

Verkündet am 27.09.2024

Urkundsbeamter/in der Geschäftsstelle



Urteil

IM NAMEN DES VOLKES

In der Sache

XXX, XXX

- Kläger -

Prozessbevollmächtigte:

Rechtsanwälte **SLD intellectual propety Rechtsanwaltsgesellschaft mbH**, Alstadt 260,
84028 Landshut, Gz.: 23-00061

gegen

XXX e. V., XXX

- Beklagter -

Prozessbevollmächtigte:

Rechtsanwälte **Heidrich Rechtsanwälte Partnerschaftsgesellschaft mbB**, Prinzenstraße 3,
30159 Hannover, Gz.: 126/23-NA-UM

erkennt das Landgericht Hamburg - Zivilkammer 10 - durch den Vorsitzenden Richter am Landgericht Hartmann, die Richterin Dr. Kanzler und den Richter am Landgericht Dr. Woodtli am 27.09.2024 auf Grund der mündlichen Verhandlung vom 11.07.2024 für Recht:

1. Die Klage wird abgewiesen.
2. Der Kläger hat die Kosten des Rechtsstreits zu tragen.
3. Das Urteil ist vorläufig vollstreckbar. Der Kläger kann die Vollstreckung des Beklagten durch Sicherheitsleistung in Höhe von 110 % des aufgrund des Urteils vollstreckbaren Betrags abwenden, wenn nicht der Beklagte vor der Vollstreckung Sicherheit in Höhe von 110 % des zu vollstreckenden Betrags leistet.

Tatbestand

Der Beklagte ist ein Verein, der mit Gründungsversammlung vom 07.07.2021 gegründet wurde (Protokoll in Anlage B7, Satzung in Anlage B1). Die konkrete Zweckrichtung der Tätigkeit des Beklagten steht zwischen den Parteien in Streit.

Der Beklagte stellt unter der Bezeichnung „XXX“ ein sogenanntes Dataset für Bild-Text-Paare öffentlich kostenfrei zur Verfügung. Es handelt sich dabei um eine Art Tabellendokument, das Hyperlinks zu im Internet öffentlich abrufbaren Bildern bzw. Bilddateien sowie weitere Informationen zu den entsprechenden Bildern enthält, darunter eine Bildbeschreibung (auch Alternativtext genannt), die Auskunft über den Inhalt des Bildes in Textform gibt. Der Datensatz umfasst 5,85 Milliarden entsprechende Bild-Text-Paare. Der Datensatz kann für das Trainieren sog. generativer Künstlicher Intelligenz genutzt werden.

Die Erstellung des Datensatzes erfolgte nach Gründung des Beklagten in der zweiten Jahreshälfte 2021. Dazu hatte der Beklagte auf einen bereits vorhandenen Datensatz der XXX aus den USA (www.XXX.org) zurückgegriffen, der für eine Art zufälligen Querschnitt der im Internet auffindbaren Bilder die jeweiligen URLs nebst textlicher Beschreibung des jeweiligen Bildinhalts enthielt. Der Beklagte extrahierte sodann die URLs zu den Bildern aus diesem Datensatz und lud die Bilder von ihrem jeweiligen Speicherort herunter. Im Anschluss wurden die Bilder beim Beklagten durch eine Software darauf geprüft, ob die im vorbestehenden Datensatz bereits vorhandene Beschreibung des Bildinhalts tatsächlich mit dem auf dem Bild zu sehenden Inhalt übereinstimmte. Bilder, bei denen Text und Bildinhalt nicht hinreichend übereinstimmten, wurden herausgefiltert. Für die verbleibenden Bilder wurden die Meta-Daten, insbesondere die URL des Speicherorts des Bildes und die Bildbeschreibung, extrahiert und in einen neu geschaffenen Datensatz, den XXX, aufgenommen. Ob im Anschluss die heruntergeladenen Bilddateien wieder gelöscht wurden, steht zwischen den Parteien – jedenfalls in Bezug auf die streitgegenständliche Fotografie – in Streit.

Im Rahmen des vorgenannten Prozesses wurde auch das streitgegenständliche Bild erfasst, heruntergeladen, analysiert und mit seinen Meta-Daten in den Datensatz XXX aufgenommen. Konkret heruntergeladen wurde dabei eine auf der Webseite der Bildagentur XXX (<https://www.XXX.com>) eingestellte, mit einem Wasserzeichen der Fotoagentur XXX versehene Bilddatei.

Auf der Webseite der Bildagentur XXX befand sich mindestens seit dem 13.01.2021 auf der Unterseite <https://www.XXX.com/de/usage.html> folgender Text:

„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.“

Der Kläger rügt eine Verletzung urheberrechtlicher Rechte an der streitgegenständlichen Fotografie in Form einer unzulässigen Vervielfältigung im Rahmen des Analyseprozesses durch den Beklagten.

Der Kläger behauptet, er sei Urheber des im Tenor aufgeführten Fotos. Die Firma XXX sei berechtigt gewesen, auf ihrer Internetseite XXX.com das streitgegenständliche Foto anzubieten und auch zu zeigen sowie Lizenzen für das Foto anzubieten; XXX sei insofern Inhaberin von einfachen, weiterlizenzierbaren Nutzungsrechten gewesen.

Die – unstrittig – im Rahmen des Analyseprozesses erfolgte Vervielfältigung verletze ihn, den Kläger, in seinen Rechten aus § 16 UrhG, insbesondere sei sie nicht durch die Schrankenregelungen der §§ 44a, 44b und 60d UrhG gedeckt:

Die Schrankenregelung des § 44a UrhG sei nicht einschlägig, der eigenständige Download einer Fotografie stelle insbesondere keine vorübergehende Handlung im Sinne dieser Norm dar.

Die Vervielfältigung sei auch nicht durch § 44b UrhG gedeckt. Das Zusammenziehen von Daten zum Zwecke des KI-Trainings sei kein Text oder Data Mining i.S.d. § 44b UrhG. Weder der europäische noch der deutsche Gesetzgeber hätten bei Schaffung der Schrankenregelung des Art. 4 DSM-RL bzw. des § 44b UrhG eine solche Verwendung „vor Augen“ gehabt. Beim Text und Data Mining i.S.d. § 44b UrhG sollten nur „in den Daten verborgene Informationen erschlossen“, „nicht aber der Inhalt der geistigen Schöpfung genutzt werden“. Beim hier vorliegenden sog. „KI-Web scraping“ gehe es jedoch gerade um den geistigen Inhalt der zu Trainingszwecken genutzten Werke „und letztlich um die Schaffung inhaltsgleicher oder ähnlicher Konkurrenzserzeugnisse“. Zudem sei der Datensatz nach dem eigenen Disclaimer des Beklagten (abgedruckt auf Bl. 47) „unkuratiert“. Schließlich sei die Sammlung und Speicherung zur Schaffung von Parallel-Archiven nach dem ausdrücklich erklärten Willen des Gesetzgebers von der Schrankenbestimmung des § 44b UrhG ausgenommen.

Zudem beeinträchtige „das massenhafte Einverleiben von urheberrechtlich geschützten Werken zu Trainingszwecken im Rahmen von generativer KI“ die normale Verwertung urheberrechtlich geschützter Werke, weil es die Voraussetzung dafür schaffe, Urheberinnen und Urheber in vielen Fällen zu ersetzen, oder jedenfalls die Verwertung des Werkes durch ein kostenloses Konkurrenzangebot erheblich erschwere. Dies stehe nach Art. 7 Abs. 2 DSM-RL in Verbindung mit Art. 5 Abs. 5 InfoSoc-RL aber der Anwendung der Schrankenregelung entgegen.

Jedenfalls sei die Vervielfältigung aufgrund des auf der Webseite www.bigstockphoto.com erklärten Nutzungsvorbehalts nach § 44b Abs. 3 UrhG unzulässig. Die entsprechende Erklärung

der Bildagentur sei dem Kläger zuzurechnen, da sie das streitgegenständliche Foto für ihn vertreibe. Entgegen der Ansicht des Beklagten sei der Vorbehalt auch maschinenlesbar i.S.d. § 44b Abs. 3 S. 2 UrhG. Die diesbezüglichen Anforderungen seien nicht höher als bei einer Maschinenlesbarkeit durch Menschen; der Vorbehalt sei aber in Druckschrift verfasst. Im Übrigen sei der Text auch für ein Computerprogramm als Vorbehalt erkennbar. So habe der Dienst XXX den entsprechenden Vorbehalt erkennen können, zudem könnten spezifische Tools wie WebOpt-Out Vorbehalte wie auf XXX.com erkennen.

Auch auf die Schrankenregelung des § 60d UrhG könne sich der Beklagte nicht berufen. Der Kläger bestreitet dazu, dass der Beklagte die Voraussetzungen des § 60d UrhG in tatsächlicher Hinsicht erfülle, namentlich

- dass er zum Zeitpunkt der hier gegenständlichen Vervielfältigungshandlung „ins Vereinsregister eingetragen“ gewesen sei,
- dass das seitens des Beklagten vorgelegte Dokument B1 die gültige Vereinssatzung des Beklagten darstelle oder zum Zeitpunkt der gegenständlichen Vervielfältigungshandlung dargestellt habe,
- dass die Vereinsmitglieder sowie der Vorstand ehrenamtlich tätig seien bzw. zum Zeitpunkt der gegenständlichen Vervielfältigungshandlung gewesen seien,
- dass der Beklagte ausschließlich forschend tätig sei bzw. zum Zeitpunkt der gegenständlichen Vervielfältigungshandlung gewesen sei bzw. dass der Beklagte wissenschaftliche Forschung betreibe, nicht-kommerzielle Zwecke verfolge und sämtliche Gewinne in die wissenschaftliche Forschung reinvestiere oder im Rahmen eines staatlich anerkannten Auftrags im öffentlichen Interesse tätig sei; zudem sei Zweck des Beklagten nach der von ihm vorgelegten Satzung nur die „Förderung der Forschung“ und nicht die „Forschung“ als solche, und es sei auch unklar, was an der von dem Beklagten erstellten Sammlung, die (unstreitig) anderen Unternehmen zur Verfügung gestellt werde, Forschung sein solle;
- dass durch den Beklagten eigene KI-Modelle auf Basis der Trainingsdaten erstellt und getestet würden, um die Möglichkeiten von KI-Technologie weiter zu erforschen,
- dass durch das öffentliche Zurverfügungstellen des Trainingsdatensatzes anderen Forschenden und Interessierten die Möglichkeit geboten werden solle, eigene KI-Modelle zu trainieren; nach den eigenen Bekundungen des Beklagten sei das streitgegenständliche Dataset auch zum Trainieren der Dienste „XXX“, „XXX“ und „XXX“ des Anbieters XXX verwendet worden; diese würden aber von (rein) kommerziellen Unternehmen betrieben; soweit der Beklagte ein Trainieren der ersten beiden Dienste bestreite, wäre es diesen jedenfalls möglich gewesen, den Datensatz zu benutzen.

Der Beklagte könne sich zudem nach § 60d Abs. 2 S. 3 UrhG nicht auf die Privilegierung des § 60d UrhG berufen. Der Beklagte arbeite offenbar intensiv mit kommerziellen KI-Anbietern zusammen:

- So bestehe offenbar eine Zusammenarbeit mit dem Privatunternehmen XXX, welches über die Finanzierung des gegenständlichen Datasets und die Besetzung relevanter Posten beim Beklagten durch eigene Mitarbeiter unmittelbaren Einfluss auf den Beklagten habe. So habe XXX nach einer Interview-Äußerung seines Gründers und Geschäftsführers das XXX Dataset finanziert.
- Mitglieder des „Teams“ des Beklagten seien zudem „vielerorts“ kommerziell im gleichen Bereich für große Tech-Firmen, u.a. als Mitarbeiter des Unternehmens XXX, tätig.
- In einem Chat auf der Plattform „XXX“ habe zudem der Mitbegründer des Beklagten, XXX, unter dem 28.09.2021 zur Eile bei der Fertigstellung des „XXX“ gemahnt, da man von einem „XXX“ bzw. dessen Unternehmen eine Finanzierung von 5.000,- \$ erhalten habe; dabei solle man diesem die Daten auch dann (schon) zur Verfügung stellen, wenn das Dataset der Öffentlichkeit noch nicht zugänglich gemacht werden könne. Bei dem besagten „XXX“ handele es sich um einen Mitarbeiter des kommerziellen KI-Anbieters XXX.

Schließlich könne sich der Beklagte auch nicht auf eine sog. „schlichte Einwilligung“ berufen. Er – der Kläger – habe die streitgegenständliche Fotografie nicht frei zugänglich gemacht, sondern zur Einräumung von entgeltlichen Lizenzen durch die Agentur XXX anbieten lassen.

Der Kläger hatte zunächst neben Unterlassung der Vervielfältigung auch Auskunft über den Umfang der erfolgten Nutzung der Fotografie beantragt. Diesen Auskunftsantrag haben die Parteien im Termin zur mündlichen Verhandlung am 11.07.2024 übereinstimmend für erledigt erklärt.

Der Kläger beantragt nunmehr noch,

den Beklagten zu verurteilen, es bei Meidung eines Ordnungsgeldes von je bis zu 250.000,- €, ersatzweise Ordnungshaft bis zu 6 Monate, für jeden einzelnen Fall der Zuwiderhandlung zu unterlassen, die nachfolgend abgebildete Fotografie

[folgt Einblendung der Fotografie]

zur Erstellung von KI-Trainingsdatensätzen zu vervielfältigen und/oder vervielfältigen zu lassen, wie geschehen im Rahmen der Herstellung des Datensatzes XXX.

Der Beklagte beantragt,

die Klage abzuweisen.

Er bestreitet, dass der Kläger das streitgegenständliche Bild selbst erstellt habe oder aus anderem Grunde im Hinblick auf das Bild berechtigt sei, Rechtsverletzungen im eigenen Namen geltend zu machen, sowie, dass der Kläger im Zeitpunkt der Erfassung des Bildes durch den Beklagten berechtigt gewesen sei, Rechtsverletzungen bezüglich des streitgegenständlichen Bildes im eigenen Namen geltend zu machen.

Vor allem stelle aber der im Rahmen der Erstellung des Datensatzes XXX erfolgte (einmalige) Download des streitgegenständlichen Bildes zwar eine urheberrechtsrelevante Vervielfältigung dar, diese sei aber von den Schrankenbestimmungen der §§ 44a, 44b und 60d UrhG sowie einer schlichten Einwilligung des Klägers gedeckt:

Die erfolgte Vervielfältigung sei zum einen von der Schrankenregelung des § 44a UrhG erfasst. Es sei keine dauerhafte Speicherung der Bilder durch den Beklagten erfolgt; vielmehr seien die Bilder nur kurzzeitig für die Analyse genutzt und sodann unverzüglich automatisiert und unwiderruflich gelöscht worden. Eine eigenständige wirtschaftliche Bedeutung komme der Vervielfältigung dabei gerade nicht zu.

Auch greife die Schrankenbestimmung des § 44b UrhG ein. Das Analysieren von Bilddateien und das Extrahieren von Metadaten zum Training Künstlicher Intelligenz sei nach dem Willen des Gesetzgebers ein Hauptanwendungsfall des Text und Data Mining. Es liege auch keine Schaffung digitaler Parallelarchive vor, da die heruntergeladenen Bilder gerade nicht dauerhaft gespeichert, sondern nur Hyperlinks aufgenommen worden seien. Der Ausnahmetatbestand des § 44b Abs. 3 UrhG greife nicht ein:

- Nach dem Vortrag des Klägers habe nicht er selbst als Rechteinhaber, sondern die Betreiberin der Website www.XXX.com als Dritte diesen Vorbehalt erklärt; der Kläger selbst habe in seiner E-Mail vom 13.02.2023 (Anlage B5) sogar ausdrücklich erklärt, dass er weder über die Qualifikation noch über die wirtschaftlichen Mittel verfüge, um einen Nutzungsvorbehalt auszusprechen.
- Zudem sei der Vorbehalt nicht ausdrücklich erfolgt, da der Passus auf der Website www.XXX.com allgemein gehalten sei und diverse unzulässige Handlungen aufführe. Weder erfolge eine explizite Erwähnung des Text und Data Mining noch von Vervielfältigungen.
- Zudem fehle es am Merkmal der Maschinenlesbarkeit. Eine in natürlicher Sprache verfasste Klausel sei generell nicht maschinenlesbar i.S.d. § 44b Abs. 3 UrhG. Voraussetzung für eine Maschinenlesbarkeit in diesem Sinne sei, dass die Klausel automatisch von einer Software verarbeitet werden könne. Dies setze eine Kodierung der entsprechenden Information voraus. Jedenfalls sei erforderlich, dass mindestens spezifische Stichwörter wie etwa „Data Mining“ in dem Text enthalten seien.

- Der Vorbehalt sei auch ersichtlich nicht als ein solcher nach § 44b Abs. 3 UrhG gemeint. Der Umstand, dass die Klausel nach dem Vortrag des Klägers schon am 13.01.2021 auf der Webseite so vorhanden gewesen sein solle, verdeutliche, dass die Klausel nicht „mit Blick auf die Regelung in § 44b Abs. 3 UrhG“ erstellt worden sein könne, da die gesetzliche Regelung zu diesem Zeitpunkt noch nicht in Kraft getreten gewesen sei. Es sei im Übrigen „auch nicht glaubwürdig“, dass ein u.s.-amerikanischer Anbieter sich auf einen – deutschen Gesetzen entspringenden – Nutzungsvorbehalt berufe.

Jedenfalls könne er – der Beklagte – sich auf die Schrankenbestimmung des § 60d UrhG berufen.

- Er – der Beklagte – sei ein gemeinnütziger Verein, der aus Forschenden bestehe und sich ausweislich der Vereinssatzung (Anlage B1) der Forschung verschrieben habe, insbesondere habe er es sich zur Aufgabe gemacht, selbstlernende Algorithmen im Sinne Künstlicher Intelligenz fortzuentwickeln und der breiten Öffentlichkeit zur Verfügung zu stellen. Er stelle dazu kostenfrei Datensätze und Modelle bereit, daneben erstelle und teste er eigene KI-Modelle auf Basis der Trainingsdaten.
- Seine Tätigkeit stelle auch „Forschung“ dar. Allein dadurch, dass er öffentlich im Internet transparent mache, wie die Trainingsdatensätze erstellt würden, leiste er einen Beitrag zur Gewinnung von Erkenntnissen über das Training künstlicher Intelligenz. So könnten andere Forscher die Schritte zur Erstellung der Datensätze nachvollziehen und darauf aufbauen. Zudem habe er zu dem streitgegenständlichen Dataset XXX eine wissenschaftliche Abhandlung mit dem Namen „XXX“ erstmals am 17.09.2022 veröffentlicht (Anlage B8). Bis zum 05.04.2024 sei die vorgenannte Abhandlung insgesamt 1403-mal in anderen wissenschaftlichen Werken zitiert worden und habe auch weitere Auszeichnungen erhalten.
- Zudem trainiere der Beklagte auf Basis der von ihm erstellten Datensätze auch eigene KI-Modelle, um Erkenntnisse darüber zu gewinnen, wie KI durch entsprechendes Training verbessert werden könne.
- Auch die dem Beklagten angehörenden natürlichen Personen, also die Vorstandsmitglieder und übrigen Vereinsmitglieder, seien „Forschende“. Soweit der Kläger bezüglich einzelner Teammitglieder darauf verweise, dass diese im gleichen Bereich für große Tech-Firmen tätig seien, habe dies keine Relevanz für die Frage, ob der Beklagte selbst kommerziell tätig ist. Im Übrigen seien die betreffenden Personen für den Beklagten ehrenamtlich tätig und müssten daher anderweitig ihren Lebensunterhalt verdienen.
- Der Umstand, dass die vom Beklagten öffentlich zur Verfügung gestellten Datasets auch von kommerziellen Anbietern genutzt würden, sei für das Eingreifen der

Schrankenbestimmung des § 60b UrhG irrelevant. Davon abgesehen seien die Dienste XXX und XXX tatsächlich gar nicht mit dem Dataset des Beklagten trainiert worden.

- Auch die in § 60 d Abs. 2 S. 3 UrhG geregelte Rückausnahme greife vorliegend nicht. Zwar habe das Unternehmen XXX dem Beklagten in der Tat in der Gründungsphase Rechenressourcen zur Verfügung gestellt. Gleiches sei aber auch durch das XXX geschehen. Eine finanzielle Unterstützung in Form von Geld habe der Beklagte nicht von XXX erhalten. Auch eine weitere Zusammenarbeit mit diesem Unternehmen habe es nicht gegeben. In jedem Fall erhalte XXX aber keinen bevorzugten Zugang zu den Forschungsergebnissen. Auch ein bestimmender Einfluss durch das Unternehmen XXX liege nicht vor. Weder das Unternehmen XXX selbst noch einer seiner gesetzlichen Vertreter des Unternehmens seien Mitglieder des Beklagten.

Schließlich liege eine sog. schlichte Einwilligung zugunsten der Nutzungen durch den Beklagten von Seiten des Klägers vor. Die streitgegenständliche Vervielfältigung sei als übliche Nutzungshandlung einzuordnen.

Im Übrigen sei darauf zu verweisen, dass der Kläger selbst Geld mit KI-generierten Bildern verdiene. Dies stelle sein „Rechtsschutzbedürfnis“ infrage.

Hinsichtlich der weiteren Einzelheiten des Parteivortrags wird auf die zur Akte gelangten Schriftsätze nebst Anlagen, soweit sie Gegenstand der mündlichen Verhandlung geworden sind, sowie auf die Sitzungsniederschrift vom 11.07.2024 (Bl. 120-123) verwiesen.

Die Parteien haben nicht nachgelassene Schriftsätze vom 29.08. und 11.09.2024 (Kläger) bzw. vom 20.09.2024 (Beklagter) zur Akte gereicht.

Entscheidungsgründe

I.

Die zulässige Klage hat in der Sache keinen Erfolg. Der Beklagte hat durch die Vervielfältigung der streitgegenständlichen Fotografie zwar in die Verwertungsrechte des Klägers eingegriffen. Dieser Eingriff ist aber durch die Schrankenregelung des § 60d UrhG gedeckt. Ob sich der Beklagte ergänzend auf die Schrankenregelung des § 44b UrhG berufen kann, bedarf vor diesem Hintergrund keiner abschließenden Beurteilung.

Die streitgegenständliche Fotografie ist jedenfalls als Lichtbild nach § 72 Abs. 1 UrhG geschützt. Das Gericht hat nach Inaugenscheinnahme der auf dem Laptop des Klägers befindlichen Rohdaten auch keinen Zweifel an der Lichtbildnereigenschaft des Klägers, § 72 Abs. 2 UrhG. Der Kläger ist auch zur Geltendmachung von Verletzungsansprüchen nach § 97 UrhG aktiv legitimiert,

so auch für den Unterlassungsanspruch nach Abs. 1 der Vorschrift; dass der Kläger der Bildagentur XXX weitergehende als (unterlizenzierbare) einfache Nutzungsrechte eingeräumt hat, hat der Beklagte nicht dargelegt. Die Bildagentur XXX hat das Foto mit einem Wasserzeichen versehen; das war eine unfreie Umgestaltung im Sinne des § 23 Abs. 1 S. 1 UrhG, so dass auch zu deren Verwertung grundsätzlich die Zustimmung des Klägers als des Urhebers erforderlich war. Im Rahmen des durchgeführten Downloads hat der Beklagte diese Fassung vervielfältigt i.S.d. § 16 Abs. 1 UrhG, ohne dafür eine Zustimmung des Klägers eingeholt zu haben.

Allerdings war der Beklagte hierzu aufgrund gesetzlicher Erlaubnis berechtigt. Die Vervielfältigung war zwar nicht durch die Schrankenregelung des § 44a UrhG gedeckt (im Folgenden 1.), und ob sich der Beklagte auf die Schrankenregelung des § 44b UrhG berufen kann, erscheint als zweifelhaft (im Folgenden 2.). Letzteres bedarf aber vorliegend keiner abschließenden Entscheidung, da die Vervielfältigungshandlung jedenfalls durch die Schrankenregelung des § 60d UrhG gedeckt war (im Folgenden 3.).

1.

Die erfolgte Vervielfältigung ist nicht durch die Schrankenregelung des § 44a UrhG gedeckt.

Danach sind vorübergehende Vervielfältigungshandlungen zulässig, die flüchtig oder begleitend sind und einen integralen und wesentlichen Teil eines technischen Verfahrens darstellen und deren alleiniger Zweck es ist, eine Übertragung in einem Netz zwischen Dritten durch einen Vermittler oder eine rechtmäßige Nutzung eines Werkes oder sonstigen Schutzgegenstands zu ermöglichen, und die keine eigenständige wirtschaftliche Bedeutung haben.

Die vorliegend erfolgte Vervielfältigung war bereits weder flüchtig noch begleitend.

a)

Flüchtig i.S.d. § 44a UrhG ist eine Vervielfältigung dann, wenn ihre Lebensdauer auf das für das ordnungsgemäße Funktionieren des betreffenden technischen Verfahrens Erforderliche beschränkt ist, wobei dieses Verfahren derart automatisiert sein muss, dass es diese Handlung automatisch, d.h. ohne Beteiligung einer natürlichen Person löscht, sobald ihre Funktion, die Durchführung eines solchen Verfahrens zu ermöglichen, erfüllt ist (EuGH, Urt. v. 16.07.2009, Az. C-5/08 – Infopaq/Danske Dagblades Forening, Rn. 64 (juris) zu Art. 5 Abs. 1 DSM-RL).

Soweit sich der Beklagte insoweit darauf beruft, dass im Rahmen des von ihm durchgeführten Analyseverfahrens die Dateien „automatisch“ gelöscht worden seien, vermag dies eine Flüchtigkeit der Vervielfältigung im vorgenannten Sinne nicht zu begründen. Abgesehen davon, dass der Beklagte nichts zu der konkreten Dauer der Speicherung vorgetragen hat, erfolgte die Löschung gerade nicht „nutzerunabhängig“, sondern aufgrund einer entsprechenden bewussten Programmierung des Analyseprozesses durch den Beklagten.

b)

Begleitend i.S.d. § 44a UrhG ist eine Vervielfältigung dann, wenn sie gegenüber dem technischen Verfahren, dessen Teil sie ist, weder eigenständig ist noch einem eigenständigen Zweck dient (EuGH, Urt. v. 05.06.2014, Az. C-360/13, Rn. 43 (juris)).

Im vorliegenden Fall erfolgte ein gezieltes Herunterladen der Bilddateien, um sie mittels einer spezifischen Software zu analysieren. Damit ist das Herunterladen kein bloß begleitender Prozess zu der durchgeführten Analyse, sondern ein der Analyse vorgelagerter bewusster und aktiv gesteuerter Beschaffungsprozess.

2.

Ob sich der Beklagte auf die Schrankenregelung des § 44b UrhG berufen kann, erscheint im vorliegenden Fall durchaus als zweifelhaft. Zwar unterfällt der von dem Beklagten vorgenommene Download grundsätzlich der Schrankenregelung des § 44b Abs. 2 UrhG, insbesondere erfolgte er zum Zwecke des Text und Data Mining im Sinne des § 44b Abs. 1 UrhG (im Folgenden lit. a). Allerdings spricht – ohne dass dies vorliegend einer abschließenden Entscheidung bedürfte – Einiges dafür, dass aufgrund eines wirksam erklärten Nutzungsvorbehalts im Sinne des § 44b Abs. 3 die Vervielfältigungshandlung nicht bereits nach § 44b Abs. 2 UrhG zulässig war (im Folgenden lit. b).

a)

Die streitgegenständliche Vervielfältigungshandlung unterfällt grundsätzlich der Schrankenregelung des § 44b Abs. 2 UrhG.

(1) Der streitgegenständliche Download erfolgte zum Zwecke des Text und Data Mining im Sinne des § 44b Abs. 1 UrhG. Danach ist Text und Data Mining die automatisierte Analyse von einzelnen oder mehreren digitalen oder digitalisierten Werken, um daraus Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen. Jedenfalls für die vorliegend streitgegenständliche Vervielfältigungshandlung ist dies zu bejahen (nachfolgend (a)); eine teleologische Reduktion des Schrankentatbestands kommt insofern nicht in Betracht (unten (b)).

Vorliegend keiner Entscheidung bedarf daher die weitere, im Schrifttum eingehend diskutierte Frage, ob das Training von Künstlicher Intelligenz in seiner Gesamtheit der Schrankenregelung des § 44b UrhG unterfällt oder nicht (eingehend zum Meinungsstand BeckOK UrhR/Bomhard, 42. Ed. 15.2.2024, UrhG § 44b Rn. 11a-11b m.w.N.; dazu auch eingehend die als Anlage K11 vorgelegte, im Auftrag der Initiative Urheberrecht erstellte Studie „Urheberrecht & Training generativer KI-technologische und rechtliche Grundlagen“).

(a) Der Beklagte hat die Vervielfältigungshandlung zum Zwecke der Gewinnung von Informationen über „Korrelationen“ im Wortsinn des § 44b Abs. 1 UrhG vorgenommen. Der Beklagte hat das streitgegenständliche Lichtbild von seinem ursprünglichen Speicherort heruntergeladen, um mittels einer bereits verfügbaren Software – offenbar der Anwendung XXX

von XXX – den Bildinhalt mit der zu dem Text bereits hinterlegten Bildbeschreibung abzugleichen. Diese Analyse der Bilddatei zum Abgleich mit einer vorbestehenden Bildbeschreibung stellt ohne Weiteres eine Analyse zum Zwecke der Gewinnung von Informationen über „Korrelationen“ (nämlich der Frage der Nicht-/Übereinstimmung von Bildern und Bildbeschreibungen) dar. Dass der Beklagte die in den Datensatz XXX aufgenommenen Bilder auf diese Art und Weise analysiert hat, wurde klägerseits als solches nicht bestritten.

Die Anwendbarkeit von § 44b Abs. 1 UrhG ist – entgegen der Auffassung des Klägers (Replik S. 13 f., Bl. 47 f. d.A.) – auch nicht deshalb ausgeschlossen, weil der Beklagte den von ihm erstellten Datensatz XXX ausweislich eines zu diesem Datensatz ausgesprochenen „Disclaimers“ nicht „kuratiert“ habe. Der klägerseits wiedergegebene Disclaimer bezieht sich allein auf eine Warnung, dass der Datensatz nicht nach „verstörenden Inhalten“ o.Ä. durchsucht worden sei. Eine solche – zusätzliche – Filterung des zu erstellenden Datensatzes ist aber nicht Anwendungsvoraussetzung für § 44b Abs. 1 UrhG und steht nicht der Annahme entgegen, dass die heruntergeladenen Bilder – wie ausgeführt – auf ihre Korrelation zwischen Bildinhalt und Bildbeschreibung hin analysiert wurden.

(b) Die streitgegenständliche Vervielfältigungshandlung ist auch nicht im Wege der teleologischen Reduktion der Schrankenregelung des § 44b UrhG aus dieser auszuschließen.

Soweit eine Herausnahme der Vervielfältigung von Daten zum Zwecke des KI-Trainings im Wege der teleologischen Reduktion im Schrifttum vereinzelt mit der Begründung befürwortet wird, dass § 44b UrhG nur die Erschließung „in den Daten verborgener Informationen“, nicht aber die Nutzung „des Inhalts der geistigen Schöpfung“ erfasse (*Schack*, NJW 2024, 113; in diese Richtung auch *Dormis/Stober*, Urheberrecht und Training generativer KI-Modelle, Anlage K11, S. 67 ff. mit einer Differenzierung zwischen Semantik und Syntax), bestehen Zweifel, ob dies zu überzeugen vermag; denn dabei wird nicht ausreichend deutlich, worin bei digitalisierten Werken der Unterschied zwischen „in den Daten verborgenen Informationen“ und „dem Inhalt der geistigen Schöpfung“ liegen soll.

Soweit ergänzend angeführt wird, dass es beim „KI-Web scraping“ um den geistigen Inhalt der zu Trainingszwecken genutzten Werke und „letztlich“ um die Schaffung inhaltsgleicher oder ähnlicher Konkurrenzzeugnisse gehe (*Schack*, a.a.O.), unterscheidet diese Argumentation nach Auffassung der Kammer nicht streng genug zwischen

- zum einen der (hier allein streitgegenständlichen) *Erstellung* eines – auch – für KI-Training nutzbaren Datensatzes,
- zum anderem dem nachfolgenden *Training* des künstlichen neuronalen Netzes mit diesem Datensatz und

- zum dritten der nachfolgenden *Nutzung der trainierten KI* zum Zwecke der Erstellung neuer Bildinhalte.

Diese letztere Funktionalität mag zwar bereits bei der Erstellung des Trainingsdatensatzes angestrebt sein. Jedoch ist im Zeitpunkt der Zusammenstellung des Trainingsdatensatzes weder absehbar, in welcher Weise der zweite Schritt (das Training) erfolgreich sein wird, noch, welche konkreten Inhalte im dritten Schritt (bei der KI-Anwendung) durch die trainierte KI werden generiert werden können. Die konkreten Anwendungsmöglichkeiten bei einer sich rasant entwickelnden Technologie wie der KI sind zum Zeitpunkt der Erstellung des Trainingsdatensatzes daher nicht abschließend absehbar und mithin nicht rechtssicher feststellbar. Wegen dieser Rechtsunsicherheit ist die bei der Erstellung des Trainingsdatensatzes zunächst allein bestehende nur allgemeine Absicht, zukünftige KI-generierte Inhalte erlangen zu wollen, kein taugliches Kriterium für die Beurteilung der rechtlichen Zulässigkeit der Erstellung des Trainingsdatensatzes als solchen.

Soweit für eine teleologische Reduktion der Schrankenregelung des § 44b UrhG schließlich angeführt wird, dass der europäische Gesetzgeber 2019 bei Schaffung der zugrundeliegenden Richtlinienbestimmung (Art. 4 DSM-RL) „das KI-Problem“ „schlicht noch nicht auf dem Schirm“ gehabt habe (*Schack*, a.a.O.; ebenso für das Training von KI-Modellen *Dormis/Stober*, a.a.O., S. 71 ff., 87 ff.), genügt allein dieser Befund für eine teleologische Reduktion ersichtlich nicht. Dabei ist insbesondere zu berücksichtigen, dass die technische Fortentwicklung im Bereich der sog. Künstlichen Intelligenz seit 2019 weniger Art und Umfang des (streitgegenständlichen) Data Mining zur Beschaffung von Trainingsdaten betrifft, sondern die Leistungsfähigkeit der mit den Daten trainierten künstlichen neuronalen Netze (dementsprechend gehen *Dormis/Stober*, a.a.O., S. 95 gleichfalls davon aus, dass die reine Erstellung von Trainingsdatensätzen „im Vorfeld des eigentlichen Trainings“ durchaus der TDM-Schranke unterfallen dürfte). Zu beachten wäre außerdem, dass die beklagenseits *abgerufene* Datenbank der Common Crawl Foundation bereits seit dem Jahr 2008 (!) erstellt wird, vgl. <https://commoncrawl.org/overview>.

Davon abgesehen hat jedenfalls der aktuelle europäische Gesetzgeber der KI-Verordnung (Verordnung (EU) 2024/1689 vom 13.06.2024, ABI. L v. 12.07.2024 S. 1) unzweifelhaft zum Ausdruck gebracht, dass auch die Erstellung von zum Training von künstlichen neuronalen Netzen bestimmten Datensätzen der Schrankenregelung des Art. 4 DSM-RL unterfällt. Denn nach Art. 53 Abs. 1 lit. c KI-Verordnung sind Anbieter von KI-Modellen mit allgemeinem Verwendungszweck verpflichtet, eine Strategie insbesondere zur Ermittlung und Einhaltung eines gemäß Art. 4 Abs. 3 DSM-RL geltend gemachten Rechtsvorbehalts vorzusehen.

Dass auch die Erstellung von zum Training von künstlichen neuronalen Netzen bestimmten Datensätzen der Schrankenregelung des Art. 4 DSM-RL unterfalle, entspricht im Übrigen auch

der Einschätzung des deutschen Gesetzgebers im Rahmen der Umsetzung der vorgenannten Schrankenbestimmung im Jahr 2021 (Begr. RegE BT-Drs. 19/27426, S. 60).

(c) Auch der in Art. 5 Abs. 5 InfoSoc-RL (i.V.m. Art. 7 Abs. 2 S. 1 DSM-RL) verankerte sog. 3-Stufen-Test rechtfertigt letztlich keine andere Beurteilung. Danach dürfen die normierten Ausnahmen nur in bestimmten Sonderfällen angewandt werden, in denen die normale Verwertung des Werks oder des sonstigen Schutzgegenstands nicht beeinträchtigt wird und die berechtigten Interessen des Rechtsinhabers nicht ungebührlich verletzt werden. Diese Anforderungen sind im vorliegenden Fall erfüllt.

Die vorliegend urheberrechtlich relevante Vervielfältigung ist auf den Zweck der Analyse der Bilddateien auf ihre Übereinstimmung mit einer vorbestehenden Bildbeschreibung nebst anschließender Einstellung in einen Datensatz beschränkt. Dass durch diese Nutzung die Verwertungsmöglichkeiten der jeweils betroffenen Werke beeinträchtigt werden würden, ist nicht ersichtlich und wird auch klägerseits nicht behauptet.

Zwar mag der auf diese Weise erstellte Datensatz nachfolgend zum Trainieren künstlicher neuronaler Netze genutzt werden können und die dabei entstehenden KI-generierten Inhalte mögen in Konkurrenz zu den Werken (menschlicher) Urheber treten können. Das allein rechtfertigt es jedoch noch nicht, bereits in der Erstellung der Trainingsdatensätze eine Beeinträchtigung auch der Verwertungsrechte an Werken i.S.v. Art. 5 Abs.5 InfoSoc-RL zu erblicken. Dies hat schon allein deshalb zu gelten, weil die Berücksichtigung bloß zukünftiger, derzeit noch gar nicht im Einzelnen absehbarer technischer Entwicklungen keine rechtssichere Abgrenzung zulässiger von unzulässigen Nutzungen erlaubt (vgl. ähnlich oben (b)).

Da eine Verwendung von im Wege des Text und Data Mining gewonnenen Erkenntnissen zum Trainieren künstlicher neuronaler Netze, die dann in Konkurrenz zu Urhebern treten können, auf Grundlage der aktuellen technologischen Entwicklung im Zweifel nie ausgeschlossen werden kann, würde die Gegenauffassung in letzter Konsequenz sogar dazu nötigen, das Text und Data Mining i.S.d. § 44b UrhG letztlich in seiner Gänze zu untersagen; eine solche vollständige Außerkraftsetzung der Schrankenregelung liefe aber der gesetzgeberischen Intention offenkundig zuwider und kann daher kein tragbares Auslegungsergebnis darstellen.

(2) Die von dem Beklagten heruntergeladene Bilddatei war auch – was der Kläger im Übrigen auch nicht in Abrede stellt – rechtmäßig zugänglich i.S.d. § 44b Abs. 2 S. 1 UrhG.

„Rechtmäßig zugänglich“ in diesem Sinne ist ein Werk insbesondere dann, wenn es frei im Internet zugänglich ist (Begr. RegE BT-Drucks. 19/27426, S. 88).

Hiervon ist für das von dem Beklagten heruntergeladene Bild auszugehen. Anders als klägerseits zunächst vorgetragen, hat der Beklagte nicht das in dem in der Klageschrift zunächst formulierten Unterlassungsantrag wiedergegebene „Originalbild“ – das von der Bildagentur XXX nur bei

Erwerb einer Lizenz zur Verfügung gestellt worden wäre –, sondern eine mit einem Wasserzeichen der Bildagentur versehene Fassung des Bildes heruntergeladen. Hierbei handelte es sich ersichtlich um das auf der Agenturseite quasi zu Werbezwecken eingestellte Vorschaubild. Dieses mit dem Wasserzeichen versehene Vorschaubild wurde von der Agentur aber gerade frei zugänglich ins Internet gestellt.

b)

Allerdings spricht einiges dafür, dass vorliegend die Schrankenregelung des § 44b Abs. 2 UrhG – ohne dass dies einer abschließenden Entscheidung bedürfte – nicht eingreift, da ein wirksam erklärter Nutzungsvorbehalt im Sinne von Abs. 3 der Vorschrift vorlag; insbesondere dürfte der auf der Webseite XXX.com unstreitig erklärte Nutzungsvorbehalt wohl den Anforderungen an eine Maschinenlesbarkeit i.S.d. § 44b Abs. 3 S. 2 UrhG genügen.

(1) Es spricht manches dafür, dass der auf der Webseite der Agentur ausgesprochene Nutzungsvorbehalt durch eine hierzu berechnigte Person ausgesprochen wurde und der Kläger sich hierauf auch zum Schutz seiner eigenen Rechte berufen kann.

Nach dem Wortlaut des § 44b Abs. 3 UrhG kann „der Rechtsinhaber“ den Nutzungsvorbehalt aussprechen. Zu beachten sind also nicht allein Vorbehaltserklärungen des Urhebers selbst, sondern auch nachfolgender Rechteinhaber, seien sie Rechtsnachfolger oder Inhaber von vom Urheber abgeleiteten Rechten. Nach dem ohne Weiteres schlüssigen Vortrag des Klägers (Protokoll v. 11.07.2024 S. 3, Bl. 122 d.A.) hatte er der Bildagentur XXX weiterlizenzierbare einfache Nutzungsrechte an dem Originalbild eingeräumt. Die Bildagentur war danach selbst Rechteinhaberin an den auf ihrer Seite eingestellten Bildern und konnte daher ohne Weiteres einen Nutzungsvorbehalt nach § 44b Abs. 3 UrhG aussprechen; dass insofern dinglich wirkende Vereinbarungen im Vertragsverhältnis zwischen dem Kläger und der Bildagentur entgegengestanden hätten, ist weder ersichtlich noch geltend gemacht worden.

Der Kläger ist wohl auch berechnigt, sich auf diese Vorbehaltserklärung seiner Lizenznehmerin zu berufen. Wirtschaftlich betrachtet fand hier die Auswertung des streitgegenständlichen Originalfotos über die Agentur statt. Damit lag in der Praxis die konkrete Entscheidung, welcher Dritte die Berechnigung zu welcher Nutzung erhalten sollte, bei der Agentur; Abschlusszwang bestand für diese nicht. In einer solchen Situation spricht aus Sicht der Kammer vieles dafür, dass sich der Urheber bei der Geltendmachung bei ihm verbliebener Verbotungsrechte auf einen von seinem Lizenznehmer erklärten Vorbehalt nach § 44b Abs. 3 UrhG berufen darf.

(2) Auch der Einwand des Beklagten, dass das in den AGB der Agentur gegenüber deren Kunden ausgesprochene Nutzungsverbot für Webcrawler schon in zeitlicher Hinsicht nicht „in Bezug auf § 44b Abs. 3 UrhG“ formuliert sein könne, ist unerheblich. Für die Rechtswirkungen der Erklärung

ist es keine Voraussetzung, dass sie bewusst mit Blick auf eine bestimmte Gesetzesfassung erklärt wird.

(3) Der Vorbehalt ist auch hinreichend klar formuliert. Art. 4 Abs. 3 DSM-RL verlangt eine ausdrückliche Erklärung des Nutzungsvorbehalts. Dieses Ausdrücklichkeitserfordernis ist mithin bei richtlinienkonformer Auslegung des § 44b Abs. 3 UrhG mitzuberücksichtigen (so auch Begr. RegE BT-Drs. 19/27426, 89). Der erklärte Vorbehalt muss damit sowohl *expressis verbis* (nicht konkludent) als auch so zielgenau (konkret-individuell) erklärt werden, dass er zweifelsfrei einen bestimmten Inhalt und eine bestimmte Nutzung erfasst (*Hamann*, ZGE 16 (2024), S. 134). Diesen Anforderungen genügt der auf der Webseite der Bildagentur XXX formulierte Nutzungsvorbehalt ohne Weiteres.

Soweit darüber hinaus argumentiert wird, dass ein für sämtliche auf einer Webseite eingestellte Werke erklärter Nutzungsvorbehalt dem Ausdrücklichkeitserfordernis des § 44b Abs. 3 UrhG widerspreche (so in Erweiterung der eigenen abstrakten Herleitung *Hamann*, a.a.O., S. 148), vermag dies nicht zu überzeugen. Denn auch der explizit für sämtliche auf einer Webseite eingestellte Werke erklärte Vorbehalt ist in seiner Reichweite und seinem Inhalt nach zweifelsfrei bestimmbar und damit ausdrücklich erklärt.

(4) Schließlich dürfte auch Einiges dafür sprechen, dass der Nutzungsvorbehalt den Anforderungen an eine Maschinenlesbarkeit i.S.d. § 40b Abs. 3 S. 2 UrhG genügt.

Dabei wird man zwar den Begriff der Maschinenlesbarkeit im Hinblick auf den ihm zugrunde liegenden gesetzgeberischen Willen, eine automatisierte Abfrage durch Webcrawler zu ermöglichen (vgl. Begr. RegE BT-Drucks. 19/27426, S. 89), durchaus im Sinne einer „Maschinenverständlichkeit“ auszulegen haben (eingehend zum Meinungsstand *Hamann*, a.a.O., S. 113, 128 ff.).

Die Kammer neigt allerdings dazu, als „maschinenverständlich“ auch einen allein in „natürlicher Sprache“ verfassten Nutzungsvorbehalt anzusehen (anders als die wohl überwiegende Auffassung im Schrifttum, s. *Hamann*, a.a.O., S. 131 ff., 146 ff. m.w.N. zum Meinungsstand, wobei dort auf einen Beitrag der hiesigen Beklagtenvertreter, nämlich auf *Akinci/Heidrich*, IPRB 2023, 270, 272 verwiesen wird, die offenbar ebenfalls die Auffassung der Kammer vertreten; der Beitrag war der Kammer allerdings bis zur Urteilsabfassung nicht unmittelbar zugänglich). Allerdings wird man die Frage, ob und unter welchen konkreten Voraussetzungen ein in „natürlicher Sprache“ erklärter Vorbehalt auch als „maschinenverständlich“ angesehen werden kann, stets in Abhängigkeit von der zum jeweils relevanten Werknutzungszeitpunkt bestehenden technischen Entwicklung beantworten müssen.

Dementsprechend hat auch der europäische Gesetzgeber im Rahmen der KI-Verordnung festgelegt, dass Anbieter von KI-Modellen eine Strategie insbesondere zur Ermittlung und

Einhaltung eines gemäß Art. 4 Abs. 3 DSM-RL geltend gemachten Rechtsvorbehalts „auch durch modernste Technologien“ vorzuhalten haben (Art. 53 Abs. 1 lit. c KI-Verordnung). Zu diesen „modernsten Technologien“ gehören aber unzweideutig gerade auch KI-Anwendungen, die in der Lage sind, in natürlicher Sprache geschriebenen Text inhaltlich zu erfassen (so offenbar insbesondere auch die hiesigen Beklagtenvertreter *Akinci/Heidrich* in dem der Kammer nicht unmittelbar zugänglichen Beitrag IPRB 2023, 270, 272, hier zit. nach *Hamann*, a.a.O. S. 148, dieser im Übrigen diese Möglichkeit in technischer Hinsicht grds. bejahend, a.a.O.). Es spricht insoweit alles dafür, dass der Gesetzgeber der KI-Verordnung mit seinem Hinweis auf „modernsten Technologien“ gerade solche KI-Anwendungen im Blick hatte.

Gegen eine solche Sichtweise wird teilweise eingewandt, sie führe zu einem Zirkelschluss: Wenn gefordert werde, der Betreiber des Text und Data Mining müsse mittels KI-Anwendungen überprüfen, ob ein Nutzungsvorbehalt erklärt worden sei, dann erfordere doch diese KI-gestützte Suche ihrerseits eine Musteranalyse, die bereits den Tatbestand des Text und Data Mining i.S.d. § 44b Abs. 1 UrhG erfülle; mit anderen Worten: Erst die Anwendung der Schranke entscheide über die Zulässigkeit ihrer Anwendung (so *Hamann*, a.a.O., S. 148). Die Kammer teilt diese Bewertung nicht: Die urheberrechtlich relevante, rechtfertigungsbedürftige Nutzungshandlung ist entgegen vorgenannter Ansicht nicht die Durchführung einer „Musteranalyse“ als solche, sondern die Vervielfältigung des urheberrechtlich geschützten Werks i.S.d. § 16 UrhG. Dass das vorausgehende Auffinden solcher Werke im Netz und deren Überprüfung, ob Vorbehalte i.S.d. § 44b Abs. 3 S. 2 UrhG erklärt sind, zwingend ein quasi vorgeschaltetes weiteres Text und Data Mining i.S.v. § 44b Abs. 1 UrhG erfordere, erscheint nicht als zwingend, denn zu denken ist insbesondere an Prozessierungen des Webseiteninhalts durch den Einsatz von Webcrawlern, bei denen lediglich flüchtige und beiläufige Vervielfältigungen entstehen, die ihrerseits bereits unter § 44a UrhG gerechtfertigt sind.

Ferner wird gegen das von der Kammer erwogene, weitere Verständnis des Begriffes der „Maschinenlesbarkeit“ auch eingewandt, dieser Begriff werde vom europäischen Gesetzgeber in anderem Zusammenhang enger verstanden. Verwiesen wird in diesem Zusammenhang auf den Erwägungsgrund 35 der PSI-Richtlinie (RL (EU) 2019/1024), der für eine „Maschinenlesbarkeit“ im Sinne dieser Richtlinie u.a. eine „einfache“ Erkennbarkeit verlange (so BeckOK UrhR/Bomhard, 42. Ed. 15.2.2024, UrhG § 44b Rn. 31 m.w.N.); dies könne für einen nur in natürlicher Sprache formulierten Vorbehalt nicht angenommen werden. Eine solche Argumentation setzt allerdings voraus, dass die Begrifflichkeiten beider Richtlinien in gleicher Weise verstanden werden müssen. Die Kammer hat Zweifel, ob eine solche Gleichsetzung der Begrifflichkeiten überzeugen kann, denn die Richtlinien haben unterschiedliche Zielrichtungen: Während die PSI-Richtlinie den rein einseitigen Zugang der Öffentlichkeit zu Informationen bzw. die rein einseitige Verpflichtung der öffentlichen Hand zur Veröffentlichung bestimmter Informationen zum Gegenstand hat, geht es im Rahmen von Art. 4 Abs. 3 DSM-RL um einen Ausgleich zwischen den Interessen der Nutzer

des Text und Data Mining (dieses möglichst einfach und möglichst rechtssicher betreiben zu können) und den Interessen der Rechteinhaber (ihre Rechte möglichst einfach und möglichst effektiv zu sichern). Dieser Interessenausgleich kann nach Auffassung der Kammer nicht einseitig zugunsten der Nutzer des Text und Data Mining gelöst werden, indem allein die für diese denkbar einfachste technische Lösung als ausreichend für die Wirksamkeit eines ausgesprochenen Nutzungsvorbehalts erachtet wird. Gegen ein solches Verständnis spräche auch die Wertung des Gesetzgebers der DSM-RL, der im Erwägungsgrund 18 gerade nicht die Erklärung eines Vorbehalts „in möglichst einfach auszulesender Weise“ fordert, sondern lediglich „in angemessener Weise“. Und auch der deutsche Umsetzungsgesetzgeber verlangt lediglich eine Erklärung in einer Weise, die „den automatisierten Abläufen beim Text und Data Mining angemessen“ ist (Begr. RegE BT-Drucks. 19/27426, S. 89).

Es wäre zudem aus Sicht der Kammer ein gewisser Wertungswiderspruch, den Anbietern von KI-Modellen einerseits über die Schranke in § 44b Abs. 2 UrhG die Entwicklung immer leistungsfähigerer textverstehender und -kreierender KI-Modelle zu ermöglichen, ihnen aber andererseits im Rahmen der Schranken-Schranke von § 44b Abs. 3 S. 2 UrhG die Anwendung bereits bestehender KI-Modelle nicht abzuverlangen.

Ob und in welchem Maße zum Zeitpunkt der streitgegenständlichen Vervielfältigungshandlung im Jahr 2021 eine ausreichende Technologie zur automatisierten inhaltlichen Erfassung des streitgegenständlichen Nutzungsvorbehalts bereits zur Verfügung stand, ist zwar klägerseits bislang nicht dargetan; der Kläger hat insoweit nur auf im Jahr 2023 verfügbare Dienste verwiesen (Replik S. 14 ff., Bl. 48 ff. d.A.). Allerdings bestehen Anzeichen dafür, dass der Beklagte bereits über geeignete Technologie verfügte. Denn nach eigenem Vortrag des Beklagten erforderte die im Rahmen der Erstellung des Datensatzes XXX durchgeführte Analyse in Form eines Abgleichs von Bildinhalten mit vorbestehenden Bildbeschreibungen ersichtlich auch und gerade eine inhaltliche Erfassung dieser Bildbeschreibungen durch die eingesetzte Software. Vor diesem Hintergrund spricht einiges dafür, dass – insbesondere auch dem Beklagten – bereits im Jahr 2021 Systeme zur Verfügung standen, die in der Lage waren, einen in natürlicher Sprache formulierten Nutzungsvorbehalt in automatisierter Weise zu erfassen.

3.

Der Beklagte kann sich hinsichtlich der streitgegenständlichen Vervielfältigung jedoch auf die Schrankenregelung des § 60d UrhG berufen.

Danach sind Vervielfältigungen für Text und Data Mining für Zwecke der wissenschaftlichen Forschung durch Forschungsorganisationen zulässig.

a)

Die Vervielfältigung erfolgte – wie dargelegt – für den Zweck des Text und Data Mining i.S.d. § 44b Abs. 1 UrhG. Sie erfolgte darüber hinaus auch zu Zwecken der wissenschaftlichen Forschung i.S.d. § 60d Abs. 1 UrhG.

Wissenschaftliche Forschung bezeichnet allgemein das methodisch-systematische Streben nach neuen Erkenntnissen (Spindler/Schuster/Anton, 4. Aufl. 2019, UrhG § 60c Rn. 3; BeckOK UrhR/Grübler, 42. Ed. 1.5.2024, UrhG § 60c Rn. 5; Dreier/Schulze/Dreier, 7. Aufl. 2022, UrhG § 60c Rn. 1). Der Begriff der wissenschaftlichen Forschung ist, indem er bereits das methodisch-systematische „Streben“ nach neuen Erkenntnissen ausreicht, nicht so eng zu verstehen, dass er nur die unmittelbar mit der Gewinnung von Erkenntnisgewinn verbundenen Arbeitsschritte erfassen würde; vielmehr genügt es, dass der in Rede stehende Arbeitsschritt auf einen (späteren) Erkenntnisgewinn gerichtet ist, wie es z.B. bei zahlreichen Datensammlungen der Fall ist, die zunächst durchgeführt werden müssen, um anschließend empirische Schlussfolgerungen zu ziehen. Insbesondere setzt der Begriff der wissenschaftlichen Forschung auch keinen späteren Forschungserfolg voraus.

Danach kann – entgegen der Auffassung des Klägers – auch bereits die Erstellung eines Datensatzes der streitgegenständlichen Art, der Grundlage für das Trainieren von KI-Systemen sein kann, durchaus als wissenschaftliche Forschung im vorstehenden Sinne anzusehen sein. Zwar mag die Erstellung des Datensatzes als solche noch nicht mit einem Erkenntnisgewinn verbunden sein; sie ist aber grundlegender Arbeitsschritt mit dem Ziel, den Datensatz zum Zwecke späteren Erkenntnisgewinns einzusetzen. Dass eine solche Zielsetzung auch im vorliegenden Fall bestand, kann bejaht werden. Dafür genügt es, dass der Datensatz – unstrittig – kostenfrei veröffentlicht und damit gerade (auch) auf dem Gebiet künstlicher neuronaler Netze Forschenden zur Verfügung gestellt wurde. Ob der Datensatz – wie es der Kläger hinsichtlich der Dienste XXX und XXX behauptet – daneben auch von kommerziellen Unternehmen zum Training bzw. zur Weiterentwicklung ihrer KI-Systeme genutzt wird, ist schon deshalb unerheblich, weil auch die Forschung kommerzieller Unternehmen noch Forschung – wenn auch nicht als solche nach §§ 60c f. UrhG privilegiert – ist.

Auf die zwischen den Parteien streitige Frage, ob der Beklagte über die Erstellung entsprechender Datensätze hinaus auch wissenschaftliche Forschung in Gestalt der Entwicklung eigener KI-Modelle tätigt, kommt es vor diesem Hintergrund nicht an.

b)

Der Beklagte verfolgt auch nicht kommerzielle Zwecke i.S.d. § 60d Abs. 2 Nr. 1 UrhG.

Für die Frage, ob Forschung nicht kommerziell ist, kommt es allein auf die konkrete Art der wissenschaftlichen Tätigkeit an, während Organisation und Finanzierung der Einrichtung, in der die Forschung erfolgt, unbeachtlich sind (ErwGr 42 InfoSoc-RI).

Die nicht-kommerzielle Zweckverfolgung des Beklagten in Bezug auf die streitgegenständliche Erstellung des Datensatzes XXX ergibt sich dabei bereits daraus, dass der Beklagte diesen unstreitig kostenfrei öffentlich zur Verfügung stellt. Dass die Entwicklung des streitgegenständlichen Datensatzes darüber hinaus zumindest auch der Entwicklung eines eigenen kommerziellen Angebots des Beklagten dienen würde (vgl. zu diesem Kriterium BeckOK IT-Recht/Paul, 14. Ed. 1.4.2024, UrhG § 60d Rn. 10), ist weder klägerseits vorgetragen noch sonst ersichtlich. Dass der streitgegenständliche Datensatz auch von kommerziell tätigen Unternehmen zum Training bzw. zur Weiterentwicklung ihrer KI-Systeme genutzt werden mag, ist für die Einordnung der Tätigkeit des Beklagten hingegen ohne Relevanz. Allein der Umstand, dass einzelne Mitglieder des Beklagten neben ihrer Tätigkeit für den Verein auch bezahlten Tätigkeiten bei solchen Unternehmen nachgehen, genügt nicht, die Tätigkeit dieser Unternehmen dem Beklagten als eigene zuzurechnen.

c)

Dem Beklagten ist eine Berufung auf die Schrankenregelung des § 60d UrhG auch nicht nach Abs. 2 S. 3 der Vorschrift verwehrt.

Danach können sich Forschungsorganisationen, die mit einem privaten Unternehmen zusammenarbeiten, das einen bestimmenden Einfluss auf die Forschungsorganisation und einen bevorzugten Zugang zu den Ergebnissen der wissenschaftlichen Forschung hat, nicht auf die Schrankenregelung des § 60d UrhG berufen. Die Darlegungs- und Beweislast für die tatsächlichen Voraussetzungen des Gegenausschlusses nach § 60d Abs. 2 S. 3 UrhG obliegt dabei nach dem Wortlaut der Norm dem Kläger.

(1) Soweit der Kläger mit der Replik zunächst darauf verwiesen hat, dass das Unternehmen XXX über die Finanzierung des gegenständlichen Datensatzes und die Besetzung „relevanter Posten“ beim Beklagten durch eigene Mitarbeiter unmittelbaren Einfluss auf den Beklagten habe (Replik S. 18, Bl. 52 d.A.), so entbehrt dieser Vortrag der Substanz.

Der Kläger verweist insoweit lediglich darauf, dass einer der Mitbegründer des Beklagten, Herr XXX, bei XXX als „Head of Machine Learning Operations“ angestellt sei, ferner ein Mitglied des Beklagten, Herr XXX, ebendort als „Research Scientist“ (Replik S. 4 f., Bl. 38 f. d.A.). Allein diese Tätigkeit von zwei Vereinsmitgliedern für das Unternehmen XXX belegt aber keinen bestimmenden Einfluss dieses Unternehmens auf die Forschungsarbeit des Beklagten.

Davon abgesehen hat der Kläger noch nicht einmal behauptet, dass der Beklagte dem Unternehmen XXX auch bevorzugten Zugang zu den Ergebnissen seiner wissenschaftlichen Forschung, namentlich dem streitgegenständlichen Datensatz, gewährt habe. Vielmehr wird insoweit nur vorgetragen, dass XXX seinen Dienst XXX mithilfe des streitgegenständlichen Datensatzes trainiert habe (Replik S. 8 f., Bl. 42 f. d.A.).

(2) Soweit der Kläger mit Schriftsatz vom 03.07.2024 auf einen im Jahr 2021 erfolgten Chat auf der Plattform XXX verweist, wonach sich der Mitbegründer des Beklagten, Herr XXX, dazu bereit erklärt haben soll, dem Unternehmen XXX aufgrund eines von diesem geleisteten Finanzierungsbeitrags von 5.000,- \$ einen vorzeitigen Zugang zu dem (damaligen kleineren) Datensatz zu gewähren, erfüllt auch dieser Vortrag nicht den Ausnahmetatbestand in § 60d Abs. 2 S. 3 UrhG.

Dabei kann dahinstehen, ob dieser – von dem Beklagten als solcher nicht bestrittene (vgl. Schriftsatz vom 09.07.2024 S. 3, Bl. 112 d.A.) – Chatverlauf die vom Kläger gezogene Interpretation überhaupt trägt. Gleichfalls kann dahinstehen, ob die Erklärung einer solchen Bereitschaft zur Gewährung eines vorzeitigen Zugangs – ob dieser tatsächlich gewährt wurde, hat der Kläger nicht vorgetragen – für das Innehaben eines bevorzugten Zugangs zu den Forschungsergebnissen i.S.d. § 60d Abs. 2 S. 2 UrhG ausreichen kann.

Denn es ist jedenfalls weder klägerseits dargetan noch sonst ersichtlich, dass das Unternehmen XXX einen bestimmenden Einfluss auf den Beklagten hätte. Soweit überhaupt personelle Verflechtungen zwischen dem Beklagten und Unternehmen der KI-Branche dargetan sind, handelt es sich um die Unternehmen XXX und XXX (Replik S. 4 ff., Bl. 38 ff. d.A.).

II.

Die Kostenentscheidung beruht auf § 91 Abs. 1, § 91a Abs. 1 S. 1 ZPO.

Die Entscheidung zur vorläufigen Vollstreckbarkeit ergeht nach § 708 Nr. 11, §§ 711, 709 ZPO.

Hartmann
Vorsitzender Richter
am Landgericht

Dr. Kanzler
Richterin

Dr. Woodtli
Richter
am Landgericht