AI Training and Copyright: Towards a Remuneration Right for Creators

A & K Metaxopoulos & Partners Law Firm **View In Analytics**

European Union, Greece | March 18 2026

Introduction

The rapid development of technology has led to the creation of Artificial Intelligence (AI) systems, a new technological innovation, that is increasingly shaping modern society. The evolution of *machine learning* has led to the development of a more advanced and innovative technology for the creation of digital content, known as Generative Artificial Intelligence. Machine learning constitutes a process through which new content is generated or created. This process can be compared to a form of learning in which the machine analyzes existing data and uses the knowledge contained in them.

A characteristic example today is the widely used generative AI platform known as ChatGPT, which primarily produces digital content. Its operation relies on the input of large volumes of data into the system, such as texts, information, images, and various forms of content, which are mainly obtained from publicly accessible sources on the internet. For the training of AI algorithms, techniques such as Text and Data Mining (TDM) are used, which involve large-scale analysis of data.

The main issue is whether the use of material, namely input data, that constitutes "works" under copyright law and is used to train these systems amounts to copyright infringement, and whether creators should be entitled to compensation or remuneration for such use of their works.

Text and Data Mining

At the EU level, the main regulatory framework governing Text and Data Mining (TDM) was introduced by Directive (EU) 2019/790 (CDSM Directive), which establishes two specific exceptions. These exceptions limit certain exclusive rights, in particular the right of reproduction and the right of extraction of works or other protected subject matters, for the purpose of carrying out text and data mining. First, Article 3 establishes an exception for text and data mining for the purposes of scientific research. This exception applies in favor of research organizations and cultural heritage institutions. The second exception, set out in Article 4, has a broader scope and may also cover commercial uses, provided that rightholders have not expressly reserved their rights (opt-out), in particular through appropriate and machine-readable means in the online environment. In this case, the creators must explicitly reserve their rights otherwise the TDM is considered permitted without prior authorization.

It is noted that, following the introduction of the EU AI Act, it has been confirmed that the exceptions introduced by the 2019 Directive also apply to the training of AI models. In particular, Recital 105 explicitly clarifies that *the text and data mining of data protected by copyright or related rights for the purpose of training AI models requires the authorization of rightholders, unless a relevant exception applies, such as those provided for in Articles 3 and 4 of Directive (EU) 2019/790.*

Creators' Right to Remuneration

Within the above framework, it is worth examining whether creators should be entitled to remuneration for the use of their works by AI systems. It should be noted that the need to remunerate creators for the use of their works in the training of AI systems is primarily grounded in the structural dependence of AI models on human creativity. To be more specific, creators produce works through an ongoing dialogue with society, drawing inspiration from social, political, and cultural conditions and transforming them into original expression. By contrast, AI systems operate on the basis of their pre-existing training data, without autonomous access to social experience. Therefore, their "creativity" essentially relies on the pre-existing works of human creators. As AI can produce content quickly and at low cost, content that may compete with or even replace human creations,

there is a risk that economic value will shift away from creators. From this perspective, ensuring fair remuneration for creators becomes crucial for maintaining a balanced and sustainable relationship between AI innovation and human creativity.

It is important to note that when a copyright exception applies, the exclusive nature of the author's right is limited and, under the current legal framework, creators are not entitled to remuneration or compensation for such use. In particular, in the case of text and data mining for scientific research, the exception applies broadly, meaning that no copyright infringement occurs that could lead to a right to remuneration. Moreover, Recital 17 of Directive (EU) 2019/790 expressly states that *rightholders are not entitled to remuneration, as any potential harm resulting from scientific use is considered negligible.* However, under the second form of the general exception for commercial purposes, the Directive provides for a right of reservation (opt-out), which could form the basis for a remuneration right. In particular, the opt-out right could be compared to the traditional licensing regime, under which the author may prohibit or permit the use and exploitation of their work in exchange for remuneration. Therefore, in the context of text and data mining, an argument *de lege lata* may already be derived from the existence of this opt-out right.

In practice, however, the exercise of the opt-out must be expressed through machine-readable means. Consequently, in order for the use of a work to be allowed in exchange for remuneration, specific technical systems for reserving rights would be required. These systems should enable the identification of rightholders and the automated processing of payments. Otherwise, the exception could not work in practice through traditional licensing agreements. Moreover, considering that text and data mining requires access to extremely large volumes of data, licensing at the level of individual works or databases would render the entire process impracticable.

Within this context, an alternative approach to the remuneration of rightholders has been proposed, drawing inspiration from the model of private copying exception. In order to balance the rights of creators and the enjoyment of works by users, the legislator has recognized that the absolute and exclusive nature of copyright may sometimes lead to unjustified restrictions on the use of works, leaving them unused or underexploited. A characteristic example is private copying, which is permitted without prior authorization. In such cases, a statutory license is created and the author cannot prohibit this specific use of their work. This license is based on an economic consideration, namely the right of creators to receive fair remuneration.

Thus, it could be argued that a right to fair remuneration for creators could be established in the context of training AI models, due to the limitation of the absolute and exclusive nature of copyright. It may therefore be preferable to replace the opt-out mechanism provided in Article 4 par. 3 of Directive (EU) 2019/790 with a remuneration right. In this way, similarly to the system of private copying, works could be used freely: AI companies would not be required to obtain individual licenses from rightholders in order to use such content for machine learning purposes, while creators would receive fair remuneration as financial compensation.

The question of how it will be possible to determine which works have been used in the training of an AI system, in order to calculate remuneration, is addressed by AI Act. The Regulation imposes transparency requirements on developers of AI systems, according to which they must provide a sufficiently detailed summary of the creative content used to train their models. Accordingly, providers of AI systems are subject to an obligation to publish summary information regarding the training data protected by copyright law. The purpose of this requirement is to facilitate cooperation between rightholders and AI system providers with regard to this new form of exploitation of works.

Another approach to the remuneration of creators has also been proposed. More specifically, this approach concerns the establishment of a lump-sum remuneration system. Under this proposal, the payment of such remuneration would depend on a "risk of substitution" criterion. In particular, it would be assessed in each case whether the creative content generated by AI could substitute human works protected by copyright. For example, if an AI system produces content resembling a literary work, remuneration would be due. This proposal is based on the doctrine of *domaine public payant*, which suggests that remuneration could be paid for the use of works that have entered the public domain, that is, works whose copyright protection has expired, in order to support creators financially. It should be noted, however, that this model has not been implemented in practice. Nevertheless, this approach could serve as a possible model for the remuneration of creators for the

use of works in AI training. Under this model, revenues from AI providers would be collected by a collective organization and distributed to rightholders as remuneration, in order to improve their living and working conditions.

Conclusion

It is evident that the level of development of AI models has reached a point where they can, to a significant extent, substitute human creativity. However, this development may lead to a substantial decline in human creative activity, potentially resulting in a form of "industrial" production of works in which emotional, interpersonal, and other forms of human intelligence are marginalized and incorporated into standardized patterns and repetitive structures.

Over time, works generated by AI may lose the creative elements they imitate from existing works, as well as their originality, since they may no longer reflect the realities of contemporary society, especially if human creators stop producing new works as AI increasingly dominates the art market. For this reason, protecting and supporting human creators is essential. Without them, the production of works risks becoming a mechanical process that could reduce the role of art to purely technical criteria and weaken its element of personal creative expression.

The solution therefore lies in a balanced regulatory approach: the development of a framework that supports technological innovation while safeguarding the value of human creativity. A fundamental principle of such a framework should be the protection of creators, which must be implemented in practice through meaningful and effective support, in particular by ensuring remuneration for the use of their works.

A & K Metaxopoulos & Partners Law Firm - Stavros Varveris