



# Generative AIs challenged by copyright and related rights

A comparative approach in  
European and US law

## Authors

### **Racine Avocats**

Charles Bouffier, Partner  
Victoire Danès, Associate  
Emma Gausson, Intern

### **Miller Nash LLP**

Delfina Homen, Special Counsel  
Brian Esler, Partner

# Introduction

Since its launch on November 30th, ChatGPT has been heralded as a major technological breakthrough, with the potential to create and speak like a human. Microsoft has said it is ready to invest \$10 billion in the company behind ChatGPT, OpenAI, in order to integrate it into its search engine Bing and compete successfully with Google. After the huge success of OpenAI's ChatGPT, Google has announced its version: Google Bard. Bard is Google's experimental AI-based conversational chat service. It is supposed to work similarly to ChatGPT, the main difference being that Google's service will exploit all the information available on the web, while ChatGPT would have been trained with the data available until 2021 only.

ChatGPT and Bard are part of the so-called "generative" artificial intelligence (AI) based on deep learning technologies, a learning method that consists of training an AI with a corpus of data composed of pre-existing texts, images and sounds. Once trained, the AI will be able to make links between the data and produce models, capable of generating new content without human intervention. Thus, ChatGPT automatically generates text, while other AIs such as Dall-E or Midjourney create images from textual descriptions. It is therefore necessary to distinguish between the corpus of data used to train the AI, which constitutes the "input data/content", and the works generated by the AI, which are the "output data/content".

These potential "AI-generated works" are created primarily by software. They are distinguished from "AI-assisted works", for which AI is only a simple tool at the service of humans, which are likely to show "even in a minimal way the originality that its creator wanted to bring"<sup>1</sup> and that the international doctrine classifies among the works protectable by copyright<sup>2</sup>. Works generated by AI are more difficult to apprehend by copyright and related rights, both from the point of view of "incoming data" and "outgoing data". They are also marked by internationality since they are mainly fed from data accessible on the Internet.

In the European Union, copyright protection is automatic from the moment the work is created, provided that it is in an original form. On the other hand, while under American copyright law a copyright exists from the moment a work is fixed in a tangible form, the copyright is only enforceable in court if it is registered with the US Copyright Office. European law has chosen to list the exceptions to copyright in a restrictive manner, whereas American copyright provides for a general exception of fair use.

The divergence of approaches between these two legal systems reflects both on the regime of input data, which feed generative AIs (I), and on that of output data, generated by these same AIs (II).

[1] Cour d'appel de Bordeaux, January 31st, 2005, RG n°03/05512.

[2] WIPO Dialogue on Intellectual Property and Artificial Intelligence, Third Session, Geneva, November 4th, 2020.



# Input data / content, copyright and related rights

## European law

### The text and data mining exception

To train an AI through *deep learning*, AI designers most often use *web scraping*. This is a technique allowing the extraction of content from one or more websites automatically, via a script or a program, in order to reuse it in another context. The extracted data can however be protected by the copyright, the related rights or by the *sui generis* right of the producers of databases.

In order to regulate this practice, the European Union has created a new exception for text and data mining provided for in Directive (EU) 2019/790 of April 17<sup>th</sup>, 2019<sup>1</sup>. This Directive meets an objective of harmonizing the legislative provisions of Member States on copyright and related rights in the context of a European digital single market.

The Directive defines text and data mining as "*any automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations*"<sup>2</sup>. Text and data mining technologies are widespread throughout the digital economy, particularly in the research and artificial intelligence sectors.

In order to protect innovation and research, Article 4 of the Directive provides that "*reproductions and extractions of lawfully accessible works and other subject matter for the purposes of text and data mining*"<sup>3</sup> are permitted, provided that the rightholder has not expressed his or her opposition (*opt-out*).

As this is a harmonization directive<sup>4</sup>, member states have very little leeway in transposing Article 4 into their national law. To date, 22 member states have already transposed Directive 2019/790 (the Netherlands, Hungary, Germany, Malta, Denmark, Croatia, Italy, Ireland, Estonia, France, Austria, Romania, Lithuania, Spain, Luxembourg, Belgium, Cyprus, Slovakia, Slovenia, Sweden, Czech Republic and Greece). The transposition procedure is still underway in Portugal, Bulgaria, Finland, Poland, Latvia and Norway<sup>5</sup>.

In France, the Ordinance n°2021-1518 of November 24<sup>th</sup>, 2021<sup>6</sup> has faithfully transposed the Directive by modifying article L.122-5 10° of the Intellectual Property Code (IPC) and by adding a new article L.122-5-3 to the IPC. The latter includes the definition of text and data mining proposed by the Directive<sup>7</sup>. The text and data mining exception applies to copyright and to objects protected by related rights<sup>8</sup> and by the *sui generis* right of the producer of databases<sup>9</sup>. The exception takes the form of two new limitations: one is for scientific research purposes, the other for all uses. Article L.122-5-3 of the IPC provides that "*copies or digital reproductions of works that have been lawfully accessed may be made*

---

<sup>1</sup> Directive (EU) 2019/790 of the European Parliament and of the Council of April 17<sup>th</sup>, 2019 on copyright and related rights in the digital single market and amending Directives 96/9/EC and 2001/29/EC.

<sup>2</sup> Article 2 of Directive (EU) 2019/790.

<sup>3</sup> Article 4 of Directive (EU) 2019/790.

<sup>4</sup> Council of State, Opinion of October 4<sup>th</sup> 2018, n°395.785: "*In certain areas, European law prohibits any difference in the law of each of the Member States and, consequently, any over-transposition. This is the case for directives that proceed, in a given sector, to a complete harmonization without opening up possibilities of option or derogation (CJEC, May 5, 1998, National Farmer's Union and others, Case C-157/96), as well as for those which, for example, list the conditions for the implementation of a rule, thus preventing national transposition measures from introducing additional conditions (ECJ, November 23<sup>rd</sup>, 1989, Kommanditgesellschaft in Firma Eau de Cologne & Parfumerie Fabrik/Provide, Case C-150/88).*

<sup>5</sup> DSM Directive implementation tracker: <https://www.notion.so/DSM-Directive-Implementation-Tracker-361cfae48e814440b353b32692bba879>.

<sup>6</sup> Ordinance No. 2021-1518 of November 24<sup>th</sup>, 2021 supplementing the transposition of Directive 2019/790 of the European Parliament and of the Council of April 17<sup>th</sup>, 2019 on copyright and related rights in the digital single market and amending Directives 96/9/EC and 2001/29/EC.

<sup>7</sup> Article L.122-5-3 al.1<sup>er</sup> of the IPC.

<sup>8</sup> Article L. 211-3 8° of the IPC.

<sup>9</sup> Article L. 342-3 6° of the IPC.

without the authorization of the authors for the sole purpose of scientific research by research organizations, libraries accessible to the public, museums, archive services or institutions holding cinematographic heritage, audiovisual or sound heritage, or on their behalf and at their request by other persons, including in the context of a non-profit partnership with private actors"<sup>10</sup> but also that "copies or digital reproductions of works that have been lawfully accessed may be made for the purpose of text and data searches carried out by any person, whatever the purpose of the search"<sup>11</sup>. The second limitation thus encompasses all purposes, including commercial<sup>12</sup>.

Several limits to this second exception for all purposes are nevertheless set by the IPC. The person who makes the copy or reproduction must have accessed the content lawfully. It is also foreseen that the author may oppose this search of texts and data. The copies and reproductions must be stored with an appropriate level of security and the person carrying out the data search must be able to provide the right holders, upon request, with all documents and evidence of the text and data mining. Furthermore, the exception allows for only temporary copying or reproduction: once the text or data has been searched, the copy or reproduction must be destroyed and the person carrying out the search must be able to justify this on request<sup>13</sup>.

Most other European countries have also faithfully transposed the Directive<sup>14</sup>, with the exception of:

- Sweden, where the transposition text extends the scope of the data mining exception to "photographic images" which do not meet the originality criterion under Swedish law, and are therefore not protected by copyright, but by a related right under Swedish law<sup>15</sup>;
- Denmark, where the transposition of the Directive has so far only covered Articles 15 and 17. Transposition of the remaining articles of the Directive is expected to be proposed to Parliament in January 2023 and included in subsequent legislation on July 1<sup>st</sup>, 2023<sup>16</sup>.

The UK government, on the other hand, has stated that it does not intend to implement Directive 2019/790, as it is no longer required to do so since the Brexit<sup>17</sup>. Although an equivalent exception was proposed, the UK government halted the project after severe criticism from creative industries<sup>18</sup>.

The text and data mining exception is therefore applicable in most member states and should cover the entire territory of the European Union in the short term, which the European Commission is scrupulously monitoring<sup>19</sup>.

### The authors' opt-out

Directive 2019/790 therefore authorizes text and data mining without the need to obtain the authorization of rights holders, their consent being presumed. However, as the Conseil Supérieur de la Propriété Littéraire et Artistique (CSPLA) points out in its report, text and data mining cannot be used "as a Trojan horse for the dissemination of protected objects without the authorization of the right holders"<sup>20</sup>.

Thus, the text and data mining exception applies, outside of scientific research, only if the use of the works "has not been expressly reserved by their rightholders in an appropriate manner, such as machine-readable means in the case of content made publicly available online"<sup>21</sup>, such as a mention in

---

<sup>10</sup> Article L.122-5-3 I. of the IPC.

<sup>11</sup> Article L122-5-3 III. of the IPC.

<sup>12</sup> A. BENSAMOUN, G. LOISEAU, Droit de l'intelligence artificielle, LGDJ, 2<sup>ème</sup> édition, 2022, §466.

<sup>13</sup> Article R.122-27 of the IPC.

<sup>14</sup> A. MATAS, "A missed deadline: the state of play of the Copyright Directive", Europeana Pro, June 14<sup>th</sup>, 2021.

<sup>15</sup> L. BORCK, E. LUTH, "Copyright directive series - a closer look at Sweden", Europeana Pro, January 11<sup>th</sup>, 2023.

<sup>16</sup> <https://www.notion.so/Denmark-494a7f08c24d45588b3810aa1b1b7212>

<sup>17</sup> P. WIENAND, "Status update on the proposed new copyright law exception to permit text and data mining for commercial use", Farrer&Co, December 7<sup>th</sup>, 2022.

<sup>18</sup> Communications and Digital Committee, House of Lords, November 22<sup>nd</sup>, 2022.

<sup>19</sup> European Commission Press Release "Copyright: Commission urges Member States to fully transpose EU copyright rules into national law", May 19<sup>th</sup>, 2022.

<sup>20</sup> CSPLA Report, « Mission Intelligence artificielle et culture », January 27<sup>th</sup>, 2020, p. 57.

<sup>21</sup> Article 4 al.3 of the Directive (EU) 2019/790.

the metadata of the file containing the protected work. In case of an *opt-out* by the right holder, text and data mining is no longer allowed for the works concerned.

The Directive leaves it up to the member states to determine the possible modalities of the *opt-out*. In France, the opposition "*does not have to be motivated and may be expressed by any means. In the case of content made available to the public online, this opposition may in particular be expressed by machine-readable means, including metadata, and by recourse to the general conditions of use of a website or service*"<sup>22</sup>.

### In practice

Online content sharing platforms have been the first to implement solutions to prevent data mining. For instance, the online platform DeviantArt dedicated to artists has created a "meta tag" (HTML tag) that authors can use to warn AI researchers/developers not to use their content for text and data mining<sup>23</sup>.

Following the example of Shutterstock, a royalty-free image, video and music bank, other content platforms are considering setting up compensation funds for authors whose work would be sold to AI companies (see below)<sup>24</sup>.

It should also be noted that a protest movement of artists against works generated by AI has recently taken place on social networks. It resulted in the creation of an "AI" logo on a black background with a red stripe, used massively on art sharing platforms to materialize the will of some artists not to see their works used by AI creators/developers<sup>25</sup>.

Another possible way of *opting out* is to include the opposition directly in the General Terms of Use (GTU) of the website, as is already the case in database law<sup>26</sup>.

Finally, the creation of databases dedicated to the training of artificial intelligences is envisaged, governed by open-source licenses and withdrawal procedures.<sup>27</sup>

The disadvantage of the different solutions presented is that they are not retroactive. Protected works that have already been used to feed AI cannot be deleted from the systems. It remains to be seen whether authors would be able to act in defense of their rights and to win the case for these past integrations.

## US law

### Copyright Protection Under US Law

To understand the copyright implications of AI under US law, it is helpful first to understand the basics of US copyright law.

Article I, Section 8, Clause 8 of the US Constitution, referred to as the Intellectual Property Clause ("IP Clause"), provides the basis for copyright protection in the United States by giving to Congress the power "[t]o promote the . . . useful Arts, by securing for limited Times to Authors . . . the exclusive Right to their respective Writings." The concept of the IP Clause is that exclusive rights are necessary to incentivize creation of new artistic works. Without legal protection, the rationale goes, competitors could freely create copies, preventing original creators from recouping their invested time and effort and disincentivizing new creation. The IP Clause thus reflects an "economic philosophy" that

---

<sup>22</sup> Article R. 122-28 of the IPC.

<sup>23</sup> K. WIGGERS "DeviantArt provides a way for artists to opt out of AI art generators", Techcrunch, November 2022.

<sup>24</sup> M. SPARKES, "Shutterstock will sell AI-generated art and 'compensate' human artists," NewScientist, October 25<sup>th</sup>, 2022.

<sup>25</sup> P. GODON, " L'intelligence artificielle va-t-elle tuer les artistes ? ", Franceinfo, January 21<sup>st</sup>, 2023.

<sup>26</sup> CJEU, Ryanair LTD, 15 January 2015, C-30/14: "*if the creator of a database protected by Directive 96/9 decides to authorize the use of his database or a copy thereof, it is open to him, as confirmed by recital 34 of that Directive, to regulate that use by means of an agreement concluded with the legitimate user, which specifies, in compliance with the provisions of that Directive, the 'purposes and [...] manner' of using the database or its copy*".

<sup>27</sup> BigCode Project: <https://www.bigcode-project.org/docs/about/the-stack/>.

“encouragement of individual effort by personal gain is the best way to advance public welfare through the talents of authors.”<sup>28</sup>

The Copyright Act is Congress’s exercise of the power granted to it under the IP Clause. Under the Copyright Act, a “copyright” is an intangible property right that protects original works of authorship fixed in any tangible medium of expression.<sup>29</sup> The requisite originality simply means the work was created independently by the author and possessed some minimal or slight degree of creativity, a low threshold that is usually met.<sup>30</sup> Fixation is met where the work is in a form sufficiently permanent or stable to permit it to be perceived, reproduced, or otherwise communicated for more than a transitory duration<sup>31</sup>; for example, taking a picture with a smartphone suffices. Finally, the requisite expression simply means the work is a tangible thing that can be identified and possessed by the owner; ideas, concepts, facts, and themes are not copyrightable.

Under the Copyright Act, authors are granted a bundle of exclusive rights, including the exclusive right to, or authorize others to make copies; prepare derivative works; distribute copies to the public by sale, rental, or lease; and perform or display the work, depending on the type of work.<sup>32</sup> Copyrights protect a wide scope of creative works, including literary works (which include software, whether in source or object code); musical works; dramatic works; choreographic works; pictorial, graphic, and sculptural works, such as fine art, photographs, and prints; motion pictures; sound recordings; compilations, which are new works that bring together and arrange preexisting material or information; and derivative works, which are a transformation or adaptation of one or more preexisting works.<sup>33</sup>

Copyright registration is not required for a copyright to exist under US law. Rather, copyrights exist from the moment a work is fixed in a tangible form.<sup>34</sup> Turning back to the earlier smartphone example, then, a copyright in a photo snapped on the smartphone exists from the moment the photo is digitally captured and stored on the device. However, unlike most countries, registration with the Copyright office is required in the US to bring a lawsuit for copyright infringement.<sup>35</sup>

Under the Copyright Act, the owner of a copyright is the work’s author—the person who reduced an idea to an original expression and committed that expression to a tangible medium—unless the work is properly assigned or is “work-made-for-hire.”<sup>36</sup> Under a typical US employer-employee situation, works created by an employee for her employer in the standard scope of her employment, utilizing resources of her employer and made during normal work hours, will typically qualify as “work-made-for-hire,” such that the employer owns any works created by the employee automatically.<sup>37</sup> Savvy employers, or companies working with persons who could be categorized as independent contractors rather than employees, often require persons who might create copyrightable materials to execute an agreement with the employer acknowledging that works created within the scope of work for the employer are to be treated as a work made for hire under the Copyright Act.

### Copyright in Input Data / Content

In the context of generative AI based on deep learning technologies, copyright is certainly implicated by the “training” stage, though potentially in different ways.

---

<sup>28</sup> *Mazer v. Stein*, 347 U.S. 201, 219 (1954).

<sup>29</sup> 17 U.S.C. § 102.

<sup>30</sup> See *Feist Publ’ns, Inc. v. Rural Tel. Serv. Co., Inc.*, 499 U.S. 340, 345-46 (1991) (holding the selection and arrangement of a phone book could be sufficiently original to be copyrightable, but that alphabetic list of phone subscribers was not sufficiently original to merit protection).

<sup>31</sup> 17 U.S.C. § 101.

<sup>32</sup> See *id.* § 106.

<sup>33</sup> See *id.* § 102.

<sup>34</sup> See *id.*

<sup>35</sup> See *Fourth Estate Pub. Benefit Corp. v. Wall-Street.com, LLC*, 139 S. Ct 881, 887 (2019) (holding that “although an owner’s rights exist apart from registration” under 17 U.S.C. § 408(a), “registration is akin to an administrative exhaustion requirement that the owner must satisfy before suing to enforce ownership rights”).

<sup>36</sup> See 17 U.S.C. § 201.

<sup>37</sup> See, e.g., *U.S. Auto Parts Network, Inc. v. Parts Geek, LLC*, 692 F.3d 1009, 1015 (9th Cir. 2012). The Copyright Act itself does not define either “employee” or “scope of employment”; instead, courts analyze these concepts under the general law of agency, as set forth by the Supreme Court in *Community for Creative Non-Violence v. Reid*, 490 U.S. 730 (1989).

ChatGPT was reportedly initially “trained” by human trainers providing the system with conversations where the trainers played both sides. Those conversations could be considered a “literary work” under the Copyright Act and hence copyrightable, assuming the requisite (low) originality was present and that the conversations were fixed in some tangible form. Assuming the trainers were employed by OpenAI or working as independent contractors under appropriate agreements, the content provided to ChatGPT in the “training” stage would itself be owned by OpenAI, ChatGPT’s creator. OpenAI’s use of its own copyrighted works to train ChatGPT would not implicate any infringement concerns under the Copyright Act.

The analysis is different with respect to generative AI systems trained on materials scraped from the internet, such as Dall-E. Many of the images Dall-E was trained on may be protected under the Copyright Act as pictorial, graphical, and sculptural works. But unlike the “conversations” used to train ChatGPT, the copyrights in these third-party-sourced images are *not* owned by OpenAI, but rather by the images’ authors. By virtue of Creative Commons and other similar licensing platforms, there are a plethora of images available online that could be used (with proper attribution, where appropriate) without copyright concerns so long as the user abides by the licensing terms. But the question becomes, what sorts of images were scraped and used to train Dall-E—and are the authors of those images being appropriately recognized and their licenses respected?

While not involving Dall-E in particular, a lawsuit filed in November 2022 against OpenAI, GitHub, and Microsoft challenges the legality of generative AI trained on open source, copyrighted works.<sup>38</sup> The lawsuit, brought by a coder/lawyer on behalf of a class of coder plaintiffs, challenges an AI system called Copilot. Copilot is an AI assistant tool that suggests code snippets for programmers while they are coding. Code is copyrightable under the Copyright Act as a literary work. However, the code at issue in the lawsuit was provided by its original creators via an open-source license through GitHub, a code repository that was allegedly scraped to provide code to Copilot during Copilot’s training. The lawsuit is not premised on copyright infringement, likely because the code at issue is open source. Instead, it alleges use of the code to train Copilot “violates the licenses that open-source programmers chose and monetizes their code despite GitHub’s pledge never to do so.” By providing the code via an open-source license through GitHub, the coder plaintiffs themselves do not receive compensation for the code, but they ensure that no one else who incorporates the code may do so either. As such, they allege, OpenAI, Microsoft (which owns OpenAI), and GitHub are all engaging in unfair competition by monetizing for themselves code that was supposed to be free for all.

It is easy to see how this first-of-its-kind challenge to AI’s use of others’ material for training purposes can spill over into generative AI. Indeed, visual artists have cried foul about platforms like Dall-E, which take existing images and “repurpose” them into something new based on a user’s prompts. Arguably such AI is creating derivative works—an exclusive right reserved to the authors of the images on which the AI was trained—without the authors’ authorization, without compensating the authors, and in a manner that threatens to put visual artists out of a job.

### The fair use exception

When legal challenges to generative AI that is trained on works owned by someone other than the AI system’s owner start to percolate in the US courts, it is likely the AI system’s owners will assert fair use as a defense. It is questionable, however, whether fair use will ultimately excuse unauthorized use of copyrighted materials to train AI.

Fair use is one of a number of exceptions to a copyright owner’s bundle of exclusive rights. When it is established, the use at issue is statutorily “not an infringement of copyright.”<sup>39</sup> The fair use exception therefore allows persons other than the copyright owner and its licensees to use a copyrighted work without authorization—essentially a golden ticket to unauthorized (and un-compensated) use of a copyrighted work.

---

<sup>38</sup> See Julie Reed, “Machine Learning Is Not Your Copilot: AI System Accused of Violating Open Source Copyright Licenses,” *IP & Technology Law Trends*, Miller Nash LLP (Jan. 9, 2023), <https://www.millernash.com/industry-news/machine-learning-is-not-your-copilot-ai-system-accused-of-violating-open-source-copyright-licenses> (last visited February 6, 2023) (discussing *Doe v. GitHub, Inc.*, No. 3:22-cv-06823, filed November 3, 2022 in the United States District Court for the Northern District of California).

<sup>39</sup> 17 U.S.C. § 107.

Under the Copyright Act, fair use includes criticism, commentary, news reporting, teaching, scholarship, and research.<sup>40</sup> However, simply pointing to one of those broad conceptual buckets is not enough to establish a use of a copyrighted work as a *fair* use. Rather, to determine whether a particular use is a fair use, US courts look to the following factors:<sup>41</sup>

- The purpose and nature of the use. Under this factor, non-commercial uses, such as educational uses, are more likely to be considered fair use.<sup>42</sup>
- The nature of the copyrighted work. Under this factor, courts look to whether the original work is creative or factual, with factual reporting given less protection than works of fiction since copyright does not protect facts or ideas.<sup>43</sup>
- How much and how substantial the portion of the copyrighted work used is relative to the work as a whole. Under this factor, which looks at both the qualitative and quantitative portions used, the less of the original work that is taken, the more likely the taking will be considered a fair use.<sup>44</sup>
- The effect of the purported fair use on the potential market for the copyrighted work or the work's value. Under this factor, courts examine the market harm from the claimed fair use, including the direct market harm to the particular copyright owner, and the harm to the market generally that may result from similar infringing uses.<sup>45</sup>

While once free-to-use, many generative AI systems are unsurprisingly shifting to fee-based systems. Using copyrighted materials to train a fee-based AI platform is a commercial use, undercutting a claim of fair use. With respect to Dall-E, the creative works used to train Dall-E are given more protection than, say, a list of facts, further undercutting fair use. Dall-E and other such generative AI are typically “trained” using the entirety of the image, a qualitatively and quantitatively large portion of the copyrighted work, even further undercutting fair use. Finally, as visual artists recognize, AI systems like Dall-E are a substantial threat both to the market for the original work and to the market for original visual works in general, even further undercutting fair use.

But US courts have previously found fair use in even widespread and wholesale copying when technology allowed a use of the works that was novel, transformative, and arguably furthered the goals of the IP Clause. For instance, at the beginning of this century, Google undertook a massive effort to scan and digitize millions of printed books (many still under copyright) to create a searchable online database. Although Google only made available snippets of the works still under copyright, it was sued for copyright infringement by publishers and authors. After years of litigation and failed settlement attempts, the trial court ultimately found Google’s use of these works to be fair use because it provided significant public benefits that advanced the progress of the arts and sciences, and the Court of Appeals agreed.<sup>46</sup> While there are significant differences between Google’s project (which continued to provide author attribution and was not a market substitute for the original works) and generative AI systems (which fail to provide author attribution and may generate market substitutes), undoubtedly an argument

---

<sup>40</sup> *Id.*

<sup>41</sup> The general factors are set forth in § 107 of the Copyright Act. Decades of jurisprudence applying these factors are followed in determining whether fair use exists in a particular situation.

<sup>42</sup> See, e.g., *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 585 (1994) (explaining that while a non-commercial purpose weighs in favor of a finding of commercial use, there is no “bright-line approach”).

<sup>43</sup> See, e.g., *Stewart v. Abend*, 495 U.S. 207, 237-38 (1990) (explaining that the more creative a work, the closer it is to the core scope of copyright protection and hence, the less likely that its use without authorization is “fair”). Note that this factor “has rarely played a significant role in the determination of a fair use dispute,” since “[t]he mere fact that the original is a factual work . . . should not imply that others may freely copy it. Those who report the news undoubtedly create factual works. It cannot seriously be argued that, for that reason, others may freely copy and re-disseminate news reports.” *Authors Guild v. Google, Inc.*, 804 F.3d 202, 220 (2d Cir. 2015).

<sup>44</sup> See, e.g., *Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 926 (2d Cir. 1994) (finding that while copying entire works “does not preclude a finding of fair use, it militates against such a finding,” and weights this factor in favor of the copyright owner, i.e., against a finding of fair use). Qualitatively, courts examine whether the portion of the original work used, even if quantitatively small, inappropriately uses the heart of the original work. See, e.g., *Video Pipeline Inc. v. Buena Vista Home Entm’t, Inc.*, 342 F.3d 191, 201 (3d Cir. 2003) (finding use of one minute of a two-hour long movie to be qualitatively unreasonable where it consisted of the climax or resolution of the movie).

<sup>45</sup> See, e.g., *Am. Geophysical Union*, 60 F.3d at 927-28 (finding market harm from lost or diverted sales of copies of the plaintiff’s works); *Zomba Enters., Inc. v. Panorama Records, Inc.*, 491 F.3d 574, 583-84 (6th Cir. 2007) (finding market harm from lost royalties or licensing revenue, including from markets the plaintiff is already in and markets the plaintiff is not in but has the right to exploit).

<sup>46</sup> *Authors Guild*, 721 F.3d 132.



can be made that what AI systems do is something like what Google did in making large quantities of copyrighted work more easily searched and digested. In effect, owners of AI systems would argue their technology is just a technology-aided acceleration of ordinary scholarship and research, which has long been considered fair (and even expected) use under US law.<sup>47</sup> Whether such arguments can prevail will depend on a much greater factual analysis than can be provided here.

### In practice

Short of repositories of training materials that are owned by the AI system's creator or that are dedicated to use in AI training (with the appropriate copyright permissions), engaging in copyright infringement by using training materials scraped from the internet could be a real risk under US law. While it remains a novel issue, US courts could hold use of copyrighted images to "train" AI systems is not a fair use, such that use of non-open source copyrighted images in training constitutes copyright infringement. For generative AI systems trained on pre-existing copyrighted works to succeed under US law—be they pictorial, literary, musical, or other works—it may be that changes need to be made to the Copyright Act to codify "training" as an exception to copyright (like the European text and data mining exception), or the free market needs to come up with a solution.

On this latter point, Shutterstock recently announced its intention to launch a first-of-its-kind "Contributor Fund," under which Shutterstock contributors whose works are used to develop AI will be compensated. Shutterstock also announced its desire to compensate contributors with royalties when their creations are used by Dall-E.<sup>48</sup> These sorts of free market adjustments, previously seen with such "new" technology as streaming audio, are likely to prove the most effective (and expedient) way to address the copyright issues with AI under US law.

---

<sup>47</sup> The Copyright Office's Fair Use Index provides numerous examples. See <https://copyright.gov/fair-use/fair-index.html> (last visited Feb. 7, 2023).

<sup>48</sup> See James Vincent, "Shutterstock Will Start Selling AI-Generated Stock Imagery With Help From OpenAI," *The Verge* (Oct. 25, 2022), <https://www.theverge.com/2022/10/25/23422359/shutterstock-ai-generated-art-openai-dall-e-partnership-contributors-fund-reimbursement> (last visited Feb. 6, 2023).

# Output data / content, copyright and related rights

## European law

### Substantive law

#### *The condition of originality*

French law has a subjective conception of originality, according to which a work is original if and only if it reflects the personality of its author. However, this conception of originality is not found in all European countries<sup>1</sup>. Germany conditions originality on the demonstration of "individuality"<sup>2</sup>. In Belgium, it is the expression of an intellectual effort that has been retained<sup>3</sup>. Finally, the Netherlands considers that an original work is one that is distinct from existing works and results from human creative work<sup>4</sup>.

In European law, a work is only protectable by copyright if it is in an original form. The "Infopaq" judgment of the Court of Justice of the European Union (CJEU) defines "original work" as "*the author's own intellectual creation*"<sup>5</sup>. The CJEU has evolved its concept of originality over time to adapt to the development of digital technology in society. Thus, the CJEU has clarified in several decisions<sup>6</sup> that "*free and creative choices*" by the author on "*the choice, sequence and combination of elements*" can express the author's creativity in an original way and lead to a result that is an intellectual creation. One of these decisions is the Painer case<sup>7</sup>, in which the CJEU stated that it is entirely possible to create original works with the aid of a machine or device. In this case, it was not a question of artificial intelligence but of photography. As a reminder, photography, like artificial intelligence for some, has long been considered by the doctrine as not protectable by copyright, notably because of the mechanical nature of the process it requires.

However, this conception of originality does not allow to qualify as original all the works involving an AI. It is necessary to measure the human involvement in the creation process. If the author uses the AI as a tool to achieve a certain result, the human author retains full creative freedom over the work and it can therefore be protected. Similarly, if a work is created independently by an AI, but the AI is still "controlled" by humans, the human authors do play a relatively passive role, but if they can prove that their creative choices had an essential impact in the creation process, the resulting work will be potentially protectable. On the other hand, if a work is entirely generated by an AI and the human intervention is limited to giving the order to run the program, then in principle there should be no room for free and creative choice on the part of the author.

Thus, AI-generated works do not seem to meet the condition of originality as defined by the CJEU. This is in fact what emerges from several judgments. In its "Football Dataco" decision, the Court states that originality is not characterized when the creation of the work is "*dictated by technical considerations, rules or constraints which leave no room for creative freedom*"<sup>8</sup>. In the "Cofemel" decision<sup>9</sup>, it recalls that the aesthetic criterion is not sufficient to characterize originality. Therefore, despite all the qualities attributed to Midjourney's paintings and ChatGPT's poems, this aesthetic criterion should not be sufficient to grant them protection under European copyright law.

#### *Ownership*

Most European Union member states consider that copyright protection can only be granted to a human intellectual creation. If it is granted to a work generated by an AI, this implies in principle a connection

---

<sup>1</sup> A. LUCAS, A. LUCAS-SCHLOETTER, C. BERNAULT, *Traité de la propriété littéraire et artistique*, LexisNexis, 2017, 5<sup>ème</sup> édition.

<sup>2</sup> German law on personal intellectual creations, September 9<sup>th</sup>, 1965.

<sup>3</sup> Cass., Pasirisie, April 27<sup>th</sup>, 1989, F-19890427-11.

<sup>4</sup> Hoge Raad, Zonen Endstra v. Nieuw Amsterdam, May 30<sup>th</sup>, 2008, RNL2008-285120.

<sup>5</sup> CJEU, Infopaq International A/S, July 16<sup>th</sup>, 2009, -C5/08.

<sup>6</sup> CJEU, Funke Medien NRW GmbH, July 29<sup>th</sup>, 2019, -C469/17; CJEU, Painer, December 1<sup>st</sup>, 2011, -C145/10.

<sup>7</sup> CJEU, Painer, December 1<sup>st</sup>, 2011, -C145/10

<sup>8</sup> CJEU, Dataco Football, March 1<sup>st</sup>, 2012, -C604/10.

<sup>9</sup> CJEU, Cofemel v. G-Star Raw CV, September 12<sup>th</sup>, 2019, C-683/17.

to a human author according to his impact on the result and his involvement in the creation process. It is then necessary to identify the human author behind the work generated by AI.

In French law in particular, an AI cannot be the owner of rights in a work, for the simple reason that one of the criteria for protection by copyright is the consciousness of creating<sup>10</sup> - a prerogative of the human being, of which an artificial intelligence is devoid. An AI cannot therefore be qualified as an author because it is not a natural person<sup>11</sup>.

The solution seems to be the same in European law, being specified that to our knowledge there is not yet any case law - French or European - allowing to determine who of the designer and/or the user of the AI should be the owner(s) of the rights on a work created through an AI and judged original.

A parallel can be drawn with industrial property law: the European Patent Office (EPO) has confirmed that the inventor named in a patent application must necessarily be a human being<sup>12</sup>. The UK Court of Appeal has also denied inventor status to an artificial intelligence<sup>13</sup>. Patent rights cannot therefore be granted to a machine either.

### Prospective law: the assumptions of legal protection

By copyright

#### *The condition of originality of output data*

Copyright does not seem to be able to protect works generated by AIs, because they do not meet the criterion of originality. The only solution for these works to be, one day, protected by copyright would be to adapt this criterion to include works generated by AI. This possibility is not totally ruled out since originality is a flexible notion that has already been made to evolve over time. Moreover, the European Union is in favor of the development of research and innovation, and more particularly the development of artificial intelligence. Finally, the exclusion of AI-generated works from the scope of copyright protection would prevent the emergence of possible related rights for the benefit of their performers<sup>14</sup>.

It is a question of legal policy based on the willingness - or not - to protect works generated by AIs by copyright.

#### *The issue of ownership of output data*

✓ The AI designer

The CSPLA report on artificial intelligence and culture<sup>15</sup> proposes to qualify the AI designer as the author of the works generated by the AI, in consideration of his or her nature as the "indirect" creator of the work. This solution is inspired by the regime of the collective work, in which the ownership of the rights is granted to the person who initiates, promotes and discloses under his name the work to the elaboration of which several persons have collaborated<sup>16</sup>. This model reflects a more economic vision of copyright, less focused on the person of the author, which could be adapted to works generated by AI. Contrary to the classical rules of ownership, the rights would arise independently of any transfer, on the head of the instigator who, in a vertical creative process, directs the creation and publishes it under his name. The CSPLA considers that such a devolution is logical since "*the AI only executes the creative framework set by the creator*"<sup>17</sup>.

---

<sup>10</sup> Cass. 1<sup>st</sup> Civil, November 13<sup>th</sup>, 2008, n°06-16.278.

<sup>11</sup> FAQs on copyright for consumers : What does copyright and related rights mean and cover, and is it the same all over the world? : <https://euiipo.europa.eu/ohimportal/en/web/observatory/faq-el>

<sup>12</sup> Board of Appeal of the European Patent Office (EPO), December 21<sup>st</sup>, 2021, J 0008/20.

<sup>13</sup> England and Wales Court of Appeal, Civil Division, Thaler v. Comptroller, September 21<sup>st</sup>, 2021.

<sup>14</sup> A. BENSAMOUN, G. LOISEAU, Droit de l'intelligence artificielle, LGDJ, 2nd edition, 2022, §483 et seq.

<sup>15</sup> CSPLA Report: Artificial Intelligence and Culture Mission, January 27<sup>th</sup>, 2020, p.37 and following.

<sup>16</sup> Article L.113-2 of the IPC.

<sup>17</sup> CSPLA report, Mission Intelligence artificielle et culture, January 27<sup>th</sup>, 2020, p.37.

Other authors also argue for taking preparatory acts into account in determining the protection of the AI-generated work (choice of incoming data, software, algorithms, settings, etc.)<sup>18</sup>.

Finally, the mechanism of accession, which comes from the law of property<sup>19</sup>, is also used to retain the status of author of the designer of the artificial intelligence<sup>20</sup>. This mechanism allows the owner to acquire the accessories produced by the thing he owns. Applied to AI, accession would allow the AI designer to acquire the rights to the works generated by the AI.

✓ The AI user

The qualification of the user as an author of AI-generated works would rest on his financial legitimacy. Indeed, most users will purchase the rights to a license in order to exploit the AI. The CSPLA report<sup>21</sup> also invokes the custody exercised by the user over the AI to justify his possible authorship.

However, in order to be able to assimilate the user to the author, the doctrine considers that it is necessary to be able to establish a link between the user and the work, or "connecting factor". The latter is difficult to demonstrate. In the particular case of ChatGPT, for example, it seems difficult to retain a minimum creative choice on the part of the user who only enters a sentence in the software. The creative involvement of the user in the determination of the work generated by the machine should in any case be assessed on a case-by-case basis.

A part of the doctrine proposes to turn to the American theory of "*work made for hire*"<sup>22</sup>, so that the AI is considered as the "author in fact" of the work, but that only the user is apprehended as the "author in law" provided that the AI has produced the work at its request<sup>23</sup>.

*By creating a special regime*

The CSPLA report proposes to legally impose a system of ownership or to provide for a condition of involvement or investment (material, human or financial) to determine the owner of the work<sup>24</sup>, following the example of the *sui generis* right of the producer of databases. The implementation of a *sui generis* right would indeed encourage developers to invest in the creation of AI by ensuring them a financial compensation for their investments. It could also allow for a shorter duration of protection than for "traditional" works. The disadvantage of providing for a *sui generis* regime rather than applying copyright lies, however, according to the authors of the CSPLA report, in the fact of denying "*the obvious kinship between traditional creations and intelligent creations*"<sup>25</sup>.

Another hypothesis could be to create a new related right inspired by the regime of posthumous works. As for the owner of the posthumous work, the person who would be at the origin of the disclosure of the work generated by the AI would benefit from less important patrimonial prerogatives than those attributed by the copyright<sup>26</sup>.

English law already provides for a derogatory regime for computer-generated works. The author is the one who has taken the necessary steps for the creation of the work. In the case of data generated by AI, the author could alternatively be the designer or the user of the AI, depending on their respective involvement in the process of creating the final work.

Finally, in a more or less near future, the owner of the intellectual property rights could be the AI itself, the European Parliament having envisaged in a 2017 resolution "*the creation, in the long run, of a*

---

<sup>18</sup> Y. GAUBIAC, « Objet du droit d'auteur - Intelligence artificielle et création artistique », JurisClasseur Lexis 360, Fasc. 1164,

<sup>19</sup> Articles 546 and following of the Civil Code; P.-Y. Gautier, "De la propriété des créations issues de l'intelligence artificielle", *Revue Pratique de la prospective et de l'innovation*, LexisNexis, October 2018.

<sup>20</sup> CSPLA report, Mission Intelligence artificielle et culture, January 27<sup>th</sup>, 2020, p.38.

<sup>21</sup> CSPLA report, Mission Intelligence artificielle et culture, January 27, 2020, p.40.

<sup>22</sup> Section 201, US Copyright Act - The American theory of "*work made for hire*" is a legal fiction according to which the "author" of works is not their creator, but their sponsor, vested *ab initio* with the copyright.

<sup>23</sup> CSPLA Report, Mission Intelligence artificielle et culture, January 27, 2020, p. 40.

<sup>24</sup> CSPLA Report, Mission Intelligence artificielle et culture, January 27, 2020, p.41 and following.

<sup>25</sup> CSPLA report, Mission Intelligence artificielle et culture, January 27, 2020, p.45.

<sup>26</sup> A. BENSAMOUN, G. LOISEAU, *Droit de l'intelligence artificielle*, LGDJ, 2nd edition, 2022, §495 et seq.

*specific legal personality for robots, so that at least the most sophisticated autonomous robots could be considered as responsible electronic persons, obliged to repair any damage caused to a third party" and, why not, to benefit from rights on their creations.*

However, there is a lot of reticence today, and this is understandable, on the part of those for whom "*artistic invention, the artistic project, is a matter for humans, not for machines, even in the context of AI*"<sup>27</sup>.

### The risks of infringement

Since AI-generated works are created from existing content that may be protected by an intellectual property right, they may constitute infringements. However, it is important to remember that ideas and genres are not protectable and that works generated "in the style of" or "in the manner of" a particular artist are not automatically infringements unless they reproduce protected works or objects without authorization.

However, legal actions have already been initiated in Europe, as in the rest of the world. The image bank Getty Images recently filed a lawsuit against Stability AI, the company behind the AI Stable Diffusion, in the High Court of Justice in London. Stable Diffusion is an AI that generates images from text entered by the user. Getty Images accused Stability AI of violating its copyright and terms of use. Getty Images claims that Stability AI illegally copied and processed millions of copyrighted images and associated metadata from their image bank in order to train its AI, even though no license was granted.

Failure to respect the rights of the owners on incoming data has repercussions on outgoing data which are likely to constitute infringements. Right holders should be able to act on the basis of infringement or unfair competition to defend their rights.

The question of a liability regime specific to AI arises. On September 28<sup>th</sup>, 2022, the European Commission published two proposals for directives to develop the law on civil liability for artificial intelligence systems<sup>28</sup>. The first one proposes to revise the directive on liability for defective products, the second one adapts the rules of extra-contractual civil liability to artificial intelligence by introducing various presumptions. A system of extra-contractual liability for AIs will thus soon see the light of day, but for the time being, none of these texts envisage the infringement of intellectual property rights and the associated liability system.

## US law

### Authorship and ownership in AI output

As previously discussed, under US law copyright protection is extended to original works of authorship fixed in any tangible medium of expression, and the owner of the copyrighted work is the author—the person who reduced an idea to an original expression and committed that expression to a tangible medium—unless the work is properly assigned or is "work-made-for-hire." Decades of jurisprudence are available for assessing whether a human-created work is sufficiently original to be copyrightable, and whether a human (or his employer) is the author of a given work. But what happens when the work at issue was created not by a human, but rather by an AI system?

That is the core issue in a pending lawsuit by Dr. Stephen Thaler, the creator of an AI system called DABUS (short for Device for the Autonomous Bootstrapping of Unified Sentience).<sup>29</sup> DABUS differs from

<sup>27</sup> Y. GAUBIAC, *Objet du droit d'auteur - Intelligence artificielle et création artistique*, JurisClasseur Lexis 360, Fasc. 1164.

<sup>28</sup> EU Comm. on Revision of the Product Liability Directive, Sept. 28<sup>th</sup>, 2022, COM (2022) 495 final, 2022/0302 (COD); EU, Prop. on Adaptation of the Rules on Non-contractual Civil Liability to the Field of Artificial Intelligence (AI Liability Directive), Sept. 28<sup>th</sup>, 2022, COM (2022) 496 final, 2022/0303 (COD).

<sup>29</sup> For an overview of Dr. Thaler's pending lawsuit, see Delfina Homen, "Paradise Denied: Copyright (Or Not) in AI-Generated Images," *IP & Technology Law Trends*, Miller Nash LLP (Jan. 17, 2023), <https://www.millernash.com/industry-news/paradise-denied-copyright-or-not-in-ai-generated-images> (last visited Feb. 6, 2023). Dr. Thaler's attempt to claim that DABUS could be an inventor for purposes of patent law was rejected last year by the Court of Appeals for the Federal Circuit. See Julie Reed, "CAFC Confirms That Artificial Intelligence Cannot Be An Inventor," *IP & Technology Law Trends*, Miller Nash LLP (Aug. 11, 2022), <https://www.millernash.com/industry-news/cafc-confirms-that-artificial-intelligence-cannot-be-an-inventor> (last visited Feb. 7, 2023). Ms. Homen has since interviewed Dr. Thaler, who generously provided the additional information concerning DABUS that is featured in this article.

traditional generative AI like ChatGPT or Dall-E. While DABUS was initially presented with certain stock images (mostly from Dr. Thaler’s smartphone) and text, DABUS’s “training” has largely been based off of its own creations; it is learning from its own experiences. Moreover, unlike AI systems which “create” works in response to prompts from users, DABUS is “sentient”—it reacts to its environment and forms new ideas based on what it sees and what it remembers, essentially operating as a synthetic brain. In effect, ideas autonomously form and ripen over time, just as in the human case. As Dr. Thaler has explained, DABUS’s “sentience” is not completely equivalent to human sentience; its memories are not of typical human experiences and so its resulting “emotion” is not akin to human emotion. But DABUS does, according to Dr. Thaler, “enable[] a form of sentient machine intelligence whose perception, learning, and imagination are keyed to its subjective feelings, all encoded as sequential chains of memories whose shapes and topologies govern the release of simulated neurotransmitters.”<sup>30</sup>

In 2012, Dr. Thaler’s DABUS created an image it self-titled “A Recent Entrance to Paradise” in reaction to a “hot button” activated by such a memory chain—namely, a zone plate image of concentric circles—which it saw in a synthetic neural chain generated in response to an environmental cue. The image is a rather fascinating and aesthetically pleasing composition of several source images—everything from a picture of Dr. Thaler’s office building to images of purple and red flowers, tunnels, and arches. The resulting composition, which can readily be deemed either a pictorial or derivative work under the Copyright Act, looks like train tracks passing through a series of flower-covered tunnels. Dr. Thaler has described *Paradise* as a “simulated near-death experience” because “it was the product of DABUS’ gradual, internal destruction.”



Diagram of “Paradise” (center) surrounded by source images combined by DABUS in response to “hot button” zone plate image. The system offered explanatory text to title and explained the image. Courtesy of Dr. Stephen Thaler.

In 2016, Dr. Thaler applied for a copyright registration in *Paradise*, identifying the author as “Creativity Machine” and listing himself as the claimant via “ownership of the machine.” The application included a note stating *Paradise* “was autonomously created by a computer algorithm running on a machine” and that he was “seeking to register this computer-generated work as a work-for-hire to the owner of the

<sup>30</sup> Dr. Stephen Thaler, “DABUS in a Nutshell,” *Philosophy and Computers*, vol. 19 no. 1, The American Philosophical Association (Fall 2019).

Creativity Machine.” The Copyright Office denied his registration twice, concluding the work “lacks the human authorship necessary to support a copyright claim.”<sup>31</sup>

Championed by Professor (and lawyer and doctor) Ryan Abbott, Dr. Thaler appealed the Copyright Office’s denial to the United States District Court for the District of Columbia.<sup>32</sup> They recently filed a motion for summary judgment on the issues of DABUS’s authorship in, and Dr. Thaler’s ownership of, *Paradise*, challenging the Copyright Office’s “human authorship requirement” as improperly relying on cases “from the 19th century [that] greatly predate even the invention of the first modern computers.”<sup>33</sup>

The motion argued *Paradise* and other AI-generated images are copyrightable because AI qualifies as an “author” under the Copyright Act since, under the plain and ordinary meaning of the word “author,” an author need not be human. Indeed, corporations can be “authors” under the work-made-for-hire doctrine. The motion further argued that finding AI to be an “author” under the Copyright Act promotes the arts not by motivating AI to create additional works, but rather by motivating “individuals like Dr. Thaler, and businesses like music and movie studios, . . . to develop and use AI to generate new works, thereby achieving the purpose of the [A]ct.”<sup>34</sup>

As to Dr. Thaler’s ownership of *Paradise*, the motion first argued standard property principles apply to make him the work’s owner. For example, analogizing to a fruit tree, Dr. Thaler argued if he owned the tree, he would also own its fruit; “[t]his does not require the tree to execute a written agreement to transfer the fruit, the fruit belongs to Dr. Thaler by virtue of his relationship to the tree.” Alternatively, Dr. Thaler argued ownership in the work automatically vested in him under work-made-for-hire.<sup>35</sup>

The Copyright Office responded to Dr. Thaler’s motion for summary judgment with a cross-motion for summary judgment, arguing it acted “reasonably and consistently with the law when it refused to extend copyright protection to a visual work that [Dr. Thaler] represented was created without any human involvement.”<sup>36</sup> It argued “the human authorship requirement is a longstanding requirement of copyright law” and that the Copyright Act “assumes that authors are humans.”<sup>37</sup> The Office pointed to Supreme Court decisions from the 1800s suggesting “human creativity” is required for a work to be copyrightable, as well as appellate court decisions holding that works created by animals, divine spirits, and nature are not copyrightable.<sup>38</sup> Noting that one of these latter decisions stated “that ‘if Congress and the President intended to take the extraordinary step of authorizing animals’ to sue, the statute would need to state so clearly,” the Copyright Office argued, “Copyright protection for works created entirely by machines would be even more extraordinary.”<sup>39</sup>

The Copyright Office’s cross-motion also argued against finding Dr. Thaler to be the owner of copyright in *Paradise* under common law principles or the work-made-for-hire doctrine. It argued Dr. Thaler’s reliance on common law property ownership principles “is irrelevant” because they “involve physical rather than intangible property. It is a fundamental principle of intellectual property, confirmed in the Act, that ownership of a material object is distinct from ownership of intangible rights embodied in that object.”<sup>40</sup> As for work-made-for-hire, the Office argued DABUS “is not a person, is not [Dr. Thaler’s] employee, and is not [Dr. Thaler’s] agent. The work made for hire doctrine does not apply here.”<sup>41</sup> Indeed, it noted that the “argument that computers can be employees for copyright purposes is extraordinary and could have broad implications for employment and tax law. The Court should not

---

<sup>31</sup> See Copyright Review Board, U.S. Copyright Office, “Second Request for Reconsideration for Refusal to Register a Recent Entrance to Paradise (Correspondence ID 1-3ZPC6C3; SR # 1-7100387071)” (Feb. 14, 2022), available at <https://www.copyright.gov/rulings-filings/review-board/docs/a-recent-entrance-to-paradise.pdf>.

<sup>32</sup> Thaler v. Perlmutter, No. 1:22-cv-01564-BAH (filed June 2, 2022).

<sup>33</sup> See Homen, n.29.

<sup>34</sup> See *id.*

<sup>35</sup> See *id.*

<sup>36</sup> Def.’s Resp. to Pl.’s Mot. Summ. J. & Cross-Mot. Summ. J. at 1, Thaler v. Perlmutter, No. 1:22-cv-01564-BAH (D.D.C. Feb. 7, 2023).

<sup>37</sup> *Id.* at 13.

<sup>38</sup> *Id.* at 14-18.

<sup>39</sup> *Id.* at 17-18 (quoting *Naruto v. Slater*, 888 F.3d 418, 425 (9th Cir. 2018)).

<sup>40</sup> *Id.* at 20.

<sup>41</sup> *Id.* at 22.

construe the term ‘employee’ in a way that would disrupt other established areas of law, such as including inanimate machines in the definition of ‘employee.’”<sup>42</sup>

Finally, the Copyright Office brushed aside the “policy arguments in favor of copyright protection for AI created works” in Dr. Thaler’s motion.<sup>43</sup> It argued that “[r]egardless of [Dr. Thaler’s] own views, the Constitutional purpose of copyright is to incentivize *humans* to create expressive works. . . . Unlike humans, machines do not have rights of free expression, and do not need economic incentives to create and disseminate expressive content.”<sup>44</sup> “In any event,” the Office concluded, “this is not the forum to resolve [Dr. Thaler’s] policy arguments.” Rather, “the Court here is limited to applying the law as it exists now, not as [Dr. Thaler] might wish it to be.”<sup>45</sup> Dr. Thaler filed a response to the Copyright Office’s cross-motion on March 7, 2023, arguing the Office was “employ[ing] smoke and mirrors to attempt to obfuscate plain [statutory] language,” and that “[t]he bottom line is that nothing in the Act’s language limits authorship to human beings.”<sup>46</sup> The court has yet to decide either motion.

*Paradise* is not the only AI-created work to come under scrutiny from the Copyright Office. On February 21, 2023, the Copyright Office issued a letter to Kristina Kashtanova, the artist behind the graphic novel “Zarya of the Dawn,” notifying Ms. Kashtanova that it is cancelling the AI-generated portion of the registration it previously granted for the work.<sup>47</sup> While the Copyright Office found that the text, written by Ms. Kashtanova, is copyrightable and will remain registered, it is cancelling the portion of the registration that extends to the images themselves, which were made with an AI text-to-image tool called Midjourney. The Copyright Office will issue a new registration that “will explicitly exclude ‘artwork generated by artificial intelligence.’”<sup>48</sup>

In arguing for the copyrightability of the images, Ms. Kashtanova claimed her text input to Midjourney provided sufficient human creativity to make the images copyrightable. The Office disagreed, concluding that “[b]ecause of the significant distance between what a user may direct Midjourney to create and the visual material Midjourney actually produces, Midjourney users lack sufficient control over generated images to be treated as the ‘master mind’ behind them.”<sup>49</sup> While the Office did “not question” Ms. Kashtanova’s arguments that she “expended significant time and effort working with Midjourney,” it noted that courts reject such “sweat of the brow” arguments.<sup>50</sup>

Interestingly, the Copyright Office was careful to cabin its decision to “the specific facts provided about Ms. Kashtanova’s use of Midjourney to create the Works’ images. It is possible that other AI offerings that can generate expressive material operate differently than Midjourney does.”<sup>51</sup> However, just as it did with *Paradise*, the Copyright Office noted that the images created with Midjourney “are not the product of human authorship,” relying on the same cases it did for *Paradise* as establishing that so-called requirement.<sup>52</sup> As long as the Copyright Office maintains this “human authorship requirement”—and the federal courts do not find otherwise—it seems that copyright registration for the output of AI in the United States will remain elusive.

It is interesting to consider whether Dr. Thaler’s ownership arguments with respect to DABUS translate to all things AI, particularly to AI systems that rely on input “prompts” from the system’s users to generate text, images, and other potentially copyrighted output. While DABUS, relying on its own perceptions rather than human input, arguably is acting with the requisite creativity copyright is supposed to reward, can the same be said for AI systems relying on human prompts? Statutorily, does it even matter when

---

<sup>42</sup> *Id.* n.17.

<sup>43</sup> *Id.* at 24.

<sup>44</sup> *Id.* at 25.

<sup>45</sup> *Id.* at 25-26.

<sup>46</sup> Pl.’s Combined Opp’n to Def.’s Mot. Summ. J. & Reply in Supp. of Pl.’s Mot. Summ. J. at 3-4, *Thaler v. Permuter*, No. 1:22-cv-01564-BAH (D.D.C. Mar. 7, 2023).

<sup>47</sup> See Copyright Review Board, U.S. Copyright Office, “Zarya of the Dawn (Registration # VAu001480196)” (Feb. 21, 2023), available at <https://www.copyright.gov/docs/zarya-of-the-dawn.pdf>.

<sup>48</sup> *Id.* at 12.

<sup>49</sup> *Id.* at 8-9.

<sup>50</sup> *Id.* at 10 (citing *Feist Publ’ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 352-53 (1991) (refusing to use copyright protection as “a reward for the hard work that went into” creating an otherwise unprotectable work, because “sweat of the brow” would permit copyright to extend further than the author’s original contribution)).

<sup>51</sup> *Id.* at 10.

<sup>52</sup> *Id.* at 1, 3-4.



under US copyright law, “[c]opyright protection subsists . . . in original works of authorship fixed in any tangible medium of expression, now known or later developed, from which they can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device”<sup>53</sup>—language that is completely agnostic to human involvement?<sup>54</sup> The Copyright Office’s cross-motion in the *Paradise* suit telegraphs its likely treatment of prompt-driven AI creations: in rejecting Dr. Thaler’s new arguments on appeal in support of his ownership of *Paradise* that he “created an AI that he directed to create artwork,” the Copyright Office stated, “[T]hat does not mean that he directed the specific contents of any work, which is what copyright protection requires.”<sup>55</sup> From this statement, the Copyright Office appears likely to reject applications to register copyrights in AI creations filed by users of prompt-driven AI systems, too.

Ownership in AI output is also up for debate. Even if Dr. Thaler succeeds in convincing the federal courts to find *him* to be the owner in DABUS’s works as the creator of DABUS, what about AI systems reliant on human prompts? Will the system creators be the copyright owners—in which case OpenAI would own the copyrights in anything generated by ChatGPT—or will the users supplying the prompts be the copyright owners? And it cannot be forgotten that the traditional rule is the *author* is the owner. If the AI is the author and ownership cannot be legally found to vest in a human or corporation, copyrights in AI-created works will be legally meaningless. Even Dr. Thaler’s motion recognizes that if AI were the copyright owner it would lack standing to sue for copyright infringement based on a federal court decision holding a monkey lacked statutory standing to bring a copyright infringement claim.<sup>56</sup> Even if machines can create copyrighted works, they cannot sue in federal court. Absent changes to US copyright law addressing ownership in AI works, an entire body of AI-generated works could effectively become orphaned, unenforceable, and completely up for grabs.

While there are not a lot of definitive statements to be made about copyright and AI under US law at this time, efforts such as those by Dr. Thaler and the coders challenging Copilot will likely shed light on these issues in the years to come. AI is not going away, and either US copyright law or its application will have to adjust accordingly. Congress apparently agrees; in October 2022, members of the US Senate wrote to the heads of the US Patent & Trademark Office and the Register of Copyrights acknowledging Congress’s agreement that “under existing intellectual property laws AI generated inventions are not eligible for protection,” calling it “the correct interpretation and understanding of **current law**,” but noting “we are equally interested in **what the law should be** in the future. In other words, we are considering what changes, if any, may need to be made to our intellectual property laws in order to incentivize future AI related innovations and creations.”<sup>57</sup> In asking for the agencies’ input, the senators suggested changes to US intellectual property laws “to continue encouraging the robust development of AI and AI generated inventions and creations . . . could include creating new forms of protection, like *sui generis* rights,” or other appropriate measures.<sup>58</sup> For the Copyright Office’s part, it acknowledged in the *Paradise* suit that “[t]he rapid development of AI technology, particularly systems capable of generating expressive material, raises many questions about its interplay with copyright laws,” issues the Office “will be addressing . . . in the coming year.”

---

<sup>53</sup> 17 U.S.C. § 102(a).

<sup>54</sup> *But see* note 57, *infra* (discussing *Naruto*, which held a monkey lacked statutory standing under the Copyright Act because the Act does not expressly authorize animals to file suit for copyright infringement).

<sup>55</sup> *See* Def.’s Resp. at 19, *supra* note 36.

<sup>56</sup> *See Naruto*, 888 F.3d at 420 (“We must determine whether a monkey may sue humans, corporations, or companies for damages and injunctive relief arising from claims of copyright infringement. Our court’s precedent requires us to conclude that the monkey’s claim has standing under Article III of the United States Constitution. Nonetheless, we conclude that this monkey—and all animals, since they are not human—lacks statutory standing under the Copyright Act.”). The *Naruto* case came to be when *Naruto*, a crested macaque living in Indonesia, took several photographs of himself using a camera left unattended by a wildlife photographer. The wildlife photographer published the images, claiming to be the copyright owner in *Naruto*’s “selfies.” People for the Ethical Treatment of Animals (PETA) sued the photographer for copyright infringement on *Naruto*’s behalf. The Ninth Circuit affirmed dismissal of PETA’s complaint because despite finding *Naruto* to have constitutional standing to sue on his own behalf, “[t]he Copyright Act does not expressly authorize animals to file copyright infringement suits under the statute.” *Id.* at 426. The Copyright Act also does not expressly authorize AI to file copyright infringement suits.

<sup>57</sup> Letter from Senators Thom Tillis and Chris Coons to Kathi Vidal, Under Secretary of Commerce for Intellectual Property and Director of the U.S. Patent and Trademark Office, and Shira Perlmutter, Register of Copyrights and Director of the U.S. Copyright Office (Oct. 27, 2022), *available at* <https://www.copyright.gov/laws/hearings/Letter-to-USPTO-USCO-on-National-Commission-on-AI-1.pdf> (emphasis in original).

<sup>58</sup> *Id.*

## Infringement by AI output

As discussed above, generative AI is “trained” on copyrighted material. Not only is using the copyrighted material in training potentially infringement, but the works generated by the AI based on the training materials potentially infringe the training materials.

To prove copyright infringement under US law, the plaintiff must show ownership of a valid copyright, copying of original elements, and substantial similarity between the infringing work and the copyrighted work. Whether the original work and the allegedly infringing work are “substantially similar” differs depending on the test applicable in a given court. For example, in the Ninth, Fourth, and Eighth Circuits, courts examine whether the works are substantially similar under both an extrinsic test, which analyzes the similarity of specific expressive elements and the intrinsic test, which subjectively analyzes whether an ordinary, reasonable person would think both works’ total concept and feel are substantially similar.<sup>59</sup> If either of these tests favor the defendant, there is no infringement.<sup>60</sup>

Dall-E provides an interesting conversation piece when it comes to potentially infringing AI output. On an information page for Dall-E 2, OpenAI explains that Dall-E 2 “can expand images beyond what’s in the original canvas, creating expansive new compositions.” The page displays a video of Dall-E 2 doing just that: taking the entirety of Johannes Vermeer’s painting “Girl with a Pearl Earring” (1665) and adding a body and a complete scene around the famed girl.<sup>61</sup> OpenAI further explains that Dall-E 2 “can take an image and create different variations of it inspired by the original,” showing exemplary AI-created variations next to Vermeer’s “girl” as well as Gustav Klimt’s “The Kiss” (1908) and Georges Seurat’s “Sunday Afternoon on the Island of La Grande Jatte” (1884).<sup>62</sup>

These works are all, of course, long out of any copyright protection. But Dall-E 2’s ability to create “variations” is not limited to aged masterpieces; to the contrary, the very same page shows variation examples made using contemporary images. It is plain to see how Dall-E 2’s capabilities, while entertaining, run the risk of generating AI-created images substantially similar to copyrighted originals, infringing US copyright law in the process. But there is a real question as to whether any individual user could be held liable for copyright infringement based on the output if the user did not know that the program was copying the protected work, as the Copyright Act requires a plaintiff to show “access” to the plaintiff’s work, which generally requires the alleged infringer to have actually viewed or heard the work.<sup>63</sup>

AI systems do not stop at mere variations on preexisting works. In late 2022, US social media became flooded with “A.I. selfies” created by an app called Lensa AI, which uses user-input “selfies” and AI to create portraits in a variety of styles.<sup>64</sup> The app, which users must pay for, generates 50 to 200 AI-generated images in themes such as “cosmic,” “fairy princess,” and “anime.”<sup>65</sup> Lensa uses Stable Diffusion, an AI-based image generator reportedly “trained on the creations of many artists who did not explicitly consent to the use of their work for Prisma Lab’s [the owner of Lensa AI] profit.”<sup>66</sup> The issue, of course, is that while one can copyright a work of art, like Andy Warhol’s iconic brightly-colored portraits of Marilyn Monroe, one cannot copyright “in the style of” Andy Warhol. But for artists who consider their style to be part of their identity, Lensa AI’s creations feel like an affront. Should those creations get “close enough” to an original work, they could cross the line into an infringement.

---

<sup>59</sup> See *Shaw v. Lindheim*, 919 F.2d 1353, 1356-57 (9th Cir. 1990); *Towler v. Sayles*, 76 F.3d 579, 583 (4th Cir. 1996); *Designworks Homes, Inc. v. Thomson Sailors Homes, L.L.C.*, 9 F.4th 961, 963-64 (8th Cir. 2021).

<sup>60</sup> *Kouf v. Walt Disney Pictures & Television*, 16 F.3d 1042, 1045 (9th Cir. 1994).

<sup>61</sup> See DALL-E 2, OpenAI, <https://openai.com/dall-e-2/> (last visited Feb. 6, 2023).

<sup>62</sup> *Id.*

<sup>63</sup> *Sid & Marty Krofft Television Prods., Inc. v. McDonald’s Corp.*, 562 F.2d 1157, 1172 (9th Cir. 1977).

<sup>64</sup> See Madison Malone Kircher & Holtermann, Callie, “How Is Everyone Making Those A.I. Selfies?,” *The New York Times* (Dec. 7, 2022), <https://www.nytimes.com/2022/12/07/style/lensa-ai-selfies.html> (last visited Feb. 6, 2023).

<sup>65</sup> *Id.*

<sup>66</sup> *Id.*

## Conclusion

In conclusion, while European and US law may both allow generative AI creators to train their systems, via deep learning technologies, using content protected by copyright and/or related rights thanks to the exceptions of (i) text and data mining in one case and - to a lesser extent - (ii) fair use in the other, neither of these two systems seems to allow generative AIs to own copyrights on the generated works. Moreover, these outgoing works are likely to give rise to claims of infringement each time they borrow too much from the incoming works, which raises the question of the related liabilities in such a case.

As the problems of generative AI go beyond borders, a worldwide harmonization of the exceptions authorizing web scraping and deep learning could prove to be relevant, as well as a harmonization of the regime applicable to output data in order to determine the respective rights of AI designers and users on the latter. A reflection in this sense is already underway at the WIPO level<sup>67</sup>. It will be interesting to see if it succeeds.

---

<sup>67</sup> Proposal for harmonization of copyright rules at the World Intellectual Property Organization by Sean Flynn and 17 other researchers. Cf. M. CLAVEY, "The Legal Complexity of Automatic Text and Data Mining," Next Impact, December 6, 2022

# Authors

## On the French side (Racine Avocats)



**Charles Bouffier**  
Partner  
cbouffier@racine.eu



**Victoire Danès**  
Associate  
vdanes@racine.eu

**Emma Gausson**  
Intern

## On the US side (Miller Nash LLP)



**Delfina Homen**  
Special Counsel  
delfina.homen@millernash.com



**Brian Esler**  
Partner  
brian.esler@millernash.com