

Kluwer Copyright Blog

TDM: Poland challenges the rule of EU copyright law

Paul Keller (Institute for Information Law (IViR)) · Tuesday, February 20th, 2024

When life gives you lemons, make lemonade. This must have been the key insight at the Polish Culture and National Heritage Ministry when the new administration took over and discovered that more than 2.5 years after the implementation deadline, Poland still had to implement the provisions of the [2019 Copyright in the Digital Single Market Directive](#) into national law. So how do you make lemonade out of the fact that you are [the only EU Member State without an implementation](#)? You claim that the delay allows you to propose a better implementation.



Basia ?abaj, CC BY-SA 4.0, via Wikimedia Commons

In this particular case, the government claims that the delay allowed it to properly consider the impact of generative AI on copyright and come to the conclusion that training generative AI systems on copyrighted works does not in fact fall within the scope of the text and data mining exceptions contained in the directive. From the explanatory memorandum accompanying [the draft implementation law published on Thursday last week for public consultation](#) (all quotes below are own translations from the polish original):

The implementation of the directive now, in 2024, dictates that we refer here to the issue of artificial intelligence and the question of whether text and data mining within the meaning of the directive also includes the possibility of reproducing works for the purpose of machine learning. Undoubtedly, at the time the directive was adopted in 2019, the capabilities of artificial intelligence were not as recognizable as they are today, when “works” with artistic and commercial value comparable to real works, i.e., man-made, are beginning to be created with the help of this technology. Thus, it seems fair to assume that this type of permitted use was not conceived for artificial intelligence. An explicit clarification is therefore introduced that the reproduction of works for text and data mining cannot be used to create generative models of artificial intelligence.

This “explicit clarification” can be found in the text of the proposed implementation for both articles 3 and 4 of the CDSM directive. The article 3 implementation states that cultural heritage institutions and academic research organizations...

may reproduce works for the purpose of text and data mining for scientific research, with the exception of the creation of generative models of artificial intelligence, if these activities are not performed for direct or indirect financial gain.

The same exception to the exception can also be found in the implementation of the general text and data mining exception:

It is allowed to reproduce distributed works for the purpose of text and data mining, except for the creation of generative artificial intelligence models, unless otherwise stipulated by the authorized party.

It is worth stressing that the language quoted above is contained in the public consultation version of the implementation law and thus not final. It also seems clear that this language has not been widely consulted within the Polish government as it [clearly contradicts efforts undertaken by other parts of the government](#). Still it is worth taking a closer look at the rationale behind this implementation and to assess the conformity with the provisions of the directive and the overall impact of the proposed approach.

A flawed rationale

First of all, while it is understandable that lawmakers seek more clarity about the relationship between the EU copyright framework and the use of copyrighted works for training AI models, the assumption that the TDM exceptions were “not conceived for artificial intelligence” is simply wrong. While there is little publicly available documentation of what lawmakers had in mind when they agreed on the structure of the TDM exceptions, what is available makes it clear that the development of artificial intelligence was explicitly factored into the discussions. Both the [European Parliament statement](#) and the [European Commission’s explainer of the directive](#), published after the adoption of the directive in March 2019 specifically highlight that the TDM exception in Article 4 was introduced “in order to contribute to the development of data analytics and artificial intelligence”.

If there was any doubt if the exception was conceived in order to facilitate the development of generative Artificial Intelligence, this relationship was further clarified in March 2023 (at a time when the impact of Generative AI was widely recognized). In response to [a Parliamentary question](#) that suggested that “The [CDSM] Directive does not address this particular matter”, [Commissioner Breton pointed out](#) that TDM exceptions do in fact “provide balance between the protection of rightholders including artists and the facilitation of TDM, **including by AI developers**”.

Finally [the upcoming Artificial Intelligence Act](#) — which has been supported by the Polish government — contains a provision that points out that developers of generative AI systems must “put in place a policy to respect Union copyright law in particular to identify and respect, including through state of the art technologies, the reservations of rights expressed pursuant to Article 4(3) of [the CDSM] Directive”. In addition, the AI act also contains a recital (60i) that explains the interaction between the training of generative Ai systems and the exceptions contained in article 3

& 4 of the copyright directive.

All of this makes it clear that “now, in 2024” the TDM exceptions as introduced in 2019 do in fact provide the framework for the use of copyrighted works for the purpose of training generative AI systems, even though some stakeholders would much prefer that this was not the case.

Compliance with the Directive

It is also clear that any attempt to exclude from the scope of the TDM provision the reproductions made in the context of training generative AI models would, *prima facie*, result in a non-compliant implementation. Defined in Article 2(2) as “any automated analytical technique aimed at analyzing text and data in digital form in order to generate information which includes, but is not limited to, patterns, trends and correlations”, the term must be considered as an autonomous concept of EU law that cannot be modified by Member States in line with political considerations. As outlined above, there is a broad consensus that the concept of text and data mining includes the training of AI models. Even if the Polish Ministry of Culture and National Heritage does not wish this to be the case, it must still implement the Directive without changing a core concept introduced in the Directive.

Expected impact

While we are waiting for the consultation process to play out, it is instructive to consider what would be the consequences should the TDM exception be implemented as proposed by the Ministry of Culture and National Heritage. By excluding the creation of generative artificial intelligence from the scope of both TDM exceptions, the Polish copyright law would remove any statutory basis for the use of copyrighted works in the context of building generative AI models. This would require AI developers to obtain permission from all rightsholders whose works are included in their training data. Given the amounts of copyrighted works that are required to train the current generation of AI models (often measuring in the billions of individual works) this would likely be impossible for anyone but the most well-resourced companies making it virtually impossible for smaller companies or public efforts (such as the [Polish open PLLuM language model](#)), as they would lack the resources to undertake the effort required to obtain the required permissions.

What is especially stunning in the Polish implementation proposal is that it not only excludes the creation of AI models from the scope of the Article 4 exception (which applies to commercial AI developers) but also from the scope of the Article 3 exception (which is designed to enable non-profit scientific research) which seems especially short sighted. The implementation proposal should therefore be read like a misguided attempt to hinder any development or use of generative AI models in Poland.

At this point, it seems useful to recall the key balances inherent in the EU’s regulatory framework for the use of copyrighted works in AI training. They form the basis of claims by the Commission and others that the EU has a uniquely balanced approach to this thorny issue. Taken together, the TDM provisions address 4 key concerns: (1) They limit permission to use copyrighted works for training data to those works that are lawfully accessible. They (2) privilege non-profit scientific

research, (3) they ensure that creators and other rights holders can exclude their works from being used to train generative AI systems, and (4) they ensure that works that are not actively managed by their rights holders can be used to train AI models.

Excluding the training of generative AI from this balanced arrangement may please some creators and rights holders, but it also pushes AI back into a legal gray area. It also seems incompatible with the provisions of the AI Act, which situates the training of generative AI models within the broader concept of TDM, and which will be directly applicable in Poland.

What is needed, instead of efforts to undermine the existing framework, are measures to ensure that the current approach can work in practice. [The new copyright provisions in the AI act are an important step into this direction](#), but they need to be complemented by the creation of a public infrastructure to facilitate opt-outs and measures aimed at ensuring fair licensing arrangements between rights holders and AI developers.

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

Kluwer IP Law

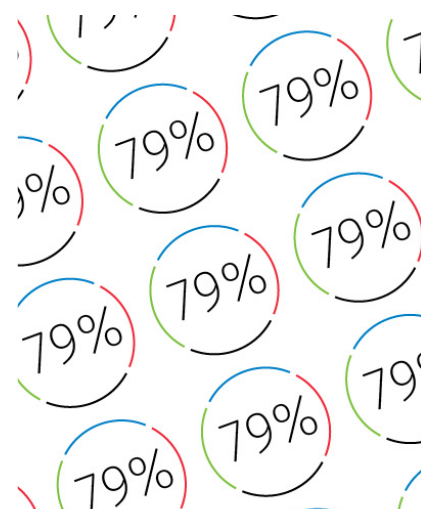
The **2022 Future Ready Lawyer survey** showed that 79% of lawyers think that the importance of legal technology will increase for next year. With Kluwer IP Law you can navigate the increasingly global practice of IP law with specialized, local and cross-border information and tools from every preferred 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 Tuesday, February 20th, 2024 at 2:07 pm and is filed under [CDSM Directive](#), [Legislative process](#), [Poland](#)

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