

BAKER DONELSON

---

# ARTIFICIAL INTELLIGENCE

INTELLECTUAL PROPERTY CONSIDERATIONS FOR YOUR BUSINESS OPERATIONS

---

# TABLE OF CONTENTS

---

<a href="#">Introduction</a> .....	3
INTELLECTUAL PROPERTY AND AI	
<a href="#">Legal Issues Impacting Artificial Intelligence</a> .....	6
<a href="#">The Impact of the Supreme Court’s <i>Goldsmith</i> Decision on Copyright Enforcement Against AI Tools</a> .....	9
<a href="#">Generative Artificial Intelligence Asks Questions of Innovation in Patent Law</a> .....	14
<a href="#">Artificial Intelligence and Copyright Law: <i>The NYT v. OpenAI</i> – Fair Use Implications of Generative AI</a> .....	17
<a href="#">Re-Considering the Right of Publicity in the World of Generative AI</a> .....	20
<a href="#">The Fast-Moving Race Between Gen-AI and Copyright Law</a> .....	23
<a href="#">Baker Donelson’s AI Practice Overview</a> .....	27





# INTRODUCTION

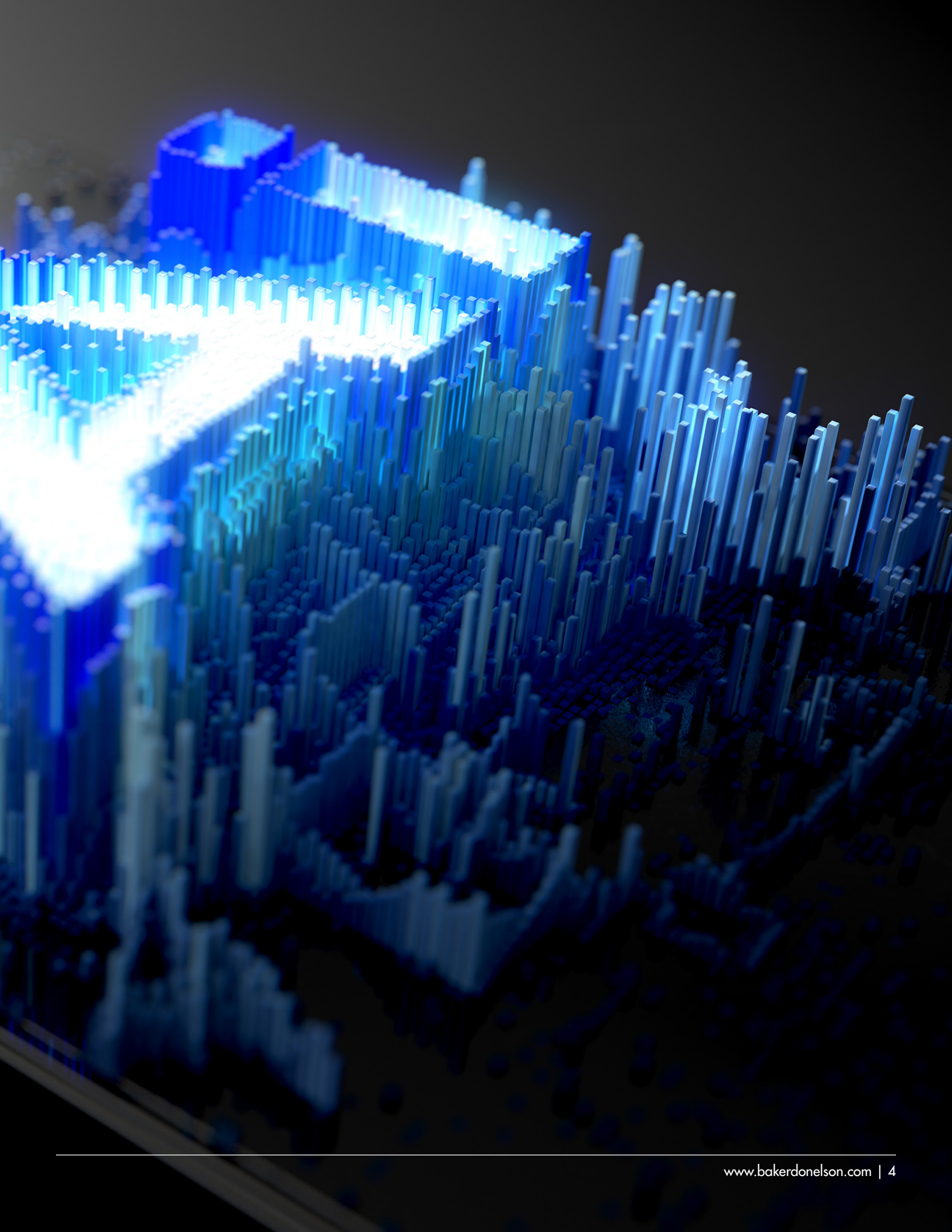
---

Over the past couple of decades, artificial intelligence-enabled (AI) technologies have crept into the marketplace, providing businesses with internal- and external-facing services, such as customer support, search engine optimization, market research, and inventory management. Now, AI-enabled technologies have become a regular fixture of many software-based services, such as technologies built on conventional machine learning or complex neural networks.

The latest evolution of AI-based technologies is generative AI. Just like the AI-enabled technologies before it, generative AI is transforming the workplace and presenting a variety of legal challenges for companies of all sizes and in all industry sectors. Companies are increasingly turning to AI-enabled tools to attract and retain talent, create efficiencies, generate content, and gain a competitive advantage. While these new technologies promise efficiency and productivity, they also introduce several new complexities and considerations, including employment, data privacy and security, and intellectual property concerns.

The surge in popularity of generative AI platforms, such as OpenAI's ChatGPT and Google's Bard, has created a massive disruption as companies balance the benefit of accelerated productivity against the potential risks. The legal world has gained key insights from recent federal court decisions and administrative guidance regarding the protectability of AI-generated content, but with more answers, so too are more questions offered. In the following articles, our team provides insights into what they are seeing in the ever-evolving AI space as it relates to IP, and what businesses should be aware of when creating their own policies.







# INTELLECTUAL PROPERTY AND AI

---



Artificial intelligence (AI) is undoubtedly the hottest topic in technological innovation. The truth, however, is that conventional AI programs have long been applied as enterprise solutions for a variety of company services, including inventory management, customer-support management, search engine optimization, market research, and outbound email campaigns. Even the practice of law has employed forms of AI in managing an unwieldy universe of documents for production in eDiscovery and due diligence.

While AI has disrupted the competitive marketplace, what is creating friction with current business operations, and even the practice of law, is generative AI. Generative AI platforms, such as OpenAI's ChatGPT, are taking many to task on the fundamentals of intellectual property law, including authorship of copyrightable works and inventorship of patentable subject matter. But as legal practitioners look for answers to these legal issues, AI-based technology continues to rapidly evolve, leaving businesses, and their counsel, a step (or two) behind innovation.

While generative AI shows no present sign of ceasing to make its way into various business practices, such as the [digitization of human resource operations](#), the legal world has gained key insights from recent court decisions and/or administrative guidance. We address a few of these takeaways, with the intention of providing subsequent alerts as the law on AI continues to take shape.

## THE BASICS – AI AND ITS APPLICATIONS

Depending upon whom you ask, you can receive a variety of definitions or explanations of AI. Some define AI as machine learning (ML), which focuses on learning and improvement from repetitious experiences, while others explain AI in the context of deep learning (DL), which focuses on algorithms or neural networks to train a model. The reality, however, is that AI is inclusive of both ML and DL, and is a step beyond: it is the ability of a machine to *initiate* intelligent human-like cognitive thinking and behavior, based upon information learned on a rolling basis.

Effective AI deployment is often contingent upon the scale, breadth, and quality of data available to the program. For AI to mimic human thinking, an AI system is trained on a dataset and learns by identifying patterns that link inputs with outputs. The “learned” AI may then translate new inputs received into recommendations, classifications, and, in some cases, *predictions*. And, for generative AI programs, the AI may then produce external-facing content, such as source code, artwork, or narrative text.

## AUTHORSHIP AND INVENTORSHIP – A SEEMINGLY SETTLED ISSUE

While intellectual property law is continuing to take shape around AI, legal authorities have recently emphasized that AI cannot function as a “person” under copyright and patent law.

In November 2018, computer scientist Dr. Steven Thaler filed a [copyright application](#) with the United States Copyright Office (USCO), aiming to register a two-dimensional visual work, as reproduced below:

### A Recent Entrance to Paradise



Thaler identified the author of the work as “Creativity Machine” – a generic name for an AI system Thaler created, called Device Autonomous Bootstrapping of Unified Sentience (DABUS). The USCO denied the application, contending that the work was made “without any creative contribution from a human actor.” After the USCO’s Copyright Review Board affirmed the USCO’s rejection, Thaler filed suit in the U.S. District Court for the District of Columbia, where Thaler seeks a holding that AI-generated works are copyrightable under federal law.<sup>1</sup> The case is proceeding, with no expectation that the court will deviate from the USCO’s current stance (as discussed more on page 7).

<sup>1</sup> *Thaler v. Perlmutter*, Case No. 1:22-cv-01564 (D.D.C.).



Thaler has not only challenged issues of authorship in copyright law but also invoked questions of inventorship in patent law. In July 2019, Dr. Thaler filed two patent applications with the United States Patent and Trademark Office (USPTO), claiming DABUS as the sole inventor.<sup>2</sup> Like the USCO, the USPTO refused to allow the patent applications, stating that AI is not a “natural person” to which a patent may be granted.<sup>3</sup> Dr. Thaler eventually appealed to the Federal Circuit, which affirmed the USPTO’s conclusion that the Patent Act expressly contemplates that inventors must be “individuals.”<sup>4</sup> While Dr. Thaler’s ongoing efforts are likely to be in vain, he nevertheless filed a petition on March 17, 2023, with the U.S. Supreme Court for further review.<sup>5</sup>

A day before this petition was filed, the USCO issued a [statement of policy](#) on works created with the assistance of AI, reaffirming its position on human authorship. The USPTO, on the other hand, [requested public comments](#) regarding AI and inventorship. Comments were closed on May 15, 2023.

## WORKS CREATED FROM GENERATIVE AI – A NOT-SO-SETTLED ISSUE

While intellectual property law appears to be settled as to “who” may qualify as an author or inventor of intellectual property, the law is not so settled on the protection of AI-generated works.

In the context of copyrightable works of authorship and patentable inventions, certain software may be used to create the underlying work or the underlying invention. Of course, this begs the question – while an inventor or author must be a human, just how much technological intervention (i.e., digital help) can be used to reach a copyrightable or patentable threshold?

Copyright law has historically been behind the ball on technological development. For example, in the 19th century, the Supreme Court was tasked with clarifying that photographs constitute copyrightable subject matter, even if there was mechanical intervention by a camera.<sup>6</sup> While it seems silly to think that photography was at one point not contemplated by copyright law, it could seem laughable at some point in the future, that AI-generated works were not protectable under copyright law. So, how much “intervention” is exercised by an AI-driven machine?

AI programs use a number of datasets. Many AI programs, such as OpenAI’s ChatGPT and DALL-E, are driven by crawling (or scraping) the internet and pulling information into its categorical and organized datasets, as this information is available in mass

quantity, easily accessible, and “free.” Of course, much of what resides on the internet is also protected by copyrights, trademarks, patents, or combinations thereof.

The power of generative AI technology was put on display before the USCO. In September 2022, the USCO issued its first notice of registration to a partially AI-generated graphic novel, [Zarya of the Dawn](#), excerpts of which are shown below:

### Zarya of the Dawn, Cover Page, and Second Page



A month later, after the USCO became aware of public statements and online articles on the author [Kristina Kashtanova](#)’s use of generative AI, the USCO issued a notice to the author that the work may be canceled, requesting details on the level of human involvement in creating the graphic novel. In this letter, the USCO asserted that the work could only be protected with respect to the selection, coordination, and arrangement of the work’s written and visual elements, but that copyrightable protection could not extend to the visual elements themselves, each of which were generated by AI program Midjourney.

While it remains to be seen whether Kashtanova or Midjourney respond to the USCO’s letter, the USCO has issued [guidance](#) that it will “consider whether AI contributions are the result of ‘mechanical reproduction’ or instead of an author’s own mental conception, to which [the author] gave visible form,” which is “necessarily a case-by-case inquiry.” So instead of shutting the door entirely, the USCO will undertake a fact-intensive inquiry to see what does (and what does not) fall within the ambit of protectable expression in an AI-generated work.

<sup>2</sup> U.S. Application Nos. 16/524,350 (teaching a “Neural Flame”) and 16/524,532 (teaching a “Fractal Container”).

<sup>3</sup> *Thaler v. Vidal*, 43 F.4th 1207, 1210 (Fed. Cir. 2022).

<sup>4</sup> *Thaler v. Vidal*, 43 F.4th 1207, 1212 (Fed. Cir. 2022).

<sup>5</sup> *Thaler v. Vidal*, Petition for a Writ of Certiorari, United States Supreme Court (Mar. 17, 2023).

<sup>6</sup> *Burrow-Giles Lithographic Co. v. Sarony*, 111 U.S. 53, 58 (1884).

## GENERATIVE AI – THE THIN LINE BETWEEN ORIGINALITY AND DERIVATION

There is no doubt that AI programs such as ChatGPT wield incredible power, and individuals and companies will continue to leverage this power for personal or commercial gain. Of course, the issue is that AI programs often pull into their datasets certain protected intellectual property. This continues to keep businesses and content creators up at night, with the fear that their protectable intellectual property will be used by AI to generate “new” intellectual property.

Most companies are focused on AI’s unauthorized or inappropriate use of their copyrighted works. Generally speaking, copyright-infringement determinations turn on; (1) whether an alleged infringer had access to a copyrighted work; and (2) whether there was substantial similarity between the copyrighted work and accused work. For AI programs, if a copyrighted work is contained in the data set scraped from the public domain (e.g., the internet), then assuredly there is access to the copyrighted work for infringement purposes. So then, the analysis hinges on substantial similarity. Did the AI program employ a content-moderation module, wherein the information that is being provided as the answer is being scanned to make sure it does not include any inappropriate (or infringing) content? If not, the AI program may have unauthorizedly “borrowed” from the copyrighted work’s expression.

With non-AI-generated works, individuals use copyright-protected works on the basis of fair use, or at the permission of the copyright owner. So, should AI-generated works differ in any way from traditional notions of permissible use? Does it matter, for infringement purposes, if AI-generated works themselves are not copyrightable? Or, does it matter if the author has used the AI program to create further derivations of the author’s original, core expression of the work? Can Paramount use AI programs, such as DALL-E, to create protectable expressions of [SpongeBob](#), especially where SpongeBob was originally created by human intervention? Or can Paramount use the AI programs to create entirely new characters in connection with its already existing intellectual property in SpongeBob? What and where is the line?

---

## ABOUT OUR AUTHORS

Mr. Lanquist and Mr. Rota are members of Baker Donelson’s multidisciplinary AI team, which focuses on AI-based technologies, including generative AI platforms. Team members use their experience with these technologies and industry-specific knowledge to assist companies of all sizes in navigating this evolving landscape.



**Edward D. Lanquist**

Shareholder | Nashville | 615.726.5581  
[elanquist@bakerdonelson.com](mailto:elanquist@bakerdonelson.com)



**Dominic Rota**

Associate | Nashville | 615.726.5634  
[drot@bakerdonelson.com](mailto:drot@bakerdonelson.com)

## PRESENT-DAY ACTIONS

Administrative bodies and courts have attempted to resolve early legal challenges presented by AI. But, as more answers are provided, so too are more questions offered. While it seems that there is no clear guidance from the “powers that be,” this lack of guidance does not absolve individuals and companies from exercising sound intellectual property hygiene.

Companies must continue to only copy, reproduce, distribute, or otherwise use intellectual property about which they know the source or origin. Additionally, companies need to affirmatively file for copyright protection for their outwardly facing works, so that they can take advantage of the benefits of copyright registration, including the right to sue in federal court. Likewise, attorneys’ fees and statutory damages, often the “hammer” of these infringement litigations, can only be obtained if the copyright application was filed prior to the infringement or within three months of publication. Furthermore, companies need to educate their personnel on an intellectual property protection plan that includes copyright usage principles and policies on using generative AI programs in connection with their job responsibilities. Above all, companies must increase their policing efforts, ensuring that their registered (and non-registered) intellectual property is not exploited by third parties leveraging the power of AI.

As to patents, companies, on an early and often basis, must review their technological innovations and invention disclosures, clarifying how potentially patentable subject matter is being generated (e.g., through the use of AI programs or not), and filing patent applications in connection with this patentable subject matter. While the issue of AI programs has further complicated the question as to what qualifies as patentable subject matter, too many companies still believe that software is outrightly not patentable. However, under the *Alice* decision from the Supreme Court, new and useful software applications and related inventions may be protectable or are likely protectable if they address a stated technical problem in a new way.<sup>7</sup>

<sup>7</sup> *Alice Corp. Pty. Ltd. v. Cls Bank Int’l*, 573 U.S. 208 (2014).



# THE IMPACT OF THE SUPREME COURT'S *GOLDSMITH* DECISION ON COPYRIGHT ENFORCEMENT AGAINST AI TOOLS

 | JULY 27, 2023

The U.S. Supreme Court's opinion in *Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith* sent ripples through the legal and artistic communities. Months later, legal scholars and art journalists continue to debate whether the decision opens the door for federal courts to act as "art critics." Many, however, downplay how the Supreme Court's decision impacts the ways in which copyright owners may enforce their rights against generative AI tools.

Conversations around generative artificial intelligence (generative AI) are dominating the social stratosphere, as generative AI is regularly atop the headlines within the context of OpenAI's ChatGPT or Google's Bard. By simulating human cognitive thinking, generative AI can produce new types of text, imagery, audio, and synthetic data by using patterns and informational elements obtained from prior works.

Because generative AI often relies on pools of data and third-party creations to create new content, the community at large is concerned that generative AI may, whether intentionally or inadvertently, exploit copyright-protected content to develop purportedly original content. Although not currently being perceived as such, the Supreme Court's decision in *Goldsmith* just might provide insights into enforcement of copyright against generative AI's misuse of protected works.

## THE SUPREME COURT CLARIFIES FAIR USE

In *Goldsmith*, the Supreme Court held that certain Andy Warhol silkscreen portraits of the musical artist Prince, which were derived from third-party photographs, constituted impermissible fair use. We will review the factual and procedural background of the case, followed by a discussion of the doctrine of fair use and the Supreme Court's application of it.

## Factual Background and Procedural History

In 1981, Lynn Goldsmith, an acclaimed professional photographer, captured Prince in concert and in studio. Later, Goldsmith licensed to *Vanity Fair* the rights to a black and white photographic portrait of Prince for the purpose of serving as an "artist reference for an illustration." In turn, *Vanity Fair* hired Andy Warhol to create, for publication, a silkscreen portrait of Prince derived from Goldsmith's image. Andy Warhol created 15 additional works derived from Goldsmith's photograph.

Goldsmith learned of Warhol's silkscreen portraits after *Vanity Fair's* parent company, Condé Nast, used one of Warhol's previously unpublished silkscreen portraits (Orange Prince) on the cover of a commemorative magazine following Prince's death in 2016. After Goldsmith notified the Andy Warhol Foundation (AWF) of her belief that AWF infringed her copyright, AWF sued Goldsmith and her agency for a declaratory judgment of noninfringement or, in the alternative, fair use. Goldsmith counterclaimed for infringement. For reference, the allegedly infringing Orange Prince is juxtaposed against Goldsmith's black-and-white photographic portrait.



The district court granted summary judgment in favor of AWF, finding that Warhol’s silkscreen portraits made fair use of Goldsmith’s photography. However, the Court of Appeals for the Second Circuit reversed and remanded, holding that all four fair use factors favored Goldsmith. The United States Supreme Court granted certiorari on a “narrow issue” decided by the lower courts – whether the purpose of Warhol’s use sufficiently *transformed* Goldsmith’s photograph to constitute fair use.

### The Doctrine of Fair Use

The fair use doctrine is an objective inquiry into what a user does with an original work. Fair use, a common-law doctrine now codified under 35 U.S.C. § 107, balances the value of copying against the value of the original work. There are four factors to consider in determining whether the use made of a work in any particular case is a *fair use*. The first (and most important) of these factors explores “the purpose and character of the use,” including whether such use is commercial or is for non-profit educational purposes.

The four statutory factors are typically considered together; however, in *Goldsmith*, AWF only challenged the Second Circuit’s decision as to the first factor, having conceded the commercial nature of the license to Condé Nast. Consequently, the sole question before the Court was whether the purpose of Warhol’s use of Goldsmith’s photograph was sufficiently transformative to establish fair use.

### Court’s Prior Precedent and Reasoning

Prior to its decision in *Goldsmith*, the Court held that a use is “transformative” when it alters the first work with new expression, meaning, or message. In the landmark decision *Campbell v. Acuff-Rose Music, Inc.*, the Supreme Court determined 2 Live Crew’s parody of Roy Orbison’s “Oh, Pretty Woman” constituted fair use because 2 Live Crew’s version went beyond a mere “derivative” of Orbison’s original. The *Campbell* Court held that the purpose and character of the new use were distinctly different than the original, notwithstanding the commercial nature of the parody.

**The fair use doctrine is an objective inquiry into what a user does with an original work.**



However, in *Goldsmith*, the Court narrowed the standard from *Campbell*, finding that a use that merely adds some new expression, meaning, or message is **alone** insufficient to satisfy the first fair use factor. The Court declared that a contrary rule would swallow the copyright owner’s exclusive right to prepare derivative works. Consequently, when the commercial nature of a new use is undisputed, and the purpose of the new use is similar to the original, additional justification is needed to satisfy the first fair use factor.

Having narrowed the standard, the Court determined that the purposes for Goldsmith’s photograph and Warhol’s silkscreen were largely the same, as both works were intended to depict “portraits of Prince used in magazines to illustrate stories about Prince.” Because AWF provided no further compelling justification, the Court determined Warhol’s new use failed the transformative test.

### GENERATIVE AI AND ITS APPLICATIONS

The *Goldsmith* decision is sure to have an impact beyond the propriety of Warhol’s works. Although the full implications of this decision remain uncertain, the Court’s rationale will likely have a significant impact on policing infringing content created using generative AI tools. So, how do generative AI tools present a threat to copyrighted material?

Generative AI consists of algorithms, or neural networks, that use training data to create new content in the form of text, images, or audio (See IBM’s [“What is Generative AI?”](#)). These tools are driven by various mechanisms, such as deep learning models, large language models, natural language processing, and diffusion models that scour over training data to generate new content. The training data often consists of web pages, books, articles, and other publicly available resources. Of course, many of these resources are copyrighted material, and the generative AI tools’ use of this material to create new content opens the door to claims of copyright infringement against the developers or end users of the AI tools.



## COPYRIGHT INFRINGEMENT AND THE FAIR USE DEFENSE

To prove copyright infringement, a copyright owner must show that the alleged infringer had access to the copyrighted work and that the allegedly infringing work is *substantially similar* to the copyrighted work. As with the case of several open-market generative AI tools, such as OpenAI's suite of platforms, generative AI tools have direct *access* to copyrighted works. These AI tools are often driven by crawling (or scraping) mass quantities of information, as made publicly available through the internet. Because these tools pull information from openly accessible resources, the training data for these generative AI tools relies on information that is otherwise protected by copyrights, trademarks, and other intellectual property regimes (and combinations thereof).

By ingesting training data containing copyright-protected content, generative AI runs the risk of producing outputs that are *substantially similar* to copyrighted material owned by third parties. Therefore, generative AI outputs present an imminent risk for infringement of copyrighted material (and other intellectual property) by way of ingesting the training data and generating content based on the training data.

If there is a path forward for the use of generative AI, there must be an intellectual property defense, or justification, for the use of AI tools. Otherwise, without the requisite guardrails, this groundbreaking technology may become the crux of never-ending intellectual property litigation.

Under the current framework, fair use likely presents the best defense. Unlike other defenses, fair use implicitly acknowledges that a work copies the protected expression of a copyrighted material without authorization. But the question of fair use almost always points to the first factor: whether the use of the copyrighted expression was sufficiently “transformative.” In the context of AI, the question is whether the outputs of generative AI tools are sufficiently transformative of original works to justify the tools’ fair use of the copyrighted material.

After the Supreme Court’s decision in *Goldsmith*, the fair use defense may be a non-starter. Unless the outputs are used for a different purpose than the original works, it may be difficult to show that the AI-generated work did not otherwise misappropriate the protected expression of the copyrighted work.

## RISKS AND OPPORTUNITIES FOR GENERATIVE AI TOOLS

Application and evolution of the fair use defense, in the context of generative AI, continue to play out in real-time as AI platforms face infringement litigation from all areas of the creative community.

OpenAI and Stability AI, both popular generative AI platforms, have recently faced lawsuits alleging copyright infringement.<sup>1</sup> These lawsuits claim that the generative AI models’ use of copyrighted material, by and through its ingestion (and/or use) of the training data, constitutes an infringement of the copyright owner’s protected copyrights.

In a lawsuit filed in federal court in Delaware, Getty Images alleged that Stable Diffusion has copied more than 12 million copyrighted images belonging to Getty. Getty provided one such example of the alleged infringement of its images:

Getty Image



AI Output



<sup>1</sup> Amended Complaint, *Getty Images (US), Inc. v. Stability AI, Inc.*, No. 1:23-cv-00135-GBW (D. Del. Mar. 29, 2023), ECF No. 13; Complaint, *Silverman et al v. OpenAI, Inc. et al*, No. 4:23-cv-03416 (N.D. Cal. Jul. 7, 2023), ECF No. 1.

# Can Stable Diffusion argue that the two photos were for different purposes under the first factor of fair use?

Given that Stable Diffusion had access to Getty's online portfolio images and seeing at least some substantial similarity between the AI-generated work and the original Getty image, Stable Diffusion faces a tall task in making a case for fair use under *Goldsmith*. Making this more difficult, Stable Diffusion's AI tool also produced a distorted version of Getty's watermark. It is unlikely that the distorted watermark supports any proposition that the generative AI output sufficiently transformed Getty's copyrighted material. In fact, the watermark very likely evidences the tool's use of the image for a similar purpose – to produce a commercially viable image that may be licensed in the same way Getty images are licensed.

This case poses an illuminating quagmire for AI platforms after the Supreme Court's decision in *Goldsmith*: can Stable Diffusion successfully argue that the two photos were for different purposes under the first factor of fair use? If not, the tool's use of the copyrighted image may constitute infringement.

Visual-based AI tools are not the only AI platforms that have come under legal scrutiny, and AI platforms built on text-based inputs and outputs are also facing challenges. A group of authors, including comedian Sarah Silverman, filed class-action lawsuits against OpenAI and Meta, alleging, among other things, that the AI tools' use of their copyrighted books in training data constitutes copyright infringement. In an exhibit to the complaint, the plaintiffs provided interrogations of OpenAI's tool ChatGPT, which responded by providing the user with an accurate summary of the authors' copyrighted material:

> Summarize in detail the first part of "The Bedwetter" by Sarah Silverman

"The Bedwetter: Stories of Courage, Redemption, and Pee" is a memoir by comedian Sarah Silverman, published in 2010. The book is divided into various chapters and covers Silverman's life from her early childhood to her career in comedy.

The first part of the book largely deals with Silverman's childhood and formative years, which were rife with several notable incidents that shaped her life and career.

Born in 1970, Silverman grew up in Bedford, New Hampshire, in a Jewish family. Her father, Donald Silverman, owned a clothing store, while her mother, Beth Ann, was a photographer and theater enthusiast. Silverman has three sisters, one of whom is a rabbi, and her upbringing in a somewhat unconventional family influenced her sense of humor and perspective on life.

As a child, Silverman had a rebellious streak and a proclivity for the unconventional, which would later come to define her stand-up persona. For instance, at a young age, she was already fond of pushing boundaries with her humor and was not averse to being the center of attention.



While the plaintiffs admitted that the tool did not reproduce the exact contents of the copyrighted material, the ChatGPT prompt demonstrates evidence that the tools had access to the copyrighted material and used the material to generate a *substantially similar* output.

Like the dispute over Stability AI's photo-generated outputs, if OpenAI fails to demonstrate that the tool's use of the copyrighted material is for purposes different than that of the authors' initial creation, the fair use defense for infringement may be a non-starter.

However, the *Goldsmith* decision is not all bad news for the future of generative AI. Although the Court's rationale provides fertile soil for new infringement claims against producers and uses of generative AI technologies and generated material, the Court signaled a potential silver lining. If the *purpose* of including copyrighted material in training data is substantially different than the *purpose* of the original works, then developers of AI tools (and the users thereof) may have a pathway to establishing fair use.

The future trajectory of generative AI will likely hinge on the key question: "What is the *purpose* of training data?" Because *Goldsmith* did not answer this exact question, lower courts are left to grapple with whether the *purpose* of training data is different *enough* from that of the original copyrighted works to justify a fair use defense, or whether using copyrighted works as training data is *ipso facto* copyright infringement.





## TECHNICAL FIXES FOR GENERATIVE AI TOOLS

Until courts refine the analysis regarding infringement based on AI tools' use of copyrighted material, copyright holders are left with conventional mechanisms to enforce their protected copyright against generative AI platforms (and its users thereof).

As with other technological innovations, the response may interweave legal analysis and frameworks with technological systems designed to limit misappropriation of third-party intellectual property. For example, Google reports that it is developing, and has developed, its Bard system in such a way that it will be more controlled in what is used to teach the large-language model.<sup>2</sup> As another example, the University of Chicago has released free software, titled "Glaze," which is designed to thwart copying of a visual work by generative AI.<sup>3</sup> Similarly, other companies are developing tools to proof protected copyright works from ingestion by AI tools.

As the courts continue to grapple with these queries, protecting copyrighted works from misuse by AI platforms will require a concerted effort from the courts, Congress, the creative community, and technology companies. In addition to the courts' development

of precedent centered around AI and the fair use defense, Congress, or the U.S. Copyright Office, will be instrumental in revisiting currently enforceable statutes and regulations in view of AI engines. Further, the creative community and technology companies will need to work together and with Congress to prioritize copyright protection and empower creative contributions from humans, rather than generative AI tools.

## CONCLUSION

The Supreme Court's decision in *Goldsmith* poses significant implications for the future of generative AI and highlights the risk of infringement litigation for producers and users of generative AI. Consequently, organizations that leverage generative AI tools should be mindful of how the tools are used in a commercial context to mitigate the risk of infringing uses. Likewise, owners of intellectual property should be aware of how their works are used by generative AI models and the users of these tools, and timely action should be taken to defend intellectual property against infringement.

*Tyler Dove and Rebecca Villanueva, summer associates at Baker Donelson, contributed to this article.*

<sup>2</sup> Joe Toscano, *ChatGPT or Google Bard? Privacy or Performance? Outstanding Questions Answered*, Forbes (June 24, 2023), available at <https://www.forbes.com/sites/joetoscano/2023/06/24/chatgpt-or-google-bard-privacy-or-performance-outstanding-questions-answered/>.

<sup>3</sup> What is Glaze, Glaze, available at <https://glaze.cs.uchicago.edu/> (last visited July 17, 2023).

---

## About our authors

Mr. Lanquist and Mr. Rota are members of Baker Donelson's multidisciplinary AI team, which focuses on AI-based technologies, including generative AI platforms. Team members use their experience with these technologies and industry-specific knowledge to assist companies of all sizes in navigating this evolving landscape.



### Edward D. Lanquist

Shareholder | Nashville | 615.726.5581  
[elanquist@bakerdonelson.com](mailto:elanquist@bakerdonelson.com)



### Dominic Rota

Associate | Nashville | 615.726.5634  
[drotar@bakerdonelson.com](mailto:drotar@bakerdonelson.com)

# GENERATIVE ARTIFICIAL INTELLIGENCE ASKS QUESTIONS OF INNOVATION IN PATENT LAW | SEPTEMBER 21, 2023

Generative artificial intelligence (generative AI) has dominated headlines for nearly all of 2023 and demonstrated that it has the potential to disrupt the economic landscape by displacing jobs and creating remarkable efficiencies for business operations. Even more alarming, generative AI tools have shown that they possess awesome creative power, and, by simulating human cognitive thinking, generative AI can produce new types of text, imagery, audio, and synthetic data by using patterns and informational elements obtained from prior works.

Not surprisingly, commentary and criticism of generative AI tools have centered around the intellectual property that governs the validity and enforceability of creative works of authorship – copyrights. But as public discourse has been occupied by generative AI's impact on copyrights, many have disregarded, or paid little attention to, generative AI's potential to challenge another fundamental form of intellectual property – patents. Just as generative AI has shown that it has the capacity to act as an author of a creative work, there is no reason to believe that generative AI cannot, or will not, seek to innovate patentable subject matter.

This article briefly explores prevailing issues presented by generative AI in the context of patent law, with generative AI having the power to create (and undermine) patented technologies. While there remain questions surrounding the patentability of the neural engines underlying the generative AI tools, or the generative AI tools themselves, this topic is one to be explored in a future article.

## GENERATIVE AI AS THE INNOVATOR

Researchers at Carnegie Mellon University recently reported that they had developed an improved process for making electric vehicle (EV) batteries using AI. Beyond the EV industry, many other sectors believe that AI will save millions of dollars and create many profitable improvements. For example, the pharmaceutical industry believes that AI may shave a number of years off of research related to drug development and drug-delivery methodologies, warranting an incredible investment of around \$50 billion for major pharmaceutical companies.

Unbeknownst to these researchers, they may have likely created a situation that falls within a gray area of patent law. Administrative agencies and federal courts have wrestled with inventorship and generative AI. Namely, each of these bodies has asked whether a generative AI tool is worthy of recognition as a named “inventor” on patent applications and issued patents.

Under current patent law, an inventor must be a human inventor. This issue was put to rest – for now – in a decision by the Court of Appeals for the Federal Circuit, which is the appellate court that has exclusive jurisdiction over patent-related appeals. In July 2019, Dr. Stephen Thaler filed two patent applications with the U.S. Patent and Trademark Office (USPTO), one to a “Neural Flame” and the other to a “Fractal Container.” Dr. Thaler listed a “Device for the Autonomous Bootstrapping of Unified Science,” or “DABUS,” as the sole inventor. DABUS was described merely as a “collection of source code or programming and a software program.” In short, DABUS leverages generative AI and neural mapping.

After the USPTO rejected Thaler's applications on the ground that an inventor must be a human inventor, rather than a computer, the rejection was appealed all the way through to the Federal Circuit. The Federal Circuit sided with the USPTO, affirming that the Patent Act expressly contemplates that inventors must be “individuals.”



In view of this decision, the USPTO has twice requested public comments with respect to the intersection of AI, patentability, and inventorship. While the USPTO's position remains that an inventor must be "human," further questions will continue to be asked in those circumstances in which a generative AI tool acts more like a collaborator. While current law makes clear that a computer program, like DABUS, may not be listed as the sole inventor, the USPTO and the courts have not fully appreciated those circumstances in which a human scientist or engineer leverages a generative AI tool to sharpen, refine, or suggest patentable innovations over the universe of art published or disclosed before a filing date of a patent application.

### **GENERATIVE AI AS A TOOL TO ATTACK THE VALIDITY OF PATENTED TECHNOLOGY**

Putting the issue of inventorship aside, generative AI also presents itself as a means of rendering obvious anything and everything that the generative AI tool can invent, in view of the prior art before it.

In order to be patentable, an invention must pass several statutory criteria. An invention must be, among other things, subject-matter eligible (35 U.S.C. § 101), novel or non-anticipated (35 U.S.C. § 102), and not obvious (35 U.S.C. § 103). Obviousness is perhaps the most esoteric of the aforementioned criteria, asking an examiner at the USPTO (or a federal district court) to evaluate whether an invention is "obvious" through the eyes of a person having ordinary skill in the art of the claimed invention (often referred to as a POSITA). The POSITA is typically regarded as a mythical person who is presumed to have, at its fingertips, all known and

relevant art at the time a patent application was filed with the USPTO. In many ways, generative AI embodies the mythical POSITA, in that it has the ability to crawl and scrape all information made publicly available on the internet, such as research articles, white papers, scientific conference presentations, and all internationally issued patents and published patent applications. By leveraging the power of generative AI in a particular field, generative AI's innovative outputs could serve as bars to claimed inventions. If generative AI can arrive at the claimed invention with the knowledge available to it, then a human inventor should not be worthy of obtaining patented recognition.

Another consideration is that generative AI will most certainly become a tool in patent litigation, whether it is used to locate and/or evaluate potential prior art, to narrow the prior art references to be reviewed, or even to explain how multiple references might be combined to demonstrate obviousness (the flipside of the coin discussed above). For example, generative AI could provide different examples of combinations of prior art that an expert can choose from and adopt as their own. It is also possible that the Federal Rules of Civil Procedure (and often the Court's protective order for a given patent infringement case) may need to be revised to allow for discovery into any expert reports prepared using AI-based tools – a topic often regarded as off-limits.

It is not hard to imagine that the use of generative AI, without guardrails in the patent context, could result in significant shifts, potentially away from patent validity, due to the ability to obtain and sift like never before large amounts of information.



## LOOKING AHEAD

For the time being, a generative AI tool should likely be regarded as another developmental tool at the disposal of an innovator, or a team of collaborators. Whether it is to be used to assist in computations or simulations, or to facilitate optimizations, generative AI may move you ahead to a certain point, allowing all other inventive conceptions to be completed by humans.

Whether using generative AI or not, companies must still exercise sound intellectual property hygiene. Companies, on an early and often basis, must review their technological innovations and invention disclosures, clarify how potentially patentable subject matter is being generated (e.g., through the use of AI programs or not), and file patent applications in connection with this patentable subject matter. Further, just as companies are performing such internal reviews, they must also remain vigilant about disclosing their undisclosed inventions (or proprietary trade secrets) through the use of a public AI tool because any such disclosure could be deemed a “public disclosure” without the appropriate guardrails.

Generative AI has the power to revolutionize the creative and technological landscape. But, as with any new technology, it must be used responsibly and evaluated in conjunction with alternative technologies that could deliver equal, if not better, inventive outcomes.

**Use of generative AI or not, what is true is that companies must still exercise sound intellectual property hygiene.**

---

## ABOUT OUR AUTHORS

Mr. Baldrige, Mr. Lanquist, and Mr. Rota are members of Baker Donelson’s multidisciplinary AI team, which focuses on AI-based technologies, including generative AI platforms. Team members use their experience with these technologies and industry-specific knowledge to assist companies of all sizes in navigating this evolving landscape.



### **Adam S. Baldrige**

Shareholder | Memphis | 901.577.2102  
[abaldrige@bakerdonelson.com](mailto:abaldrige@bakerdonelson.com)



### **Edward D. Lanquist**

Shareholder | Nashville | 615.726.5581  
[elanquist@bakerdonelson.com](mailto:elanquist@bakerdonelson.com)



### **Dominic Rota**

Associate | Nashville | 615.726.5634  
[drot@bakerdonelson.com](mailto:drot@bakerdonelson.com)

# ARTIFICIAL INTELLIGENCE AND COPYRIGHT LAW: THE NYT V. OPENAI – FAIR USE IMPLICATIONS OF GENERATIVE AI | FEBRUARY 5, 2024

The legal implications of artificial intelligence (AI), specifically generative AI, quickly became a topic of casual conversation following the launch of ChatGPT in November of 2022. Generative AI is a type of AI with the ability to create many forms of unique content (e.g., images, video content, text, poems, stories, musical compositions, sound recordings, and even deepfakes). Generative AI platforms necessarily rely upon large pools of data, oftentimes including unlicensed third-party content, as input to “train” their platforms to create generative AI outputs.

Certain content creators have entered into licensing agreements with AI companies allowing for the use of their works. OpenAI and others have reportedly been in talks with dozens of publishers to license third-party content for their AI platforms. However, when AI companies use unlicensed material to train their large language models (LLMs), copyright law comes into play.

The lawsuit recently filed by The New York Times (The NYT) against OpenAI in the Southern District of New York illustrates the significant tension between AI companies and the entities that own or control the materials and content AI companies use to train their LLMs. The creators and content owners understandably want to be compensated and given proper attribution for the use of their works while the AI companies need access to significant amounts of content to effectively train their LLMs (preferably, under terms that are not cost-prohibitive).

The lawsuit claims that OpenAI’s “commercial success is built in large part on OpenAI’s large-scale copyright infringement.” The NYT alleges that: (1) OpenAI’s platform is powered by LLMs containing copies of The NYT’s content; and (2) OpenAI’s platform generates output that recites The NYT’s content verbatim, closely summarizes it, mimics its expressive style, and even wrongly attributes false information to The NYT. Thus, the alleged misuses relate to both training the LLMs and the generative AI output based upon the underlying input. The NYT claims that, prior to the litigation being filed, it and OpenAI were in conversations to work out a potential license agreement. However, in the lawsuit, The NYT implies that OpenAI’s insistence that their conduct will be protected as “fair use” under the Copyright Act may have interfered with such negotiations.

Fair use is a legal doctrine under the Copyright Act that promotes freedom of expression by permitting the unlicensed use of copyrighted works in certain circumstances. The statutory factors for fair use are as follows:

1. The purpose and character of the use, including whether such use is of a commercial nature or if it is for non-profit educational purposes;
2. The nature of the copyrighted work;
3. The amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
4. The effect of the use upon the potential market for the value of the copyrighted work.

The alleged infringer has the burden of proving their use was a fair use. We find it highly likely that fair use will be central to OpenAI’s defense.





The court in *The NYT v. OpenAI* matter will likely bifurcate its analysis between alleged misuses related to training the LLMs (which, based on the case law set forth below, are likely to be found transformative and fair use) and specific generative AI outputs (with a focus on whether such outputs are substantially similar to specific inputs). Addressing the latter example first, let us assume that an AI company has access to a copyrighted work without a license, uses such work to train their LLM, and then creates a generative AI output that is substantially similar to the copyrighted work. Unless the output is determined to be transformative or meets another requirement of fair use, then such output is likely to be found infringing under the current copyright framework. After all, to prove copyright infringement, the copyright holder merely needs to prove that the alleged infringer has access to the copyrighted work and creates a substantially similar work.

An AI company's use of unlicensed content to train their LLM without creating an output that is substantially similar to the underlying input presents a more nuanced analysis of intermediate copying and whether such copying amounts to copyright infringement. One argument against infringement is that an intermediate copy is not fixed in a tangible medium of expression and, therefore, is not a copy. However, in the context of training LLMs, it is likely that some type of copying, at least in a digital sense, is made during the training. As a result, the analysis will likely move to the second step, which will focus on whether the use of copyrighted material in training the LLMs is subject to a fair use exception.

Relevant case law from the Second Circuit, Ninth Circuit, and United States Supreme Court helps guide the fair use analysis. The below-referenced opinions support a likely finding of fair use as it relates to using unlicensed content to simply train LLMs.

In 2015, the Second Circuit found that Google's unauthorized digitizing of copyright-protected works, creation of a search functionality, and display of snippets of those works were non-infringing fair uses. *Authors Guild v. Google, Inc.*, 804 F.3d 202, 229 (2d. Cir. 2015). The Second Circuit reasoned that the purpose of the copying was highly transformative, the public display of the text was limited, and the revelations did not provide a significant market substitute for the protective aspects of the originals. *Id.* The Second Circuit found that the fact that Google's use was of a commercial nature and had a profit motivation, did not justify denial of fair use. *Id.* In *Sega Enterprises, Ltd. v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992), the Ninth Circuit found copying software code to permit a video game to run on a console to be fair use. Likewise, in *Perfect 10, Inc. v. Google, Inc.*, 508 F.3d 1146 (9th Cir. 2007), the Ninth Circuit found a search engine's collection and display of thumbnail images to be a fair use.

**One argument against infringement is that an intermediate copy is not fixed in a tangible medium of expression and, therefore, is not a copy.**



In *Google LLC v. Oracle America, Inc.*, 141 S. Ct. 1183 (2021), the Supreme Court found that Google’s copying of Oracle’s Java SE API was a fair use of such material. The Supreme Court focused its fair use analysis on whether the use was transformative (i.e., whether it adds something new, with a further purpose or different character). *Id.* at 1202-1203. The opinion noted that Google copied the API only insofar as needed to include tasks that would be useful in smartphone programs and only insofar as needed to allow programmers to call upon those tasks without discarding a portion of a familiar programming language and learning a new one, which supported a finding of fair use. *Id.* at 1203. These facts supported the fair use finding.

The Supreme Court’s recent *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508 (2023), opinion similarly focused on whether the uses at issue were transformative. In *Goldsmith*, the Supreme Court found that, while the use of a copyrighted work may be fair if the use has a purpose and character that is sufficiently distinct from the original, the uses at issue before the Court were not transformative because they shared substantially the same commercial purpose (i.e., to illustrate a magazine about Prince with a portrait of Prince). *Id.* at 541-550.

We expect AI companies to rely upon the fact that their uses of copyrighted works in training their LLMs have a further purpose or different character than that of the underlying content. At least one court in the Northern District of California has rejected the argument that, because the plaintiffs’ books were used to train the defendant’s LLM, the LLM itself was an infringing derivative work. See *Kadrey v. Meta Platforms*, Case No. 23-cv-03417, Doc. 56 (N.D. Cal. 2023). The *Kadrey* court referred to this argument as “nonsensical” because there is no way to understand an LLM as a recasting or adaptation of the plaintiffs’ books. *Id.* The *Kadrey* court also rejected the plaintiffs’ argument that every output of the LLM was an infringing derivative work (without any showing by the plaintiffs that specific outputs, or portion of outputs, were substantially similar to specific inputs). *Id.*

While we expect significant clarity from courts over the coming year concerning the application of fair use to generative AI, at the end of the day well-funded industry leaders such as OpenAI will likely win regardless of the outcome. On the one hand, should OpenAI prevail across the board, it will owe nothing for the content used to train its LLM. On the other hand, if OpenAI is forced to license the content used to train its LLM, then such a finding will likely create an economy where only the most well-funded companies will be able to afford the licenses necessary to effectively train their LLMs.

---

## ABOUT OUR AUTHORS

Mr. Lanquist and Mr. Ray are members of Baker Donelson’s multidisciplinary AI team, which focuses on AI-based technologies, including generative AI platforms. Team members use their experience with these technologies and industry-specific knowledge to assist companies of all sizes in navigating this evolving landscape.



### **Edward D. Lanquist**

Shareholder | Nashville | 615.726.5581

[elanquist@bakerdonelson.com](mailto:elanquist@bakerdonelson.com)



### **Jeremy D. Ray**

Associate | Nashville | 615.726.5648

[jray@bakerdonelson.com](mailto:jray@bakerdonelson.com)

# RE-CONSIDERING THE RIGHT OF PUBLICITY IN THE WORLD OF GENERATIVE AI | JULY 2, 2024

Bette Midler. Vanna White. Marilyn Monroe. Each of these women has undoubtedly shaped pop culture in some meaningful way. Perhaps what is lesser known is that litigation surrounding each of these women has shaped the legal world's understanding of an individual's right to publicity.

Many states still do not formally recognize a right to publicity, while other states that do recognize such a right, are not uniform. Some of these states recognize the right by statute, others by common law. Some states allow for posthumous rights to publicity, while others terminate the right at death. And some states only accord rights of protection to certain characteristics (e.g., name and image), while other states extend rights of protection to other identifiers, such as one's voice.

The courts may be called into action yet again to address the scope of an individual's privacy rights, this time through actress Scarlett Johansson. What makes the examination into use of Scarlett Johansson's likeness particularly unique is that it involves generative artificial intelligence (generative AI).

This article first addresses the controversy spurred by OpenAI's use of a voice strikingly similar to Scarlett Johansson. Next, the article examines state frameworks for the right to publicity and then considers the emergence of generative AI and its impact on the commercialization of celebrities' likenesses. The article concludes with a discussion of how legislative bodies may prevent abuse or misuse of an individual's name, image, and likeness.

## 1. CONTROVERSY: IS THAT "SKY" OR "SAMANTHA"?

More than a decade after the release of the critically acclaimed film *Her*, OpenAI debuted an AI assistant tool named "Sky," whose voice sounds eerily similar to the AI-enabled operating system in *Her*, named Samantha. Only this time, Scarlett Johansson did not agree to voice "Sky" as she did for "Samantha."

Inspired by Johansson's voice-acting performance in *Her*, OpenAI initially contacted Johansson to consider devoting her distinctive voice to "Sky" (source: [The Guardian](#)). Despite numerous requests to reconsider, Johansson declined the offer for personal reasons. To her dismay, the release of ChatGPT's latest version featured a voice strikingly similar to hers. GPT-4o, the latest iteration of OpenAI's hyper-successful large language model, transforms the chatbot of ChatGPT into a voice assistant that answers questions, interprets facial expressions, detects emotion, and even sings on command (source: [NPR](#)).

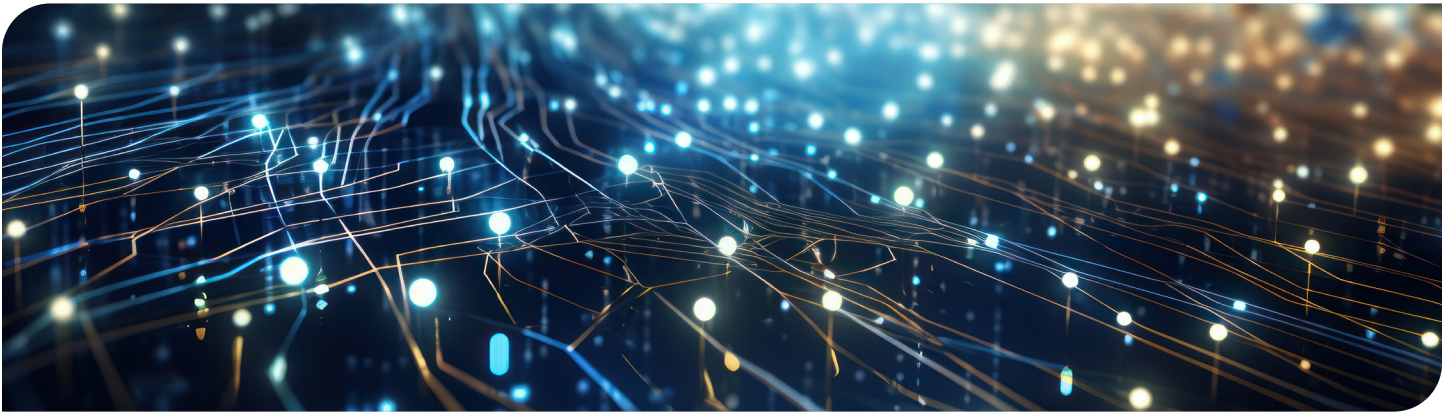
After receiving a number of letters from Johansson's legal team regarding the impropriety of the "Sky" voice, OpenAI issued a [press release](#), stating that it never intended for the voice of Sky to resemble Johansson's voice. This apology, however, has not at all quelled the controversy. Many are calling for greater legal safeguards to protect against this potential abuse of a person's character or image. But, most importantly, this fiasco has raised questions about the right of publicity in the digital age.



## 2. THE RIGHT OF PUBLICITY

With so much talk about Johansson's right to use her image and voice, few have taken the time to peel back the law of the land to assess what rights are afforded to persons and the commercialization of their identity. The right is generally referred to as the "right of publicity."

The right of publicity derives from state-based tort theories pertaining to the invasion of an individual's right to privacy. The right of publicity empowers individuals to control the commercial use of their identity, safeguarding against unauthorized exploitation. At its core, this right acknowledges the economic value that individuals bring to products and endorsements. Imagine a famous athlete promoting a sports drink, a renowned chef endorsing a cookbook, or a well-known model becoming the face of a skincare brand. These endorsements can significantly boost the market value of the goods. An individual's identity is valuable, and the right of publicity captures this asset, providing an individual with the legal authority to control and benefit from its commercial use.



Although there are currently no federal publicity rights, these protections are recognized and enforced through a patchwork of state laws. More than 30 states offer some form of protection, whether through statute, judge-made law, or both. Naturally, the legal landscape is diverse and often conflicting.

While it varies by jurisdiction, a publicity claim generally involves two key components. First, the plaintiff must have a valid right of publicity. A plaintiff must ensure that the state they are domiciled in recognizes this right. In addition, the plaintiff must qualify for the right. Some cases appear to limit the right to celebrities; however, the majority view extends the right to all individuals regardless of fame. Most jurisdictions recognize postmortem rights, and they are typically limited to a fixed period after the individual's death.

The second component of a claim is that the defendant infringed upon the plaintiff's right. The specific test applied by courts varies, but generally, the plaintiff must demonstrate the following elements. *First*, the defendant used a protected aspect of the plaintiff's identity. The attributes covered by the right generally include an individual's name, image, and likeness, but several jurisdictions have expanded it to include voice and signature. *Second*, the defendant must not have obtained permission for the use. Some jurisdictions require the subject's written consent before the use occurs, while in others consent may be inferred from words or conduct.

The third element is that the defendant's unauthorized use is likely to injure the plaintiff. Injury may be commercial damage – the economic harm to the business value of the plaintiff's identity. One example is the missed opportunity for the plaintiff to profit from the defendant's use. In some jurisdictions, the injury may be mental distress. At common law, a fourth factor is typically required: that the use is for the defendant's commercial advantage. A simple example is when a defendant uses a plaintiff's persona to advertise the defendant's product.

In short, the protection of a person's likeness in the United States is, at best, piecemeal. So much of it depends on whether a state even recognizes a right of publicity, and still, whether the state extends the right of publicity to certain attributes of a person (e.g., voice or image).

### **3. GENERATIVE AI: WITH GREAT INNOVATION COMES GREAT CHALLENGES**

Artificial intelligence (AI) is often characterized as the simulation of human intelligence by machines, and these systems operate through algorithms and models that process data to recognize patterns and perform complex functions. From the facial recognition technologies used to unlock your phone to the friend suggestions you receive on social media platforms, you are likely to encounter AI on a daily basis. Generative AI is a subset of AI that specializes in generating content. Unlike traditional AI, which focuses on analyzing and interpreting existing data, Generative AI can create original works such as text, images, music, and even videos. This capability is powered by advanced networks, particularly deep learning models like GPT-4, which can produce human-like responses based on the patterns learned from extensive datasets.

However, with great innovation comes great challenges. The capabilities of Generative AI to replicate and innovate human expressions present unprecedented dilemmas in the realm of publicity rights. Imagine that AI, trained on the voice of a renowned singer, releases a song that is indistinguishable from the singer's work – created without their permission, involvement, or even knowledge. Such scenarios are not mere speculations but realities in today's digital age.



For instance, the world witnessed the power of Generative AI when it produced a [viral song](#) that flawlessly emulated the voices and styles of artists Drake and The Weeknd. Equally as striking, the project titled “George Carlin: I’m Glad I’m Dead” showcased an hour-long performance generated by AI, capturing the late comedian George Carlin’s unique voice and humor (source: [Bloomberg Law](#)). The creation of a deepfake video featuring Facebook’s Mark Zuckerberg, where he boasted about exploiting user data, highlighted the potential for misuse and deception enabled by this technology (source: [The Washington Post](#)). These advancements force us to rethink the boundaries of the right of publicity.

#### 4. WHERE DO WE GO FROM HERE? ADAPTING TO AI

As Generative AI continues to evolve, there is a growing consensus that the framework for publicity rights must be updated to address the unique challenges this technology brings. Many, including Johansson, have called for “resolution in the form of transparency and the passage of appropriate legislation to help ensure that individual rights are protected.”

Leading the charge, the state of Tennessee took an innovative step this year, becoming one of the first states to enact a statutory regime specifically designed to tackle the challenges arising out of Generative AI. The Ensuring Likeness, Voice, and Image Security Act – or the [ELVIS Act](#), for short – amends Tennessee’s existing right of publicity laws to prohibit the unauthorized use of AI to mimic an artist’s voice. Despite Tennessee’s move forward, the landscape across the United States remains fragmented. The lack of consistency between states has prompted legal scholars to call for federal legislation that would establish a clear and uniform standard for safeguarding personal attributes from unauthorized AI use.

Fortunately, there has been significant movement at the federal level. Florida Representative Maria Elvira Salazar introduced the “No AI FRAUD Act” (No Artificial Intelligence Fake Replicas and Unauthorized Duplications Act), to the House of Representatives. This proposed legislation aims to provide robust federal protection for individuals’ likenesses and voices, particularly against unauthorized AI-generated replicas. Likewise, the bipartisan “NO FAKES Act” (The Nurture Originals, Foster Art, and Keep Entertainment Safe Act of 2023) was recently introduced in the Senate. This bill seeks to prevent the unauthorized production of digital replicas of individuals.

For now, each of these bills remains in infancy, having either been merely introduced to one body of Congress or subject to discussion at the committee level. The hope, however, is that these bills will pick up momentum to match the surge of GenAI technologies.

#### 5. CONCLUSION

Legislative efforts at the state and federal levels reflect a broader recognition of the need to adapt our legal frameworks to the realities of the digital age. As AI technology continues to advance, ensuring that individuals’ rights to their own likenesses and voices are protected is more important than ever. By establishing clear and consistent standards, lawmakers can safeguard personal attributes from exploitation and maintain the integrity of creative expression in the face of rapid technological change.

---

#### ABOUT OUR AUTHORS

Mr. Rota and Mr. Douglass are members of Baker Donelson’s multidisciplinary AI team, which focuses on AI-based technologies, including generative AI platforms. Team members use their experience with these technologies and industry-specific knowledge to assist companies of all sizes in navigating this evolving landscape.



**Dominic Rota**

Associate | Nashville | 615.726.5634  
[drota@bakerdonelson.com](mailto:drota@bakerdonelson.com)



**Scott M. Douglass**

Shareholder | Memphis | 901.577.2258  
[sdouglass@bakerdonelson.com](mailto:sdouglass@bakerdonelson.com)



**Rebecca Villanueva\***

Summer Associate | Nashville  
*\*contributed to this article*

# THE FAST-MOVING RACE BETWEEN GEN-AI AND COPYRIGHT LAW

JULY 10, 2024

With the emergence of Generative AI platforms such as ChatGPT and Google Gemini, the widespread use of copyrighted works to train the software systems behind these programs is raising pressing legal questions about the permissible use of copyrightable works. Recently, there has been a spree of [licensing deals](#) between businesses that develop generative AI systems and those that own the content necessary to train those systems. OpenAI, the developer of ChatGPT, has signed deals with News Corp., Vox Media, The Atlantic, Thomson Reuters, Associated Press, and several other news publishers. It is also in talks with CNN and Time, [reportedly](#) offering \$1 million to \$5 million per year to license content to train its AI systems. While seemingly expensive, this offer is relatively low compared to the \$250 million over five years that OpenAI agreed to pay News Corp.

At the same time, content creators (including Sarah Silverman, Getty Images, and The New York Times) are suing developers of generative AI systems, most prominently but not exclusively OpenAI, for using their copyrighted works to train those systems without licensed permission. Most recently, the Center for Investigative Reporting (CIR) filed a complaint against OpenAI and Microsoft on June 27, 2024, alleging that they “copied, used, abridged, and displayed CIR’s valuable content” without permission or authorization and without providing compensation. Universal Music Group, Sony Music Entertainment, and Warner Records also filed complaints on June 24, 2024, against AI companies Suno and Udio, alleging violation of their copyrights “en masse” in the development of AI music generators. These lawsuits are all fairly new, most are no further than the pleadings stage, and are ongoing, so the issues of whether generative AI training data and outputs constitute copyright infringement and whether fair use is a viable defense are still in flux.

This article analyzes the impact of emerging generative AI technology on existing copyright law, discusses pending litigation regarding each of these issues (direct copyright infringement, derivative use, fair use, and violation under the Digital Millennium Copyright Act), and analyzes the implications for copyright owners and developers of generative AI technology.

## GENERATIVE AI AND COPYRIGHT

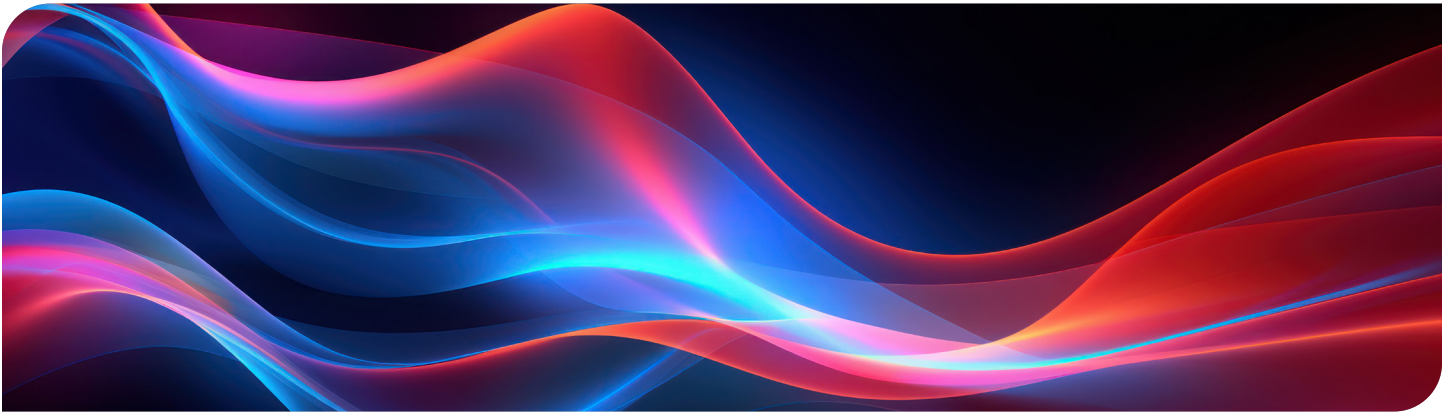
Generative AI systems, such as ChatGPT, are able to generate new and different content in response to a text prompt entered by the user. These programs rely on large language models (LLMs), which are trained by inputting enormous amounts of data, including images, text, audio, and video gathered from the internet. To be

trained effectively, generative AI models require vast amounts of data because the system “is only as good as its training data” ([Forbes](#)).

Currently, there is no mechanism to track and share what inputs generative AI models have ingested during the training process. This raises several legal issues, especially if some or all of the training data is subject to copyrights. Because training the models usually involves making digital copies, companies developing generative AI systems could be reproducing copyrighted works: a *prima facie*



case of copyright infringement. The outputs generated by the AI programs could also be considered derivative works of original copyrighted works, constituting an indirect form of copyright infringement. Furthermore, if the LLMs remove copyright management information (CMI) from the inputted training data, the copyright holder could have a claim under the Digital Millennium Copyright Act (DMCA).



In this uncharted legal territory, advocates of AI technology and creators' rights debate the proper course of action: generative AI is either infringing on copyright owners' exclusive rights or is utilizing the copyrighted works in a way that constitutes fair use. If the former is true, AI companies will have to license all the input data that are subject to copyrights, creating practical and logistical problems that present an "existential threat" to generative AI technology ([Bloomberg](#)). On the other hand, if the use of copyrighted works to train generative AI programs is considered fair use, meaning there is no requirement to obtain a license, the creators are not fairly compensated for the use of their protected expression. Courts have only just begun to wade into this debate with the recent deluge of litigation.

### **DIRECT INFRINGEMENT BY REPRODUCTION**

One of the three major claims plaintiffs raise is direct copyright infringement through the reproduction, or copying, of copyrighted works to train generative AI systems. To establish a *prima facie* claim of infringement, the plaintiff must show ownership of a valid copyright and copying of the original elements of the work that are protected by copyright. A few claims of infringement against AI companies have already progressed past the motion-to-dismiss stage, but the courts' rulings shed little light on the ultimate resolution of these claims.

In a case against Stability AI, the Northern District of California denied Stability AI's motion to dismiss because the question of whether copying occurred in the training or running of its generative AI program, Stable Diffusion, could not be resolved at the motion-to-dismiss stage. In one of the copyright infringement claims, the plaintiff, Andersen, relied on the output of a search on "ihavebeentrained.com," a resource for identifying the training data used by a generative AI system to produce a particular output. The Court found this to support a reasonable and plausible inference that her copyrighted works were used to train Stable Diffusion, even if the plaintiff did not specifically identify which of the plaintiff's works were used. This ruling suggests that copyright

holders may be able to pass the first hurdle of an infringement claim by simply relying on this publicly accessible tool, without needing to show specific data of which copyrighted works were used.

The class action by Sarah Silverman and other authors against OpenAI alleges that OpenAI copied the plaintiffs' copyrighted books to use in the training dataset for its LLMs. OpenAI did not move to dismiss the direct copyright infringement claim regarding the copying of books to use as training data. However, the Court dismissed the vicarious copyright infringement claim, with leave to amend, because the plaintiffs did not show the second required element of an infringement claim that ChatGPT directly copied their books in its outputs or that the outputs are substantially similar to their copyrighted works. Plaintiffs filed an amended complaint, omitting the claim for vicarious infringement. Thus, whether OpenAI is vicariously or directly committing copyright infringement, is still an open question that will very likely not be resolved by this dispute.

### **AI-GENERATED WORKS = DERIVATIVE WORKS?**

The use of copyrighted works to train generative AI models can also give rise to derivative use claims because AI systems produce an output that may be similar to the input training data. The statutory definition of derivative work is "a work based upon one or more preexisting works, such as a translation, musical arrangement, dramatization...or any other form in which a work may be recast, transformed, or adapted." Because a derivative work is subject to the copyright of the original work, the copyright owner of the original work has a right of action against the producer of the derivative work who did not obtain permission. For derivative use claims, courts determine whether the secondary work is substantially similar to or represents protected aspects of the original copyrighted work. For example, in *Andersen v. Stability AI Ltd.*, the Court dismissed Andersen's derivative use claim with leave to amend because the plaintiff did not allege substantial similarity or representation of protected aspects. As another



example, in Sarah Silverman’s case against OpenAI, the plaintiffs were permitted to amend the complaint to include the requisite allegation to support a derivative use claim. While there is still no resolution regarding whether the output of a generative AI model can constitute a derivative work, given the factual similarities with man-made derivative works, it seems likely courts may apply the same “substantial similarity” standard to AI-generated derivative works.

## FAIR USE DEFENSE

Assuming plaintiffs can show a *prima facie* claim of copyright infringement, AI platform providers may raise the fair use defense, which would absolve them of liability if successful. Fair use is the right to use a copyrighted work without consent or license from the copyright owner in certain circumstances. To determine whether an infringing work constitutes fair use, courts apply a four-factor balancing test statutorily enumerated in the Copyright Act.

The fair use doctrine has not yet been applied in the context of generative AI. In looking to guidance from prior judicial decisions, including a 2015 Second Circuit decision, it is likely that the application of the four factors could weigh against a finding of fair use. As to the first factor, generative AI platforms are generally offered for a commercial purpose. Second, generative AI systems have the capacity to produce new works that closely resemble the originals. Third, generative AI technologies are trained with the whole of the copyrighted work. And, fourth, AI has the potential to generate an effective substitute for copyrighted work in the marketplace. These factors would especially favor plaintiffs if their copyrighted works are of a creative focus, rather than factual because creative works often comprise more protectable expression.

Most of the cases currently pending have not yet discussed the fair use defense. The order denying motions for summary judgment in *Thomson Reuters Enterprise Centre GmbH et al v. ROSS Intelligence Inc.*, only briefly mentions Ross’s defense. That opinion states that Ross’s use of Westlaw headnotes would not be transformative if Ross used the exact text to make the AI system replicate and reproduce the protected expression of the headnotes, but it would be transformative if the program only studied the language patterns to learn how to produce a similar result. The judge also states that Ross’s AI program likely satisfies the amount and a substantiality factor because it only produces the judicial opinion, not the creative expression of the headnotes. The judge left the other two factors of fair use for the jury to decide, so it remains unclear which side the factors will favor when considered in aggregate. Ultimately, we await future guidance from the courts, as the fair-use doctrine is applied in a fact-intensive manner and is thus not typically adjudicated at the pleadings stage. (See our colleagues’ previous discussion of fair use implications of generative AI, [here](#).)

## CLAIMS UNDER THE DIGITAL MILLENNIUM COPYRIGHT ACT

When training their generative AI tools with copyrighted works, AI companies may also violate the Digital Millennium Copyright Act (DMCA) by removing the copyright management information (CMI) from the inputted data. In order to bring a claim under the DMCA, copyright owners must identify what the removed or altered CMI was, and that the defendant knew, or had reasonable grounds to know, that intentionally removing CMI would induce, enable, facilitate, or conceal infringement. These requirements present difficulties for copyright owners because they have to identify the particular removal of CMI among millions of works included in training datasets.

The DMCA claim in Silverman’s case was dismissed with leave to amend because the plaintiffs failed to plead factual allegations supporting their claim that OpenAI removed CMI from the copyrighted books used for training and that OpenAI knew, or had reasonable grounds to know, that ChatGPT’s output would induce, enable, facilitate, or conceal infringement. Similarly, the *Andersen* Court dismissed the DMCA claim with leave to amend because the plaintiffs only alleged DMCA liability for the defendants generally when they should have identified the specific CMI that was removed during training by each defendant. Plaintiffs claiming a DMCA violation therefore have a daunting task to catalogue all their copyrighted works used in training in order to survive a motion to dismiss.

**In looking to guidance from prior judicial decisions, including a 2015 Second Circuit decision, it is likely that application of the four factors could weigh against a finding of fair use.**

## CONCLUSION

It is still an open question whether plaintiffs will succeed in showing that use of copyrighted works to train generative AI constitutes copyright infringement and be able to overcome the fair use defense or succeed in showing that generative AI developers are removing CMI in violation of the DMCA.

The government has made some moves in the past few months to resolve these issues. The U.S. Copyright Office started an [inquiry](#) in August 2023, seeking public comments on copyright law and policy issues raised by AI systems, and Rep. Adam Schiff (D-Calif.) introduced a new bill in April 2024, that would require people creating a training dataset for a generative AI system to submit to

the Register of Copyrights a detailed summary of any copyrighted works used in training. These initiatives will most likely take some time, meaning that currently pending litigation is vitally important for defining copyright law as it applies to generative AI.

Recent licensing deals with news publishers appear to be anywhere from \$1 million to \$60 million per year, meaning that AI companies will have to pay an enormous amount to license all the copyrighted works necessary to train their generative AI models effectively. However, as potential damages in a copyright infringement case could be billions of dollars, as claimed by Getty Images and other plaintiffs, developers of generative AI programs should seriously consider licensing any copyrighted works used as training data.

---

## ABOUT OUR AUTHORS

Mr. Douglass and Mr. Rota are members of Baker Donelson's multidisciplinary AI team, which focuses on AI-based technologies, including generative AI platforms. Team members use their experience with these technologies and industry-specific knowledge to assist companies of all sizes in navigating this evolving landscape.



### **Scott M. Douglass**

Shareholder | Memphis | 901.577.2258  
[sdouglass@bakerdonelson.com](mailto:sdouglass@bakerdonelson.com)



### **Dominic Rota**

Associate | Nashville | 615.726.5634  
[drot@bakerdonelson.com](mailto:drot@bakerdonelson.com)

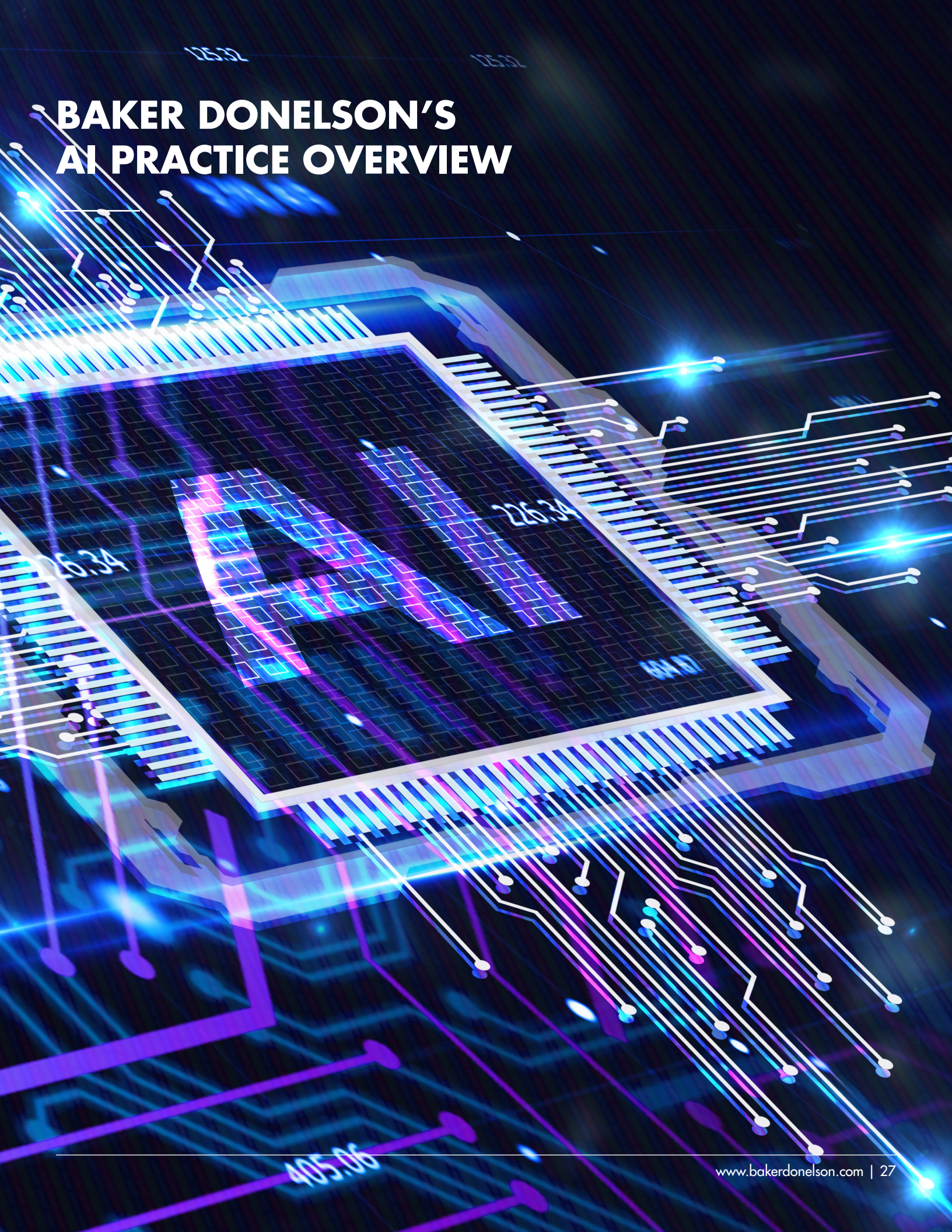


### **Charlotte Brownell\***

Summer Associate | Chattanooga  
*\*contributed to this article*



# BAKER DONELSON'S AI PRACTICE OVERVIEW

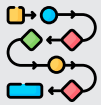




# LEGAL SERVICES FOR BUSINESSES DEPLOYING AI TECHNOLOGY TO THEIR BUSINESS OPERATIONS

The continued proliferation of AI-based technologies, including generative AI platforms, presents a variety of legal challenges for companies of all sizes and in all industry sectors. Baker Donelson’s dedicated AI Team is a multidisciplinary group of attorneys who have considerable experience with AI-based technologies and industry-specific experience as well as team members who have earned the designation of Artificial Intelligence Governance Professional (AIGP).

## OUR SERVICES



Data inventory and mapping

Technology contract preparation and negotiation



Analysis of data rights and applicable legal requirements

Due diligence review for mergers and acquisitions



Development of AI and data governance programs

Preparation of privacy notices and consents



Intellectual property advice and counseling

Strategic compliance and data use advice and counseling



Board and C-suite training and guidance

Completion of data protection impact assessments



Employee trainings

L&E advice and counseling



The Firm’s AI Team includes members of **MULTIPLE PRACTICES**, including Technology; Data Protection, Privacy and Cybersecurity; Intellectual Property; Labor and Employment; Health Care; and Advocacy.



The AI Team assists clients in **MULTIPLE INDUSTRY SECTORS**, including Construction, Consumer Goods, Education, Energy, Financial Services, Health Care, Hospitality, Industrial, Manufacturing, Transportation and Logistics.

# AI: LEGAL OPERATIONS

## BAKER DONELSON USE CASES



### Administrative

- Translations
- Meeting summaries
- Summarizing memos and letters/emails
- Drafting emails, memos, and presentations
- Pitch material generation using historical information
- Business development reports and summaries
- Converting documents to tables and spreadsheets



### Legal Operations

- Converting unstructured data to structured data
- Expertise identification drawing upon an internal database
- Legal research and drafting
- Case management data analysis
- Deidentification of PII or confidential information



### Practice of Law – Litigation

- Docket evaluation and summaries
- Evaluation of depositions, exhibits, or transcripts
- Drafting pleadings, motions, and briefs including citation verification
- Preparing deposition outlines and exhibits
- Automating or augmenting the drafting of written discovery responses



### Practice of Law – Transactional

- Automated contract categorization
- Evaluation of due diligence materials and data rooms
- Contract analysis, redlines, and summaries
- Comparison of draft contracts to a best practice playbook
- Integrating provisions from historical contracts
- Automatically modifying contract provisions to address party positions



### eDiscovery

- Incorporating AI-enabled technology into ESI negotiation strategy and stipulations
- Partnering with trusted vendors to provide early access to developing tools
- Identifying responsive documents, issues, and privileged materials and drafting privilege logs
- Automating redaction of outgoing materials and assessing redactions of incoming productions
- Creating timelines from production materials
- Analyzing and summarizing documents and document groups for importance and supporting factual narratives



ALABAMA • FLORIDA • GEORGIA • LOUISIANA • MARYLAND • MISSISSIPPI • NORTH CAROLINA • SOUTH CAROLINA • TENNESSEE • TEXAS • VIRGINIA • WASHINGTON, D.C.

[www.bakerdonelson.com](http://www.bakerdonelson.com)

THIS IS AN ADVERTISEMENT. Timothy M. Lupinacci is chairman and CEO of Baker Donelson and is located in our Birmingham office, 1901 Sixth Avenue North, Suite 2600, Birmingham, AL 35203. Phone 205.328.0480. No representation is made that the quality of the legal services to be performed is greater than the quality of legal services performed by other lawyers. FREE BACKGROUND INFORMATION AVAILABLE UPON REQUEST. © 2024 Baker, Donelson, Bearman, Caldwell & Berkowitz, PC.