

# Kluwer Copyright Blog

## The Jungle Bird, El Diablo, and the Zombie or Machine Learning Models, Computer Programs and Copyright put to the test

Begoña González Otero (Max Planck Institute for Innovation and Competition) · Monday, December 13th, 2021

European and international policymakers have raised how artificial intelligence (AI) interacts with intellectual property (IP) law on several occasions. Nonetheless, before any policy and law-making endeavour can be undertaken, a fitness test of the existing IP framework is indispensable. Recent discussions have focused on AI-aided and AI-generated output, concentrating on whether an AI system can be a creator or an inventor (see, among others, [here](#) and [here](#)). Nonetheless, a more holistic view accounting for the role of IP law across the AI innovation cycle would be beneficial and elemental. Against this backdrop, I have in a recent [paper](#) developed such a fitness analysis to the so-called “core-components” of Machine Learning (ML) systems.



Image by [Dorothee QUENNESSON](#) from Pixabay

Those versed in the intersection between IP rights and technology know that computer programs protection has always caused inevitable frictions IP law. Even if an international agreement were reached to protect computer programs (code) as literary works of copyright (See [WIPO Copyright Treaty, 1996](#)), code protection would be at the centre in a world increasingly dominated by digitisation and “[softwarization](#)”. Nevertheless, when one considers ML models, the first issue is conceptualising them. ML models have been called ‘learning algorithms’, ‘AI computer programs’, and ‘software 2.0’. However, there is no unanimity about how to technically defined them (see [here](#)). This is pertinent from a copyright perspective because the regime of protection granted by copyright will be different depending on whether the ML model qualifies as a computer program, as a mathematical method, or as another type of work. Additionally, all proprietary and open

source software licensing rely on copyright protection. The license is not triggered in most open licenses if applied to subject matter not protected by copyright (or related rights). Thus, questioning whether EU copyright law provides adequate protection for ML models is not a naïve endeavour.

The paper begins with an overview of computer programs' justifications for copyright protection, finding similarities with ML models. However, in machine learning, the input is training data, and the desired result comes from the ML model, based on a trial-error process. ML models are learners that generate an output by making inferences from data. This leads to a first preliminary conclusion, questioning whether copyright protection of computer programs has or may have any practical relevance in the world of ML models.

The second part delves into the legal framework for copyright protection of computer programs. This is followed by technological explanations of ML model architectures and the intermediate stages as a part of the ML system to ensure a sufficient understanding of the legal challenges, focusing on the differences with traditional software. It also classifies the different ML techniques and training methods used therein to frame the type of work that an ML model would be (or, as engineers would call it, "a simple ML workflow"). Based on the results of this conceptualisation, where ML models are considered as computer programs and as independent works, the final part puts each of them to the copyright test.

The fitness test considers the [Software Directive](#) and all relevant case-law of the CJEU, mainly focusing on the requirements of originality and authorship. To the first test, the result is that for ML models, even if we might have a computer program in the sense of copyright, it will be excluded from protection to the extent that a valid authorship claim cannot be articulated. Moreover, in those cases where authorship could be claimed, the scope of protection would be limited to that concrete formulation or specific arrangement of the algorithms in the ML model, to its inner structure. Accordingly, the most valuable part of the model – the functionality of such an arrangement – would not be protected and would remain in the public domain.

Assessing ML models as independent works does not bring much better results. The purposive interpretation of the Software Directive by the CJEU in *SAS Institute Inc. v World Programming Ltd* may appear to offer ground for concluding that choices for interfaces concerning the implementation of abstract ideas contained in the source code could be sufficiently original, as those concerning their languages or formats. The same could be said of ML models if replacing languages and formats with programmed backpropagations or programmed optimisation procedures such as stochastic gradient descent.

However, things get slipperier when we need to identify an ML model as a work. Bear in mind that an ML model is a learner that, on the one hand, follows the instructions provided by a programmer and does what is told to do, but on the other hand (and according to the same learning principles) has a specific capacity for optimisation that goes beyond given instructions. In this context, identifying an ML model as work would entail a general contradiction for the copyright system. In other words, identifying when an ML model has reached a sufficient creative level would imply that the originality requirement is objective, i.e. it can be appreciated compared to pre-existing creations. This opens the door for questioning whether the criterion of minimum creative effort to protect human intellect works should not be subject to revision and re-evaluation. Moreover, it brings us back to the debate of protecting functional works under the copyright regime.

The fitness legal test results are clear. Even if it were possible in some individual cases to “express” an ML model as a computer program, the scope of copyright protection would not cover its most valuable feature: its functionality. Thus, most cases would remain in the public domain under impossible authorship claim articulation. However, the absence of copyright protection for ML models does not seem to affect a potential risk of misappropriation or piracy. The reason is that the factual control over many other elements, components, and parameters of the ML system is crucial for potential re-use. Access to APIs, such as in the case of OpenAI’s model GPT-3, the use of technical protection measures and access to training data limited by database rights are some already existing examples. These topics are in need of further research.

On the whole, it seems there is no justification for the creation of a sui generis right or ancillary right for the protection of ML models.

For more details, have a look at the paper, open access, [here](#).

---

*To make sure you do not miss out on regular updates from the Kluwer Copyright Blog, please subscribe [here](#).*

## Kluwer IP Law

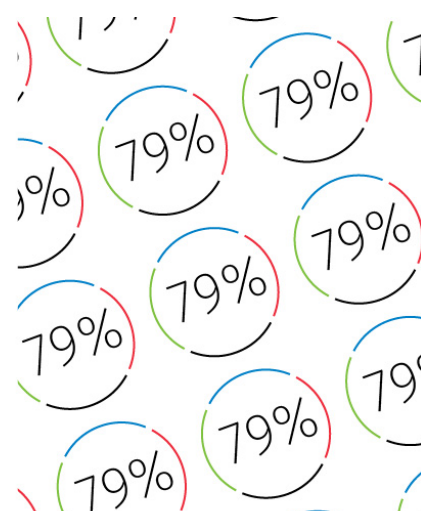
The **2022 Future Ready Lawyer survey** showed that 79% of lawyers think that the importance of legal technology will increase for next year. With Kluwer IP Law you can navigate the increasingly global practice of IP law with specialized, local and cross-border information and tools from every preferred location. Are you, as an IP professional, ready for the future?

Learn how **Kluwer IP Law** can support you.

---

79% of the lawyers think that the importance of legal technology will increase for next year.

**Drive change with Kluwer IP Law.**  
The master resource for Intellectual Property rights and registration.



2022 SURVEY REPORT  
The Wolters Kluwer Future Ready Lawyer  
Leading change

---

This entry was posted on Monday, December 13th, 2021 at 8:06 am and is filed under [Artificial Intelligence \(AI\)](#), [Authorship](#), [Originality](#), [Software](#)

You can follow any responses to this entry through the [Comments \(RSS\)](#) feed. You can leave a response, or [trackback](#) from your own site.