided design program, or laboratory instrument. A rule that punished artists or inventors for using advanced tools would be both impractical and hostile to innovation. Generative AI, however, does more than carry out a fully specified human instruction. It may determine expressive form, suggest technical structures, generate near-complete designs, produce branding concepts, imitate styles, summarize markets, and shape consumer choices. The other common response moves too far in the opposite direction and treats the system as if it should become an author, inventor, or legal person. That step would create far more problems than it solves. Authorship and inventorship are not decorative labels. They connect legal rights to accountability, ownership, transfer, moral and economic interests, priority, disclosure, and liability. The D.C. Circuit reached that conclusion for copyright in *Thaler v. Perlmutter*⁴⁾, and the Federal Circuit reached the parallel

1) OECD, "Intellectual Property Issues in Artificial Intelligence Trained on Scraped Data", OECD Artificial Intelligence Papers, No. 33, OECD Publishing, 2025, pp. 1-49. <https://doi.org/10.1787/d5241a23-en>

2) World Intellectual Property Organization (WIPO), "Generative Artificial Intelligence: Patent Landscape Report", WIPO, 2024, pp. 1-110. <https://doi.org/10.34667/tind.49740>

3) United States Copyright Office, "Copyright and Artificial Intelligence, Part 3: Generative AI Training (Pre-Publication Version)", United States Copyright Office, 2025, pp. 1-108.

conclusion for patent inventorship in *Thaler v. Vidal*⁵).

The more serious legal question lies between these positions. The central task is to decide how much human judgment, control, selection, arrangement, validation, or supervision is needed before the law attributes a work or invention to a person. That question is familiar, yet generative AI sharpens it. A novelist who uses an AI system to test dialogue, reorganize scenes, or revise prose may remain the author of the final manuscript. A user who enters a short prompt and accepts a generated image with no expressive intervention has a weaker claim. A scientist who uses a model to identify candidate molecules may still be an inventor if she defines the research problem, constrains the search, evaluates the candidates, and validates the claimed compound. A person who merely asks a model to invent something useful and files the result as a patent application presents a very different case. The law needs a vocabulary for these differences.

The argument developed here is that generative AI requires reconstruction of intellectual property law around the lifecycle of AI systems. The legal analysis should begin by identifying the stage at which the dispute arises. Training involves data collection, scraping, copying, filtering, tokenization, embedding, and statistical optimization. Model development involves architecture, weights, fine-tuning, reinforcement learning, evaluation, safety mitigations, and deployment choices. Output generation involves prompting, inference, selection, editing, dissemination, and possible resemblance to protected works or marks. Market deployment involves recommendation systems, advertising, licensing, consumer interaction, product ranking, and contractual allocation of risk. Each stage raises different legal questions. A training use may be lawful while a particular output infringes. A generated output may lack copyright protection while the human arrangement of that output is protected. An AI-assisted invention may be patentable while the applicant still fails to show significant human contribution. A firm may protect model weights as trade secrets while still owing regulators, courts, or auditors targeted access to information under appropriate safeguards. This lifecycle matrix is the organizing contribution of the present article. While prior scholarship has examined training, authorship, inventorship, and trade secrets in productive isolation, it has not offered a unified stage-sensitive framework that evaluates copyright, patent, trademark, and trade secret questions within a single analytical architecture.

The four-stage lifecycle approach proposed here makes the interactions between IP regimes visible and provides practitioners, legislators, and courts with a coherent method for situating any new dispute at the correct point of analysis. Copyright cannot resolve patent inventorship. Patent disclosure cannot solve false endorsement. Trademark confusion cannot decide whether text and data mining should be licensed. Trade secret law cannot provide public accountability for foundation models on its own. Generative AI therefore makes intellectual property

4) *Thaler v. Perlmutter*, 130 F.4th 1039 (D.C. Cir. 2025).

5) *Thaler v. Vidal*, 43 F.4th 1207 (Fed. Cir. 2022).

more interdependent. A realistic governance model must combine copyright, patent, trademark, trade secret, contract, competition, consumer protection, privacy, and AI-specific regulation. That combination may appear untidy, yet the technology itself is untidy from a legal perspective. It does not respect the doctrinal boundaries on which twentieth-century IP analysis often relied.

This article proceeds in eight further parts. Section 2 explains the doctrinal, comparative, and normative methodology. Section 3 examines copyrightability and human authorship. Section 4 turns to training data, fair use, text and data mining, and licensing. Section 5 addresses AI-assisted invention in patent law. Section 6 considers trademarks, algorithmic markets, and synthetic endorsement. Section 7 analyzes trade secrets and the transparency dilemma. Section 8 proposes a layered governance framework. Section 9 concludes that intellectual property law should remain human-centered while becoming more stage-sensitive, transparent, and institutionally coordinated.

2. Methodology

This article uses doctrinal, comparative, and normative legal analysis. The doctrinal part examines how copyright, patent, trademark, and trade secret rules apply to generative AI under existing law⁶⁾. The analysis draws on statutes, cases, administrative guidance, and regulatory materials, with particular attention to the U.S. Copyright Office reports on copyrightability⁷⁾ and training⁸⁾, the revised USPTO guidance on AI-assisted inventions⁹⁾, the European Commission General-Purpose AI Code of Practice¹⁰⁾, the EU AI Act provisions for general-purpose AI models¹¹⁾, and recent litigation concerning AI training, outputs, and marks, including *Thomson Reuters v. Ross Intelligence*¹²⁾ and *Getty Images v. Stability AI*¹³⁾.

The comparative part focuses on the United States, the European Union, and the United Kingdom. The United States is central because fair use has become one of the main battlegrounds for AI training disputes, and because U.S. agencies have taken influential positions on human authorship and AI-assisted invention. The European

6) Peter S. Menell et al., *Intellectual Property in the New Technological Age 2025 Vol. I Perspectives, Trade Secrets and Patents*, Clause 8 Publishing, 2025, pp. 1-533; Peter S. Menell et al., *Intellectual Property in the New Technological Age 2025 Vol. II Copyrights, Trademarks and State IP Protections*, Clause 8 Publishing, 2025, pp. 1-813.

7) United States Copyright Office, "Copyright and Artificial Intelligence, Part 2: Copyrightability", United States Copyright Office, 2025, pp. 1-41.

8) United States Copyright Office, "Copyright and Artificial Intelligence, Part 3: Generative AI Training (Pre-Publication Version)", United States Copyright Office, 2025, pp. 1-108.

9) United States Patent and Trademark Office, "Revised Inventorship Guidance for AI-Assisted Inventions (Docket No. PTO-P-2025-0014)", Federal Register, Vol.90 No.227, 2025, pp. 54636-54637.

10) European Commission, "The General-Purpose AI Code of Practice", European Commission, <<https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai>>, 작성일: 2025. 7. 10.

11) European Union, Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), OJ L, 2024/1689, 12.7.2024, Article 53.

12) Thomson Reuters Enterprise Centre GmbH and West Publishing Corp. v. Ross Intelligence Inc., No. 1:20-cv-613-SB, 2025 WL 458520 (D. Del. Feb. 11, 2025).

13) Getty Images (US), Inc. and others v. Stability AI Ltd, [2025] EWHC 2863 (Ch).

Union has combined text-and-data-mining rules with AI-specific transparency duties, a relationship examined in recent scholarship on the AI Act's copyright provisions¹⁴⁾ and the practical limits of training-data transparency¹⁵⁾. The United Kingdom retains a distinctive statutory rule for computer-generated works while reconsidering the place of that rule in its copyright and AI consultation¹⁶⁾. Though the three jurisdictions do not exhaust the global debate, they offer useful contrasts of litigation-centered adjudication, regulatory transparency, and statutory experimentation.

The normative analysis evaluates whether existing doctrines still serve the values they were designed to protect. Copyright should encourage human creativity and dissemination without enclosing machine-generated material that lacks meaningful human expression. Patent law should reward technical disclosure and human inventive contribution without granting rights over results that routine AI tools make obvious. Trademark law should preserve source identification and consumer trust in markets where algorithmic agents increasingly shape attention. Trade secret law should protect legitimate commercial know-how without allowing secrecy to defeat accountability where AI systems affect public welfare. These values are familiar, but generative AI changes the practical conditions under which they operate. The methodology therefore treats AI as a stress test for IP's institutional purposes rather than as a reason to discard those purposes.

The article also relies on a lifecycle matrix. Rather than treating AI output as the single object of analysis, it asks where in the technical and commercial process a legal claim arises. Training involves acquisition, copying, preprocessing, tokenization, embedding, storage, filtering, and optimization. Model development involves architecture, weights, fine-tuning, reinforcement learning, safety layers, evaluation, and deployment strategy. Output generation involves prompting, inference, selection, editing, dissemination, and possible reproduction of protected expression or marks. Market deployment involves recommendation, ranking, advertising, product integration, licensing, consumer interaction, and contractual allocation of risk. The legal status of one stage does not automatically determine the legal status of another. A training use might be lawful while a particular output infringes. A model might generate unprotected output while the human arrangement of those outputs receives copyright protection. AI assistance might not defeat patentability, yet an application may fail if the human contribution to conception is not shown. A provider might validly protect model weights as trade secrets while still owing regulators or courts targeted disclosures under protective conditions.

This methodological choice also avoids a common weakness in AI-law writing

14) João Pedro Quintais, "Generative AI, Copyright and the AI Act", *Computer Law & Security Review*, Vol.56(2025), Article No. 106107. <https://doi.org/10.1016/j.clsr.2025.106107>

15) Adam Buick, "Copyright and AI Training Data—Transparency to the Rescue?", *Journal of Intellectual Property Law & Practice*, Vol.20 No.3(2025), pp. 182-192. <https://doi.org/10.1093/jiplp/jpae102>

16) Intellectual Property Office et al., "Copyright and Artificial Intelligence", UK Government, 2024, paras. 131, 133, 141-143.

treating legal uncertainty as if it were a purely doctrinal problem. Some of the uncertainty is doctrinal, but much of it is evidentiary and institutional. Rightsholders may not know whether their works were used for training. Courts may not be able to see the relation between training data, model weights, and outputs without expert assistance. Patent examiners may not know how much of an alleged inventive step was generated by a model rather than conceived by a human. Consumers may not know whether a recommendation is a neutral response, a sponsored result, or an AI hallucination. Regulators may not know whether a model provider has actually respected copyright reservations or merely described a compliance policy in general terms. A lifecycle method forces the analysis to ask who has information, who needs it, and what form of access is proportionate.

3. Human Authorship and the Copyright Status of AI-Assisted Works

Copyright law has become the first battlefield because generative systems produce outputs that look like the ordinary objects of copyright, including text, images, music, video, software, and design. The problem is easy to state and difficult to administer. When does an AI-assisted work remain a human-authored work? When does it become an unprotected machine output? The answer should depend on the authorial contribution embodied in the material claimed as protected, rather than on the mere use of AI.

The U.S. Copyright Office's 2025 copyrightability report provides the clearest administrative account of this approach. The Office accepts that copyright may protect works containing AI-generated material when a human author contributes protectable expression through selection, coordination, arrangement, modification, or other creative acts. It also concludes that material generated by AI without sufficient human control over expressive elements is not protected.¹⁷⁾ The D.C. Circuit's decision in *Thaler v. Perlmutter* is consistent with that view. The case involved an application that identified the machine as the sole creator. It did not require the court to resolve the harder and more common question of how much human involvement in an AI-assisted process is enough.¹⁸⁾

The human-authorship baseline has deep roots in copyright theory. Long before contemporary generative models, Samuelson examined how law might allocate rights in computer-generated works and emphasized that ownership questions cannot be answered simply by treating the computer as an author.¹⁹⁾ Ginsburg and Budiardjo later returned to the relationship between authors, tools, assistants, and machines, showing that copyright authorship depends on more than causal involvement in the production of a work.²⁰⁾ Burk added that causation, volition, and

17) United States Copyright Office, "Copyright and Artificial Intelligence, Part 2: Copyrightability", United States Copyright Office, 2025, pp. 23-33.

18) *Thaler v. Perlmutter*, 130 F.4th 1039 (D.C. Cir. 2025).

19) Pamela Samuelson, "Allocating Ownership Rights in Computer-Generated Works", *University of Pittsburgh Law Review*, Vol.47 No.4(1986), pp. 1185-1228.

intention help explain why some outputs can be attributed to a human author while others cannot.²¹⁾ These accounts prevent the analysis from becoming a simple technological analogy. The relevant question is not whether a machine was involved. The relevant question is whether the human contribution supplies the expressive authorship that copyright protects.

This distinction matters in practice. A filmmaker who uses AI to generate background images, then edits, arranges, scores, and integrates those images into a film, may have a strong copyright claim in the film as a whole. A poet who uses a model to test alternative phrasings and then chooses, revises, and orders the final text may be the author of the poem. A designer who generates dozens of images and then substantially redraws or recomposes them may have copyright in the resulting design. By contrast, a user who enters a prompt and accepts the output without meaningful expressive intervention has a much thinner claim. Prompting may involve skill, taste, persistence, and imagination, yet those qualities do not automatically translate into control over the fixed expression. That is why the Copyright Office treats prompts alone as generally insufficient under current technology.²²⁾

The issue should not be confused with aesthetic value. A machine-generated image may be striking, useful, or commercially valuable. It may look more sophisticated than a human sketch. Copyright, however, does not protect artifacts because they are attractive or costly to produce. It protects original expression attributable to an author. A legal system that grants rights to any attractive output would detach copyright from its authorial foundation and convert it into a generalized investment protection regime. That shift would be especially dangerous in generative AI because outputs can be created at enormous scale. Without a meaningful human-authorship threshold, copyright could be flooded with thin claims over machine variations, increasing clearance burdens for human creators and downstream users.

The temptation to create a new right in AI-generated outputs should therefore be resisted. Such a right would add exclusivity where the policy case remains weak. Major AI developers already have substantial incentives through subscriptions, enterprise contracts, API pricing, patents, trade secrets, data advantages, computing infrastructure, and brand recognition. Users who make human creative contributions can rely on ordinary copyright for those contributions. A broad right in machine output would create clearance burdens, increase the volume of low-cost proprietary claims, and risk reducing the public domain without a convincing showing that protection is necessary. Lemley is right that generative AI puts

20) Jane C. Ginsburg & Luke Ali Budiardjo, "Authors and Machines", *Berkeley Technology Law Journal*, Vol.34 No.2(2019), pp. 343-448. <https://doi.org/10.15779/Z38SF2MC24>

21) Dan L. Burk, "Thirty-Six Views of Copyright Authorship, by Jackson Pollock", *Houston Law Review*, Vol.58 No.2(2020), pp. 263-321.

22) United States Copyright Office, "Copyright and Artificial Intelligence, Part 2: Copyrightability", United States Copyright Office, 2025, pp. 18-21.

pressure on copyright's basic architecture, especially the idea-expression divide and substantial similarity analysis. That pressure calls for doctrinal care, not a new monopoly over machine production.²³⁾

The United Kingdom's approach illustrates the risk of legal fiction. The Copyright, Designs and Patents Act contains a special rule for computer-generated works, attributing authorship to the person who made the necessary arrangements for creation. The U.K. government's copyright and AI consultation recognizes that the rule's practical value and justification are uncertain in the generative AI setting.²⁴⁾ The uncertainty is telling. A legal system can always assign a right to someone. The harder question is whether the assignment advances the purposes of copyright. Where no human author supplies protectable expression, the case for exclusivity becomes fragile.

Human-authorship doctrine should also remain sensitive to collaboration. In many creative industries, generative AI will not replace human authors but will be embedded into workflows involving teams, editors, designers, programmers, sound engineers, translators, and marketing professionals. The law should not demand an unrealistic account of a single isolated author. It should ask whether the claimed expression is the product of human creative choices. In works that combine human and AI material, registration and infringement analysis may need to identify protected and unprotected elements with more care than before. That is not a radical departure from ordinary copyright. Courts already separate ideas from expression, facts from original selection, public-domain material from protected arrangement, and scènes à faire from protectable details. Generative AI simply makes that analytic separation more common and more important.

4. Training Data, Fair Use, Text and Data Mining, and Licensing

The copyrightability of outputs is only one side of the copyright problem. The larger economic fight concerns training. Generative models are built through large-scale ingestion and processing of data. Some of that data consists of copyrighted books, images, music, films, journalism, code, legal materials, scientific writing, and other protected works. The legal question is not whether training may involve copying. In many technical settings, it does. The hard question is whether that copying is authorized, excused, fair, covered by a text-and-data-mining exception, or subject to compensation.

A careful analysis must separate training from output. Training-stage copying involves collection, storage, preprocessing, filtering, tokenization, embedding, and model optimization. Output-stage infringement asks whether a generated result

23) Mark A. Lemley, "How Generative AI Turns Copyright Law Upside Down", *Science and Technology Law Review*, Vol.25 No.2(2024), pp. 190-212. <https://doi.org/10.52214/stlr.v25i2.12761>

24) Intellectual Property Office et al., "Copyright and Artificial Intelligence", UK Government, 2024, paras. 142-143.

reproduces protected expression, creates an unauthorized derivative work, removes rights-management information, or otherwise violates an exclusive right. The two questions are related, yet one does not decide the other. A model may be trained under a lawful exception and still generate an infringing output. A model may produce few visibly infringing outputs and still face legal challenge over the materials used to train it. Rosati's work treats liability for AI-generated outputs as a distinct inquiry under international, EU, and U.K. copyright law.²⁵⁾

In the United States, fair use remains unsettled. AI developers often rely on analogies to search indexing, intermediate copying, software reverse engineering, and computational analysis. Rightsholders answer that generative AI is different because the models may produce substitute expressive goods and because training has sometimes relied on unauthorized or pirated sources. The U.S. Copyright Office's training report sensibly avoids a categorical answer. It treats training uses along a spectrum. Uses aimed at non-expressive analysis may have stronger fair use claims, especially where outputs do not substitute for the works used. Uses involving expressive works taken from unauthorized sources to build systems that generate competing outputs face a more difficult case, particularly where licensing markets are available or developing.²⁶⁾

Recent litigation confirms that facts matter. In *Thomson Reuters v. Ross Intelligence*, the District of Delaware rejected a fair use defense in a dispute involving Westlaw headnotes used to train a competing legal research tool.²⁷⁾ The decision should not be treated as a universal rule for all AI training. It involved a curated database, a direct market relationship, and a product that competed in legal research. Its broader significance is more modest and more important. Courts may be unwilling to describe training as transformative in the abstract when the defendant uses protected material to build a substitute product in a market where licensing already exists. Other AI training disputes, including book-training litigation against major AI companies, have raised different factual questions about transformative purpose, source legitimacy, piracy, market harm, and the availability of licensing. A serious doctrine must leave room for those differences.

European law has taken a different path by linking text-and-data-mining exceptions with AI-specific transparency duties. The DSM Directive created text-and-data-mining exceptions, including a general exception subject to rights reservation. The AI Act then added obligations for providers of general-purpose AI models, including a duty to maintain technical documentation, adopt a policy to comply with EU copyright law, respect rights reservations, and publish a sufficiently detailed summary of training content according to an AI Office template.²⁸⁾ The

25) Eleonora Rosati, "Infringing AI: Liability for AI-Generated Outputs under International, EU, and UK Copyright Law", *European Journal of Risk Regulation*, Vol.16 No.2(2025), pp. 603-627. <https://doi.org/10.1017/err.2024.72>

26) United States Copyright Office, "Copyright and Artificial Intelligence, Part 3: Generative AI Training(Pre-Publication Version)", United States Copyright Office, 2025, pp. 73-107.

27) Thomson Reuters Enterprise Centre GmbH and West Publishing Corp. v. Ross Intelligence Inc., No. 1:20-cv-613-SB, 2025 WL 458520 (D. Del. Feb. 11, 2025).

European Commission General-Purpose AI Code of Practice, published in 2025, gives providers a voluntary route for demonstrating compliance with transparency, copyright, safety, and security obligations.²⁹⁾ Quintais identifies the importance of this new framework while showing that its relationship with EU copyright law remains legally complex.³⁰⁾ Buick adds a practical concern that training-data summaries may help only if they are detailed enough to be meaningful.³¹⁾

The transparency turn is essential because ordinary litigation cannot easily govern industrial-scale training on its own. Rightsholders often do not know whether their works were used. Developers may claim trade secrecy over datasets and model design. Courts may lack a reliable factual record. Regulators need documentation to assess compliance. Downstream users need to understand the legal and technical risks of models they integrate into products. Public summaries, rights-reservation systems, audit trails, and confidential regulator-facing disclosures serve different functions. They should not be confused with one another.

Text and data mining deserves separate attention because it shows how old copyright categories can struggle with new forms of knowledge extraction. A system may copy expressive works not to communicate the expression to the public but to extract statistical regularities. That purpose may look less threatening than ordinary reproduction. Yet generative AI can use those regularities to generate outputs that compete with the works from which patterns were learned. Tyagi's account of text and data mining and the innovation dimension of generative AI treats text and data mining not merely as an exception but as part of a larger innovation policy debate.³²⁾ The policy question is not only whether AI developers should be allowed to mine works. It is also whether the legal system can distinguish socially valuable computational analysis from unlicensed substitution in markets that support human creation.

Licensing will develop unevenly. Concentrated rights markets such as news archives, stock photography, music catalogs, legal databases, and academic publishing may support direct or collective licensing. Highly fragmented open web materials, orphan works, and mixed datasets pose more difficult transaction-cost problems. The Copyright Office is right to caution against premature government intervention while licensing markets continue to develop, while also leaving room for targeted intervention if market failure appears in particular sectors.³³⁾ A mature

28) European Union, Regulation (EU) 2024/1689, Article 53.

29) European Commission, "The General-Purpose AI Code of Practice", European Commission, <<https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai>>, 작성일: 2025. 7. 10.

30) João Pedro Quintais, "Generative AI, Copyright and the AI Act", *Computer Law & Security Review*, Vol.56(2025), Article No. 106107. <https://doi.org/10.1016/j.clsr.2025.106107>

31) Adam Buick, "Copyright and AI Training Data—Transparency to the Rescue?", *Journal of Intellectual Property Law & Practice*, Vol.20 No.3(2025), pp. 182-192. <https://doi.org/10.1093/jiplp/jpae102>

32) Kalpana Tyagi, "Copyright, Text & Data Mining and the Innovation Dimension of Generative AI", *Journal of Intellectual Property Law & Practice*, Vol.19 No.7(2024), pp. 557-570. <https://doi.org/10.1093/jiplp/jpae028>

framework should therefore combine contextual fair use, sector-specific licensing, machine-readable rights reservations, lawful dataset practices, provenance tools, and liability for outputs that reproduce protected expression.

News and journalism provide an especially useful illustration. News organizations have filed high-profile claims and negotiated licenses with AI companies, but news also plays a democratic role that complicates simple proprietary analysis. Licensing may support journalism by creating new revenue streams. Excessive restriction could impede search, summarization, accessibility, or public-interest uses. Complete unlicensed ingestion could weaken markets for journalism and dilute incentives to produce costly reporting. The same tension appears in academic publishing, where access to scholarship may support research and innovation, yet unlicensed extraction from curated databases can undermine the economic structure that supports editorial selection, peer review, and dissemination. These examples show why copyright policy for AI training should be sector-sensitive rather than abstract.

Source legitimacy should also matter. A model trained on lawfully accessed licensed materials presents a different legal and moral profile from one trained on pirated books or unauthorized copies. Some fair use analysis focuses on purpose and transformation, but the manner of acquisition should not disappear. A legal system that treats pirated training corpora as functionally equivalent to licensed or lawfully accessible materials would weaken the basic architecture of copyright compliance. At the same time, courts should avoid treating every technical copy made during computational analysis as if it were an ordinary market-facing reproduction. The hardest cases are precisely those in the middle, including lawful access, transformative analytical purpose, expressive outputs, uncertain substitution, and emerging licensing markets.

Drawing these threads together, the present article proposes the following criteria for assessing whether training on infringing or unauthorized sources can still qualify as fair use. First, under the purpose and character factor, courts should ask whether the training use is genuinely non-expressive and analytical, that is, whether the model is extracting statistical patterns rather than reproducing expression for communicative effect. A system whose commercial value lies in generating outputs that substitute for the works ingested will struggle to demonstrate transformation under this factor. Second, the nature of the copyrighted work weighs more heavily against the developer where the corpus consists primarily of creative rather than factual works, since such works sit closer to the core of copyright protection. Third, the amount and substantiality factor should account for the scale of ingestion as a whole, not merely the proportion of any single work reproduced. Industrial-scale ingestion of entire libraries differs materially from intermediate copying of limited extracts for interoperability purposes. Fourth, and most critically, the effect on the

33) United States Copyright Office, "Copyright and Artificial Intelligence, Part 3: Generative AI Training(Pre-Publication Version)", United States Copyright Office, 2025, pp. 96-107.

market should examine not only whether a particular output substitutes for a given work but also whether the availability of the AI system as a whole undermines licensing markets for the training corpus. Where rights-holders had established or were developing licensing programs for computational uses, as several news organizations and academic publishers had before major model releases, unlicensed ingestion of pirated copies forecloses those markets and weighs strongly against fair use. Source legitimacy thus enters the analysis not as an independent fifth factor but as a condition that shapes the weight of all four statutory factors. Training on pirated corpora tends to darken purpose, exploit creative works at scale, rely on wholesale copying, and foreclose licensing markets simultaneously.

5. Patent Law and AI-Assisted Invention

Patent law faces a less visible challenge than copyright, although its long-term significance may be equally great. Generative AI already assists with software engineering, drug discovery, materials science, electronics, design optimization, claim drafting, and experimental planning. WIPO's patent landscape report shows rapid growth in GenAI-related patenting and scientific publication, with tens of thousands of GenAI inventions identified in the decade through 2023.³⁴⁾ As AI becomes an ordinary research instrument, patent law must decide how to treat inventions produced through human-machine collaboration.

The formal rule remains human-centered. The Federal Circuit held in *Thaler v. Vidal* that an inventor under U.S. patent law must be a natural person.³⁵⁾ The U.K. Supreme Court rejected the DABUS inventorship claims under U.K. patent law.³⁶⁾ The European Patent Office has reached the same basic conclusion in the DABUS proceedings.³⁷⁾ The revised USPTO guidance preserves the natural-person rule and clarifies that AI systems themselves are not inventors.³⁸⁾ These positions are sound. Naming a machine as inventor would not clarify ownership, disclosure, assignment, priority, or accountability. It would unsettle the legal machinery that makes patent rights administrable.

The difficult question is how to evaluate human contribution when AI plays a substantive role in the inventive process. Levy describes this problem through the idea of augmented inventorship, emphasizing that generative systems may participate in tasks that look close to conception while the law continues to require human inventors.³⁹⁾ Zech similarly stresses that inventorship serves legal allocation

34) World Intellectual Property Organization(WIPO), "Generative Artificial Intelligence: Patent Landscape Report", WIPO, 2024, pp. 1-110. <https://doi.org/10.34667/tind.49740>

35) *Thaler v. Vidal*, 43 F.4th 1207 (Fed. Cir. 2022).

36) *Thaler v. Comptroller-General of Patents, Designs and Trade Marks*, [2023] UKSC 49.

37) European Patent Office, Boards of Appeal, Technical Board 3.2.07, T 0528/25 (Designation of inventor/DABUS), decision of 5 February 2026, ECLI:EP:BA:2026:T052825.20260205.

38) United States Patent and Trademark Office, "Revised Inventorship Guidance for AI-Assisted Inventions(Docket No. PTO-P-2025-0014)", Federal Register, Vol.90 No.227, 2025, pp. 54636-54637.

39) Haim V. Levy, "Augmented Inventorship: Rethinking Patent Law in the Age of Generative AI",

and accountability functions, even as AI may affect inventive step, disclosure, and institutional practice.⁴⁰⁾ These accounts point toward a practical solution. Courts and patent offices should ask whether a human researcher made a meaningful intellectual contribution to the claimed invention. That contribution may involve defining the technical problem, choosing or creating the dataset, selecting the model, setting constraints, designing the objective function, evaluating outputs, identifying the operative solution, validating the result, and drafting claims that reflect the actual inventive concept.

Patent practice may need better records of human-machine collaboration. A human-machine contribution statement could identify the human acts that shaped conception. A traceability record could document how candidate solutions were generated, filtered, tested, and selected. Such records should be proportionate. They need not disclose every model parameter or trade secret. Their purpose would be to support inventorship, enablement, written description, and nonobviousness analysis. Patent law should not punish the use of AI. It should prevent applicants from claiming exclusive rights over technical results without showing the human contribution and technical teaching that justify the grant.

AI also changes the nonobviousness inquiry. If advanced generative tools become routine in a field, the capabilities of a person having ordinary skill in the art may change. A solution that once required unusual creativity may become accessible through standard AI-assisted search. Patent law already adjusts to changing technical knowledge and available tools. AI should be treated as part of that evolving baseline. This does not mean that every AI-assisted invention is obvious. It means that courts and examiners should take seriously the possibility that generative systems alter what ordinary experimentation can achieve.

Disclosure deserves equal attention. The patent bargain requires public teaching in exchange for exclusivity. If an applicant's invention depends on hidden datasets, proprietary model behavior, undisclosed constraints, or inaccessible optimization processes, enablement and written description may become more difficult. Firms may wish to patent downstream inventions while keeping discovery pipelines secret. That is understandable, but secrecy cannot swallow the disclosure obligation. Where an AI-assisted invention depends on technical information needed to practice the claim, patent law should require enough disclosure to allow the skilled person to reproduce the invention without undue experimentation.

The relationship between patenting and trade secrecy is especially important in AI. A company may seek patents on methods, applications, or downstream products while retaining core model weights, datasets, and optimization procedures as trade secrets. That combination can be legitimate. Patents and trade secrets have always

Journal of Law and the Biosciences, Vol.13 No.1(2026), Article No. lsag014.
<https://doi.org/10.1093/jlb/lsag014>

40) Herbert Zech, "AI and Patents: The Role of Inventorship", Edit. by Philipp Hacker et al., *The Oxford Handbook of the Foundations and Regulation of Generative AI*, Oxford University Press, 2025, pp. 536-544. <https://doi.org/10.1093/oxfordhb/9780198940272.013.0027>

coexisted. Yet the cumulative effect may become problematic if firms claim broad patent rights while withholding the technical knowledge that made the invention possible. Patent law should therefore be attentive to the difference between using AI as a tool and relying on an undisclosed AI system as the real source of the technical teaching. The former should be allowed. The latter may undermine the disclosure bargain.

Patent offices will also face capacity challenges. Examiners may need to evaluate prior art generated or discovered by AI, assess whether claims are obvious in light of AI-assisted design tools, and determine whether specifications disclose enough about computational workflows. Applicants may use generative systems to draft longer and more complex applications. Competitors may use AI to search for invalidating prior art or design around claims. These practical changes could alter the economics of patent prosecution and enforcement. They also reinforce the need for patent institutions to develop technical literacy without treating AI as mysterious. AI-assisted invention should be governed through ordinary patent values such as human contribution, novelty, nonobviousness, enablement, and public disclosure.

6. Trademark Law, AI-Mediated Markets, and Synthetic Endorsement

Trademark law has received less attention than copyright and patent law in the AI debate, yet generative AI may alter the practical environment in which trademarks function. Trademark law assumes that signs help consumers identify source, reduce search costs, and make reliable purchasing decisions. AI-mediated markets complicate that picture. Consumers increasingly encounter products through ranking systems, recommendation engines, voice assistants, conversational shopping tools, generated descriptions, and automated replenishment. In that environment, source information may be filtered, summarized, rearranged, or displaced by an AI intermediary.

Farley argues that trademark law's search-cost justification becomes unstable when retail is virtual, marketing is data-driven, and purchasing decisions are increasingly automated.⁴¹⁾ Pane similarly describes the rise of the algorithmic marketplace, where AI-driven search, visual recognition, conversational interfaces, and autonomous agents reshape consumer behavior.⁴²⁾ These accounts suggest that trademark law should look beyond the traditional question of whether two marks appear similar to a human consumer at point of sale. Courts may also need to ask how the AI system ranked products, whether sponsored or affiliated results were disclosed, whether a brand was suppressed or substituted, and whether generated descriptions created an impression of affiliation or equivalence.

41) Christine Haight Farley, "Trademarks in an Algorithmic World", *Washington Law Review*, Vol.98 No.4(2023), pp. 1123-1186.

42) Meghan Pane, "Navigating the Algorithmic Marketplace: Reassessing Trademark Law in the Age of the Artificially Intelligent Consumer", *Rutgers University Law Review*, Vol.77 No.1(2024), pp. 197-232.

Generative AI also changes brand creation. A business can now generate hundreds of candidate names, slogans, logos, product mockups, and advertising concepts in minutes. This may reduce costs for small firms, yet it may also increase the number of close variants and near-conflicting signs. Trademark offices may use AI to improve classification, image similarity search, and examination. Applicants may use AI to generate marks that are intentionally or accidentally close to existing brands. The result could be a noisier trademark environment, with greater pressure on clearance, opposition, and enforcement systems. The special issue of *The Trademark Reporter* on artificial intelligence captures this shift across theory, practice, and enforcement.⁴³⁾

The litigation in *Getty Images v. Stability AI* illustrates how AI output disputes may move from copyright into trademark. The public debate around the case focused heavily on training and copyright. Yet the case also involved allegations that generated images could include Getty watermarks or marks, creating trademark and source-identification concerns.⁴⁴⁾ Even when a model does not store training works in a conventional database-like manner, outputs may still reproduce marks, logos, watermarks, trade dress, or other source cues. Trademark law is therefore likely to play a larger role in AI disputes than early commentary assumed.

Synthetic endorsement creates another layer. Generative systems can produce convincing images, voices, videos, and personas that suggest sponsorship, affiliation, or approval. The resulting harm may fall partly under trademark false endorsement, partly under rights of publicity, partly under privacy law, and partly under unfair competition. Murray argues that deepfakes can be addressed through a combination of publicity, privacy, and trademark doctrines, especially where synthetic content creates commercial deception.⁴⁵⁾ Trademark law should remain focused on source and sponsorship confusion, while publicity and privacy regimes address unauthorized identity exploitation more directly. The boundary will not always be clean. A synthetic advertisement using a celebrity's voice to sell a product may involve both false endorsement and unauthorized commercial appropriation of identity.

The policy response should update trademark doctrine without turning it into a general AI-misrepresentation law. Trademark law should address confusion, false affiliation, dilution, counterfeiting, and misleading uses of marks in generated content. Consumer protection law should address deceptive AI interfaces and undisclosed sponsored recommendations. Platform governance should address repeat generation of infringing marks, counterfeit product listings, and synthetic endorsements. Provenance and labeling tools may help where generated content is

43) International Trademark Association, "Artificial Intelligence Issue, *The Trademark Reporter*", International Trademark Association, Vol.114 No.6, 2024, p. 691.

44) *Getty Images (US), Inc. and others v. Stability AI Ltd*, [2025] EWHC 2863 (Ch).

45) Michael D. Murray, "Deceptive Exploitation: Deepfakes, the Rights of Publicity and Privacy, and Trademark Law", *IDEA: The Law Review of the Franklin Pierce Center for Intellectual Property*, Vol.65 No.2(2025), pp. 124-210.

likely to mislead. The central concern is source integrity in a market where consumers increasingly rely on systems that present conclusions rather than raw options.

The algorithmic consumer also requires a more subtle account of confusion. Classical confusion analysis asks what an ordinary purchaser is likely to believe after encountering the defendant's sign. In an AI-mediated transaction, the relevant question may be what the consumer would have believed if the AI assistant had not transformed the informational environment. A recommendation engine may conceal alternatives, summarize quality, compare prices, or infer preferences before the consumer sees a mark. A voice assistant may respond with one product rather than a list. A visual-search tool may identify a product from an image and send the user to a marketplace listing. In each setting, the consumer's exposure to trademarks is mediated by ranking, omission, and framing. Courts should therefore examine not only the marks themselves but also the interface through which source information is delivered.

Trademark law must also preserve space for lawful expression. Generative AI will be used for parody, commentary, criticism, comparative advertising, fan works, artistic experimentation, and education. Not every generated use of a mark should be treated as infringement. The challenge is to distinguish expressive and referential uses from uses that cause source confusion, false sponsorship, dilution, counterfeiting, or commercial deception. Existing doctrines, including nominative fair use, parody, functionality, and First Amendment-sensitive analysis, remain important. The AI setting does not eliminate these doctrines. It requires courts to apply them to generated content and algorithmic interfaces with attention to context.

7. Trade Secrets, Confidentiality, and the Transparency Dilemma

Trade secret law may become the most strategically important IP regime for generative AI. Many valuable features of AI systems are not patented or published. Model weights, training-data mixtures, filtering methods, fine-tuning data, safety techniques, benchmark strategies, system prompts, inference optimizations, and deployment processes may all be held as trade secrets if they derive economic value from secrecy and are protected by reasonable measures. AI development is expensive, competitive, and vulnerable to copying. Firms have legitimate reasons to protect sensitive technical and commercial information.

The difficulty is that the same secrecy can obstruct accountability. Creators may need to know whether their works were used for training. Regulators may need to evaluate safety, bias, security, and copyright compliance. Courts may need evidence about datasets, model behavior, and output generation. Downstream firms may need documentation to assess their own legal risks. Public researchers may need enough information to test claims about reliability and social impact. Keller and Aplin frame

this as a conflict between trade secret protection and public transparency, while arguing that the better path lies in structured access, third-party participation, and public-interest stewardship rather than broad destruction of trade secret rights.⁴⁶⁾

The EU AI Act reflects a similar compromise. Article 53 requires general-purpose AI providers to maintain documentation and provide information to downstream providers and regulators while protecting intellectual property rights, confidential business information, and trade secrets.⁴⁷⁾ This structure is sensible because transparency is not a single thing. Public training summaries, regulator access, downstream documentation, independent audits, court-supervised discovery, and emergency disclosures serve different audiences. The law should therefore calibrate access. Public summaries can identify categories of training content and general compliance policies. Regulators can receive more detailed information under confidentiality obligations. Courts can use protective orders. Auditors can inspect sensitive systems without publishing trade secrets. This model respects legitimate secrecy while preventing secrecy from becoming a complete answer to public accountability.

AI also creates new risks for trade secret owners. Employees may paste confidential code, business plans, customer lists, research data, or legal strategies into public AI systems. Vendors may train on submitted data unless contracts forbid it. Models may memorize and regurgitate sensitive information. Attackers may use prompt injection, model extraction, or membership inference techniques to infer protected content. NIST's generative AI risk profile identifies confidentiality and security risks associated with training data, model outputs, code, and model weights.⁴⁸⁾ These risks will affect the reasonable-measures requirement in trade secret law. A company that permits uncontrolled employee use of public AI tools may have difficulty proving that it took reasonable steps to preserve secrecy. Effective trade secret governance now requires AI-use policies, employee training, access controls, vendor review, prompt restrictions, logging, data-loss prevention, and contractual limits on model training.

Trade secrecy also raises competition concerns. A foundation model provider may combine proprietary datasets, model weights, cloud partnerships, API contracts, patents, brand recognition, and secrecy over safety and fine-tuning methods. Each element may be lawful in isolation. Together, they may create durable control over an infrastructure used by downstream firms. IP law has always tolerated some exclusivity to encourage investment. The risk in generative AI is cumulative enclosure with patents, contracts, data advantages, trade secrets, and platform effects reinforcing one another. Competition law, interoperability policy,

46) Perry Keller & Tanya F. Aplin, "Reconciling Trade Secrets and AI Public Transparency", *The Cambridge Law Journal*, Vol.85 No.1(2026), pp. 151-185. <https://doi.org/10.1017/S0008197325101086>

47) European Union, Regulation (EU) 2024/1689, Article 53.

48) Chloe Autio et al., "Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile (NIST AI 600-1)", National Institute of Standards and Technology, 2024, pp. 1-60. <https://doi.org/10.6028/NIST.AI.600-1>

procurement standards, and data-access rules may therefore become necessary companions to trade secret law.

A balanced regime should resist two mistakes. The first mistake is to treat trade secret claims as inherently suspect whenever AI is socially important. Trade secrecy can protect real innovation and security interests. The second mistake is to allow trade secrecy to defeat every request for information. That approach would leave rightsholders, regulators, courts, and affected communities unable to evaluate systems that shape markets, culture, and public decision-making. The better solution is tiered access. Public disclosure should be limited where security and commercial interests are legitimate. Independent oversight, regulator access, confidential discovery, and certified audits can provide deeper accountability without full public exposure. This is not a perfect solution, but it is more realistic than either total secrecy or total transparency.

It is important to distinguish the existing legal landscape from the normative proposal advanced here. Article 53 of the EU AI Act and the GPAI Code of Practice already require providers to maintain technical documentation and publish training summaries, and Keller and Aplin advocate structured third-party access as the preferred path over wholesale disclosure. The present article builds on but goes beyond these positions by proposing a four-tier access hierarchy calibrated to audience and risk: (i) public training summaries sufficient to enable rights-holders to identify whether their works were ingested, (ii) confidential regulator-facing documentation that discloses model weights, safety evaluations, and dataset provenance under non-disclosure obligations, (iii) court-supervised discovery subject to protective orders in active litigation, and (iv) certified independent audits for high-risk deployments, where auditors operate under professional secrecy duties comparable to those governing financial auditors. The novelty of this proposal lies not in any single tier, since each has analogues in existing law or scholarship, but in their integration as a coherent and escalating system governed by proportionality, rather than as ad hoc responses to individual disputes.

Trade secret law also intersects with professional duties. Lawyers, doctors, engineers, researchers, and public officials may use AI systems in contexts involving confidential client information, patient data, technical designs, or government records. The question is not merely whether the AI provider owes duties. It is also whether the professional user has taken reasonable steps to prevent disclosure. Contract terms, data-processing settings, enterprise controls, and internal policies all become part of the trade secret and confidentiality analysis. As generative AI becomes normalized, the standard of care for confidentiality will rise. What seemed cautious in 2022 may appear negligent in 2027.

8. A Layered Governance Framework for AI and IP

A workable framework for generative AI and intellectual property should begin

with human attribution. Copyright should protect human expression reflected in the claimed work. Patent law should protect inventions with meaningful human contribution. AI systems should not be named as authors or inventors. That rule is not hostility to technology. It preserves accountability and keeps the legal system connected to persons and institutions capable of owning, transferring, licensing, disclosing, and answering for rights. The D.C. Circuit's copyright decision denies authorship to the machine-generated work⁴⁹, the Federal Circuit's patent decision rejects machine inventorship⁵⁰, the U.K. Supreme Court takes the same view under U.K. patent law⁵¹, and the revised USPTO guidance treats AI systems as tools used by human inventors rather than inventors themselves⁵²).

The second element is stage-sensitive analysis. Training, output, and deployment should be evaluated separately. Training involves questions of copying, exceptions, fair use, rights reservation, licensing, and documentation. Outputs involve questions of protectability, substantial similarity, derivative works, rights-management information, trademark use, and false endorsement. Deployment involves recommendation design, consumer deception, platform accountability, contractual allocation of risk, and market power. Collapsing these stages creates confusion. A lawful training practice does not guarantee lawful outputs. An unprotectable output may still infringe another work. An AI-generated mark may be registrable in some circumstances while still creating confusion or dilution in others.

The third element is meaningful transparency. Transparency should be treated as legal infrastructure rather than as a slogan. Article 53 of the EU AI Act points in this direction through documentation, copyright-policy, and training-summary obligations⁵³), while the GPAI Code gives providers a more operational route for demonstrating compliance with the Act's transparency and copyright expectations⁵⁴). Their success will depend on specificity, enforcement, and trusted access. Vague descriptions of training data will not support licensing, enforcement, or public accountability. Complete public disclosure of sensitive model details may endanger security and legitimate commercial interests. A layered access system is more promising, including public summaries for general accountability, confidential regulator access for compliance, protected discovery for litigation, and independent audits for high-risk systems.

The fourth element is sector-sensitive licensing. No single licensing model will work for all creative and technical materials. News archives, stock images, music catalogs, legal databases, academic publications, open-source code, and ordinary

49) *Thaler v. Perlmutter*, 130 F.4th 1039 (D.C. Cir. 2025).

50) *Thaler v. Vidal*, 43 F.4th 1207 (Fed. Cir. 2022).

51) *Thaler v. Comptroller-General of Patents, Designs and Trade Marks*, [2023] UKSC 49.

52) United States Patent and Trademark Office, "Revised Inventorship Guidance for AI-Assisted Inventions(Docket No. PTO-P-2025-0014)", Federal Register, Vol.90 No.227, 2025, pp. 54636-54637.

53) European Union, Regulation (EU) 2024/1689, Article 53.

54) European Commission, "The General-Purpose AI Code of Practice", European Commission, <<https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai>>, 작성일: 2025. 7. 10.

web text differ in ownership structure, substitution risk, transaction costs, and market norms. Direct licensing may work for concentrated catalogs. Collective mechanisms may help where rights are fragmented. Machine-readable rights reservations may support prospective control. Courts should account for licensing feasibility when applying fair use, a point emphasized by the U.S. Copyright Office in its training report⁵⁵⁾, while the OECD's work on scraped data shows why licensing questions intersect with broader IP and data-access problems⁵⁶⁾.

The fifth element is accountability for outputs and interfaces. Providers, deployers, and users may all matter, depending on the system design and the facts. A user who requests infringing output may bear responsibility. A provider that designs a service to generate close substitutes, ignores repeated infringing behavior, or markets the system for imitation may also face liability in appropriate circumstances. Rosati shows that output liability can extend beyond end users depending on the applicable legal framework and factual control.⁵⁷⁾ In trademark and consumer settings, the interface itself matters. An AI assistant that ranks, summarizes, or recommends products may shape consumer understanding more strongly than any single mark displayed on a page.

The final element is integration with adjacent law. Copyright cannot fully address deepfakes. Patent law cannot resolve market concentration around foundation models. Trademark law cannot regulate training datasets. Trade secret law cannot by itself produce public accountability. The legal response should therefore connect IP with AI regulation, privacy law, publicity rights, consumer protection, cybersecurity, competition law, and professional responsibility. This does not weaken IP law. It allows each regime to do the work for which it is suited.

This integrated framework also suggests a more modest role for courts. Courts will remain essential because they decide concrete disputes and give legal meaning to open-textured standards such as fair use, authorship, inventorship, likelihood of confusion, and reasonable secrecy measures. Yet litigation is slow, expensive, and dependent on the facts presented by the parties. It cannot alone build licensing infrastructure, technical standards, or regulatory documentation systems. Copyright offices, patent offices, AI regulators, standards bodies, competition authorities, professional associations, and private contracting networks will shape the practical settlement. The strongest governance structure will therefore be plural. It will combine adjudication with ex ante documentation, market-making, and institutional oversight. What distinguishes the framework proposed in this article from the substantial existing scholarship—including the governance proposals

55) United States Copyright Office, "Copyright and Artificial Intelligence, Part 3: Generative AI Training(Pre-Publication Version)", United States Copyright Office, 2025, pp. 96-107.

56) OECD, "Intellectual Property Issues in Artificial Intelligence Trained on Scraped Data", OECD Artificial Intelligence Papers, No. 33, OECD Publishing, 2025, pp. 1-49. <https://doi.org/10.1787/d5241a23-en>

57) Eleonora Rosati, "Infringing AI: Liability for AI-Generated Outputs under International, EU, and UK Copyright Law", *European Journal of Risk Regulation*, Vol.16 No.2(2025), pp. 603-627. <https://doi.org/10.1017/err.2024.72>

canvassed by Quintais, Keller and Aplin, and Rosati, and the regulatory structures of the EU AI Act and GPAI Code—is its cross-regime character. Prior proposals have typically operated within a single IP branch or a single regulatory instrument. The layered framework developed here is the first to integrate the four classical IP regimes with AI-specific regulation around a common lifecycle matrix, thereby enabling the analysis to track legal risk from the moment of data collection through to market deployment and to assign responsibility to the appropriate actor at each stage.

9. Conclusion

Generative artificial intelligence has made intellectual property law more important and less self-sufficient. Copyright remains essential because training and outputs affect the markets and dignity of human creators. Patent law remains essential because AI-assisted research will shape future technical innovation. Trademark law remains essential because source identification and trust are vulnerable in algorithmic markets. Trade secret law remains essential because much of the value of AI systems lies in information firms do not disclose. Yet none of these regimes can govern the field alone.

The law should keep human authorship and inventorship at the center. That baseline protects the coherence of copyright and patent systems while leaving ample room for AI-assisted work. The law should also avoid a broad new right in machine-generated outputs. The stronger path is to protect human creative and inventive contributions, regulate training through contextual doctrine and licensing where appropriate, require meaningful transparency, impose responsibility for infringing or deceptive outputs, and prevent trade secrecy from defeating accountability. The distinctive contribution of the present article is its integration of these prescriptions within a lifecycle matrix that runs across all four IP regimes. Rather than proposing reforms within copyright, patent, trademark, or trade secret law in isolation, the article shows how the same four stages of training, model development, output generation, and market deployment generate different legal questions in each regime, and how a coherent governance response must therefore operate across regimes rather than within them.

The future of IP in the AI era will depend on careful institutional design. Courts will decide important disputes, but litigation alone is slow, expensive, and fragmented. Copyright offices, patent offices, AI regulators, standards bodies, firms, creators' organizations, and competition authorities will all shape the practical settlement. The best framework will be human-centered, lifecycle-based, transparent enough to be accountable, respectful of legitimate secrecy, attentive to market power, and flexible across creative and technical sectors. Generative AI does not make intellectual property obsolete. It asks intellectual property law to become clearer about the human, economic, and public values it is meant to protect.

참고문헌

단행본(서양)

- Herbert Zech, “AI and Patents: The Role of Inventorship”, Edit. by Philipp Hacker et al., *The Oxford Handbook of the Foundations and Regulation of Generative AI*, Oxford University Press, 2025.
- Peter S. Menell et al., *Intellectual Property in the New Technological Age 2025 Vol. I Perspectives, Trade Secrets and Patents*, Clause 8 Publishing, 2025.
- Peter S. Menell et al., *Intellectual Property in the New Technological Age 2025 Vol. II Copyrights, Trademarks and State IP Protections*, Clause 8 Publishing, 2025.

학술지(서양)

- Adam Buick, “Copyright and AI Training Data—Transparency to the Rescue?”, *Journal of Intellectual Property Law & Practice*, Vol.20 No.3(2025).
- Christine Haight Farley, “Trademarks in an Algorithmic World”, *Washington Law Review*, Vol.98 No.4(2023).
- Dan L. Burk, “Thirty-Six Views of Copyright Authorship, by Jackson Pollock”, *Houston Law Review*, Vol.58 No.2(2020).
- Eleonora Rosati, “Infringing AI: Liability for AI-Generated Outputs under International, EU, and UK Copyright Law”, *European Journal of Risk Regulation*, Vol.16 No.2(2025).
- Haim V. Levy, “Augmented Inventorship: Rethinking Patent Law in the Age of Generative AI”, *Journal of Law and the Biosciences*, Vol.13 No.1(2026).
- Jane C. Ginsburg & Luke Ali Budiardjo, “Authors and Machines”, *Berkeley Technology Law Journal*, Vol.34 No.2(2019).
- João Pedro Quintais, “Generative AI, Copyright and the AI Act”, *Computer Law & Security Review*, Vol.56(2025).
- Kalpna Tyagi, “Copyright, Text & Data Mining and the Innovation Dimension of Generative AI”, *Journal of Intellectual Property Law & Practice*, Vol.19 No.7(2024).
- Mark A. Lemley, “How Generative AI Turns Copyright Law Upside Down”, *Science and Technology Law Review*, Vol.25 No.2(2024).
- Meghan Pane, “Navigating the Algorithmic Marketplace: Reassessing Trademark Law in the Age of the Artificially Intelligent Consumer”, *Rutgers University Law Review*, Vol.77 No.1(2024).
- Michael D. Murray, “Deceptive Exploitation: Deepfakes, the Rights of Publicity and Privacy, and Trademark Law”, *IDEA: The Law Review of the Franklin Pierce Center for Intellectual Property*, Vol.65 No.2(2025).
- Pamela Samuelson, “Allocating Ownership Rights in Computer-Generated Works”, *University of Pittsburgh Law Review*, Vol.47 No.4(1986).
- Perry Keller & Tanya F. Aplin, “Reconciling Trade Secrets and AI Public Transparency”, *The Cambridge Law Journal*, Vol.85 No.1(2026).

판례

- European Patent Office, Boards of Appeal, Technical Board 3.2.07, T 0528/25 (Designation of inventor/DABUS), decision of 5 February 2026, ECLI:EP:BA:2026:T052825.20260205.
- Getty Images (US), Inc. and others v. Stability AI Ltd, [2025] EWHC 2863 (Ch).
- Thaler v. Comptroller-General of Patents, Designs and Trade Marks, [2023] UKSC 49.
- Thaler v. Perlmutter, 130 F.4th 1039 (D.C. Cir. 2025).
- Thaler v. Vidal, 43 F.4th 1207 (Fed. Cir. 2022).
- Thomson Reuters Enterprise Centre GmbH and West Publishing Corp. v. Ross Intelligence Inc.,

No. 1:20-cv-613-SB, 2025 WL 458520 (D. Del. Feb. 11, 2025).

인터넷 자료

European Commission, “The General-Purpose AI Code of Practice”, European Commission, <<https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai>>, 작성일: 2025. 7. 10.

연구보고서

Chloe Autio et al., “Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile (NIST AI 600-1)”, National Institute of Standards and Technology, 2024.

OECD, “Intellectual Property Issues in Artificial Intelligence Trained on Scraped Data”, OECD Artificial Intelligence Papers, No. 33, OECD Publishing, 2025.

United States Copyright Office, “Copyright and Artificial Intelligence, Part 2: Copyrightability”, United States Copyright Office, 2025.

United States Copyright Office, “Copyright and Artificial Intelligence, Part 3: Generative AI Training (Pre-Publication Version)”, United States Copyright Office, 2025.

World Intellectual Property Organization (WIPO), “Generative Artificial Intelligence: Patent Landscape Report”, WIPO, 2024.

기타 자료

Intellectual Property Office et al., “Copyright and Artificial Intelligence”, UK Government, 2024.

International Trademark Association, “Artificial Intelligence Issue, The Trademark Reporter”, International Trademark Association, Vol.114 No.6, 2024.

United States Patent and Trademark Office, “Revised Inventorship Guidance for AI-Assisted Inventions(Docket No. PTO-P-2025-0014)”, Federal Register, Vol.90 No.227, 2025.