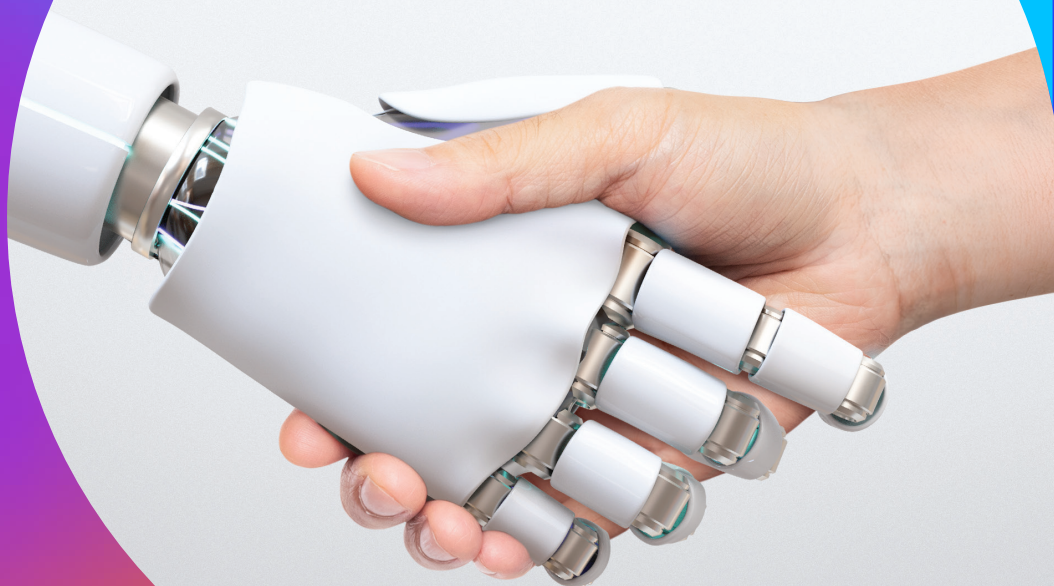


AI and copyright: Unveiling the legal challenges



Intro

The potential of AI seems unlimited, and the hype surrounding it is ever-growing. With the launch of generative AI-powered tools like ChatGPT, Stable Diffusion, Bard, DALL-E or Speechify, to name only a few, businesses worldwide are embracing this major technological innovation and leveraging creativity, efficiency and productivity gains among their teams. However, we are only about to see the tip of the iceberg regarding the potential disruption triggered by generative AI.

Bill Gates recently wrote in an open-ed that “the development of AI is as fundamental as the creation of the microprocessor, the personal computer, the Internet, and the mobile phone”. No matter what comparisons one might favour, it is undeniable that generative AI is meant to stay and will continue to develop at an unbelievable speed. For businesses, and agencies in particular, this presents an unlimited potential of opportunities but also comes with certain risks.

What the following pages entail

In the subsequent lines, we will distinguish between the implications around the legitimacy of the reuse of the input data on which algorithms are trained before focussing on the protection that can be granted to AI-generated output data. After that, we will present the specificities of UK law on ‘computer-generated works’ and highlight how they differ from EU and US law. Finally, we will briefly summarise what these findings mean for agencies and provide a few recommendations when it comes to using generative AI tools within an agency environment.

We invite you to flip through this publication and uncover a few of the most pressing legal challenges associated with using generative AI tools and how agencies can best mitigate risks in this context.

Contents

Contents

Definitions.....	3
The legal challenges around Generative AI.....	4
The legal implications around the input data.....	5
The legal challenges around the output data.....	7
What are the current criteria for AI-assisted output to be granted copyright protection?.....	8
The UK position on computer-generated works.....	8
What does this mean for agencies regarding the possibility of using generative AI tools?.....	10
Conclusion.....	11

Definitions

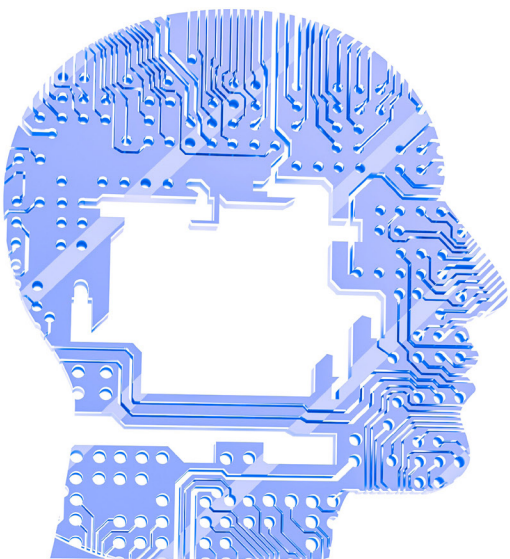
What is Artificial Intelligence (AI)?

Artificial intelligence is a field, which combines computer science and robust datasets, to enable problem-solving. It also encompasses sub-fields of machine learning and deep learning, which are frequently mentioned in conjunction with artificial intelligence. These disciplines are comprised of AI algorithms which seek to create expert systems which make predictions or classifications based on input data¹.

What is Generative AI?

Generative AI is a branch of artificial intelligence which focuses on creating models capable of generating new content that resembles existing data. Unlike traditional machine learning algorithms that can only analyse or act on existing data, generative AI can learn from existing artefacts to produce new ones that reflect the characteristics of the training data while avoiding repetition. It can produce novel content, such as images, video, music, speech, text, software code and product designs.

Generative AI has many different use cases for agencies, ranging (without being limited to) from various possibilities in graphic and audiovisual content creation to copywriting, developing chatbots for customer service or generating personalised marketing content. It will, without any doubt, redefine and offer new options for improvement in the agency-client relationship.

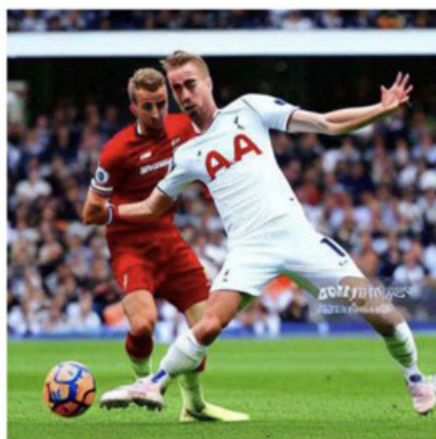


¹What is Artificial Intelligence (AI)?, IBM

The legal challenges around Generative AI

Generative AI relies on extensive datasets to train its algorithms, which poses various legal challenges, ranging from intellectual property rights infringements to privacy and data protection issues, media quality guarantees, etc. In this paper, we will primarily focus on challenges relating to intellectual property rights, and especially to copyright. Copyright laws serve as a legal framework to promote the progress of the arts and sciences by incentivising authors and protecting works from being reproduced without permission². With the rise of generative AI technologies, existing copyright frameworks have come under increased scrutiny over the last few years and will likely evolve to level the technological evolutions in this field.

Indeed, several recent court decisions have highlighted the challenges surrounding generative AI and copyright. In February 2023, Getty Images filed a lawsuit³ against Stability AI in the United States for copying over 12 million photographs from its collection. In this lawsuit, the claim identified some of the output delivered by Stability AI to include a modified or distorted version of a Getty Images watermark, underscoring the clear link between the copyrighted images and the final product. Similarly, in a case⁴ filed in late 2022, three artists formed a class to sue multiple generative AI platforms based on the AI using their original works without a license to train their AI. The plaintiffs claim that this allows users to generate pieces that may be insufficiently transformative from their existing, protected works, and, as a result, would be unauthorised derivative works⁵.



*This is an illustration from Getty Images' lawsuit, showing an original photograph and a similar image (complete with Getty Images' watermark) generated by Stable Diffusion.
Image: Getty Image*

² Christopher May, 'The Venetian Moment: New Technologies, Legal Innovation and the Institutional Origins of Intellectual Property', 2002

³ Getty Images (US), Inc v Stability AI, Inc, Case 1:23-cv-00135-UNA

⁴ Andersen et al v. Stability AI Ltd. et al, Case 3:23-cv-00201

⁵ [Generative AI Has an Intellectual Property Problem](#), G. Appel, J. Neelbauer & D. Schweidel, Harvard Business Review, 2023

The legal implications around the Input data

To best analyse the legal challenges of data reuse and copyright protection in generative AI tools, we will first focus on the implications around the input data (i.e. the data used to train AI models) before addressing implications regarding the output data (i.e. the legal protection that can be granted to works that are produced using generative AI tools).

Starting with input data, employing generative AI for image applications, such as text-to-image, image-to-image, or semantic image-to-photo, raises significant legal questions regarding licensing AI-generated content for commercial reuse, and differs depending on the applicable legal framework between the EU, US and the UK.

In the EU and the UK, and as a general principle, AI tools demand input data suitable for commercial reuse, meaning it must not be protected by copyright to be lawfully trained. In this context, commercial reuse refers to the purpose for which the algorithms are trained⁶.

In the US, the legal framework entails one notable exception: the fair-use doctrine.

The fair-use doctrine permits the reuse of copyrighted material without prior permission from the rightsholder “for purposes such as criticism (including satire), comment, news reporting, teaching (including multiple copies

for classroom use), scholarship, or research,” and for transformative use of the copyrighted material in a manner for which it was not intended⁷. It is one of the limitations of copyright, which is intended to balance the interests of copyright holders with the public interest.

In this context, it is not a coincidence that most generative AI tool algorithms find their training grounds in the US. Many tech companies are justifying the legitimacy of reusing copyright-protected input data to train their algorithms primarily by referring to the fair use doctrine. However, the extent to which fair use can be applied to generative AI input data, especially for commercial purposes, remains to be determined at this stage.

In Europe, a copyright exception on text and data mining exists under the 2019/790 ‘DSM’ Directive⁸. The DSM directive admits the possibility of text and data mining purposes for commercial use under two conditions:

1. Provided that the content could be accessed legitimately for text and data mining.
2. Provided that the owner of the copyright and related rights has not expressly reserved the extraction of text and data and thus relied on the so-called opt-out mechanism.

⁶ Recital 42 of the Infosoc Directive 2001, [Consolidated TEXT: 32001L0029 — EN — 06.06.2019 \(europa.eu\)](#)

⁷ 17 U.S. Code § 107 (US Copyright Act 1976) - Limitations on exclusive rights: Fair use

⁸ Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market [L 2019130EN.01009201.xml \(europa.eu\)](#)

In practice, it can be difficult for rightsholders to take efficient advantage of the opt-out mechanism, as the latter may choose among different options to opt-out, which are not necessarily all limited to a machine-readable format and thus can be challenging to respect regarding the automated training of vast amount of data to feed AI algorithms.

The legal implications around the Input data

Moreover, many rightsholders claim not to be aware that their works are used for text and data mining purposes. Thus, they cannot efficiently take advantage of the opt-out mechanism, which limits the legal certainty around the legitimacy of reusing such data.

How is the EU trying to address copyright infringements in the training of AI models ?

The EU flagship Regulation Proposal on artificial intelligence, the so-called AI Act⁹, tries to address this problem by requiring tech companies to set up public repositories that encompass copyright-protected datasets on which relevant algorithms are trained.

To this end, article 28b of the consolidated draft AI Act¹⁰ states that: “Providers of foundation models used in AI systems specifically intended to generate (...) content shall (...) document and make publicly available a sufficiently detailed summary of the use of training data protected under copyright law. It appears, however, questionable to what extent these disclosures will prove effective as many legal experts claim that the low threshold of originality, the territorial fragmentation of copyright and the poor state of rights ownership metadata demonstrate the impossibility of such a transparency obligation to prove efficient¹¹ .



⁹ [Proposal](#) for a regulation of the European Parliament and the Council laying down harmonised rules on artificial intelligence

¹⁰ Amendments adopted by the European Parliament on 14 June 2023 on the AI Proposal [TA \(europa.eu\)](#)

¹¹ [Generative AI, Copyright and the AI Act](#), KluwerCopyrightBlog, J. Quintais, 2023

The legal challenges around the Output data

Besides the copyright protection issues of the input data, the question also arises as to whether AI-generated output can be granted copyright protection. A first important distinction to establish in this context is the one of AI-assisted versus AI-generated output¹².

AI-generated output refers to the generation of an output by AI without any human intervention, while AI-assisted output is generated with material human intervention and/or direction¹³. Except for the case of the UK, AI-generated output cannot as of today be granted copyright protection in the US and the EU as it does not qualify as 'work'¹⁴.

The degree to which AI-assisted output can be granted copyright protection varies in jurisdictions and relies on a case-by-case assessment. In this context, the US Copyright Office made in February 2023 a groundbreaking decision, further tightening the boundaries between AI-generated and AI-assisted content, particularly in terms of copyright protection.

The decision arose from a case¹⁵ involving a comic book, "Zarya of the Dawn," by Kris Kashtanova. The author sought copyright protection for the book, which included images generated by the AI tool 'Midjourney.' Kashtanova argued that while AI played a supportive role in the creative process, the book was not solely created by the AI.

Indeed, the author asserted that she conceived and structured the story, designed the layout of each page, and performed the art direction.

Nevertheless, while the textual elements were being granted copyright protection, the US Copyright Office declared that the images generated through AI were not. If this case law should be confirmed in the future, it could pave the way for AI-generated content, which relies on little to no human input, to systematically end up in the public domain. It is however too early at this stage to make any definite assumptions.



"Zarya of the Dawn", written by Kris Kashtanova. Image generated with Midjourney

¹² [The qualification of AI creations as "works" under EU copyright law](#), P. De Grauwe & Sacha Gryspeerd, 2022

¹³ Ibid

¹⁴ Ibid

¹⁵ US Copyright Office (USCO), [Zarya of the Dawn case](#), 2023

What are the current criteria for AI-assisted output to be granted Copyright protection?

The current legal framework appears to be increasingly ill-adapted to address the challenges AI-generated content poses, because of the historical human-centric approach of copyright protection and the potential difficulties distinguishing human-made from computational acts of creation. However, if we apply the current copyright protection criteria to generative AI tools, the main elements to be taken into consideration are briefly summarised hereafter.

For AI-assisted output to be granted copyright protection, it must first qualify as 'work'. Article 2 of the Berne Convention on Copyright holds that literary and artistic works include every production in the literary, scientific and artistic domain¹⁶. Further, AI-assisted output must be the result of human intellectual effort. The EU Court of Justice has clarified in a notable case that the use by the author of their creative freedom must be perceptible in their expression¹⁷. This does however not exclude the possibility of creating works of authorship with the aid of a machine or device¹⁸. Last, the EU Copyright Directive (2001/29/EC), as interpreted by well-established case law, requires the concept of originality¹⁹. In other words, for copyright to subsist in a "work", it must be original, meaning that it must represent the author's own intellectual creation, and it must be an expression of that intellectual creation²⁰. Ideas lacking a specific shape or form cannot qualify as works that can be copyright-protected. challenges will likely keep experts busy for the months and years to come.

The UK position on computer-generated works

As mentioned before, in the EU and the US, copyright thus relates in practice to a human creator. This has been established in the above-mentioned EU case law with the elaboration on the concepts of originality and creative choices²¹, and confirmed in the US in the now famous "monkey selfie" dispute²², during which both the Copyright Office and the courts found that animals could not hold copyright.



One of the monkey selfies at issue in the dispute Public Domain

¹⁶ [Berne Convention for the Protection of Literary and Artistic Works, 1886, as amended 1979](#)

¹⁷ CJEU 1 December 2011, Painer v Standard, C-145/10

¹⁸ *Ibid*

¹⁹ CJEU 16 July 2009, Infopaq v Danske Dagblades Forening, C-5/08

²⁰ *Ibid*

²¹ *Supra* 18

²² *Naruto v. Slater*, USCA, Ninth Circuit, April 23, 2018

The UK position on computer-generated works

Moving away from the US and the EU, in some copyright laws of the British tradition –including in the UK, Ireland, New Zealand, and South Africa– the requirement of human authorship has been circumvented by establishing the authorship of “computer-generated works” in cases where no human authorship can be established²³. Under the Copyright, Designs and Patents Act 1988 (“CDPA”), artistic works created by a computer can benefit from copyright protection. Section 178 of the CDPA provides that a “computer-generated” work is generated by a computer in circumstances where there is no human author of the work. In those cases, the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken²⁴. The CDPA thus creates a legal fiction as it constructs an artificial author for computer-generated works, which by definition are authorless²⁵.

However, this appears to be problematic from a legal point of view because the CPDA is not equipped to deal with the complexities of AI-generated works²⁶. In 1988, when the provision was enacted, a human author could always be identified due to their clear involvement in programming the computer outputs, which is different with AI-generated works, for they do not rely on creative human input²⁷. Despite drastic technological advancements, the CDPA has not been updated since its enactment and thus does not provide for more clarity regarding copyright protection.

Returning to the earlier example of the comic book “Zarya and the Dawn”, assuming it qualifies as an original work of intellectual creation, as required for copyright protection, if Kris Kashtanova, a UK citizen or working for a UK company, had created the work, different legal experts claim it would likely have had high chances to be granted copyright protection in the UK, including for the images generated²⁸.

However, the legal landscape in the UK is also likely to change in the future. With the increased use of generative AI tools, the House of Commons Science, Innovation and Technology Select Committee held an evidence session in May 2023 in the UK parliament on the impact of AI in the creative industry. Experts suggested that the CDPA’s approach to “computer generated” works is no longer appropriate as AI is less of a tool that aids in the creation of works, but rather, is what creates the works²⁹.

²³ AIPPI German Delegation: Copyright in artificially generated works national [report](#), J. Osha Et al., 2019

²⁴ UK Copyright, Design and Patent Act 1998, Section 9(3).

²⁵ E. Bonadio & L. McDonagh, ‘AI as Producer and Consumer of Copyright Works: Evaluating the Consequences of Algorithmic Creativity’, 2020

²⁶ E. Chaw, Algorithmic Creativity: [How Should the UK Copyright Regime Accommodate Autonomous AI-generated Works?](#), 2023

²⁷ J. Lee, ‘Computer-generated Works under the CDPA 1988’, Artificial Intelligence and Intellectual Property (Oxford Academic), 2021

²⁸ C. Daniel, J. Grasser, J. Collis, Copyright protection for AI works: UK vs US, Global IP & Privacy Law Blog, Squire Patton Boggs, 2023

²⁹ Ibid

What does this mean for agencies regarding **The possibility of using generative AI?**

Using generative AI tools for production purposes with clients appears to be often unsafe and needs to be assessed on a case-by case basis, as it cannot be generally assumed that the algorithms have been trained on data fit for reuse.

A notable exception to this rule represents Adobe's Firefly. Adobe claims to be training its algorithms on data for which it owns the right to reuse or provides financial compensation to the authors³⁰. Nonetheless, from an operational perspective, Firefly is often considered to have limitations regarding integrations and creative capabilities compared to other generative AI tools in the market. Consequently, as long as the status quo prevails, agencies will most likely not be unable to unlock the full potential of many of the existing tools in creative production processes. Besides, most contracts with clients entail warranties and indemnities that deliverables cannot infringe copyright (or third-party rights generally).

Thankfully, many use cases involving generative AI do not pose the same legal uncertainty as 'for image applications', and prove very useful daily, such as some use cases involving different language models (ChatGPT, Bard etc..), campaign optimisation and content creation, the development of client-specific chatbots, etc. And again, the most significant potential of generative AI is likely still ahead of us.



³⁰[Adobe Firefly vs. Midjourney: How Firefly can help speed up creative workflows, 2023](#)

Conclusion

Establishing an internal agency policy on AI usage is highly recommended for the inherent risk of working with such tools in the production phase with clients. It is essential to understand in which cases generative AI tools are safe to deploy and in which case they are not.

Before using AI tools, it is helpful to review the AI service's terms and conditions, mainly its terms around copyright, warranties, and any exclusion or limitation of liability. This provides valuable insights into the legal risks of relying on these tools. Likewise, it is recommended to review specific contracts with clients and IPR-related terms and conditions. These agreements often specify that deliverables need to be free of copyright infringements.

Whenever feasible, it is recommended to work with AI tools trained on data that is safe for reuse. As innovation in AI tools is fast-paced, generative AI tools that are safe for reuse should be expected to expand significantly in number in the coming months.

Please note that these recommendations remain subject to change, as is the evolving legal framework. This publication does not constitute legal advice.





Alexis Bley
Public Affairs Manager, EACA
alexis.bley@eaca.eu

Images from unsplash.com and vecteezy.com

www.eaca.eu