

Hamburg court rules LAION's dataset building is lawful research text and data mining

Source copy kept by wheresthe.ai for the story "Hamburg court rules LAION's dataset building is lawful research text and data mining" (<https://wheresthe.ai/d/de-hamburg-court-laion-tdm-ruling-2024-09/>).

Source: <https://lausen.com/en/lg-hamburg-kneschke-v-laion-e-v/>

Text saved 18 Oct 2026 (the page as published).

SHA-256 of the saved text: d749c4c51b6dd946aeeb06b8bf8fd949e6e884c8ba220cfb07b9d63f60455b65

Author

Dr. Ursula Feindor-Schmidt, LL.M.

Lawyer, Partner

Specialised Lawyer for Copyright and Media Law

[View profile](#)

[Share blog post via](#)

[Xing](#)

[LinkedIn](#)

The use of copyrighted works in AI training datasets – Created for science, used for profit?

The decision of the Hamburg Regional Court (judgment of September 27, 2024, file no. 310 O 227/23 ;in German) in the proceedings brought by photographer Robert Kneschke against the non-profit organisation LAION e.V. ("LAION") has attracted a great deal of media attention worldwide. According to many commentators, this is the first decision in Germany (possibly also in Europe) on the use of copyrighted works in connection with the training of (generative) artificial intelligence ('AI').

In fact, however, this decision is not about the actual use of a copyrighted work for AI training. On a closer look, the proceedings were rather limited to acts preceding a possible AI training. The limitation resulted from the plaintiff's specific narrow complaint as well as the plaintiff's factual and provable submissions.

The case is nevertheless extremely interesting: the court clarifies first issues in connection with the copyright exception for "text and data mining" (TDM), in particular TND for the purposes of scientific research. In doing so, it becomes clear that rightsholders will be facing massive challenges if they want to effectively protect their works from being included in AI training. In particular, the possibilities of a collaboration between science and commercial companies as well as the performance of actions in different jurisdictions (which might be considered possibly advantageous for AI operators) can make it difficult to enforce copyrights.

Rightsholders as well as providers of AI systems should therefore plan their steps carefully. However, due to the legal uncertainties, sustainable AI systems will always require strategic cooperation between the content and technology industries.

The case

The photographer Robert Kneschke is challenging the use of a photograph he took. This photograph was contained in a data set of over 5 billion image-text pairs, which the defendant LAION e.V. made publicly available free of charge on its website under the name "LAION 5B". The (still retrievable) dataset does not consist of the works themselves, but of a table containing hyperlinks to images or image files publicly available on the internet as well as further information on the corresponding images, including an image description (also known as 'alternative text'), which provides information on the content of the image in text form.

To create the dataset, in particular for the automated comparison of image and description, LAION downloaded the works, but ultimately only published the original links (accessible on the internet). To create the dataset, LAION used a specific software to analyze the correctness of the textual image description associated with an image. The data set is suitable for training of generative AI and was presumably also used for such training.

However, LAION itself did not use the dataset for AI training. Therefore, the use of the works for AI training was not the subject of the decision.

The disputed photograph by Robert Kneschke was marketed on the internet by an stock photo agency. A preview of the photograph (in low resolution and with a watermark) was made publicly available on the agency's website. The agency had included the following rather general note in its terms of use (as of 2021) on the website:

“RESTRICTIONS YOU MAY NOT:

18. use automated programs, applets, bots or the like to access the XXX.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.”

The photographer filed complaint for an injunction against the reproduction of his work for the creation of AI training data sets.

The decision

The court dismissed the action on the grounds that the reproduction at issue by LAION was covered by the exception for text and data mining for scientific purposes under Section 60d UrhG and was therefore permitted.

According to the opinion of the court the reproductions were made by LAION (only) for the purpose of text and data mining within the meaning of Section 44b (1) UrhG , namely for the automated analysis of individual or several digital or digitized works in order to obtain information, in particular about patterns, trends and correlations. In the present case, the specific purpose was to analyze the correlation between image and text (i.e. the question of non-conformity/conformity between images and image descriptions). A further purpose of the reproduction (in addition to the purpose of the TDM) was possible, but was not yet sufficiently specific at the time the reproduction was made. It was not foreseeable for LAION who in particular would subsequently use the datasets created by LAION in what way – i.e. whether and in what way they would be used by third parties in the context of the training of AI.

The court found LAION to be privileged pursuant to Section 60d UrhG. The TDM was carried out for the purposes of scientific research. The term scientific research was to be understood broadly; the methodical and systematic “pursuit” of new knowledge was sufficient; it was also sufficient that the work step in question was aimed at a (later) gain in knowledge. This is the case, for example, with numerous data collections that must first be compiled for subsequent empirical conclusions. In particular, the concept of scientific research does not presuppose any subsequent research success. For this later research objective, it is sufficient that the dataset was – indisputably – published free of charge and thus also made available to researchers in the field of artificial neural networks. By making the data set available free of charge, the non-commercial purpose required under the exception was held to be complied with.

The applicability of the TDM exception was also not excluded by the exception in Section 60d (2) sentence 3 UrhG. Whether private companies had exerted a decisive influence on the research or whether private companies were granted preferential access to the research results could not be proven by the plaintiff.

Marginal notes of the court (obiter dictum)

In its decision, the Hamburg Regional Court took the opportunity to make some interesting comments on the subject of text and data mining and AI training in general. However, these only represent the current legal opinion of the Regional Court of Hamburg, but were neither relevant to the present case nor do they have any binding effect on other courts.

This includes the classification of the general reservation of rights in relation to automated crawling and scraping in the terms of use of the website on which the preview of the image was available. The court considered this general notice to be sufficiently clear and also ‘machine-readable’ within the meaning of Section 44b (2) sentence 3 UrhG. In the opinion of the Regional Court, the latest state of the art – including the latest AI technology – must be taken into account in this respect. It remains to be seen whether this view will be upheld by other courts with regard to technologies available in 2021 or whether the term ‘machine-readable’ will be understood more as a standard in the sense of robots.txt. Certainly, however, highly technological companies or institutions must definitely make use of technologies available to clarify a reservation of rights. We would still not recommend relying solely on such general, above all human-readable reservations. Nevertheless, this reasoning of the Hamburg Regional Court could prove helpful for rights holders in the future. It is also important in this context that

the court considered it sufficient that the holder of simple rights of use is (also) entitled to declare the TDM reservation.

In addition, the Regional Court differentiated between various steps in connection with the use of works for the training and application of generative AI. It was (in any case) necessary to differentiate between

- 1) the creation of a data set (which is the sole subject of dispute in these proceedings) that can also be used for AI training;
- 2) the subsequent training of the artificial neural network with this dataset;
- 3) the subsequent use of the trained AI for the purpose of creating new image content.

The Regional Court took the view that the application of the text and data mining exceptions under Section 44b UrhG or 60d UrhG for step 1), i.e. the creation of a dataset under the conditions of the exception provisions, does not preclude the fundamental possibility of steps 2 and 3. This would also follow from the explanatory memorandum to the German transposition of EU Directive 2019/790 and the explanatory memorandum to the AI ACT (Regulation (EU) 2024/1689) (Recital 105).

On the other hand, this does not mean that step 2) or even 3) also fall under the text and data mining exceptions under Section 44b or 60d UrhG. It was precisely these steps that were not the subject of the Regional Court's decision or, in any case, could not be proven by the plaintiff.

With respect to steps 2) or 3) it would be necessary to assess independently who carries out which technical steps for such training and where, and whether local law or local exceptions apply here.

Take away

As a result, the rightsholders – as is often the case with new technological developments – are faced with the problem that the technical processes and, above all, possibly even a planned collaboration are not transparent for the rightsholders concerned. In addition, most cases are international and cross-jurisdictional. In order to be able to enforce copyrights in court, these new constellations involving the training and use of AI systems require comprehensive technical and legal knowledge (or expert opinions) that examine the technical background, the connecting factors of an infringing act and the applicable law and exceptions in different jurisdictions. Against this backdrop, effective legal enforcement is practically impossible for individual authors or smaller rightsholders.

In this context, the AI ACT is intended to make things a little easier for rights holders. It is currently unclear how the requirement in the AI Act to create and publish a “sufficiently detailed summary of the content used for the training of the general purpose AI model” (Article 53 (1) (c)) will actually prove the use of a specific work. The AI Office set up is currently working on a template for creating those summaries.

In light of the many unanswered questions and legal uncertainties, however, it is clear that sustainable AI will always require strategic cooperation between the content and technology industries.

More posts

On 31 July 2026, the Munich Regional Court I (Landgericht München I) handed down its second landmark AI judgment in GEMA v Suno (Case No. 42 O 763/25). Following GEMA v OpenAI (judgment of 11 November 2025, Case No. 42 O 14139/24), which concerned song lyrics, the 42nd Civil Chamber now had to rule on ...

Read more

Re-named a Top Commercial Law Firm We are pleased to have once again been recommended by FOCUS Business: LAUSEN is once again among Germany's top commercial law firms in the media and press sector for 2026. The survey considers commercial law firms based in Germany. Respondents can recommend up to 20 firms in each of ...

Read more

We are pleased to announce that three attorneys from LAUSEN have been recommended in the latest Handelsblatt ranking, “Germany's Best Lawyers 2026.” The ranking is compiled annually in collaboration with Best Lawyers and is considered one of the most prestigious accolades in the German legal market. These recommendations confirm our firm's exceptional expertise in the ...

[Read more](#)