

# Kluwer Copyright Blog

## German Regional Court (Landgericht) of Hamburg paves the way to treat the reproduction of works as AI training data under the EU text and data mining exceptions

Jonathan Pukas, Jan Bernd Nordemann (NORDEMANN) · Friday, October 25th, 2024

The German Regional Court (Landgericht) of Hamburg handed down its judgment in the LAION case on 27 September 2024 (file no. 310 O 227/23, published in German [here](#)).

The key points of the decision are as follows:

- The reproduction of works for the purpose of creating URL lists that can be used for artificial intelligence training does not fall under the temporary reproduction exception under Section 44a German Copyright Act (Art. 5(1) EU InfoSoc Directive 2001/29 [InfoSoc Directive]).
- The reproduction as mentioned above does, however, serve the purpose of text and data mining within the meaning of Section 44b (1) German Copyright Act (Art. 2 No. 2 EU DSM Directive 2019/790 [DSM Directive]) where the works are examined for



Image from [pixabay](#)

**correlations during the data preprocessing stage.**

- **The term “scientific research” as defined in Section 60d German Copyright Act (Art. 3 (1) EU DSM Directive 2019/790) also includes preparatory work aimed at obtaining knowledge at a later date.**
- **When assessing the non-commercial purpose within the meaning of Section 60d (2) No. 1 of the German Copyright Act (Art. 2 No. 1 (a) EU DSM Directive 2019/790), the character of the specific scientific activity is the sole determining factor and not how the institution is funded.**

For an official English translation of the German Copyright Act (UrhG), please see [here](#)).

## **Facts**

The Plaintiff is the author of the photograph in dispute, which enjoys protection as a photo under the German Copyright Act. The photo had been legally made available to the public on the website of a stock-photo provider who held non-exclusive exploitation rights. The photograph was downloaded from this website by the defendant for the purpose of creating an AI training data set and was thus reproduced, despite a usage restriction against web scraping of the published content being declared, in natural language, on a subpage of the website. The purpose of the reproduction was to compare the photograph with a description of the image created by a third-party provider. The actual training data set later contained only URLs to the photos, as well as the associated image descriptions and other metadata.

The plaintiff sought an injunction against the defendant preventing the defendant from any future reproductions of the plaintiff's photograph for the creation of AI training data sets.

## **Decision**

The Regional Court (Landgericht) of Hamburg dismissed the action. While the court rejected the applicability of the temporary reproduction exception (Section 44a German Copyright Act/Art. 5

(1) InfoSoc Directive), it was of the view that the reproduction was covered by the exception for text and data mining (TDM) for the purposes of scientific research (Section 60d German Copyright Act/Art. 3 DSM Directive).

The reason given was firstly that the defendant's comparison with the image description in preparation for its URL-based training data set is classified as TDM within the meaning of Sections 60d (1) and 44b (1) of the German Copyright Act (Art. 2 (2) DSM Directive). It was therefore not necessary to rule on whether the actual *training* of artificial intelligence falls under the TDM exceptions in Section 44b German Copyright Act (Art. 4 DSM Directive) and Section 60d German Copyright Act (Art. 3 DSM Directive).

Moreover, the court found that the association against which action was taken fulfilled the other requirements of Section 60d German Copyright Act (Art. 3 EU DSM Directive 2019/790). One particularly important factor is that the data set has been made publicly available free of charge.

Therefore, the decision did not depend on whether the general TDM exception under Section 44b (2) German Copyright Act (Art. 4 DSM Directive) applied.

The court did consider, however, the comprehensive usage restriction, declared in natural language (English) on a subpage, to be machine-readable and thus effective (see Section 44b (3) German Copyright Act/Art. 4 (3) DSM Directive). The reservation on a subpage of the website read as follows:

*“RESTRICTIONS – YOU MAY NOT:*

*(...)*

18. *Use automated programs, applets, bots or the like to access the \*\*\*.com website or any content thereon for any purpose, including, by way of example only, downloading Content, indexing, scraping or caching any content on the website.“*

## **Implications**

The court's rejection of the applicability of the temporary reproduction exception (Section 44a German Copyright Act/Art. 5 (1) InfoSoc Directive) seems compelling.

The significance of the judgment for the application in practice of the TDM exceptions for placing works in AI training data sets is huge. This is the first decision of a court within an EU Member State on the question of whether the TDM exceptions in Article 3 and Article 4 DSM Directive also apply to AI training. Accordingly, there is also great interest in the decision from other EU countries. These EU provisions have been implemented in Section 60d and Section 44b of the German Copyright Act. The court tries to minimise the landmark nature of its statements with a decision that ostensibly focuses on the specific individual case. Nevertheless, the judgment provides important practical guidance on the reproduction of works as AI training data and the application of the relevant exceptions to this.

The first point to note is that the judgment strengthens the position of all those who want to include AI training under the concept of TDM in Section 44b (1) German Copyright Act (Art. 4 DSM

Directive). In its reasoning for this subsumption, the court addresses important arguments that have been repeatedly invoked in the debate surrounding Section 44b of the German Copyright Act or Art. 4 DSM Directive and AI training: It rejects a teleological reduction of the scope, even if the EU legislature “had not yet considered” the “AI problem” when the DSM Directive was adopted. The Regional Court also convincingly supports its view by pointing out that Article 53 (1) (c) and (d) of the new AI Act (Regulation (EU) No. 1689/2024) explicitly refers to the reservation of rights in Article 4 (3) DSM Directive.

The Regional Court also rejected a restrictive interpretation on the basis of the three-step test (Article 5 (5) InfoSoc Directive). The reproduction to be examined is, according to the court, limited to the purpose of analysing the image files for their conformity with a pre-existing image description, along with then adding them to a data set. In the court’s opinion, this would not have a relevant effect on the ability to exploit the work in question. This opinion of the Court does not prejudice a stricter application of the three-step-test to the actual AI training. It has to be borne in mind that the case was only about the *collection* of training data and not the reproduction of the work during AI training. Later on, the collection had been made publicly available via a URL list. For the actual AI training, another download (i.e. reproduction) would be necessary. Therefore, the court did not accept that the later training reproductions and their effects must be considered when applying the three-step test. Given this background, for these later training reproductions there could indeed be direct effects, for example by creating a competitive situation in which the output of the trained AI will compete with the training works that were used.

Intermediary companies for training data collection will benefit from the decision, especially the court’s broad interpretation of Section 60d German Copyright Act (Art. 3 DSM Directive). It is also interesting for them that their data preprocessing steps can already be covered by the concept of TDM. It is then no longer a matter of subsuming the actual AI training itself under the legal requirements.

Another important point in practice is that the Regional Court of Hamburg was leaning towards considering a reservation of rights declared in natural language to be machine-readable within the meaning of the second sentence of Section 44b (3) German Copyright Act (Art. 4 (3) DSM Directive), although it did not make a conclusive ruling on this. In addition, the declaration of reservation must be explicit. This is not set out in Section 44b (3) German Copyright Act directly, but in Art. 4 (3) DSM Directive. Therefore, it must be considered when applying the German TDM exception. In principle, both points are plausible. However, a number of important practice-related questions remain, including whether the somewhat complex wording in the present case is appropriate in the meaning of Art. 4 (3) DSM Directive, as the facts of the case require under an interpretation in conformity with the directive. It would also be necessary to look at whether a rights reservation in English would meet the legal requirements for a German author and a German user of the work. Due to its widespread use on the internet, however, there is much to suggest that it would. The right of reservation of holders of non-exclusive rights, like the stock photo provider, also remains unclear. The curtain has therefore fallen for now on the matter of reservation of rights, leaving many questions unanswered. Due to these unresolved questions, the declaration in question should only be used with caution as a template. It seems likely that the courts will establish stricter standards in this regard. For the Regional Court of Hamburg, it was not necessary to examine the rights reservation in detail.

The judgment of the Regional Court of Hamburg should also give pause for thought as to which applications for relief make sense for rightsholders when taking action against collectors of

training data. It might be just as appropriate to choose an application that, for example, includes the establishment of joint responsibility on the part of the training data collector for copyright infringements committed by third-party commercial AI developers who train AI using the collector's link lists and reproduce works for which rights reservations have been declared. Since it is not necessary for the collector to keep any copies of works after the link list has been produced, this might be even more effective to prohibit AI training with the work in question.

*The authors would like to thank Adam Ailsby, Belfast ([www.ailsby.com](http://www.ailsby.com)), for co-authoring the English translation. An edited German version of this article was first published in the law journal GRUR-Prax 2024, page 639.*

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

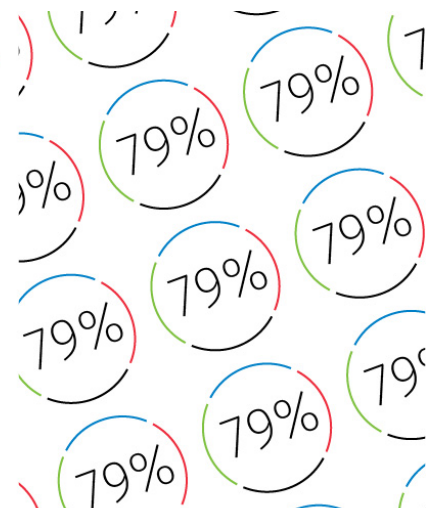
## Kluwer IP Law

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 Friday, October 25th, 2024 at 1:17 pm and is filed under [Artificial Intelligence \(AI\)](#), [Exceptions and Limitations](#), [Germany](#), [Text and Data Mining \(TDM\)](#)

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