

IMPLIED LICENSES IN THE AGE OF AI: CAN SILENCE SPEAK?

ESTHER POMERANTZ*

As generative artificial intelligence (AI) technology advances, the legality of using publicly available online content for AI training has been the subject of much legal debate. While existing scholarship focuses on the applicability of the fair use defense, this Column explores an alternative defense: implied licenses. By examining the implied license copyright doctrine and drawing parallels to existing jurisprudence in the online context, this Column elucidates how, in the near future, failure to deploy opt-out mechanisms might be construed as the grant of an implied license for AI training. Embracing this framework could hold substantial implications for both AI companies and online content businesses, redefining how training permissions and liability are construed in an increasingly AI-driven society.

* J.D. *cum laude*, Harvard Law School, 2025. Thank you to the *Harvard Business Law Review* editors for their edits and feedback.

TABLE OF CONTENTS

INTRODUCTION	789
I. IMPLIED LICENSES: THE CURRENT LEGAL LANDSCAPE	790
A. <i>What is an Implied License?</i>	790
B. <i>Standard Test and Shift Toward Flexibility</i>	791
II. APPLICATION TO GENERATIVE AI TRAINING	792
A. <i>Background</i>	792
B. <i>Implied Licenses in the Online Context</i>	794
C. <i>Applicability to Generative AI</i>	796
III. POTENTIAL IMPLICATIONS OF THE IMPLIED LICENSE DEFENSE	798
A. <i>Implications for Generative AI Companies</i>	798
1. Stronger Legal Shield Against Copyright Infringement Claims	798
2. Fewer Licensing Deals	798
3. Increased Innovation	799
B. <i>Implications for Businesses with Online Content</i>	799
1. Onus to Opt-Out of AI Training	799
2. New Content Management Strategies	800
3. Impact on M&A Practices	800
CONCLUSION	800

INTRODUCTION

Implied licenses have long played a vital role in business, particularly in the realm of intellectual property (IP). While explicit agreements are the gold standard, commercial interactions often function without formalized contracts, relying instead on agreements inferred from conduct. An implied license is the copyright parallel to an implied contract, allowing for the use of a copyrighted work in situations where a rights holder's actions indicate that they have granted permission, even where no explicit agreement exists.¹ In the evolving digital landscape, implied licenses have become increasingly relevant for companies seeking to justify their use of IP that was not explicitly licensed to them.² While implied licenses provide commercial flexibility, they also introduce legal uncertainty—particularly as technology continues to reshape how content is created, accessed, and used.

With the rise of generative artificial intelligence (AI), a new frontier for this uncertainty is emerging. AI models rely on vast amounts of data to improve their outputs, often scraping publicly available content from the internet.³ Generative AI companies take the position that publicly available data is fair game,⁴ but there is significant debate over whether that position is legally supported.⁵ While discussions around the legality of AI training data have largely focused on fair use,⁶ implied licenses may soon become a part of this debate.

Implied licenses could provide an alternative defense for generative AI companies facing copyright infringement claims. Unlike traditional implied licenses, which arise from direct interactions between parties,⁷ AI training involves no negotiated relationship at all—only the passive inaction of data owners who may not even realize their content is being harvested. Yet, courts have, in some instances, treated the failure to implement access controls, such as a robots.txt file to prevent web scraping, as evidence of implied consent.⁸ This raises an interesting possibility: perhaps a content owner's failure to opt out of AI training could be interpreted as the grant of an implied license to use

¹ See *infra* Part I(A).

² See *e.g.*, *Bitmanagement Software GmbH v. United States*, 989 F.3d 938 (Fed. Cir. 2021) (asserting implied license defense in the copyright infringement case); *AP v. Meltwater U.S. Holdings, Inc.*, 931 F. Supp. 2d 537 (S.D.N.Y. 2013) (same); *Righthaven LLC v. Klerks*, 2010 U.S. Dist. LEXIS 105307 (same).

³ See Matt Burgess and Reece Rogers, *How to Stop Your Data from Being Used to Train AI*, WIREd (Oct 12, 2024, 9:30 AM), <https://www.wired.com/story/how-to-stop-your-data-from-being-used-to-train-ai/> [https://perma.cc/6KT4-MUG3].

⁴ Ina Fried, *For AI Firms, Anything “Public” is Fair Game*, AXIOS (Apr. 5, 2024), <https://www.axios.com/2024/04/05/open-ai-training-data-public-available-meaning>.

⁵ See *infra* Part II(A).

⁶ See *id.*

⁷ See *infra* Part I(B).

⁸ See *infra* Part II(B).

that content for training. Although this defense has not yet been the subject of major litigation, with the exponential growth of generative AI platforms, and as mechanisms allowing content owners to signal restrictions on data use become more widespread, it may gain traction.

This Column explores the potential role of implied licenses in the AI training legality debate, along with its business implications. It begins by examining the current legal landscape of implied licenses. Drawing on implied license case law in the online context, it then analyzes the applicability of this doctrine to the AI training context. Finally, this Column considers the potential implications of this defense for both AI companies and businesses with online content.

I. IMPLIED LICENSES: THE CURRENT LEGAL LANDSCAPE

A. *What is an Implied License?*

An implied license is an unwritten license—or grant of permission—where the legal right to use someone’s IP is inferred from the conduct of the parties involved, as opposed to a formal agreement.⁹ Implied licenses “are an outgrowth of basic contract law principles,” where courts recognize that an agreement can exist even in the absence of express terms, if the parties’ actions suggest they intended to be bound.¹⁰ A classic example of an implied contract is dining at a restaurant: when you order a meal, it is understood that you are expected to pay for the food upon receiving it.¹¹ Even without a specific verbal agreement to pay, the actions and circumstances create an implied contract that obligates you to settle the bill before leaving.¹² With implied licenses, copyright law has embraced a similar concept, allowing for the use of a copyrighted work in situations where a rights holder’s actions indicate that they have granted permission, even where no explicit agreement exists.¹³

An implied license functions as an affirmative defense to a claim of copyright infringement.¹⁴ Practically, this means that if Company A uses Company B’s intellectual property and Company B sues Company A for

⁹ See 1 GRAY MARKETS: PREVENTION, DETECTION AND LITIGATION § 16.03 (2024); 2 MILGRIM ON LICENSING § 15.48 (2024).

¹⁰ Jacqui Lipton, *Implied Licenses in Copyright Law*, AUTHORS ALLIANCE (May 27, 2020), <https://www.authorsalliance.org/2020/05/27/implied-licenses-in-copyright-law/> [<https://perma.cc/L4DL-5XWR>].

¹¹ See Wil Kenton, *Implied Contract: Definition, Example, Types, and Rules*, Investopedia (Feb. 10, 2025), https://www.investopedia.com/terms/i/implied_contract.asp [<https://perma.cc/L48A-KAP7>]; see also RESTATEMENT (SECOND) OF CONT. § 4.

¹² See *id.*

¹³ See Lipton, *supra* note 10.

¹⁴ *Latimer v. Roaring Toyz, Inc.*, 601 F.3d 1224, 1235 (11th Cir. 2010) (“an implied license is an affirmative defense to a claim of copyright infringement”).

copyright infringement, Company A can avoid liability by demonstrating, among other defenses against infringement, that an implied license to use the intellectual property exists.

Under U.S. copyright law, an exclusive license—where only one person is given permission to use the work—must be in writing.¹⁵ An implied license, which is by definition unwritten, must therefore be non-exclusive.

An implied license can be found when a “copyright holder engages in conduct ‘from which [the] other [party] may properly infer that the owner consents to his use.’”¹⁶ Consent to use the copyrighted work is often inferred when one party creates a work at the request of another and delivers it with the intention that it will be copied and shared.¹⁷ However, this consent need not be implied from an action—consent can sometimes be inferred from a copyright owner’s silence or *lack* of objection to the ongoing use of their copyrighted material.¹⁸

B. Standard Test and Shift Toward Flexibility

To determine whether an implied license exists, courts have traditionally employed a three-factor test, finding a nonexclusive implied license when (i) a person (the licensee) requested the work at issue; (ii) the creator (the licensor) created the work and delivered it to the licensee who requested it; and (iii) the licensor intended that the licensee would copy and distribute the work.¹⁹ This framework, largely derived from the Ninth Circuit’s decision in *Effects Associates, Inc. v. Cohen*,²⁰ has been widely adopted by the Third, Fourth, Fifth, Sixth, Seventh, Eleventh, and D.C. Circuits as the standard test for implied licenses.²¹

¹⁵ See 17 U.S. Code § 204(a) (explaining that a “transfer of copyright ownership” must be in writing).

¹⁶ *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1116 (D. Nev. 2006) (citing *De Forest Radio Tel. Tel. Co. v. United States*, 273 U.S. 236, 241 (1927)).

¹⁷ See 1 INTELLECTUAL PROPERTY COUNSELING AND LITIGATION § 3.07 (2024).

¹⁸ See 3 NIMMER ON COPYRIGHT § 10.03 (2024); 1 GRAY MARKETS: PREVENTION, DETECTION AND LITIGATION § 16.03 (2024); 2 MILGRIM ON LICENSING § 15.48 (2024); 1 INTELLECTUAL PROPERTY COUNSELING AND LITIGATION § 3.07 (2024); *Field*, 412 F. Supp. 2d at 1116.

¹⁹ See 3 NIMMER ON COPYRIGHT § 10.03 (2024) (citing *I.A.E., Inc. v. Shaver*, 74 F.3d 768, 776 (7th Cir. 1996); *Atkins v. Fischer*, 331 F.3d 988, 992 (D.C. Cir. 2003)). See also *Estate of Hevia v. Portrio Corp.*, 602 F.3d 34, 41 (1st Cir. 2010); *Beholder Prods., Inc. v. Catona*, 629 F. Supp. 2d 490, 494 (E.D. Pa. 2009) (citing non-precedential Third Circuit opinion as well); Scott J. Sholder and Sara Gates, *This Is Not Another Fair Use Article: The Implied License and De Minimis Use Copyright Defenses*, AMERICAN BAR ASSOCIATION (Dec. 1, 2020), https://www.americanbar.org/groups/intellectual_property_law/publications/landslide/2020-21/november-december/this-not-another-fair-use-article-implied-license-de-minimisuse-copyright-defenses/ [https://perma.cc/H7GE-8GWU].

²⁰ *Effects Assocs., Inc. v. Cohen*, 908 F.2d 555, 558–59 (9th Cir. 1990).

²¹ See Sholder & Gates, *supra* note 19; 3 NIMMER ON COPYRIGHT § 10.03 (2024).

More recently, however, with the advent of digital media and online copyright cases—where the first element of someone requesting the work generally is not at play—there has been a jurisprudential shift towards flexibility.²² While the three-factor test is still used, many courts have moved towards an *intent*-based contextual inquiry, focusing on the “totality of the parties’ conduct” and the third element of intent.²³ Courts including the First, Fifth, Ninth, Eleventh, and Federal Circuits have found the existence of implied licenses where the first two *Effects Associates* elements were not present.²⁴

The Supreme Court has not directly addressed this issue in the copyright context, but it has provided relevant guidance in a patent case, noting that “[a]ny language used by the owner . . . or any conduct on his part exhibited to another from which that other may properly infer that the owner consents to his use . . . upon which the other acts, constitutes a license.”²⁵ Many circuits have applied this reasoning to copyright cases, finding implied licenses even when the three *Effects Associates* factors are not met.²⁶ The emerging legal standard therefore seems to be that an implied license exists when the conduct of both parties, “taken as a whole, supports the intent to grant a license.”²⁷

II. APPLICATION TO GENERATIVE AI TRAINING

With an understanding of what an implied license is, and how courts have approached the determination of its existence, we now turn to the application of implied licenses in the generative AI training context.

A. Background

Generative AI companies rely on vast amounts of publicly available data to train their models, scraping content from across the web to refine their outputs.²⁸ This practice has sparked significant debate over whether AI developers, by using online content for training purposes without permission, are infringing on the content owners’ copyright.²⁹ “Under U.S. copyright

²² *See id.*

²³ *See e.g.*, *Estate of Hevia v. Portrio Corp.*, 602 F.3d at 41; *Baisden v. I’m Ready Prods., Inc.*, 693 F.3d 491, 501 (5th Cir. 2012), cert. denied, 558 U.S. 1229 (2013); *Falcon Enters., Inc. v. Publishers Serv., Inc.*, 438 F. App’x 579, 581 (9th Cir. 2011); *Midlevelu, Inc. v. ACI Info. Grp.*, 989 F.3d 1205, 1216 (11th Cir. 2021), cert. denied, 141 S. Ct. 2863 (2021); *Bitmanagement Software GmbH v. United States*, 989 F.3d 938, 947 (Fed. Cir. 2021).

²⁴ *See id.*

²⁵ *De Forest Radio Tel. Co. v. United States*, 273 U.S. 236, 241 (1927).

²⁶ *See e.g.*, *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1116 (D. Nev. 2006); *Parker v. Yahoo!, Inc.*, No. 07-2757, 2008 U.S. Dist. LEXIS 74512 at *10–11 (E.D. Pa. Sep. 25, 2008).

²⁷ *See Lipton, supra* note 10.

²⁸ *See Burgess & Rogers, supra* note 3.

²⁹ *See Negar Bondari, AI, Copyright, and the Law: The Ongoing Battle Over Intellectual Property Rights*, SITES AT USC (Feb. 4, 2025), <https://sites.usc.edu/iptls/2025/02/04/>

law, copying entire copyrighted works and storing them for more than a transitory period,”—which is precisely what is being done to train generative AI—clearly constitutes copyright infringement.³⁰ “However, the AI developer’s infringing conduct is excused if either the copyright owner has licensed [the content to the AI developer], or if an applicable exception in the Copyright Act excuses the infringing conduct.”³¹ The fair use exception, codified in Section 107 of the Copyright Act, is the relevant exception that could apply to AI training.³²

Given that most content owners have not explicitly licensed their works for AI training, generative AI companies have primarily relied on the “fair use” doctrine as their legal justification.³³ Fair use is a copyright infringement defense that permits limited use of copyrighted material without permission, based on whether the “purpose and character of use,” the “nature of the copyrighted work,” the amount used, and the effect on the potential market, weigh in favor of the use.³⁴ While some AI training practices may be deemed fair use, the issue remains unsettled, with multiple lawsuits currently challenging generative AI companies over their use of copyrighted material on this basis.³⁵ The outcome of ongoing cases will shape the legal landscape, but fair use remains uncertain and fact-specific,³⁶ making it an imperfect shield for generative AI developers.

However, implied licenses could provide an alternative defense. It is possible that content owners, by failing to take steps to prevent their data from being used, might be deemed to have granted AI companies an implied license to use that content for training.

At first glance, this application of implied licenses may seem novel. AI developers generally do not interact directly with content owners before using their works for large language model (LLM) training, making this unlike a typical implied license case. Yet, a core principle of an implied license is that permission can be inferred from conduct. Courts have recognized that an implied license can arise not only from affirmative actions but *also* from a rights holder’s silence or lack of objection when a reasonable

ai-copyright-and-the-law-the-ongoing-battle-over-intellectual-property-rights/ [https://perma.cc/7Z9S-UMGL].

³⁰ Keith Kupferschmid, *Insights from Court Orders in AI Copyright Infringement Cases*, COPYRIGHT ALLIANCE (Dec. 12, 2024), <https://copyrightalliance.org/ai-copyright-infringement-cases-insights/> [https://perma.cc/K6ZU-YZYV].

³¹ *Id.*

³² *See id.*

³³ *See id.*; Roomy Khan, *AI Training Data Dilemma: Legal Experts Argue For ‘Fair Use’*, FORBES (Oct. 4, 2024, 12:24 PM) <https://www.forbes.com/sites/roomykhani/2024/10/04/ai-training-data-dilemma-legal-experts-argue-for-fair-use/> [https://perma.cc/62C9-YR8R].

³⁴ *What Is Fair Use?*, COPYRIGHT ALLIANCE (last visited Mar. 30, 2025), <https://copyrightalliance.org/faqs/what-is-fair-use/> [https://perma.cc/JJM9-W85V].

³⁵ *See* Khan, *supra* note 33.

³⁶ *See* Kupferschmid, *supra* note 30.

person would expect them to speak up.³⁷ In copyright disputes, this principle has been applied in cases where content owners knowingly distributed their works without imposing explicit restrictions.³⁸

Arguably, a rights holder's failure to object to usage of a work in circumstances where objection is reasonably expected could be interpreted as granting an implied license, particularly in an environment where industry norms or technological tools exist to signal restrictions. In fact, case law applying implied licenses in the online context may lead us to believe that, under certain circumstances, the implied license defense could apply to LLM training.

B. *Implied Licenses in the Online Context*

Courts have repeatedly recognized that an implied license may arise where a copyright holder's inaction is reasonably interpreted as granting permission to use their content. This principle has been particularly relevant in cases involving search engines and web crawlers, where website owners could have implemented technical measures to prevent unauthorized use of their content, but choose not to do so.

A fundamental case in this area is *Field v. Google, Inc.*, where Blake Field, a blogger, sued Google for copyright infringement.³⁹ Field alleged that Google's practice of caching his website posts and making them available through its search engine violated his exclusive rights to reproduce and distribute his copyrighted works.⁴⁰ Google uses an automated program, "Goog-lebot," to crawl the Internet, locating and analyzing available website pages and cataloging those pages into its search engine.⁴¹ As part of this process, website owners have the option to opt out of Google's caching by utilizing "no-archive" HTML meta-tags.⁴² Despite being aware of this option, Field chose not to implement these "no-archive" tags.⁴³ Google, in its defense, argued that, by knowingly choosing not to use the "no-archive" meta-tags, Field had granted Google an implied license to cache his site.⁴⁴

The court ruled in Google's favor, holding that Field's failure to utilize the opt out meta-tags constituted an implied license for Google to cache and display his content.⁴⁵ In its decision, the court noted that usage of meta-tags to prevent caching is a "well-known industry standard," and highlighted the impracticality of Google contacting each website owner individually to

³⁷ *Effects Assocs., Inc.*, 908 F.2d at 558–59.

³⁸ *See supra* Part I(B).

³⁹ *Field v. Google Inc.*, 412 F. Supp. 2d 1106 (D. Nev. 2006).

⁴⁰ *See id.* at 1116.

⁴¹ *Id.* at 1110.

⁴² *Id.* at 1112.

⁴³ *See id.* at 1116.

⁴⁴ *See id.* at 1115–16.

⁴⁵ *See id.* at 1116.

determine their caching preferences.⁴⁶ Field’s decision not to use those tags was thus “reasonably interpreted” by Google as consent, and thereby the grant of an implied license, to index his posts.⁴⁷

Parker v. Yahoo! Inc. reaffirmed this holding.⁴⁸ Parker, an author who published his work online, sued Yahoo! and Microsoft for copyright infringement, alleging that their search engines created and republished cached copies of his work without permission.⁴⁹ Much like in *Field*, Parker was aware of the opt-out mechanisms—“robots.txt” files—that both Yahoo! and Microsoft provided to prevent caching, and that his failure to employ these mechanisms would mean that “the search engines would display a copy of his works.”⁵⁰ Pointing to the reasoning in *Field*, the court held that Parker’s failure to utilize the “robots.txt” protocols established an implied license to cache his work.⁵¹ This reinforced the idea that industry norms and the availability of opt-out mechanisms play a crucial role in determining whether an implied license can be inferred.

However, courts have also set limits. In *MidlevelU, Inc. v. ACI Info. Grp.*, the Eleventh Circuit declined to find an implied license when a content aggregator republished blog articles sourced from an open “really simple syndication” (RSS) feed.⁵² The court distinguished the case from *Field*, noting that while allowing web crawlers to index a site might imply permission for search engines, it did not automatically extend to third-party republication.⁵³ The court drew an analogy: “Implied permission to enter through a front door (web crawler) does not also imply permission to enter through a back window (RSS feed).”⁵⁴ This ruling highlights that an implied license does not arise merely because content is accessible online; the determination depends on industry norms and the specific nature of the use.

A further limitation was articulated in *Associated Press v. Meltwater U.S. Holdings, Inc.*, where the Associated Press (AP) sued Meltwater, a media monitoring service, for scraping and republishing excerpts of its news articles.⁵⁵ Meltwater argued that AP’s failure to block crawlers using robots.txt protocols signaled an implied license.⁵⁶ The court rejected this defense, reasoning that implied licenses require some indication of mutual intent—mere

⁴⁶ See *id.* at 1112, 1116.

⁴⁷ See *id.*

⁴⁸ See *Parker v. Yahoo!, Inc.*, No. 07-2757, 2008 U.S. Dist. LEXIS 74512 (E.D. Pa. Sep. 25, 2008).

⁴⁹ See *id.*

⁵⁰ See *id.* at *15.

⁵¹ See *id.*

⁵² See *Midlevelu, Inc. v. ACI Info. Grp.*, 989 F.3d 1205 (11th Cir. 2021).

⁵³ See *id.* at 1217.

⁵⁴ *Id.*

⁵⁵ See *AP v. Meltwater U.S. Holdings, Inc.*, 931 F. Supp. 2d 537, 541 (S.D.N.Y. 2013).

⁵⁶ See *id.* at 563.

failure to prevent access does not suffice.⁵⁷ The court also emphasized that shifting the burden onto copyright holders to implement technical restrictions would be unfair, as it would effectively force all rights holders to take affirmative steps to prevent unauthorized use rather than requiring the alleged infringer to seek permission.⁵⁸

Notably, the court distinguished the facts at hand from those in *Field* and *Parker*, explaining that the website protocols at issue in those cases functioned differently than the robot.txt protocols here.⁵⁹ Unlike in those cases, where the protocols at issue were for website owners to instruct search engines not to “cache” their sites, the robot.txt files here would prevent web crawlers from accessing the website.⁶⁰ Additionally, unlike in those cases, where the plaintiffs knew that if they had utilized those meta-tags the search engines would honor the tags’ signal not to “cache,” Meltwater “reserve[d] the right to disregard certain robots.txt instructions, and ha[d] not suggested that it [would] remove content from its system [upon] request.”⁶¹

Taken together, these decisions underscore that in the online-content context, an implied license may be recognized when: (i) there is an industry practice that clearly informs a content owner that inaction, such as failure to implement a specific technical signal, might be interpreted as a grant of permission to use their work; and (ii) the content owner was aware of the preventative measures they could have taken to restrict use, but chose not to take them.

C. Applicability to Generative AI

This framework can be applied to the context of LLM training. Just as website owners have the ability to prevent search engines from caching their pages using meta-tags or robots.txt files, new tools have emerged that enable content owners to restrict AI developers from scraping their data for training. These tools include:

- **Robots.txt files:** Robots.txt files provide instructions to web crawlers and bots about which portions of a website they are allowed to visit.⁶² They can be used to indicate that a website’s content should not be included in AI training datasets.⁶³

⁵⁷ See *id.* at 562.

⁵⁸ See *id.* at 563.

⁵⁹ See *id.* at 564.

⁶⁰ See *id.* at 563–64.

⁶¹ See *id.* at 564.

⁶² See Neil Clarke, *Block the Bots that Feed “AI” Models by Scraping Your Website*, NEIL CLARKE (Aug. 23, 2023), <https://neil-clarke.com/block-the-bots-that-feed-ai-models-by-scraping-your-website/> [<https://perma.cc/7E2B-SVZC>].

⁶³ See *id.*

- **AI.txt files:** A more recent initiative, ai.txt files are similar to robots.txt, but are designed specifically for AI training.⁶⁴ By adding an ai.txt file to a website, content owners can explicitly indicate whether their data may be used for commercial AI model training.⁶⁵
- **NoAI tags:** NoAI tags, including “noai” and “noimageai,” are meta tags that content owners can embed in their website’s HTML to signal web crawlers and bots that they do not consent to their content being used to train AI models.⁶⁶
- **Platform-specific opt-out mechanisms:** Several major AI developers, including OpenAI, Adobe, Amazon, and Google, have implemented opt-out mechanisms that allow users to exclude their data from AI training datasets.⁶⁷ For some platforms, this can be done via the platform settings, while others may require users to send a specific request or follow detailed procedures outlined in their support pages.⁶⁸

While these mechanisms are still evolving, they suggest an industry-wide move toward standardizing the ability to opt out of AI training. Although there is currently no universally accepted standard requiring companies to implement such signals in the U.S., the increasing adoption of these tools raises the possibility that they may soon become standard practice.

If the use of robots.txt or ai.txt files, “noai” tags, or similar mechanisms becomes an industry norm, their absence may carry legal significance. Just as the courts in *Field* and *Parker* found that failure to implement standard opt-out measures constituted the grant of an implied license,⁶⁹ a content owner’s choice not to use a ‘do not train’ signal may soon be seen as the grant of an implied license for AI models to train on their content.

Of course, there are distinctions. Unlike search engines, which generally display indexed content in a way that benefits the original content owner, AI training involves repurposing content for an entirely different use—one that may directly compete with the content owner’s business model. Courts may be hesitant to extend the *Field* and *Parker* reasoning to AI training, particularly given the potential economic harm to rights holders.

Nonetheless, as generative AI platforms become ubiquitous and opt-out mechanisms become increasingly standardized, the likelihood of courts finding implied licenses in this context will grow. If courts recognize that

⁶⁴ See SPAWNING AI.TXT, <https://site.spawning.ai/spawning-ai-txt> [https://perma.cc/76AQ-UDRF] (last visited Mar. 30, 2025).

⁶⁵ See *id.*

⁶⁶ Miriam Tremelling, *NoAI Meta Tag FAQs*, RAPTIVE (Jun. 21, 2023), <https://help.raptive.com/hc/en-us/articles/13764527993755-NoAI-Meta-Tag-FAQs> [https://perma.cc/6886-SN2B].

⁶⁷ See Burgess & Rogers, *supra* note 3.

⁶⁸ See *id.*

⁶⁹ See *supra* Part II(B).

widespread industry norms place content owners on notice of their ability to opt out, they may begin to interpret the absence of an opt-out signal as implicit permission, potentially giving AI companies a new line of defense in copyright disputes over generative AI training data.

III. POTENTIAL IMPLICATIONS OF THE IMPLIED LICENSE DEFENSE

If courts recognize an implied license defense for generative AI training, it could have significant implications for both AI companies and content-owning businesses.

A. *Implications for Generative AI Companies*

1. Stronger Legal Shield Against Copyright Infringement Claims

For generative AI companies, an implied license defense could reduce the risk of copyright infringement liability. If courts begin recognizing implied licenses in the AI training context, where applicable, companies could argue that by failing to implement an industry standard opt-out mechanism, a content owner has granted permission for them to train on that work. This defense could supplement fair use as an additional, and potentially more predictable, legal basis for AI companies to continue training their models without explicit permission from content owners.

Moreover, unlike fair use, which requires a nuanced, multi-factor analysis,⁷⁰ an implied license argument could be more straightforward: If a content owner knowingly fails to implement an AI training opt-out mechanism, they have thereby granted permission for AI training. This could make it more difficult for content owners to bring successful legal challenges, potentially resulting in fewer lawsuits or more favorable litigation outcomes for AI companies.

2. Fewer Licensing Deals

As a result of publishers starting to push back against their data being used for AI training, there has been a surge of licensing deals between AI companies and major publishers and news organizations.⁷¹ Under these deals, “publishers let the AI companies use their content” for training, while

⁷⁰ See *supra* note 34.

⁷¹ See Sara Guaglione, *2024 in Review: A Timeline of the Major Deals Between Publishers and AI Companies*, DIGIDAY (Dec. 27, 2024), <https://digiday.com/media/2024-in-review-a-timeline-of-the-major-deals-between-publishers-and-ai-companies/> [https://perma.cc/MC63-GU46].

in exchange, publishers receive benefits like attribution and access to AI technology.⁷² This enables AI companies to access high-quality data while avoiding legal disputes.

If the implied license defense gains traction, however, AI companies may feel less pressure to negotiate such deals, arguing that their use of some publicly available content is already permissible. This could reduce the overall frequency of licensing agreements, prompting publishers to seek new strategies to monetize and protect their content effectively. Additionally, it may lead to fragmented licensing practices, with some companies relying on implicit permissions, while others opting for explicit deals to mitigate legal risk. AI companies may still engage in licensing agreements for premium content or where content owners have opted out of training, but the overall demand for these deals could decrease.

3. Increased Innovation

With a clearer legal framework allowing broader access to publicly available data, generative AI companies may feel more confident in channeling resources into research and development. This could foster growth and innovation, leading to advancements in AI capabilities and the creation of tools that enhance business efficiency and productivity. Additionally, easier access to data could prompt AI companies to explore new applications and markets, sparking competition and driving technological breakthroughs. Ultimately, the adoption of the implied license defense could create an innovative landscape that further transforms AI solutions across various business sectors.

B. Implications for Businesses with Online Content

1. Onus to Opt-Out of AI Training

For businesses that publish content online, the most immediate implication of an implied license defense is the shift in responsibility to affirmatively restrict AI training. Currently, the burden of permission falls on AI companies. However, if courts recognize that failing to implement an opt-out signal constitutes implied consent, the burden will shift to content owners. Businesses will need to proactively use industry standard opt-out tools (whether it is robots.txt, ai.txt, NoAI tags, etc.) to prevent their content from being used for training.

This could create logistical and technical challenges, particularly for smaller businesses or independent creators who may lack the resources to

⁷² *Id.*

easily implement such opt-out mechanisms. Larger corporations with legal and technical teams may adapt quickly, but smaller businesses could find themselves at a disadvantage.

2. New Content Management Strategies

To ensure that their IP is protected, businesses may want to adopt new content management strategies. A robust approach could involve implementing automated opt-out mechanisms, to prevent unwanted scraping and training of proprietary data. Companies may also consider incorporating explicit licensing terms within their terms of service, to clearly prohibit AI training on their content without consent. In addition, businesses could consider deploying watermarking and detection tools to provide an extra layer of security. This would enable businesses to track and manage how their content is utilized within AI datasets, helping to ensure that they are used only according to agreed terms, whether explicit or implicit.

3. Impact on M&A Practices

Emerging recognition of an implied license defense for AI training could potentially impact mergers and acquisitions (M&A) practices in the context of publisher deals. Acquirers may place greater emphasis on evaluating a target company's approach to content rights management during due diligence. The absence of clear opt-out mechanisms may be viewed as a potential liability, leading to adjustments in how these businesses are valued. Acquirers might reassess their offers to factor in the risks associated with unauthorized content use and potential litigation.

To navigate these changes and remain competitive targets, content businesses may need to adopt robust opt-out measures and clearly document their data rights strategies. Proactively managing content use and maintaining exclusivity will become increasingly important for ensuring attractive valuations and favorable negotiation terms. To safeguard their value and better position themselves for successful acquisitions, online content businesses will need to align their operations with evolving legal frameworks and industry standards.

CONCLUSION

As courts grapple with the legal boundaries of AI training, the concept of implied licenses offers an alternative framework for assessing whether AI companies have lawfully accessed publicly available content. While much of the current debate has centered on fair use, implied licenses could provide an additional defense for using online content for AI training, particularly as technical opt-out mechanisms become more widely embraced.

However, adopting this implied license defense is not without risks. Doing so would shift the burden onto content owners to affirmatively opt out, raising concerns about consent, control, and the broader implications for copyright enforcement. If courts begin to recognize implied licenses for AI training, online content owners may face increased pressure to implement restrictive measures to prevent unauthorized use. This could lead to a fragmented digital landscape where content is either heavily restricted or freely available for AI consumption, depending on the technical sophistication of its owners.

At the same time, adoption of an implied license defense would align with broader trends in copyright law towards encouraging innovation. As AI-generated content becomes further integrated into business and creative industries, the legal system will need to strike a balance between fostering innovation and protecting content owners. Ultimately, the way in which the boundaries of implied consent are defined—whether through litigation, legislative action, or industry standards—could influence the trajectory of the AI revolution.

