LG Hamburg, judgement of 27.09.2024 - 310 O 227/23

Reference	openJur 2024, 9199	Rkr: ? AmtlSlg: ?
Civil procedural law	IT and media law	Civil law

Tenor

1 1 The action is dismissed.

2 2 Orders the plaintiff to pay the costs.

3 3 The judgement is provisionally enforceable. The plaintiff may avert the defendant's enforcement by providing security amounting to 110% of the amount enforceable on the basis of the judgement, unless the defendant provides security amounting to 110% of the amount to be enforced prior to enforcement.

Facts of the case

4 The defendant is an association that was founded with a founding meeting on 7 July 2021 (minutes in Annex B7, articles of association in Annex B1). The specific purpose of the defendant's activities is in dispute between the parties.

5 The defendant makes a so-called dataset for image-text pairs publicly available free of charge under the name "...". This is a type of table document that contains hyperlinks to publicly accessible images or image files on the Internet as well as further information on the corresponding images, including an image description (also known as alternative text) that provides information on the content of the image in text form. The data set comprises 5.85 billion corresponding image-text pairs. The data set can be used to train so-called generative artificial intelligence.

6 The data set was created after the defendant was founded in the second half of 2021. For this purpose, the defendant had used an existing data set of ... from the USA (www...org), which contained the respective URLs along with a textual description of the respective image content for a kind of random cross-section of the images that could be found on the Internet. The defendant then extracted the URLs to the images from this data set and downloaded the images from their respective storage locations. The defendant then used software to check the images to see whether the description of the image content already in the existing data set actually matched the content to be seen in the image. Images where the text and image content did not match sufficiently were filtered out. For the remaining images, the metadata, in particular the URL of the image storage location and the image description, were extracted and included in a newly created dataset, the .... Whether the downloaded image files were subsequently deleted again is in dispute between the parties - at least with regard to the photograph in dispute.

7 As part of the aforementioned process, the image at issue was also captured, downloaded, analysed and included with its metadata in the data set ... . Specifically, an image was downloaded from the website of the image agency ... (https://www...com) with a watermark of the photo agency ... watermarked image file.

8 On the website of the picture agency ... the following text has been on the subpage https://www...com/de/usage.html since at least 13 January 2021:

9 "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."

10 The plaintiff alleges an infringement of copyright in the photograph at issue in the form of unauthorised reproduction by the defendant as part of the analysis process.

11 The plaintiff claims that he is the author of the photograph listed in the operative part. The company ... was entitled to offer and also display the photo in dispute on its website ...com and to offer licences for the photo;
... was in this respect the owner of simple, sublicensable rights of use.

12 The - undisputed - reproduction that took place as part of the analysis process infringed the plaintiff's rights under § Section 16 UrhG, in particular it is not covered by the limitation provisions of Sections 44a, 44b and 60d UrhG:

13 The limitation provision of Section 44a UrhG is not relevant; in particular, the independent download of a photograph does not constitute a temporary act within the meaning of this provision.

14 The reproduction is also not covered by Section 44b UrhG. The aggregation of data for the purpose of AI training is not text or data mining within the meaning of Section 44b UrhG. Neither the European nor the German legislator had such a use "in mind" when creating the limitation provision of Art. 4 DSM Directive or Section 44b UrhG. In the case of text and data mining within the meaning of Section 44b UrhG, only "information hidden in the data should be made accessible", "but not the content of the intellectual creation should be utilised". However, the so-called "AI web scraping" at issue here is precisely about the intellectual content of the works used for training purposes "and ultimately about the creation of identical or similar competing products". Moreover, according to the defendant's own disclaimer (printed on p. 47), the data set is "uncurated". Finally, the collection and storage for the creation of parallel archives is excluded from the limitation provision of Section 44b UrhG according to the expressly declared intention of the legislator.

15 In addition, "the mass incorporation of copyright-protected works for training purposes in the context of generative AI" impairs the normal exploitation of copyright-protected works because it creates the conditions for replacing authors in many cases, or at least makes it considerably more difficult to exploit the work by offering it to competitors free of charge. However, according to Article 7(2) of the DSM Directive in conjunction with Article 5(5) of the InfoSoc Directive, this precludes the application of the limitation rule.

16 In any case, the reproduction is inadmissible due to the reservation of use declared on the website www.bigstockphoto.com pursuant to Section 44b (3) UrhG. The corresponding declaration of the picture agency is attributable to the plaintiff, as it distributes the disputed photo for him. Contrary to the defendant's view, the reservation was also machine-readable within the meaning of Section 44b (3) sentence 2 UrhG. The requirements in this regard are not higher than for machine readability by humans; however, the reservation is written in block letters. Moreover, the text was also recognisable as a reservation by a computer program. The service ... was able to recognise the corresponding reservation, and specific tools such as WebOpt-Out can also recognise reservations such as those on ...com.

17 Nor can the defendant invoke the limitation provision of Section 60d UrhG. The plaintiff disputes that the defendant fulfils the requirements of Section 60d UrhG in fact, namely

18 - that he was "registered in the register of associations" at the time of the act of reproduction in question here,

19 - that the document B1 submitted by the defendant represented the valid articles of association of the defendant or had represented them at the time of the act of reproduction in question,

20 - that the members of the association and the board were working on a voluntary basis or were at the time of the act of reproduction in question,

21 - that the defendant is exclusively engaged in research or was engaged in research at the time of the act of reproduction in question or that the defendant conducts scientific research, pursues non-commercial purposes and reinvests all profits in scientific research or is active in the public interest within the framework of a state-recognised mandate; furthermore, according to the articles of association submitted by the defendant, the purpose of the defendant is only the "promotion of research" and not "research" as such, and it is also unclear what research is supposed to be in the collection created by the defendant, which is (indisputably) made available to other companies;

22 - that the defendant would create and test its own AI models based on the training data in order to further explore the possibilities of AI technology,

23 - that by making the training dataset publicly available, other researchers and interested parties should be offered the opportunity to train their own AI models; according to the defendant's own statements, the dataset at issue was also used to train the services "...", "..." and "..." of the provider ... However, these were operated by (purely) commercial companies; insofar as the defendant denied that the first two services had been trained, it would in any case have been possible for them to use the dataset.

24 In addition, the defendant cannot invoke the privileged status of Section 60d (2) sentence 3 UrhG pursuant to Section 60d UrhG. The defendant is apparently working intensively with commercial AI providers:

25 - There is apparently a co-operation with the private company ..., which has direct influence on the defendant through the financing of the dataset in question and the filling of relevant positions at the defendant by its own employees. Thus ... had, according to an interview with its founder and managing director, financed the ... Dataset.

26 - Members of the defendant's "team" are also "in many places" commercially in the same field for large tech companies, including as an employee of the company ....

27 - In a chat on the "..." platform on 28 September 2021, the co-founder of the defendant, ..., also urged the company to hurry up with the completion of the "...", as it had received funding of $5,000 from a "..." or his company; the data should (already) be made available to him even if the dataset could not yet be made available to the public. The "..." in question was an employee of the commercial AI provider ...

28 Finally, the defendant could not rely on a so-called "simple consent". He - the plaintiff - had not made the photograph in dispute freely accessible, but had it offered for the granting of licences against payment by the agency ... .

29 The plaintiff had initially requested information about the extent to which the photograph had been used in addition to an injunction against reproduction. The parties unanimously declared this request for information to be settled at the hearing on 11 July 2024.

30 The plaintiff now still applies,

31 order the defendant to cease and desist from using the following photograph on pain of a fine of up to € 250,000.00 each, or alternatively imprisonment for up to 6 months, for each individual case of infringement

...

to reproduce and/or have reproduced for the creation of AI training datasets, as happened in the context of the production of the dataset ...

32 The defendant claims that the Court should,

33 dismiss the action.

34 He denies that the plaintiff created the disputed image himself or that he is otherwise entitled to assert rights infringements in his own name with regard to the image, and that the plaintiff was entitled to assert rights infringements with regard to the disputed image in his own name at the time the image was captured by the defendant.

35 Above all, however, the (one-off) download of the disputed image in the context of the creation of the data set ... (one-off) download of the disputed image in the context of the creation of the data set constitutes a copyright-relevant reproduction, but this is covered by the limitation provisions of §§ 44a, 44b and 60d UrhG as well as the simple consent of the plaintiff:

36 On the one hand, the reproduction that took place was covered by the limitation provision of Section 44a UrhG. The images had not been permanently stored by the defendant; rather, the images had only been used for a short time for the analysis and then automatically and irrevocably deleted without delay. The reproduction did not have any independent economic significance.

37 The limitation provision of Section 44b UrhG also applies. The analysing of image files and the extraction of metadata for the training of artificial intelligence is a main application of text and data mining according to the intention of the legislator. There was also no creation of digital parallel archives, as the downloaded images were not permanently stored, but only hyperlinks were included. The exception under Section 44b (3) UrhG does not apply:

38 - According to the plaintiff's submission, it was not he himself as the rights holder, but the operator of the website www...com as a third party who declared this reservation; the plaintiff himself even expressly stated in his email of 13 February 2023 (Annex B5) that he had neither the qualifications nor the economic means to declare a reservation of use.

39 - Furthermore, the reservation was not made explicitly, as the passage on the website www...com is general and lists various unauthorised actions. There is no explicit mention of text and data mining or of reproductions.

40 - In addition, the characteristic of machine readability is lacking. A clause written in natural language is generally not machine-readable within the meaning of Section 44b (3) UrhG. The prerequisite for machine-readability in this sense is that the clause can be automatically processed by software. This presupposes that the corresponding information is encoded. In any case, it is necessary that at least specific keywords such as "data mining" are contained in the text.

41 - The reservation was also clearly not intended as such in accordance with Section 44b (3) UrhG. The fact that, according to the plaintiff's submission, the clause was already present on the website on 13 January 2021 makes it clear that the clause could not have been created "with a view to the provision in Section 44b (3) UrhG", as the statutory provision had not yet entered into force at that time. Moreover, it was "also not credible" that a US provider would invoke a reservation of use based on German law.

42 In any event, he - the defendant - could invoke the limitation provision of Section 60d UrhG.

43 - It - the defendant - is a non-profit association consisting of researchers and, according to the association's articles of association (Annex B1), is dedicated to research; in particular, it has set itself the task of further developing self-learning algorithms in the sense of artificial intelligence and making them available to the general public. To this end, it provides data sets and models free of charge and also creates and tests its own AI models based on the training data.

44 - His work also constitutes "research". Simply by making it publicly transparent on the internet how the training data sets are created, he contributes to gaining knowledge about the training of artificial intelligence. In this way, other researchers could understand the steps involved in creating the data sets and build on them. In addition, he had published a scientific paper on the dataset at issue ... published a scientific paper with the name "..." for the first time on 17 September 2022 (Annex B8). By 5 April 2024, the aforementioned treatise had been cited a total of 1403 times in other scientific works and had also received further awards.

45 - In addition, the defendant also trains its own AI models on the basis of the data sets it has created in order to gain insights into how AI can be improved through appropriate training.

46 - The natural persons belonging to the defendant, i.e. the board members and other members of the association, are also "researchers". Insofar as the plaintiff refers to the fact that individual team members work in the same field for large tech companies, this has no relevance to the question of whether the defendant itself is commercially active. Moreover, the persons concerned work for the defendant on a voluntary basis and therefore have to earn their living elsewhere.

47 - The fact that the datasets made publicly available by the defendant are also used by commercial providers is irrelevant for the intervention of the limitation provision of Section 60b UrhG. Apart from that, the services ... and ... were not actually trained with the defendant's dataset.

48 - The reverse exception provided for in § 60 d para. 2 sentence 3 UrhG does not apply in the present case either. The company ... had indeed made computing resources available to the defendant during the start-up phase. However, the same had also been done by the ... . The defendant had not received any financial support in the form of money from ... . There had also been no further co-operation with this company. In any case, ... did not receive preferential access to the research results. There was also no decisive influence by the company ... did not exist. Neither the company ... itself nor any of its legal representatives were members of the defendant.

49 Finally, there is a so-called simple consent in favour of the use by the defendant on the part of the plaintiff. The reproduction at issue is to be categorised as a customary act of use.

50 Furthermore, it should be noted that the plaintiff himself earns money with AI-generated images. This calls into question his "need for legal protection".

51 With regard to the further details of the parties' submissions, reference is made to the written submissions and annexes to the file, insofar as they became the subject of the oral hearing, as well as to the minutes of the hearing of 11 July 2024 (pp. 120-123).

52 The parties have submitted unredacted written submissions dated 29 August and 11 September 2024 (plaintiff) and 20 September 2024 (defendant) to the file.

Reasons

53 I.

54 The admissible action is not successful on the merits. By reproducing the photograph at issue, the defendant has indeed infringed the plaintiff's exploitation rights. However, this encroachment is limited by the limitation provision of the § 60d UrhG is covered. Whether the defendant can additionally invoke the limitation provision of Section 44b UrhG does not need to be conclusively assessed against this background.

55 The photograph in dispute is in any case protected as a photograph pursuant to Section 72 (1) UrhG. After inspecting the raw data on the plaintiff's laptop, the court also has no doubt about the plaintiff's status as a photographer, Section 72 (2) UrhG. The plaintiff is also actively authorised to assert infringement claims pursuant to Section 97 UrhG, including the claim for injunctive relief pursuant to para. 1 of the provision; the fact that the plaintiff has granted the picture agency ... more extensive than (sub-licensable) simple rights of use, the defendant has not demonstrated. The picture agency ... added a watermark to the photo; this was a non-free transformation within the meaning of Section 23 (1) sentence 1 UrhG, so that the consent of the plaintiff as the author was also required in principle for its utilisation. As part of the download, the defendant reproduced this version within the meaning of Section 16 (1) UrhG without having obtained the plaintiff's consent.

56 However, the defendant was entitled to do so on the basis of statutory authorisation. The reproduction was not covered by the limitation provision of Section 44a UrhG (hereinafter 1.), and whether the defendant can invoke the limitation provision of Section 44b UrhG appears doubtful (hereinafter 2.). However, the latter does not require a final decision in the present case, as the act of reproduction was in any case covered by the limitation provision of Section 60d UrhG (hereinafter 3.).

57 1.

58 The reproduction that has taken place is not covered by the limitation provision of Section 44a UrhG.

59 Accordingly, temporary acts of reproduction are authorised which are transient or incidental and constitute an integral and essential part of a technological process and whose sole purpose is to enable a transmission in a network between third parties by an intermediary or a lawful use of a work or other subject-matter and which have no independent economic significance.

60 The duplication in the present case was already neither ephemeral nor concomitant.

61 a)

62 A reproduction is volatile within the meaning of Section 44a UrhG if its lifetime is limited to what is necessary for the proper functioning of the technical process concerned, whereby this process must be automated in such a way that it automatically deletes this act, i.e. without the involvement of a natural person, as soon as its function of enabling the performance of such a process has been fulfilled (ECJ, judgment of 16 July 2009, ref. 16 July 2009, Case C-5/08 - Infopaq/Danske Dagblades Forening, para. 64 (juris) on Art. 5 (1) DSM Directive).

63 Insofar as the defendant refers in this respect to the fact that the files were deleted "automatically" as part of the analysis process carried out by him, this cannot justify the volatility of the duplication in the aforementioned sense. Apart from the fact that the defendant has not stated anything about the specific duration of the storage, the deletion did not take place "user-independently", but due to a corresponding deliberate programming of the analysis process by the defendant.

64 b)

65 A reproduction is ancillary within the meaning of Section 44a UrhG if it is neither independent of the technical process of which it is a part nor serves an independent purpose (ECJ, judgement of 5 June 2014, case C-360/13, para. 43 (juris)).

66 In the present case, the image files were downloaded in a targeted manner in order to analyse them using specific software. This means that downloading is not merely an accompanying process to the analysis carried out, but a conscious and actively controlled procurement process upstream of the analysis.

67 2.

68 Whether the defendant can invoke the limitation provision of Section 44b UrhG appears to be doubtful in the present case. It is true that the download carried out by the defendant is in principle subject to the limitation provision of Section 44b (2) UrhG, in particular it was carried out for the purpose of text and data mining within the meaning of Section 44b (1) UrhG (hereinafter lit. a). However, without this requiring a final decision in the present case, there is some evidence to suggest that the act of reproduction was not already authorised under Section 44b (2) UrhG due to an effectively declared reservation of use within the meaning of Section 44b (3) UrhG (hereinafter lit. b).

69 a)

70 The act of reproduction at issue is in principle subject to the limitation provision of Section 44b (2) UrhG.

71 (1) The download at issue was made for the purpose of text and data mining within the meaning of Section 44b (1) UrhG. Accordingly, text and data mining is the automated analysis of individual or multiple digital or digitised works in order to obtain information, in particular about patterns, trends and correlations. In any case, this must be affirmed for the act of reproduction at issue in the present case (below (a)); a teleological reduction of the offence cannot be considered in this respect (below (b)).

72 In the present case, there is therefore no need to decide the further question, which has been discussed in detail in the literature, as to whether or not the training of artificial intelligence in its entirety is subject to the limitation provision of Section 44b UrhG (in detail on the state of opinion BeckOK UrhR/Bomhard, 42nd ed. Ed. 15.2.2024, UrhG § 44b Rn. 11a-11b m.w.N.; see also in detail the study "Urheberrecht & Training generativer KI- technologische und rechtliche Grundlagen" (Copyright & Training of Generative AI - Technological and Legal Foundations) submitted as Annex K11 on behalf of the Initiative Urheberrecht).

73 (a) The defendant carried out the act of reproduction for the purpose of obtaining information on "correlations" in the literal sense of Section 44b (1) UrhG. The defendant downloaded the photograph at issue from its original storage location in order to obtain information about the correlations by means of an already available software - apparently the application ... from ... - to compare the image content with the image description already stored for the text. This analysis of the image file in order to compare it with a pre-existing image description constitutes without further ado an analysis for the purpose of obtaining information about "correlations" (namely the question of the non-conformity/conformity of images and image descriptions). The fact that the defendant analysed the images included in the data set ... in this way was not disputed as such by the plaintiff.

74 The applicability of Section 44b (1) UrhG is also not excluded - contrary to the opinion of the plaintiff (reply p. 13 f., p. 47 f. of the original) - because the defendant did not "curate" the data set created by him ... according to a "disclaimer" issued for this data set. The disclaimer reproduced by the plaintiff refers solely to a warning that the data set had not been searched for "disturbing content" or similar. Such an - additional
- However, filtering the data set to be created is not a prerequisite for the application of Section 44b (1) UrhG and does not contradict the assumption that the downloaded images - as explained - were analysed for their correlation between image content and image description.

75 (b) The act of reproduction at issue is also not to be excluded from the limitation provision of Section 44b UrhG by way of teleological reduction.

76 Insofar as an exclusion of the reproduction of data for the purpose of AI training by way of teleological reduction is occasionally advocated in the literature on the grounds that Section 44b UrhG only covers the development of "information hidden in the data", but not the use of "the content of the intellectual creation" (Schack, NJW 2024, 113; in this direction also Dormis/Stober, Urheberrecht und Training generativer KI-Modelle, Annex K11, pp. 67 et seq. with a differentiation between semantics and syntax), there are doubts as to whether this is convincing, as it is not sufficiently clear what the difference is between "information hidden in the data" and "the content of the intellectual creation" in the case of digitised works.

77 To the extent that it is additionally argued that "AI web scraping" is about the intellectual content of the works used for training purposes and "ultimately" about the creation of identical or similar competing products (Schack, loc. cit.), this argumentation does not distinguish strictly enough, in the Chamber's opinion, between

78 - on the one hand, the creation of a data set (which is the sole subject of the dispute here) that can also be used for AI training,

79 - among other things, the subsequent training of the artificial neural network with this data set, and

80 - thirdly, the subsequent use of the trained AI for the purpose of creating new image content.

81 This latter functionality may already be the aim when the training data set is created. However, at the time of compiling the training dataset, it is not possible to foresee how successful the second step (training) will be, nor what specific content can be generated by the trained AI in the third step (in the AI application). The specific application possibilities for a rapidly developing technology such as AI are therefore not conclusively foreseeable at the time the training data set is created and therefore cannot be determined with legal certainty. Due to this legal uncertainty, the mere general intention to obtain future AI-generated content when the training dataset is created is not a suitable criterion for assessing the legal admissibility of the creation of the training dataset as such.

82 Finally, to the extent that it is argued in favour of a teleological reduction of the limitation provision of Section 44b UrhG that the European legislator "simply did not yet have the AI problem" "on its radar" when creating the underlying directive provision (Art. 4 DSM Directive) in 2019 (Schack, loc. cit.; likewise for the training of AI models Dormis/Stober, loc. cit., p. 71 et seq., 87 et seq.), this finding alone is clearly not sufficient for a teleological reduction. In particular, it must be taken into account that the technical development in the field of artificial intelligence since 2019 concerns less the type and scope of the (disputed) data mining for the procurement of training data, but rather the performance of the artificial neural networks trained with the data (accordingly, Dormis/Stober, loc. cit., p. 95, also assume that the mere creation of training data sets "in advance of the actual training" may well fall under the TDM threshold). It should also be noted that the Common Crawl Foundation database retrieved by the defendant has been created since 2008 (!), cf. https://commoncrawl.org/overview.

83 Apart from this, the current European legislator of the AI Regulation (Regulation (EU) 2024/1689 of 13 June 2024, OJ L of 12 July 2024 p. 1) has unequivocally expressed that the creation of data sets intended for the training of artificial neural networks is also subject to the limitation rule of Art. 4 of the GDPR. According to Art. 53 para. 1 lit. c AI Regulation, providers of AI models with a general purpose are obliged to provide a strategy, in particular to identify and comply with a legal reservation asserted in accordance with Art. 4 para. 3 DSM Directive.

84 The fact that the creation of data sets intended for the training of artificial neural networks is also subject to the restriction provision of Art. 4 of the DSM Directive also corresponds to the assessment of the German legislator in the context of the implementation of the aforementioned restriction provision in 2021 (explanatory memorandum BT-Drs. 19/27426, p. 60).

85 (c) The so-called three-step test enshrined in Art. 5 para. 5 InfoSoc Directive (in conjunction with Art. 7 para. 2 sentence 1 DSM Directive) does not ultimately justify a different assessment. According to this test, the standardised exceptions may only be applied in certain special cases in which the normal exploitation of the work or other protected subject matter is not impaired and the legitimate interests of the rightholder are not unduly infringed. These requirements are met in the present case.

86 The reproduction relevant to copyright law in the present case is limited to the purpose of analysing the image files for their conformity with a pre-existing image description and subsequent entry into a data set. It is not apparent and is not claimed by the plaintiff that this use would impair the utilisation possibilities of the works concerned.

87 It is true that the data set created in this way may subsequently be used to train artificial neural networks and the resulting AI-generated content m a y compete with the works of (human) authors. However, this alone does not justify considering the creation of the training datasets as an impairment of the exploitation rights to works within the meaning of Art. 5 (5) of the InfoSoc Directive. This must apply simply because the consideration of merely future technical developments, which are not yet foreseeable in detail, does not allow a legally certain distinction to be made between authorised and unauthorised uses (see similarly (b) above).

88 Since the use of knowledge obtained by means of text and data mining to train artificial neural networks, which can then compete with authors, can never be ruled out in case of doubt on the basis of current technological developments, the opposing view would ultimately even require text and data mining within the meaning of Section 44b UrhG to be prohibited in its entirety; however, such a complete overriding of the limitation provision would obviously run counter to the legislative intention and can therefore not represent a viable result of interpretation.

89 (2) The image file downloaded by the defendant was also - which the plaintiff also does not dispute - lawfully accessible within the meaning of Section 44b (2) sentence 1 UrhG.

90 A work is "lawfully accessible" in this sense in particular if it is freely accessible on the internet (Begr. RegE BT-Drucks. 19/27426, p. 88).

91 This must be assumed for the image downloaded by the defendant. Contrary to the plaintiff's initial assertion, the defendant did not download the "original image" reproduced in the injunction request initially formulated in the statement of claim. - which would have been made available by the picture agency ... only if a licence had been purchased - but downloaded a version of the image with a watermark from the picture agency. This was clearly the preview image posted on the agency's website for advertising purposes. However, this watermarked preview image had just been made freely accessible on the Internet by the agency.

92 b)

93 However, there are some indications that the limitation provision of Section 44b (2) UrhG does not apply in the present case - without this requiring a final decision - since there was an effectively declared reservation of use within the meaning of paragraph 3 of the provision; in particular, the reservation of use indisputably declared on the website ...com is likely to meet the requirements for machine readability within the meaning of Section 44b (3) sentence 2 UrhG.

94 (1) There is much to suggest that the reservation of use stated on the agency's website was issued by a person authorised to do so and that the plaintiff can also rely on this to protect his own rights.

95 According to the wording of Section 44b (3) UrhG, "the rightholder" can declare the reservation of use. This means that not only declarations of reservation by the author himself, but also by subsequent rights holders, be they legal successors or holders of rights derived from the author, must be taken into account. According to the plaintiff's readily conclusive submission (minutes of 11 July 2024, p. 3, p. 122 of the file), he had granted the picture agency ... simple rights of use to the original image that could be sublicensed. The picture agency was therefore itself the owner of the rights to the images posted on its website and could therefore easily declare a reservation of use in accordance with Section 44b (3) UrhG; it is neither apparent nor has it been asserted that agreements with an effect in rem in the contractual relationship between the plaintiff and the picture agency would have stood in the way of this.

96 The plaintiff is probably also entitled to invoke this declaration of reservation by its licence holder. From an economic point of view, the exploitation of the original photo in dispute took place via the agency. Thus, in practice, the specific decision as to which third party was to receive the authorisation for which use lay with the agency; the agency was not obliged to conclude a contract. In such a situation, the Chamber believes that there is much to suggest that the author may rely on a reservation declared by his licensee pursuant to Section 44b (3) UrhG when asserting the remaining prohibition rights.

97 (2) The defendant's objection that the prohibition of use for web crawlers in the agency's general terms and conditions towards its customers could not be formulated "in relation to Section 44b (3) UrhG" in terms of time alone is also irrelevant. It is not a prerequisite for the legal effects of the declaration that it is consciously declared with regard to a specific version of the law.

98 (3) The reservation is also sufficiently clearly formulated. Art. 4 (3) DSM Directive requires an express declaration of the reservation of use. This expressiveness requirement must therefore be taken into account when interpreting Section 44b (3) UrhG in conformity with the Directive (see also the explanatory memorandum to the draft bill BT-Drs. 19/27426, 89). The declared reservation must therefore be declared both expressis verbis (not implied) and so precisely (concretely and individually) that it unequivocally covers a specific content and a specific use (Hamann, ZGE 16 (2024), p. 134). These requirements are met by the reservation of use formulated on the website of the picture agency ... easily fulfils these requirements.

99 Insofar as it is also argued that a reservation of use declared for all works posted on a website contradicts the expressiveness requirement of Section 44b (3) UrhG (according to Hamann, loc. cit., p. 148, extending his own abstract derivation), this is not convincing. This is because even the reservation explicitly declared for all works posted on a website can be unequivocally determined in terms of its scope and content and is therefore expressly declared.

00 (4) Finally, there should also be some evidence in favour of the reservation of use satisfying the requirements of machine readability within the meaning of Section 40b (3) sentence 2 UrhG.

01 In view of the underlying legislative intention to enable automated searches by web crawlers (see explanatory memorandum to the draft bill BT-Drucks. 19/27426, p. 89), the term "machine readability" will certainly have to be interpreted in the sense of "machine comprehensibility" (see Hamann, loc. cit., pp. 113, 128 et seq.).

02 However, the Chamber tends to regard a reservation of use written solely in "natural language" as "machine understandable" (unlike the probably predominant view in the literature, see Hamann, loc. cit., p. 131 et seq,
146 et seq. with further references to the current state of opinion, where reference is made to a contribution by the defendant's representatives here, namely Akinci/Heidrich, IPRB 2023, 270, 272, who apparently also take the Chamber's view; however, the contribution was not directly accessible to the Chamber until the judgement was issued). However, the question of whether and under what specific conditions a reservation declared in "natural language" can also be regarded as "machine understandable" will always have to be answered depending on the technical development existing at the relevant time of use of the work.

03 Accordingly, the European legislator has also stipulated within the framework of the AI Regulation that providers of AI models must have a strategy in place, in particular to identify and comply with a legal reservation asserted in accordance with Art. 4 para. 3 of the GDPR "also by means of state-of-the-art technologies" (Art. 53 para. 1 lit. c of the AI Regulation). However, these "state-of-the-art technologies" unambiguously include AI applications that are capable of recognising the content of text written in natural language (according to the defendants' representatives Akinci/Heidrich in particular in the article IPRB 2023, 270, 272, not directly accessible to the Chamber, cited here after Hamann,
op. cit. p. 148, who, incidentally, generally affirms this possibility from a technical point of view, loc. cit.) In this respect, there is every indication that the legislator of the AI Regulation had precisely such AI applications in mind with its reference to "state-of-the-art technologies".

04 Some argue that such a view leads to a circular argument: If it is demanded that the operator of the text and data mining must use AI applications to check whether a reservation of use h a s been declared, then this AI-supported search in turn requires a pattern analysis, which already fulfils the offence of text and data mining within the meaning of Section 44b (1) UrhG; in other words, only the application of the limitation decides on the permissibility of its application (according to Hamann, loc. cit., p. 148). The Chamber does not share this assessment: contrary to the aforementioned view, the copyright-relevant act of use requiring justification is not the performance of a "pattern analysis" as such, but the reproduction of the copyrighted work within the meaning of Section 16 UrhG. It does not appear compelling that the prior discovery of such works on the Internet and their verification as to whether reservations within the meaning of Section 44b (3) sentence 2 UrhG have been declared necessarily requires a quasi upstream further text and data mining within the meaning of Section 44b (1) UrhG, because processing of the website content through the use of web crawlers, in which only fleeting and incidental reproductions are created, which in turn are already justified under Section 44a UrhG, is to be considered in particular.

05 Furthermore, it is also objected to the broader understanding of the term "machine readability" considered by the Chamber that this term is understood more narrowly by the European legislator in a different context. In this context, reference is made to recital 35 of the PSI Directive (Directive (EU) 2019/1024), which requires, among other things, "simple" recognisability for "machine readability" within the meaning of this Directive (according to BeckOK UrhR/Bomhard, 42nd ed. 15 February 2024, UrhG Section 44b para. 31 with further references); this cannot be assumed for a reservation formulated only in natural language. However, such an argument presupposes that the terms of both directives must be understood in the same way. The Chamber has doubts as to whether such an equation of the terms is convincing, as the directives have different objectives: While the PSI Directive deals with the purely unilateral access of the public to information or the purely unilateral obligation of public authorities to publish certain information, Article 4(3) of the DSM Directive deals with a balance between the interests of the users of text and data mining (to be able to operate this as simply and as legally securely as possible) and the interests of the rights holders (to secure their rights as simply and as effectively as possible). In the opinion of the Chamber, this balance of interests cannot be resolved unilaterally in favour of the users of text and data mining by considering only the simplest conceivable technical solution for them as sufficient for the effectiveness of a declared reservation of use. Such an understanding would also be contradicted by the assessment of the legislator of the DSM Directive, which in recital 18 does not require the declaration of a reservation "in the simplest possible manner", but only "in an appropriate manner". And the German transposing legislator also only requires a declaration in a way that is "appropriate to the automated processes of text and data mining" (explanatory memorandum to the draft bill BT-Drucks. 19/27426, p. 89).

06 In the Chamber's view, it would also be a certain contradiction of values to allow the providers of AI models to develop ever more powerful text-understanding and text-creating AI models via the barrier in Section 44b (2) UrhG on the one hand, but not to require them to use existing AI models within the framework of the barrier in Section 44b (3) sentence 2 UrhG on the other.

07 Whether and to what extent sufficient technology for the automated recording of the content of the disputed reservation of use was already available at the time of the act of reproduction at issue in 2021 has not yet been demonstrated by the plaintiff; in this respect, the plaintiff has only referred to services available in 2023 (replica p. 14 et seq., p. 48 et seq. of the annex). However, there are indications that the defendant already had suitable technology. According to the defendant's own submission, the analysis carried out as part of the creation of the data set ... in the form of a comparison of image content with pre-existing image descriptions clearly also and precisely required the content of these image descriptions to be recorded by the software used. Against this background, there are some indications that systems were already available in 2021 - especially to the defendant - that were capable of automatically recording a reservation of use formulated in natural language.

08 3.

09 However, the defendant can invoke the limitation provision of Section 60d UrhG with regard to the reproduction at issue.

10 Accordingly, reproductions for text and data mining are permitted for the purposes of scientific research by research organisations.

11 a)

12 As explained above, the reproduction was made for the purpose of text and data mining within the meaning of Section 44b (1) UrhG. It was also made for the purposes of scientific research within the meaning of Section 60d (1) UrhG.

13 Scientific research generally refers to the methodical and systematic pursuit of new knowledge (Spindler/Schuster/Anton, 4th ed. 2019, UrhG Section 60c para. 3; BeckOK UrhR/Grübler, 42nd ed. 1.5.2024, UrhG Section 60c para. 5; Dreier/Schulze/Dreier, 7th ed. 2022, UrhG Section 60c para. 1). The concept of scientific research, in that it already allows the methodical-systematic "endeavour" to gain new knowledge to be sufficient, is not to be understood so narrowly that it would only cover the work steps directly associated with the acquisition of knowledge; rather, it is sufficient that the work step in question is aimed at a (later) acquisition of knowledge, as is the case, for example, with numerous data collections that must first be carried out in order to subsequently draw empirical conclusions. In particular, the concept of scientific research does not presuppose any subsequent research success.

14 Accordingly, contrary to the plaintiff's view, the creation of a data set of the type at issue, which can form the basis for the training of AI systems, can certainly be regarded as scientific research in the aforementioned sense. Although the creation of the data set as such may not yet be associated with a gain in knowledge, it is a fundamental work step with the aim of using the data set for the purpose of gaining knowledge at a later date. It can be affirmed that such an objective also existed in the present case. It is sufficient that the data set was - indisputably - published free of charge and thus made available to researchers (also) in the field of artificial neural networks. Whether the data set - as the plaintiff claims with regard to the services ... and ... is also used by commercial companies for the training or further development of their AI systems, is irrelevant because the research of commercial companies is also still research - even if not privileged as such under Sections 60c et seq. UrhG - is still research.

15 Against this background, the question in dispute between the parties as to whether the defendant also carries out scientific research in the form of the development of its own AI models in addition to the creation of corresponding data sets is irrelevant.

16 b)

17 The defendant is also not pursuing commercial purposes within the meaning of Section 60d (2) no. 1 UrhG.

18 The question of whether research is non-commercial depends solely on the specific nature of the scientific activity, while the organisation and financing of the institution in which the research is carried out are irrelevant (recital 42 InfoSoc Directive).

19 The non-commercial purpose pursued by the defendant in relation to the creation of the data set at issue ... already results from the fact that the defendant indisputably makes it publicly available free of charge. The fact that the development of the data set in dispute would also at least serve the development of the defendant's own commercial offer (cf. on this criterion BeckOK IT-Recht/Paul, 14th ed. 1 April 2024, UrhG § 60d para. 10) has neither been submitted by the plaintiff nor is it otherwise apparent. The fact that the data set in dispute may also be used by commercially active companies for training or further development of their AI systems is irrelevant for the categorisation of the defendant's activity. The mere fact that individual members of the defendant also carry out paid activities for such companies in addition to their work for the association is not sufficient to attribute the activities of these companies to the defendant as its own.

20 c)

21 The defendant is also not barred from invoking the limitation provision of Section 60d UrhG under subsection 2 sentence 3 of the provision.

22 Accordingly, research organisations that collaborate with a private company that has a decisive influence on the research organisation and preferential access to the results of scientific research cannot invoke the limitation provision of Section 60d UrhG. According to the wording of the standard, the burden of presentation and proof for the actual requirements of the counter-exclusion pursuant to Section 60d (2) sentence 3 UrhG lies with the plaintiff.

23 (1) Insofar as the plaintiff initially referred to the fact that the company ... has direct influence on the defendant through the financing of the dataset in question and the filling of "relevant positions" at the defendant by its own employees (replica p. 18, p. 52 of the annex), this submission lacks substance.

24 In this respect, the plaintiff merely refers to the fact that one of the co-founders of the defendant, Mr ..., is employed at ... as "Head of Machine Learning Operations", and that a member of the defendant, Mr ..., is also employed there as a "Research Scientist" (replica p. 4 f., p. 38 f. of the file). This activity of two members of the association for the company ... does not prove that this company has a decisive influence on the defendant's research work.

25 Apart from this, the plaintiff has not even alleged that the defendant granted the company ... preferential access to the results of its scientific research, namely the data set at issue. Rather, it is only submitted in this respect that ... its service ... with the help of the data set in dispute (replica p. 8 f., p. 42 f. of the file).

26 (2) Insofar as the plaintiff, in his statement of 3 July 2024, refers to a chat that took place in 2021 on the platform ... according to which the co-founder of the defendant, Mr ..., is said to have agreed to provide the company ... on the basis of a financial contribution of USD 5,000 made by the latter, this submission does not fulfil the exception in Section 60d (2) sentence 3 UrhG either.

27 It remains to be seen whether this chat process - not disputed by the defendant as such (cf. statement of 9 July 2024 p. 3, p. 112 of the annex) - supports the interpretation drawn by the plaintiff at all. It also remains to be seen whether the declaration of such a willingness to grant early access - the plaintiff has not stated whether this was actually granted - can be sufficient for holding preferential access to the research results within the meaning of Section 60d (2) sentence 2 UrhG.

28 In any event, it has neither been demonstrated by the plaintiff nor is it otherwise apparent that the company ... would have a decisive influence on the defendant. Insofar as any personal ties between the defendant and companies in the AI sector have been shown, these are the companies ... and ... (replica p. 4 et seq., p. 38 et seq. of the file).

29 II.

30 The decision on costs is based on Section 91 (1), Section 91a (1) sentence 1 ZPO.

31 The decision on provisional enforceability is made in accordance with Section 708 No. 11, Sections 711, 709 ZPO.