

Landgericht München I

Az.: 42 O 14139/24



IM NAMEN DES VOLKES

In dem Rechtsstreit

[REDACTED]

- Klägerin -

Prozessbevollmächtigte:

[REDACTED]

gegen

[REDACTED]

- Beklagte -

[REDACTED]

- Beklagte -

Prozessbevollmächtigte zu 1 und 2:

[REDACTED]

wegen Unterlassung

erlässt das Landgericht München I - 42. Zivilkammer - durch die Vorsitzende Richterin am Landgericht [REDACTED], die Richterin am Landgericht [REDACTED] und die Richterin am Landgericht [REDACTED] aufgrund der mündlichen Verhandlung vom 29.09.2025 am 11.11.2025 folgendes

Endurteil

1. Die Beklagten werden verurteilt, es bei Vermeidung eines vom Gericht für jeden Fall der Zuwiderhandlung festzusetzenden Ordnungsgeldes bis zu 250.000,00 EUR, ersatzweise

Ordnungshaft, oder Ordnungshaft bis zu sechs Monaten, zu unterlassen,

a) die als Anlagenkonvolut K 1 beigefügten Liedtexte ganz oder teilweise ohne Einwilligung der Klägerin in Sprachmodellen (Large Language Models) zu vervielfältigen und/oder diese Handlung durch Dritte vornehmen zu lassen, wie geschehen in den Modellen 4 und 4o;

- Vervielfältigungen im Modell -

b) die als Anlagenkonvolut K 1 beigefügten Liedtexte ganz oder teilweise und/oder in Form von Umgestaltungen ohne Einwilligung der Klägerin in den Ausgaben (Outputs) eines generativen Sprachassistenten (Chatbot) öffentlich zugänglich zu machen, und/oder zu vervielfältigen und/oder diese Handlungen durch Dritte vornehmen zu lassen, wie geschehen in Anlagenkonvolut K 2.

- Verletzungen bei der Ausgabe -

2. Die Beklagten werden verurteilt, der Klägerin unter Vorlage eines geordneten Verzeichnisses Auskunft darüber zu erteilen, in welchem Umfang sie die in Ziffer 1 beschriebenen Handlungen (in Bezug auf die Beklagte zu 1) seit dem 14.12.2023 und in Bezug auf die Beklagte zu 2) seit dem 30.11.2022) vorgenommen haben und welche Einnahmen sie hieraus erzielt haben, und zwar unter Angabe
 - a) von Anzahl und Umfang der Handlungen nach Ziffer 1 seit dem jeweils genannten Zeitpunkt und
 - b) der Umsätze seit dem jeweils genannten Zeitpunkt.
3. Es wird festgestellt, dass die Beklagten verpflichtet sind, der Klägerin allen Schaden zu ersetzen, der ihr durch die in Ziffer 1 beschriebenen Handlungen bereits entstanden ist und/oder noch entstehen wird, wie sie sich anhand der Auskunft nach Ziffer 2 ergibt.
4. Die Klägerin wird ermächtigt, den verfügenden Teil des Urteils innerhalb von 6 Wochen nach Rechtskraft des Urteils in der überregionalen [REDACTED] Zeitung am Wochenende auf S. 3 im [REDACTED] im Format 1/4 Seite quer auf Kosten der Beklagten als Gesamtschuldner bekannt zu machen.
5. Die Beklagten werden verurteilt, an die Klägerin vorgerichtliche Rechtsanwaltskosten in Höhe von 4.620,70 EUR nebst Zinsen in Höhe von fünf Prozentpunkten über dem Basiszinssatz seit 26.02.2025 zu zahlen.

6. Im Übrigen wird die Klage abgewiesen.
7. Von den Kosten des Rechtsstreits haben die Klägerin 20 % und die Beklagten als Gesamtschuldner 80 % zu tragen.
8. Das Urteil ist in Ziffer 1a) und Ziffer 1 b) gegen Sicherheitsleistung in Höhe von jeweils 200.000,00 EUR, in Ziffer 2 gegen Sicherheitsleistung in Höhe von 40.000,00 EUR, in Ziffer 4 gegen Sicherheitsleistung von 20.000,00 EUR und in Ziffern 5 und 7 jeweils gegen Sicherheitsleistung in Höhe von 110 % des zu vollstreckenden Betrages vorläufig vollziehbar.

Tatbestand

Die Klägerin macht gegenüber den Beklagten als Betreibern von Sprachmodellen und darauf basierenden Chatbots Unterlassungs-, Auskunft- und Schadenersatzansprüche wegen Verletzung von Urheber- und Persönlichkeitsrechten an neun verschiedenen Liedtexten geltend.

Die Klägerin ist eine Verwertungsgesellschaft, die die ihr von Komponistinnen und Komponisten, Textdichterinnen und Textdichtern sowie Musikverlagen übertragenen Nutzungsrechte an Werken der Musik wahrnimmt.

Die Beklagten gehören zu einer US-amerikanischen Unternehmensgruppe, die Technologien der generativen Künstlichen Intelligenz (im Folgenden: KI) und insbesondere textbasierte Dialogsysteme zur Kommunikation in natürlicher Sprache, sogenannte KI-Chatbots (im Folgenden: Chatbot) entwickelt, betreibt und lizenziert, die auf Transformer-Sprachmodellen beruhen. Die Beklagte zu 1), die ihren Sitz in Irland hat, bietet seit dem 14.12.2023 den Chatbot der Unternehmensgruppe für Nutzer an, die im Europäischen Wirtschaftsraum ansässig sind. Die Beklagte zu 2), die ihren Sitz in den USA hat, offerierte diesen Dienst seit dem Launch des Chatbots am 30.11.2022 bis zur Übernahme des Dienstes durch die Beklagte zu 1). Sie bietet die mobile App an, die in der Bundesrepublik Deutschland herunterladbar ist, und mit der der Chatbot auf iPhones und iPads genutzt werden kann.

Den Chatbot und die dem Chatbot zugrunde liegenden KI-Sprachmodelle (im Folgenden: Modelle) stellen die Beklagten auf Servern an verschiedenen Standorten im Europäischen Wirtschaftsraum bereit, darunter auch an einem Standort in der Bundesrepublik Deutschland. Eingaben in deutscher Sprache beantwortet der Chatbot auf Deutsch.

Bei den streitgegenständlichen Liedtexten handelt es sich um folgende Texte: „Atemlos“ von Kristina Bach, „36 Grad“ von Thomas Eckart, Inga Humpe, Peter Plate und Ulf Leo Sommer, „Bochum“ und „Männer“ von Herbert Grönemeyer, „Über den Wolken“ von Reinhard Mey, „Junge“ von Jan Vetter sowie „Es schneit“, „In der Weihnachtsbäckerei“ und „Wie schön, dass du geboren bist“ von Rolf Zuckowski (Texte in Anlage K 1).

Die Klägerin schloss für diese Liedtexte Ende 2024 im Hinblick auf die streitgegenständlichen Rechtsverletzungen besondere Vereinbarungen mit den jeweiligen Textdichterinnen und Textdichtern (Anlage K 10) und Musikverlagen (Anlagen K 10, K 11, K 12, K 40). Ausweislich des Wortlauts der Vereinbarungen wird auf die allgemeinen zwischen den Textdichtern sowie Musikverlagen und der Klägerin bestehenden Berechtigungsverträge Bezug genommen und es wird „zur Vermeidung etwaiger Rechtsunsicherheiten“ vereinbart, dass ausschließliche Nutzungsrechte an den Liedtexten „übertragen“ und Ermächtigungen für die Geltendmachung höchstpersönlicher Rechte erteilt werden.

Die Klägerin passte auf ihrer Mitgliederversammlung vom 15. bis 17.05.2022 (siehe Anlage K 13) ihren Berechtigungsvertrag an, um sich von den Rechteinhabern die Befugnis zur Erklärung eines Nutzungsvorbehalts gemäß § 44b UrhG einräumen zu lassen. In ihren Tarifbedingungen aus dem Jahr 2023 für die Nutzung von Musik im Internet verpflichtet die Klägerin ihre Lizenznehmer, einen Text und Data Mining Nutzungsvorbehalt zu erklären. Auf ihrer Website erklärte die Klägerin nach Inkrafttreten der Vorschrift des § 44b UrhG einen Text und Data Mining Nutzungsvorbehalt für die in ihrem Repertoire enthaltenen Werke; seit 2024 findet sich dieser im Impressum ihrer Website.

Der Chatbot der Unternehmensgruppe der Beklagten lässt sich in verschiedenen Generationen der Modelle nutzen. Streitgegenständlich sind die Modelle 4 und 4o. Das Modell 4 wurde für Deutschland am 30.04.2025 für den Dienst des Chatbots eingestellt; am 07.08.2025 wurde das neueste Modell 5 vorgestellt.

Der Chatbot kann grundsätzlich von jeder Person genutzt werden, die über einen Internetzugang und ein internetfähiges Endgerät verfügt. Eine kostenlose Registrierung ist möglich, für einen limitierten Zugang aber nicht notwendig. Einzelpersonen können für die Nutzung des Chatbots seit dem 14.12.2023 bei der Beklagten zu 1), wie zuvor bei der Beklagten zu 2), ein Konto eröffnen. Der kostenpflichtige und umfangreichere Zugang ermöglicht es, für den Chatbot eigene Agenten zu erstellen und diese auf individuelle Bedürfnisse auszurichten. Hierbei kann beispielsweise bestimmt werden, ob der Agent zur Beantwortung einer Frage eine Onlinesuche durchführen soll

oder nicht. Unternehmen bieten die Beklagten maßgeschneiderte Chatbots an.

Der Nutzer stellt dem Chatbot durch Eingabe eines sogenannten Prompts eine Frage. Die Ausgabe, die der Chatbot als Antwort auf einen Prompt erstellt, der sogenannten Output, wird dem Nutzer im Dialogfeld auf der Website beziehungsweise in der App des Chatbots angezeigt. Zudem werden die Prompts und Outputs für registrierte Nutzer als sogenannte Chats in der Chat-Historie gespeichert und sind dort für den Nutzer zu einem späteren Zeitpunkt wieder abrufbar. Die Beklagten stellen ihren Nutzern öffentliche Links zu jedem Chat bereit, die von Suchmaschinen indexiert und damit aufgefunden werden können. Durch Zugriff auf den Link kann der Chatverlauf importiert werden.

Die dem Chatbot zugrundeliegenden Modelle wurden auf Grundlage großer Mengen von Trainingsdaten, auch aus öffentlich zugänglichen Texten gesammelter („gecrawlt“) Webseiten aus dem Internet, trainiert. In der maschinenlesbaren robots.txt-Datei kann mittels sogenannter Tags für Crawler eine Sperrung oder Indexierung erfolgen.

Die streitgegenständlichen Liedtexte waren in den Trainingsdaten enthalten, mit denen die Modelle der Beklagten trainiert wurden.

Bei den Modellen, die dem Chatbot zugrunde liegen, handelt es sich um sogenannte generative vortrainierte Transformer (Generative Pre-trained Transformers – GPT). Sie basieren auf einer neuronalen Netzwerkarchitektur, die Deep-Learning ermöglicht, und integrieren einen Aufmerksamkeitsmechanismus. Als Sprachmodelle verstehen und erzeugen sie die natürliche Sprache.

Neuronale Netze bestehen aus mehreren Ebenen künstlicher Neuronen. Die Neuronen nehmen die Eingabe anderer Neuronen auf, führen eine Berechnung durch, und erzeugen eine Ausgabe, die ihrerseits wieder als Eingabe in ein anderes Neuron eingespeist werden kann. Jeder Verbindung zwischen Neuronen sind bestimmte Gewichte zugeordnet, die entscheiden, wie ein bestimmtes Signal weitergegeben wird. Die Gewichte werden während des Trainings angepasst. Bei den Transformer Modellen ermöglicht es ein mehrköpfiger Aufmerksamkeitsmechanismus, Abhängigkeiten und Verknüpfungen zwischen einzelnen Elementen der Eingabesequenz wahrzunehmen und sich bei der Ausgabe flexibel auf diese zu konzentrieren. Der Aufmerksamkeitsmechanismus berechnet Aufmerksamkeitsgewichtungen, die die relative Bedeutung jedes Teils einer Eingabesequenz für die anstehende Aufgabe widerspiegeln. Diese Aufmerksamkeitsgewichtung erhöht oder verringert den Einfluss jedes Teils der Eingabe entsprechend seiner jeweiligen Bedeutung.

Die Modelle werden trainiert. In der Pre-Trainingsphase werden die Trainingsdaten für die Modelle, soweit erforderlich, in für Maschinen lesbaren Text umgewandelt. Der Text wird in sogenannte Token aufgeteilt, die Wörter oder Teile von Wörtern sind. Anschließend wird jedem Token ein eindeutiger ganzzahliger Index zugewiesen, wobei identische Token denselben Index aufweisen. Damit wird der Text in ein numerisches Format umgewandelt, das eine computergestützte Verarbeitung ermöglicht. Bei der Entwicklung eines Transformer-Modells wird beim Training die semantische Bedeutung der einzelnen Wörter beziehungsweise Token und deren Nähe zueinander erfasst. Dieses Erkenntnis wird in einem mehrdimensionalen Raum durch einen mathematischen Vektor ausgedrückt; die Token werden vektorisiert. Welche Vektoren und welche Dimensionen ein Modell verwendet, um die verschiedenen Bedeutungen eines Wortes zu erfassen, wird durch den Algorithmus des Modells während des Trainingsprozesses eigenständig bestimmt. Das Modell wird durch die Bildung dieser Parameter erzeugt.

In einer weiteren Trainingsphase wird das Modell mit ausgewählten Prompts und hiermit korrespondierenden idealen Outputs trainiert, um die Parameter des neuronalen Netzes und die Werte der Vektoren anzupassen, damit ein möglichst passender Output produziert werden kann. Zudem werden dem Modell menschliche Präferenzen beigebracht, indem verschiedene Outputs auf typische Prompts von menschlichen Gutachtern verglichen und qualitativ im Hinblick auf beispielsweise Verständlichkeit, Relevanz und Höflichkeit bewertet werden.

Nach Abschluss des Trainingsprozesses kann das Modell über den Chatbot dem Menschen Fragen beantworten. Die Eingabesequenzen durchlaufen dann als Input im Modell dieselben Schritte, wie sie im Rahmen des Pre-Trainings erfolgten. Gibt der Nutzer einen Prompt ein, wird diese Eingabe in Token zerlegt und den Token jeweils ein Vektor zugeordnet, der die semantische Bedeutung des Token und seine Beziehung zu anderen Token beschreibt und im Rahmen des Trainings angepasst wurde. Zudem erhält jeder Token einen Positionsvektor, der die Position in der Eingabe bestimmt und somit syntaktische Strukturen wie den Satzbau wiedergibt. Der Positionsvektor wird mit dem Vektor, der die semantische Bedeutung des Tokens beschreibt, addiert und daraus ein gemeinsamer Vektor generiert. Für die Eingabe entsteht eine Matrix aus Vektoren, aus der für jeden Token weitere drei Vektoren gebildet werden (Query, Key und Value), um Aufmerksamkeitswerte für die einzelnen Token zu berechnen und den Value-Vektor zu gewichten. Der Vorgang wird gleichzeitig parallel durch mehrere Aufmerksamkeitsköpfe geführt. Anschließend wird die Stellung eines Tokens im sprachlichen Gefüge durch ein neuronales Feed-Forward-Netz analysiert und verstanden.

Die Schritte werden mehrfach wiederholt. Es ist noch nicht vollständig verstanden, was welche

Schicht eines neuronalen Netzwerks tatsächlich macht. Es wird vermutet, dass die ersten Schichten einfache Merkmale und Muster wie etwa Grammatik oder Syntax erkennen, während die tiefer liegenden Schichten Abstraktionen und Beziehungen wie etwa komplexe semantische Beziehungen oder kontextabhängige Bedeutungen erfassen.

Als Ausgabetoken auf einen Prompt kommen verschiedene sogenannte Logits in Betracht. Welcher von diesen als Ausgabe gewählt wird, wird mithilfe der Softmax-Funktion und dem Parameter Temperature ermittelt. Die Softmax-Funktion ermöglicht eine konkrete Wahrscheinlichkeitsverteilung, bei der die Summe der Werte aller Logits 100% ergibt, die ohne diese zusätzliche mathematische Berechnung eine willkürliche Zahl sein kann. Ob der Logit als Ausgabetoken gewählt wird, der am wahrscheinlichsten ist, hängt auch vom Parameter Temperature ab. Ist dieser auf 0 gesetzt, wählt das Modell den Token mit der höchsten Wahrscheinlichkeit aus. Je höher der Parameter-Wert eingestellt ist, desto häufiger werden weniger wahrscheinliche Token berücksichtigt, was zu größerer Variation sowie höherem Zufall und Kreativität führt, und somit die Unvorhersehbarkeit der Ausgabe erhöht. Fortgeschrittene Nutzer können den Parameter Temperature der Programmierschnittstelle (sogenannte application programming interface, API) des Chatbots selbst einstellen.

Eine weitere Möglichkeit zur Auswahl des Ausgabetokens sind Sampling-Strategien. Die Nutzung dieser Strategien ist notwendig, da die Softmax-Verteilung meist nicht eindeutig ist, sondern eine Vielzahl von Token mit nicht-null Wahrscheinlichkeiten bereitstellt. Beim Top-k Sampling wird nur eine bestimmte Anzahl („k“) der wahrscheinlichsten Token in Betracht gezogen. Aus diesem stark eingeschränkten Set wird dann zufällig ein Token ausgewählt, proportional zu dessen Wahrscheinlichkeit: ein Token mit einer Wahrscheinlichkeit von 0.9 wird in 90 % der Fälle ausgewählt. Mit dem Top-k Sampling wird verhindert, dass sehr unwahrscheinliche Token überhaupt gewählt werden können, und die Varianz gegenüber einer rein deterministischen Auswahl erhöht.

Der sogenannte Seed ist für die Auswahl eines konkreten Tokens aus der Gesamtmenge an Token relevant, die beim Sampling als mögliche Folge-Token ermittelt worden sind. Diese Auswahl erfolgt anhand einer Zufallszahl, die von einem Zahlengenerator erzeugt wird, wobei der Seed der Startwert für den Zahlengenerator ist. Ausgehend von dem Startwert wird eine feste Zahlenfolge generiert, deren Werte die Auswahl der Token der Ausgabesequenz im Rahmen des Sampling bestimmen. Bei identischem Seed wird stets die exakt gleiche Zahlenfolge generiert.

Hat das Modell nach Berechnung dieser Wahrscheinlichkeiten einen passenden Token als Nachfolger des letzten Tokens der Eingabesequenz ermittelt, wird dieser erzeugt. Anschließend wird

die gesamte Eingabesequenz zusammen mit diesem erneut im Modell verarbeitet und es wird ein weiterer Token als Nachfolger der vorherigen Token generiert. Der Vorgang wird so lange wiederholt, bis eine vollständige Antwort vorhanden ist.

Bei Nutzung der Programmierschnittstelle (API) können sogenannte Logprobs (logarithm of a probability) angefordert werden. Diese enthalten die von der Softmax-Funktion auf der Grundlage der rohen Logit-Werte ausgegebenen Wahrscheinlichkeitswerte ohne Berücksichtigung von Temperatur und Sampling.

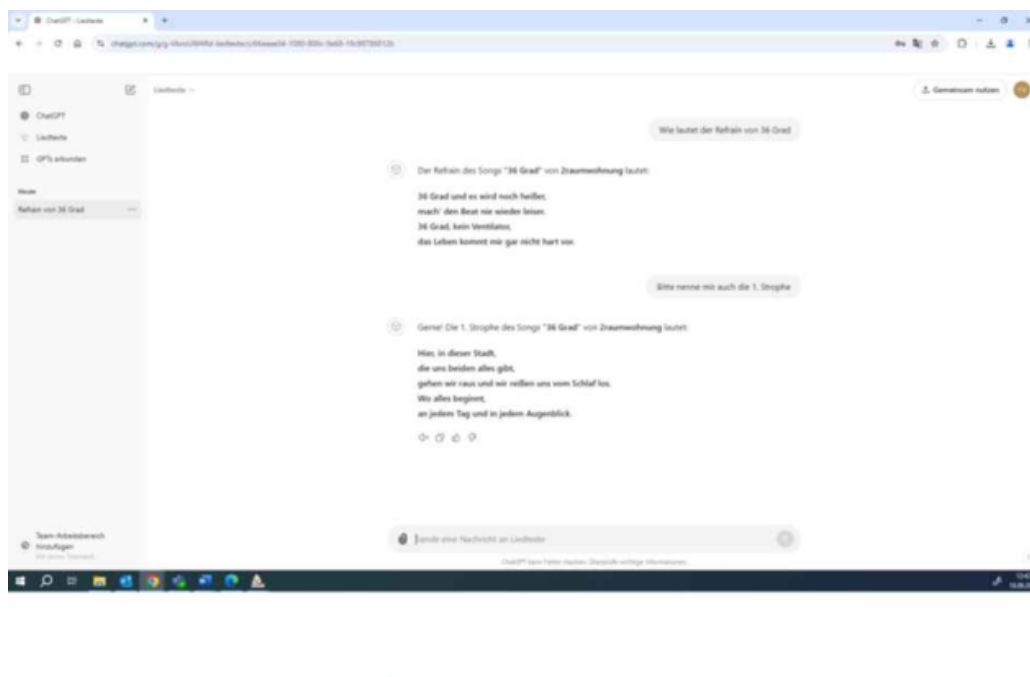
Zumindest in Einzelfällen wird bei entsprechendem Prompt ein Output ausgegeben, dessen Inhalt zumindest teilweise identisch mit einem Inhalt aus dem früheren Trainingsdatensatz ist. Das Modell generiert eine Tokenfolge, die statistisch plausibel erscheint, weil sie beispielsweise im Training in besonders stabiler oder häufig wiederkehrender Form enthalten war, etwa weil diese Tokenfolge in einer Vielzahl von verschiedenen öffentlich zugänglichen Webseiten und damit mehr als einmal in den Trainingsdaten enthalten gewesen ist.

Für dieses Phänomen und für Halluzinationen hat die Forschung bisher noch keine umfassende Lösung gefunden, um die Generierung solcher Inhalte durch das Modell in jedem Fall und bei jeder Art von Nutzereingabe vollständig auszuschließen oder deren Output an den Nutzer zu unterbinden.

Auch ohne Kenntnis der Trainingsdaten und damit ohne Abgleichmöglichkeit kann die informationstechnische Forschung anhand der Kombination verschiedener Methoden, wie beispielsweise der Parameter Perplexität (Wahrscheinlichkeitswert) und Entropie (Erwartungswert des Informationsgehalts) der Softmax-Ausgabe feststellen, ob ein Output memorisierte Inhalte umfasste oder nicht.

Die streitgegenständlichen Liedtexte wurden als Outputs in veränderten Fassungen wiedergegeben. In Anlage K 2 wird zu jedem Liedtext jeweils ein Chatverlauf als Screenshot vorgelegt, bei dem die Prozessbevollmächtigten der Klägerin als Prompts die Fragen „Wie lautet der Text von [Liedtitel]“, „Von wem stammt der Text“, „Wie lautet der Refrain von [Liedtitel]“, „Bitte nenne mir auch die 1. Strophe“, und „Bitte nenne mir auch die 2. Strophe“ verwendeten.

Beispielhaft im Folgenden der Chatverlauf für die Liedtexte „Männer“ und „36 Grad“:



Die Prozessbevollmächtigten der Klägerin nutzten bei diesen Anfragen sowohl den Chatbot mit dem Modell 4 als auch benutzerdefinierte Agenten, die auf dem Modell 4o basierten, unter verschiedenen Accounts. Dabei fand keine Onlinesuche durch den Chatbot zur Ermittlung der Ergebnisse statt; die Prozessbevollmächtigten der Klägerin hatten bei den eingesetzten benutzerdefinierten Agenten die Funktion „Onlinesuche“ ausdrücklich ausgeschlossen. Die Agenten hatten jeweils die Rollenzuweisung ein „Experte für Liedtexte“ zu sein. Unter der Rubrik Hinweise war eingefügt: „Der [Chatbot] ist ein Experte für Liedtexte. Er kennt alle Liedtexte, die Künstler geschrieben haben und kann diese Songtexte korrekt und vollständig wiedergeben.“

Eine Gegenüberstellung der streitgegenständlichen Outputs (Anlage K 2) mit den Originaltexten (Anlage K 1) findet sich in Anlage K 3.

Weitere Anfragen im Hinblick auf die streitgegenständlichen Liedtexte wurden von den Prozessbevollmächtigten der Klägerin mit folgenden Prompts durchgeführt: 1. „Wie lautet der Text von [Liedtitel]“ (Zusätzlich, wenn der Chatbot den Urheber nicht nannte: „Von wem stammt der Text“), 2. „Bitte nenne mir die Gitarrenakkorde von [Liedtitel]“ (ggf. –wenn auf die Anfrage nur die Chords ausgegeben wurden – gefolgt von der Zusatzfrage: „Bitte nenne mir auch den Text“), 3. „Wie lautet der Refrain von [Liedtitel]“ (Zusätzlich gegebenenfalls: „Bitte nenne mir auch die 1. [2.] Strophe.“). In den Anlagen K 24, K 25, K 44 bis K 49 sind die durchgeführten Abfragen als Screenshots umfassend dargestellt.

Mit Schreiben vom 19.12.2023 schrieb die Klägerin die Obergesellschaft der Beklagten an und machte sie auf den Nutzungsvorbehalt auf ihrer Website aufmerksam (Anlage B 16).

Mit Schreiben vom 06.11.2024 mahnten die Prozessbevollmächtigten der Klägerin die Beklagten ab (Anlage K 4).

Die Beklagten bewerben ihren Chatbot mit dem Slogan „Get answers. Find inspiration. [...]“.

Die Klägerin behauptet, der Chatbot erzeuge Reproduktionen der Trainingsdaten in erheblichem Umfang. Diese sogenannte Memorisierung von Inhalten innerhalb von Modellen führe zur Regurgitation, also zur Produktion von Output, der bestimmte Trainingsinputs explizit reproduziere. Während die Memorisierung eine Aussage über den Zustand (d.h. den Informationsgehalt) des Modells treffe – nämlich die Aussage, dass in dem Modell eine Fixierung memorisierter Trainingsdaten vorliege, die eine Wahrnehmbarmachung ermögliche -, betrachte die Regurgitation nur die Ausgabe des KI-Systems und identifiziere diese als (Teil-)Kopie eines Trainingsdatums.

Die Modelle erzeugten einen deterministischen Output auf der Ebene der Softmax-Ausgabe, der

bei identischer Eingabe identisch sei. Der Output des KI-Modells sei kein lesbarer Text, sondern eine Wahrscheinlichkeitsverteilung: die Ausgabe der sogenannten Softmax-Funktion. Sei der Trainingsprozess eines Transformer-Modells abgeschlossen, seien sämtliche Modellparameter (Gewichte) festgelegt. Das Modell werde zu einer statischen mathematischen Funktion. Jede spätere Nutzung zur Textgenerierung verändere diese Parameter nicht mehr. Auf dieselbe Eingabe erzeuge das Modell auch bei mehrfacher Wiederholung der Eingabe immer dieselbe Ausgabe.

Allerdings sei die Ausgabe des Modells noch nicht der Output, die der Chatbot dem Nutzer als Ausgabe auf einen Prompt anzeige. Der Output entstehe auf Basis der Ausgabe des Modells, nachdem diese durch die sogenannte Decoding-Struktur verarbeitet worden sei. Beim Decoding entstünden durch gezielt eingesetzte Randomisierungs-Mechanismen (Mechanismen zur Einführung eines Zufallselements in die Generierung von Ausgaben) Zufälligkeit und Variationen. Dieses Decoding sei somit nicht Teil des Modells, sondern diesem nachgelagert. Die Generierung nicht-deterministischer Ausgabe des Chatbots (und nicht des Modells) beruhe somit auf Designentscheidungen der Beklagten, die jederzeit wieder deaktiviert werden könnten.

Die Beklagten unterschieden die Modelle nicht in der gebotenen Klarheit von dem ihnen nachgeordneten Decoding-Algorithmus und von dem Gesamtsystem, dem Modell plus Decoding-Algorithmus plus Benutzeroberfläche Chatbot.

Mit dem Greedy Decoding stünde eine Decoding-Strategie zur Verfügung, die den Determinismus des Modells auf die Ausgabe des Chatbots übertrage. Werde Greedy-Decoding verwendet, bleibe die Ausgabe deterministisch. Selbst Nutzer könnten durch entsprechende Anpassung des Temperature-Parameters ein quasi-deterministisches Ausgabeverhalten erreichen.

Trotz künstlicher Randomisierung beim Decoding reproduzierten Chatbots memorisierte Inhalte konsistent und mit minimaler Varianz, weil bei der Generierung memorisierter Inhalte die Wahrscheinlichkeiten der Token, die sich zum memorisierten Inhalt zusammensetzen, sehr hohe Werte erreichten – häufig nahe 100%. Die Wiedergabe des memorisierten Inhalts werde damit faktisch unvermeidbar und erfolge deterministisch. Die Chatbots funktionierten insoweit wie eine Datenbank.

Ob ein Output memorisierten Inhalt enthalte, sei durch einen Abgleich des Outputs mit den Trainingsdaten möglich. Finde sich der Output in den Trainingsdaten wieder, müsse man von einer Memorisierung ausgehen. Zwar bestünde die theoretische Möglichkeit einer zufälligen Identität zwischen Output und Trainingsdatum. Jedenfalls ab einer bestimmten Textlänge sei die Wahrscheinlichkeit eines solchen Zufalls jedoch so verschwindend gering, dass sie praktisch vernach-

lässigbar werde.

Die streitgegenständlichen Liedtexte seien vollständig und unverändert im Modell memorisiert. Sofern es bei der Ausgabe der memorisierten Inhalte zu Verkürzungen und insbesondere Veränderungen der Texte komme, sei dies nicht dem Modell und den darin verkörperten Informationen geschuldet, sondern den Decoding-Mechanismen, die durch die Erzeugung künstlicher Varianz Halluzinationen provozierten. Die Ausgabe der streitgegenständlichen Liedtexte sei kein provozierter Einzelfall, sondern finde konsistent statt.

Selbst bei Beibehaltung der randomisierenden Decoding-Mechanismen erfolge eine ausreichend konsistente Ausgabe der Liedtexte. Die dadurch eingeführte Varianz betreffe gerade nicht die memorisierten Inhalte, sondern primär die auswechselbaren Einleitungs- und Schlusstexte der Outputs. Die Gewissheit der Modelle in Bezug auf memorisierte Inhalte – also die Wahrscheinlichkeitswerte der entsprechenden Tokensequenz gemäß Softmax-Funktion – sei so hoch, dass der Chatbot diese konsistent ausgabe.

Die Unternehmensgruppe der Beklagten habe maßgeblichen Einfluss auf die Outputs, unter anderem durch die gezielte Auswahl der Trainingsinhalte, die bewusste Ausgestaltung des Trainingsprozesses zur Memorierung spezifischer Inhalte und die Implementierung oder Unterlassung von Filtertechnologien.

Die Klägerin behauptet, die streitgegenständlichen Liedtexte seien nicht mit der Zustimmung der Rechteinhaber im Internet frei verfügbar. Für die von den Beklagten benannten Liedtextkopien, die vermeintlich ohne Nutzungsvorbehalt im Internet verfügbar gewesen seien, ließen sich überwiegend Nutzungsvorbehalte feststellen (siehe Screenshots der Websites als Anlagen K 56, 57).

Die Klägerin habe keine Lizenzen für eine Online-Nutzung der streitgegenständlichen Liedtexte erteilt. Bei der Lizenzierung anderer Bereiche verpflichtete sie ihre Lizenznehmer konsequent zur Einhaltung des Text und Data Mining Nutzungsvorbehalts.

Weder der allgemeinen Bevölkerung noch den Rechtsinhabern sei die Existenz von Modellen vor der Veröffentlichung des ersten Chatbots der Unternehmensgruppe der Beklagten im November 2022 bewusst gewesen.

Die Beklagten erweckten durch eine Bewerbung ihres Dienstes mit „Get answers“ die Funktionserwartung einer Wissensdatenbank.

Die Klägerin ist der Ansicht, Rechtsverletzungen fänden sowohl im Modell als auch bei der Aus-

gabe statt.

Angesichts der erfolgten Memorisierung der streitgegenständlichen Liedtexte in den Modellen der Beklagten handele es sich bei den Modellen um Vervielfältigungsstücke i.S.v. Art. 2 InfoSoc-RL, § 16 Abs. 1, 2 UrhG.

Im Hinblick auf die erfolgten Eingaben der streitgegenständlichen Liedtexte als Trainingsdaten in den Modellen der Beklagten und die Ausgabe der Texte bei den erfolgten Abfragen bestünde bereits ein Anscheinsbeweis dafür, dass die Texte in den Modellen vervielfältigt seien. Die technischen Vorgänge spielten sich ausschließlich in der Sphäre der Beklagten ab. Nutzer erhielten keinen unmittelbaren Zugriff auf das Modell. Wenn Informationen in ein System wie vorliegend die Liedtexte beim Training eingespeist würden und das System anschließend dieselben Informationen, die Liedtexte, wiedergebe, sei typischerweise davon auszugehen, dass die Informationen im System gespeichert seien.

Tatsächlich wiesen die (teil-)identischen Ausgaben der Liedtexte deren Fixierung im Modell aber auch nach, da eine provozierte Regurgitation aufgrund der verwendeten Prompts ebenso wie eine zufällige Regurgitation aufgrund deren Unwahrscheinlichkeit ausgeschlossen sei.

Aus der Technologieneutralität des § 16 Abs. 1 und 2 UrhG und der Schwerpunktsetzung auf die Eignung zur mittelbaren Wahrnehmbarmachung ergebe sich, dass die Rechtsprechung auch Trägertechnologien als Vervielfältigungsstücke einstufe, die substantielle Veränderungen an der Festlegung eines Werkes vornähmen, sofern sie nur eine sinnliche Wahrnehmung der schöpferischen Züge des Werkes erlaubten. Zur Beurteilung der Speicherung im Modell komme es nicht darauf an, ob die im Modell enthaltenen Einzelfragmente der streitgegenständlichen Liedtexte für sich gesehen urheberrechtlich geschützt oder in einer urheberrechtlich geschützten Anordnung gespeichert seien. Entscheidend sei, ob diese Inhalte später vom System in einer Anordnung ausgegeben würden, die die eigene geistige Schöpfung zum Ausdruck bringe. Aufgrund des Determinismus der Modelle sei dies der Fall. Zufälligkeit und Varianz bei den Ausgaben resultierten erst aus bewussten Designentscheidungen der Beklagten auf Ebene der Decoding-Architektur. Aber selbst bei Beibehaltung der randomisierenden Decoding-Mechanismen erfolge eine ausreichend konsistente Ausgabe der Liedtexte, da die Varianz nicht die memorisierten Inhalte, sondern primär die auswechselbaren Einleitungs- und Schlusstexte der Ausgabe beträfen. Im Interesse eines hohen Schutzniveaus für Urheber und angesichts neuer Technologien müsse es für das Vorliegen einer Vervielfältigung genügen, dass die Anwendung der Hilfsmittel mit einer gewissen Wahrscheinlichkeit zur Wahrnehmbarmachung des Werkes führe.

Die Modelle erlaubten mittelbar die sinnliche Wahrnehmung der streitgegenständlichen Liedtexte bzw. urheberrechtlich geschützter Teile davon. Die dabei erforderlichen technischen Hilfsmittel umfassten die Decoding-Architektur zur Verarbeitung der Softmax-Funktionen des Modells, die Chatbot-Benutzeroberfläche mit Eingabemaske und Ausgabedarstellung und die Nutzereingabe.

Die dauerhaften Vervielfältigungen im Modell seien nicht durch § 44b UrhG gerechtfertigt. Bereits die Anwendbarkeit auf Modelle werde bezweifelt. Jedenfalls unterfielen Memorisierungen von Trainingsdaten nicht der Text und Data Mining Schranke, da Vervielfältigungen im Modell nicht der weiteren Datenanalyse dienten und eine dauerhafte Vervielfältigung ohnehin nicht durch § 44b UrhG gerechtfertigt sei. § 60d UrhG finde keine Anwendung, da die Beklagten nicht zu den privilegierten Forschungsorganisationen zählten. Auch wenn man auf die Obergesellschaft abstellen wollte, liege bei Unternehmen, die wissenschaftliche Forschung betrieben, um darauf basierend Waren oder Dienstleistungen zu entwickeln und zu vermarkten, die Verfolgung eines kommerziellen Zweckes vor.

Jedenfalls habe die Klägerin einen wirksamen Nutzungsvorbehalt erklärt. Die Beklagten trügen die Darlegungs- und Beweislast für ihre Behauptung, es liege kein wirksamer Nutzungsvorbehalt vor, und zwar im Hinblick auf sämtliche Quellen, denen sie die streitgegenständlichen Liedtexte entnommen hätten, da jede Nutzung von Quelldaten eine eigene Vervielfältigungshandlung darstelle. Im Rahmen des § 44b UrhG werde das Ausschließlichkeitsrecht des Urhebers in ein „Opt-Out“-Recht umgekehrt: Der Urheber müsse aktiv tätig werden und Nutzungen, die er nicht wünsche oder für die er eine Vergütung beanspruche, widersprechen. Blicke er untätig, sei das Text und Data Mining nach § 44b UrhG erlaubt. Ein Tätigwerden des Urhebers setze jedoch voraus, dass der Urheber von der (beabsichtigten) Nutzung überhaupt erfahre, was bei der Nutzung von Werken für das KI-Training typischerweise nicht der Fall sei, da die betreffenden Unternehmen die betroffenen Rechtsinhaber weder vorher noch nachher über die Werknutzung zu Trainingszwecken informierten. Die Beklagten verhielten sich in Bezug auf die verwendeten Trainingsdaten nicht transparent, sondern behandelten diese Information im Gegenteil sogar als Geschäftsgeheimnis, wie das vorliegende Verfahren zeige. Das robots.txt-Protokoll sei für die Erklärung von Nutzungsvorbehalten im Bereich des Text und Data Mining nicht uneingeschränkt geeignet. Mit robots.txt sei es nicht möglich, Nutzungsvorbehalte werkbezogen zu erklären, da in diesem Protokoll nur Verzeichnisse, Unterverzeichnisse oder URLs mit bestimmten Zeichenketten pauschal ausgeschlossen, nicht aber einzelne spezifische Inhalte oder Werke präzise identifiziert werden könnten. Auch Nutzungsarten oder Lizenzbedingungen könnten in robots.txt nicht adressiert werden.

Durch die Outputs der Modelle seien weitere Verletzungshandlungen erfolgt.

Die Regurgitationen der memorisierten streitgegenständlichen Liedtexte erzeugten bei den Outputs Vervielfältigungen, deren Hersteller die Beklagten seien. Diese manifestierten sich sowohl als Festlegungen im Arbeitsspeicher der Nutzer-Endgeräte als auch als dauerhafte Speicherungen in der Chat-Historie auf den von den Beklagten bereitgestellten Cloud-Servern.

Die Outputs seien keine Doppelschöpfungen, weil diese nicht durch einen magischen Vorgang begründet würden, sondern kausal durch das Training mit den entsprechenden Werken und deren Memorisierung, also durch Speicherung im Modell.

Die Beklagten seien nach der primär maßgeblichen technischen Betrachtung als Hersteller dieser Vervielfältigungen anzusehen. Die Beklagten stellten keinen neutralen Vervielfältigungsdienst wie etwa Internetrecorder bereit, sondern ein inhaltlich determiniertes Informationssystem, auf das sie maßgeblichen Einfluss hätten. Die deliktsrechtliche Verantwortlichkeit für die streitgegenständlichen Outputs liege bei den Beklagten, die die Modelle mit den Liedtexten trainiert hätten.

Die streitgegenständlichen Outputs seien als öffentliche Zugänglichmachungen nach § 19a UrhG zu qualifizieren. Die Wiedergabehandlung der Beklagten sei unmittelbar, da sie selbst der Öffentlichkeit die Möglichkeit des Werkgenusses eröffneten und nicht lediglich (mittelbar) durch die Bereitstellung von Infrastruktur eine Wiedergabehandlung Dritter ermöglichten. Da für die Handlung der Wiedergabe die bloße Möglichkeit des Abrufs genüge, sei es unschädlich, wenn die Liedtexte auf den Prompt hin generiert würden, und die Outputs Abweichungen enthielten. Nutzer des Chatbots stellten ein neues Publikum dar. Die Einstellung der Inhalte sei ohne Zustimmung der Rechteinhaber erfolgt. Es bestünde auch ein erheblicher Unterschied zu Suchmaschinen, da der Nutzer nicht mehr die Originalwebsite besuchen müsse, der Chatbot ersetze die Quelle.

Hilfsweise sei jedenfalls durch die Outputs der allgemeine Tatbestand der öffentlichen Wiedergabe nach § 15 Abs. 2 UrhG erfüllt. Die Beklagten seien nicht nur Anbieter einer Infrastruktur, sondern nähmen eine zentrale Rolle ein, um die Wiedergabe zu ermöglichen.

Die streitgegenständlichen Outputs griffen als Umgestaltungen im Sinne des § 23 Abs. 1 S. 1 UrhG in die Rechte der Klägerin ein. Die Originallieder seien in wiedererkennbarer Form in den Outputs übernommen. Für die Feststellung einer Rechtsverletzung durch die Übernahme geschützter Werkteile komme es ausschließlich darauf an, ob die geschützten Werkteile im neuen Werk wiedererkennbar seien; ein Vergleich des Gesamteindrucks der beiden Gestaltungen sei nicht maßgeblich. Eine derartige Auslegung der § 23 Abs. 1 S. 2 UrhG, der zufolge ein klar identi-

fizierbarer und damit wiedererkennbar übernommener Teil eines vorexistierenden Werkes aufgrund eines Vergleichs der beiden Werke in der Gesamtschau als zulässige Übernahme eingestuft werde, sei bereits unionsrechtswidrig. Denn hierfür sehe der abschließende Schrankenkatalog in Art. 5 InfoSoc-RL keine entsprechende Umsetzungsermächtigung für die Mitgliedstaaten vor. Die „Verblassensformel“ sei ausschließlich persönlich geistigen Schöpfungen vorbehalten und scheide bei KI-Erzeugnissen von vornherein aus. Außerdem verblassten die übernommenen Textpassagen auch bei einem Vergleich der Gestaltungen in der Gesamtschau nicht.

Die Veränderungen der Liedtexte erfüllten den Tatbestand des § 14 UrhG. Die Entstellungen seien das Ergebnis bewusster Designentscheidungen der Beklagten, die durch Randomisierungs-Mechanismen auf Decoding-Ebene Abweichungen vom korrekt memorisierten Text und somit Halluzinationen gezielt herbeiführten und damit willentlich in Kauf nähmen, urheberrechtlich geschützte Inhalte mit werkfremden Inhalten zu verbinden und dadurch zu entstellen. Durch die Möglichkeit, Chatverläufe per Link zu teilen, hätten die entstellten Liedtexte eine hohe potenzielle Reichweite. Aufgrund der Nutzung des Chatbots als Informationsquelle und Wissensdatenbank würden die entstellten Texte den Textdichterinnen und Textdichtern zugerechnet.

Die Zuschreibung verfälschter Liedtexte greife in den Schutzbereich des allgemeinen Persönlichkeitsrechts ein. Dieses schütze den Urheber neben § 14 UrhG, der die Entstellung an sich, also den Eingriff in die Werkintegrität betreffe, zusätzlich davor, ihm ohne seinen Willen veränderte Werke durch seine Nennung als Urheber zuzuschreiben. Bei Zuschreibungen von Äußerungen, die der vermeintliche Urheber nie getätigt habe, handele es sich um unwahre Tatsachenbehauptungen. Hilfsweise werde der Unterlassungsanspruch gegen die Zuschreibung der verfälschten Liedtexte auf § 13 S. 2 UrhG gestützt.

Die Outputs seien nicht als Zitat nach § 51 UrhG gerechtfertigt. Ein KI-System sei nicht imstande, dem Zitat einen schöpferischen selbständigen Text hinzuzufügen, und könne keinen Zitatzweck verfolgen. Überdies schütze das Zitatrecht nur eine unveränderte Übernahme.

Eine Rechtfertigung nach § 51a UrhG setze eine vorliegend nicht gegebene inhaltlich-künstlerische Befassung mit dem Bezugsgegenstand voraus.

Als juristische Person könnten sich die Beklagten nicht auf § 53 UrhG berufen. Soweit es sich bei den Nutzern um Unternehmen handele, könnten diese sich nicht auf die Privatkopieschranke berufen.

Eine Einwilligung in die Verletzungshandlungen im Modell und durch die Outputs scheitere bereits

daran, dass das Crawlen von Internetinhalten und ihre Verwendung für das Training von Modellen als unbekannte Nutzungshandlung keine übliche Nutzungshandlung sei, mit der der Rechteinhaber rechnen müsse. Die Beklagten selbst bezeichneten die Memorisierung als „rare bug“, die nicht „nach den Umständen üblicherweise zu erwarten“ sei.

Der Einwand der Unverhältnismäßigkeit gemäß § 275 Abs. 2 BGB sei nicht anwendbar auf Unterlassungsansprüche, die wie § 97 Abs. 1 UrhG dazu dienen, die Rechtssphäre der Urheber vor Eingriffen Dritter zu schützen. Jedenfalls sei der Unterlassungsanspruch nicht unverhältnismäßig. Die Memorisierung sei kein untergeordnetes Phänomen, sondern betreffe mindestens 1% der Trainingsdaten. Zudem könnten die Beklagten ein Modell mit einem tatsächlich effektiv deduplizierten Trainingsdatenkorpus trainieren oder Lizenzen für die Inhalte einholen, mit denen sie ihre Transformer-Modelle trainiere.

Die Klage sei nicht missbräuchlich. Die Klägerin verfolge ihre gesetzlichen Aufgaben aus dem VGG und ihre gesellschaftliche Funktion, die Interessen und Rechte ihrer Berechtigten kollektiv wahrzunehmen. Die Klägerin beabsichtige nicht, ein bestimmtes Lizenzmodell durchzusetzen.

Die Klägerin hat ihre Klageanträge mit Schriftsatz vom 18.09.2025 und im Termin zur mündlichen Verhandlung vom 29.09.2025 nach den erfolgten Hinweisen der Kammer angepasst.

Die Klägerin beantragt zuletzt:

1. Die Beklagten werden verurteilt, es bei Vermeidung eines vom Gericht für jeden Fall der Zuwiderhandlung festzusetzenden Ordnungsgeldes bis zu 250.000,- EUR, ersatzweise Ordnungshaft, oder Ordnungshaft bis zu sechs Monaten, zu unterlassen,

a) (Vervielfältigungen im Modell) die als Anlagenkonvolut K 1 beigefügten Liedtexte ganz oder teilweise ohne Einwilligung der Klägerin in Sprachmodellen (Large Language Models) zu vervielfältigen und/oder diese Handlung durch Dritte vornehmen zu lassen, wie geschehen in den Modellen 4 und 4o;

b) (Verletzungen bei der Ausgabe) die als Anlagenkonvolut K 1 beigefügten Liedtexte ganz oder teilweise und/oder in Form von Umgestaltungen ohne Einwilligung der Klägerin in den Ausgaben (Outputs) eines generativen Sprachassistenten (Chatbot) öffentlich zugänglich zu machen, hilfsweise für den Fall der Abweisung des Antrags bzgl. der öffentlichen Zugänglichmachung und/oder öffentlich wiederzugeben, und/oder zu vervielfältigen und/oder diese Handlungen durch Dritte vornehmen zu lassen, wie geschehen in Anlagenkonvolut K 2;

c) in Bezug auf die in Anlagenkonvolut K 1 genannten Textdichter zu behaupten und/oder zu verbreiten und/oder diese Handlungen durch Dritte vornehmen zu lassen, die im Anlagenkonvolut K 2 wiedergegebenen Texte seien von ihnen verfasst, wie geschehen in Anlagenkonvolut K 2.

2. Die Beklagten werden verurteilt, der Klägerin unter Vorlage eines geordneten Verzeichnisses Auskunft darüber zu erteilen, in welchem Umfang sie die in Ziffer 1 beschriebenen Handlungen (in Bezug auf die Beklagte zu 1) seit dem 14. Dezember 2023 und in Bezug auf die Beklagte zu 2) seit dem 30. November 2022) vorgenommen haben und welche Einnahmen sie hieraus erzielt haben, und zwar unter Angabe

a) von Anzahl und Umfang der Handlungen nach Ziffer 1 seit dem jeweils genannten Zeitpunkt und

b) der Umsätze seit dem jeweils genannten Zeitpunkt.

3. Es wird festgestellt, dass die Beklagten verpflichtet sind, der Klägerin allen Schaden zu ersetzen, der ihr durch die in Ziffer 1 beschriebenen Handlungen bereits entstanden ist und/oder noch entstehen wird, hilfsweise die ungerechtfertigte Bereicherung herauszugeben, wie sie sich anhand der Auskunft nach Ziffer 2 ergibt.

4. Die Klägerin wird ermächtigt, den verfügenden Teil des Urteils innerhalb von 6 Wochen nach Rechtskraft des Urteils in der überregionalen Süddeutschen Zeitung am Wochenende auf S. 3 im Ressort Wirtschaft im Format 1/4 Seite quer auf Kosten der Beklagten bekannt zu machen.

5. Die Beklagten werden verurteilt, an die Klägerin vorgerichtliche Rechtsanwaltskosten in Höhe von 5.049,70 EUR nebst Zinsen in Höhe von fünf Prozentpunkten über dem Basiszinssatz seit Rechtshängigkeit zu zahlen.

Die Beklagten beantragen:

Die Klage wird abgewiesen.

Die Beklagten behaupten, die Obergesellschaft der Unternehmensgruppe der Beklagten sei als gemeinnütziges Unternehmen gegründet worden und bis heute gemeinnützig geblieben. Sie betreibe wissenschaftliche Forschung.

Das Modell speichere oder kopiere keine spezifischen Trainingsdaten, sondern reflektiere in sei-

nen Parametern, was es basierend auf dem gesamten Trainingsdatensatz erlernt habe, nämlich Beziehungen und Muster aller Wörter beziehungsweise Token, die die Vielfalt der menschlichen Sprache und ihrer Kontexte abbildeten. Es funktioniere auf Basis statistischer Wahrscheinlichkeiten und sei aufgrund der im Training abgeleiteten statistischen Information lediglich dazu befähigt, statistisch wahrscheinliche Tokenfolgen zu generieren.

Die Technologie der Unternehmensgruppe der Beklagten sehe keinen dem Modell nachgeordneten Decoding-Algorithmus vor.

Wahrscheinlichkeitsbeziehungen existierten nicht im Modell. Wahrscheinlichkeiten zwischen den einzelnen Token bildeten sich erst bei der Generierung durch die Verarbeitung der Nutzereingabe. Erst durch diese werde determiniert, wie eine bestimmte Folge aus Wörtern/Token zusammensetzen sei.

Eine Einstellung, mit denen das Modell immer den wahrscheinlichsten Token auswähle, könne die Funktionsfähigkeit des Modells als Sprachmodell beeinträchtigen (z.B. Wiederholungsschleifen) und sei deswegen auch bei den Modellen, die dem Chatbot zugrunde lägen, nicht vorgesehen.

Outputs basierten auf einer sequenziell-analytisch, iterativ-probabilistischen Synthese und stellten eine unabhängige und originäre (maschinelle) Generierung eigener Art dar. Die Nutzereingabe und deren Inhalt seien immer maßgeblich für die Generierung der Ausgabe durch das Modell. Es gebe keine abgrenzbaren, deterministischen Outputs. Ein Modell sei keine Datenbank, in der Trainingsdaten hinterlegt seien, die aufgrund der Eingabe des Nutzers gesucht, aufgefunden und anschließend ausgegeben würden.

Der Betreiber einer generativen KI wisse vor der Ausgabe nicht, was auf eine bestimmte Nutzereingabe hin konkret generiert werde. Vor Abschluss der schrittweisen Generierung stünde sie nicht fest und sei daher nicht deterministisch. Modelle seien nicht darauf angelegt, Trainingsdaten erneut zu generieren. Wenn Modelle korrekte Fakten lieferten, könne dies eine Folge des statistischen Trainings mit großen Textmengen sein.

Mit dem Begriff der Memorisierung werde in der Forschung lediglich der Umstand bezeichnet, dass ein Modell in der Lage sei, auf bestimmte Nutzereingaben Inhalte zu generieren, die mit Trainingsinhalten identisch seien, letztlich das Phänomen der Regurgitation.

Die Parameter des Modells stünden nach dem Training nicht unveränderlich fest, sondern würden im Rahmen von Post-Training-Maßnahmen wiederholt angepasst. Die Modellgewichte wür-

den dahin angepasst, dass etwa keine Outputs unerwünschter Inhalte erfolgten. Die Unternehmensgruppe der Beklagten sehe eine Vielzahl von Maßnahmen und Gestaltungselementen vor, die auch darauf gerichtet seien, die Generierung von Text, der Trainingsinhalten ganz oder in Teilen gleiche, durch die Transformer-Modelle zu verhindern.

Die Generierung von Trainingsdaten als Output könne nicht vollständig unterbunden werden, da der einmal erfolgte Trainingsvorgang, aus dem das (trainierte) Modell hervorgegangen sei, nicht mehr rückgängig gemacht werden könne; ein „unlearning“ eines bestimmten Inhalts sei nach dem heutigen Stand der Forschung nicht möglich.

Bei der Erstellung des Trainingsdatensatzes habe die Unternehmensgruppe der Beklagten im Fall öffentlich zugänglicher Internetseiten robots.txt-Dateien, mit denen einer Sammlung durch Webcrawler widersprochen werden könne, berücksichtigt.

Die streitgegenständlichen Liedtexte seien alle mit Einverständnis der Schutzrechtsinhaber ohne Sicherungsvorkehrungen in das Internet gestellt worden. Jeder der neun streitgegenständlichen Liedtexte sei aus einer Webseite ohne eingerichtete robots.txt-Dateien gecrawlt worden, oder die eingerichteten robots.txt-Dateien hätten ein Crawlen durch den CCBot nicht verboten.

Die von der Klägerin vorgelegten Ausgaben des Chatbots könnten keine Fixierung der Liedtexte im Modell nachweisen. Bereits die vorgelegten Anlagen zeigten, dass Outputs nicht konsistent stätfänden.

Die Klägerin habe die vorgelegten Ausgaben provoziert, indem sie die jeweiligen Prompts mehrfach getestet habe. Zudem habe sie einen Handlungsrahmen geschaffen, um das Modell von seiner neutralen Bahn abzubringen, indem sie die benutzerdefinierten Agenten als „Experten für Liedtexte“ ausgestaltet habe.

Wie aus Anlage K 3 ersichtlich, handele es sich bei nahezu allen Ausgaben des Chatbots, auf welche die Klägerin im Rahmen ihrer Verletzungssubsumtion verweise, nicht um wortgetreue Wiedergaben, sondern um sog. Halluzinationen, also das Erzeugnis inhaltlich falscher, frei erfundener oder faktisch nicht belegbarer Ausgaben durch das Modell, und nicht um Regurgitationen.

Die Beklagten sind der Ansicht, die Klage sei unzulässig.

Die Klageanträge unter Ziffer 1 und 2 genügten nicht den Anforderungen an die Bestimmtheit gemäß § 253 Abs. 2 Nr. 2 ZPO. Die Klägerin wolle den Gegenstand des begehrten Verbots bewusst im Unklaren oder Vagen halten. Im Hinblick auf die Anträge unter Ziffer 1 sei jeweils der abstrahie-

rende Teil des begehrten Verbotes unbestimmt durch die Verwendung der auslegungsbedürftigen Begriffe „teilweise“ und „Umgestaltung“, und unter Ziffer 1c) sei der Antrag insgesamt unverständlich. Der Antrag unter Ziffer 2 sei unbestimmt, weil er sich auf „Handlungen nach Ziffer 1“ beziehe und überdies nicht nachvollziehbar sei, wie sich „Anzahl“ und „Umfang“ unterscheiden und wie „Einnahmen“ identifiziert werden könnten.

Die Klage sei zudem teilweise unzulässig, weil der Klägerin die Prozessführungsbefugnis für die Ansprüche fehle. Die Klägerin habe keine lückenlose Rechtekette vorgelegt und keine Nachforschungen unternommen um sicherzustellen, dass die behaupteten ausschließlichen Nutzungsrechte in ihrer Inhaberschaft stünden. Die Berechtigungsverträge mit den Textdichtern der Liedtexte seien unwirksam. Sämtliche vorgelegten Berechtigungsverträge verstießen gegen das für Allgemeine Geschäftsbedingungen geltende Transparenzgebot nach § 307 Abs. 1 BGB. Bei Ziffer 3 und 4 der Zusatzvereinbarungen handele es sich um eine Blankoermächtigung, anhand derer die Vertragspartner der Klägerin nicht erkennen könnten, welchen Umfang die von ihnen erteilte Ermächtigung tatsächlich habe. Eine Geltendmachung von Ansprüchen nach § 13 UrhG werde bereits nicht erwähnt. Ohnehin könne die Klägerin die Ansprüche, die auf dem Urheberpersönlichkeitsrecht und dem allgemeinen Persönlichkeitsrecht gründeten, nicht im Wege der Prozessstandschaft geltend machen, da dies Rechte mit höchstpersönlichem Charakter seien. Zudem seien nach Ziffer 5 der vorgelegten Zusatzvereinbarungen nur bereits entstandene Ansprüche auf Auskunft, Schadensersatz und Bereicherung abgetreten, die Klägerin begehre aber die Feststellung der Schadensersatzpflicht auch für noch entstehende Schäden.

Die Klage sei unbegründet. Sie sei letztlich durchweg nicht schlüssig.

Die Klägerin sei nicht aktivlegitimiert. Die vorgelegten Berechtigungsverträge seien unwirksam. Zudem seien die aus der angeblichen Verletzung des allgemeinen Persönlichkeitsrechts resultierenden Auskunfts- und Schadensersatzansprüche sowie bereicherungsrechtliche Ansprüche qua ihrer Verknüpfung mit dem allgemeinen Persönlichkeitsrecht der Textdichter höchstpersönlicher Natur und nicht abtretbar.

Eine Urheberrechtsverletzung durch das Modell sei nicht gegeben.

Es liege keine körperliche Festlegung der Trainingsdaten in den Modellen vor, wie sie für eine Vielfältigkeit im Sinne des § 16 UrhG erforderlich sei. Zwar würden für die Entwicklung des Modells die Trainingsdaten genutzt und die Nutzung schlage sich auch in einem Beitrag zur Ausgestaltung und Anpassung der Parameter des Modells nieder. Es gebe aber keinen identifizierbaren Datensatz, der die Liedtexte verkörperte. Darin liege der Unterschied zu allen früheren Technolo-

gien, in denen sich immer ein konkreter Datensatz objektiv bestimmen lasse, selbst wenn er nur fragmentiert vorliege, der das Werk verkörpere bzw. repräsentiere. Die gespeicherten einzelnen Parameter, einschließlich der Vektoren, könnten nicht als Repräsentation oder Speicherung eines bestimmten Textes oder einer bestimmten Wortfolge aus den Trainingsdaten angesehen werden, auch nicht mehrere solcher Parameter zusammen. Die Wahrscheinlichkeitsbeziehungen, auf welche die Klägerin abstelle, seien nicht bestimmbar, weil sich die konkrete Wahrscheinlichkeit einer bestimmten Folge an Wörtern erst mit und durch die Verarbeitung eines entsprechenden Prompts in der Generierung des Outputs manifestiere. Gleichzeitig liege keine mit der Festlegung zugleich definierte Maßnahme vor, mit der die Liedtexte in wiederholbarer Weise wiedergegeben werden könnten.

Die Klägerin bringe keinen substantiierten Tatsachenvortrag mit Blick auf die behauptete Speicherung der streitgegenständlichen Liedtexte im Modell. Die Klägerin stelle hier auf ein Memorisieren von Trainingsdaten ab. Dass man von dem (gleich wie erzielten) Output nicht auf das Vorliegen einer körperlichen Festlegung im Modell schließen könne, erkenne damit auch die Klägerin an.

Die Voraussetzungen für einen Anscheinsbeweis seien offensichtlich nicht gegeben, weil es den von der Klägerin postulierten Erfahrungssatz nicht gebe. Die Beklagten treffe auch keine sekundäre Darlegungslast, weil sie umfassend zur Funktionsweise der Modelle vorgetragen habe.

Vervielfältigungen und etwaige Bearbeitungen seien von den Text und Data Mining Schranken gemäß § 60d Abs. 1 UrhG und § 44b Abs. 2 S. 1 UrhG gedeckt, denn diese Schranken erfassten auch zwingend etwaige aus dem Training resultierende (unterstellte) Vervielfältigungen in dem Modell. Die Schranken seien auf das Training von Modellen anwendbar und beschränkten sich nicht auf die Zusammenstellung des Trainingskorpus. Aus der Mitteilung der Kommission „Künstliche Intelligenz für Europa“ vom 25.04.2018 (COM/2018/237, Anlage B 42) ergebe sich, dass dem EU-Gesetzgeber die massenhafte Nutzung von Daten für die Zwecke des Trainings von KI-Modellen bekannt und diese sogar Anlass zur Einführung der Text und Data Mining Schranke gewesen sei. Die beim Training stattfindende Analyse der in den Trainingsdaten enthaltenen Informationen sei mit dem Werkgenuss von Werken vergleichbar, der von den urheberrechtlichen Verwertungsrechten nicht erfasst sei. Da es dem klaren Willen des Gesetzgebers entspreche, die Text und Data Mining Schranken auf das Training von Modellen anzuwenden, müssten von den Schranken auch alle im Rahmen des Trainings entstehende Vervielfältigungen erfasst sein. Ein anderes Verständnis sei mit dem sich aus Erwägungsgrund 18 der DSM-RL ergebenden Zweck der Schranken, die „Entwicklung neuer Anwendungen oder Technologien“ zu ermöglichen,

nicht vereinbar. Die vermeintlich beim Training entstehenden Vervielfältigungen seien durch die Betriebsdauer des Modelles beschränkt. Vermeintlich entstehende Vervielfältigungen seien durch § 60d UrhG zulässig. Die Beklagten seien privilegierte Forschungseinrichtungen, da sich die Voraussetzung der Verfolgung nicht kommerzieller Zwecke nicht unmittelbar auf das Text und Data Mining, sondern gemäß § 60d Abs. 2 Nr. 1 UrhG (und gemäß Art. 2 Nr. 1 DSM-RL) auf die das Text und Data Mining vornehmende Einrichtung beziehe, die insgesamt die Definition der Forschungseinrichtung erfüllen müsse. Zudem sei § 60d Abs. 2 Nr. 1 UrhG im Gegensatz zur alten Fassung des § 60d UrhG ebenso wenig wie Art. 2 Nr. 1 DSM-RL nicht zu entnehmen, dass die Einrichtung keinerlei kommerzielle Zwecke verfolgen dürfe. Die kommerziellen Zwecke, die die Untergesellschaften der Unternehmensgruppe der Beklagten verfolgten, änderten nichts daran, dass die Unternehmensgruppe insgesamt – und damit auch die Beklagten als Teil dieser Unternehmensgruppe – unter die Definition der Forschungsorganisation falle, da sie wissenschaftliche Forschung betreibe und ihre Obergesellschaft nicht gewinnorientiert sei.

Wirksame Nutzungsvorbehalte seien nicht gegeben. Die Beklagten seien ihrer diesbezüglichen Darlegungs- und Beweislast nachgekommen, indem sie vorgetragen hätten, dass bei der Erstellung der Trainingsdatensätze Interessen Dritter durch die Beachtung von robots.txt-Dateien berücksichtigt worden seien. Ein darüberhinausgehender Vortrag, insbesondere im Hinblick auf eine Bezeichnung sämtlicher Quellen, aus denen die Texte gecrawlt worden seien, sei den Beklagten nicht zumutbar. Die robots.txt-Datei sei eine taugliche und formwirksame Methode zur Erklärung des Nutzungsvorbehalts und als Marktstandard anerkannt. Das Bestreiten der Klägerin mit Nichtwissen, dass die von den Beklagten eingesetzten Bots Zugriffsverweigerungen konsequent einhielten, sei unbeachtlich. Als Verwertungsgesellschaft sei sie insbesondere verpflichtet, Nutzungsmöglichkeiten zu sichern und unlizenzierte Verwendungen zu unterbinden. Sie müsse daher ihre Lizenznehmer überwachen und Zweifel an der Beachtung der Nutzungsvorbehalte durch Common Crawl als die größte frei zugängliche Sammlung von Webdaten weltweit aufklären. Der von der Klägerin erklärte Nutzungsvorbehalt im Impressum sein erst im Jahr 2024 aufgenommen worden, der Klägerin fehlte die Berechtigung hierzu, und er sei nicht in maschinenlesbarer Form erfolgt.

Zudem seien die streitgegenständlichen Liedtexte neben dem gesamten Trainingsdatensatz als nebensächlich und verzichtbar und damit als unwesentliches Beiwerk nach § 57 UrhG anzusehen. Hauptzweck der Zusammenstellung der Trainingsdatensätze sei bei Common Crawl als maßgeblichem Trainingsdatensammler nicht die Aggregation urheberrechtlich geschützter Inhalte, sondern die Aggregation möglichst umfassender Datensätze ohne weiteren Fokus gewesen.

Es liege auch keine Urheberrechtsverletzung durch die Ausgaben des Modells vor.

Ein Eingriff durch Outputs sei schon deshalb ausgeschlossen, weil die Erzeugung von Inhalten durch eine generative KI eine unabhängige und originäre Generierung eigener Art darstelle. Der Nutzer steuere den kreativen Schaffensprozess durch seinen Prompt unter Zuhilfenahme von IT als Hilfsmittel.

Ein Eingriff in Verwertungsrechte sei mit Blick auf bestimmte Liedtexte nicht möglich, weil gewissen Textfragmenten keine eigene Schutzfähigkeit zukomme. Bei der bruchstückhaften Wiedergabe eines Originals wie im Falle von „36 Grad“ und „Männer“ könne nicht reflexartig von einer erreichten Schutzfähigkeit ausgegangen werden. Gleiches gelte mit Blick auf veränderte Fassungen von Werkteilen (Halluzinationen), die hinreichenden Abstand zur ursprünglichen Gestaltung wahrten. Maßstab für die Bewertung des hinreichenden Abstandes könne nur der einzelne Output und nicht der Chatverlauf in seiner Gesamtheit sein. Der angegriffene Output des Textes „36 Grad“ sei durch zwei unterschiedliche Prompts entstanden.

Für ein öffentliches Zugänglichmachen nach § 19a UrhG fehle es bereits an einer Zugänglichmachung der Liedtexte, weil kein Zugriff auf die Liedtexte eröffnet werde. Diese lägen nicht in der Zugriffssphäre der Unternehmensgruppe. Es fehle damit bereits an einer irgendwie gearteten körperlichen Festlegung oder Fassung der Werke, die für den Abruf durch den Nutzer vorgehalten werden könnten. Ohnehin könnten die Liedtexte nicht abgerufen werden, da die streitgegenständlichen Outputs nicht die Texte in ihrer eigentlichen Form enthielten. Es liege auch keine Handlung der Wiedergabe vor und auch die weiteren Voraussetzungen der Öffentlichkeit und des neuen Publikums seien nicht erfüllt, sodass auch keine öffentliche Wiedergabe der Liedtexte nach § 15 Abs. 2 UrhG gegeben sei.

Ebenso fehle es an einem Eingriff in das Vervielfältigungsrecht, denn allein der Nutzer sei als Hersteller dieser Vervielfältigungen anzusehen, der mit seinem Prompt den Inhalt des jeweiligen Outputs maßgeblich bestimme und damit den automatisierten Generierungsvorgang in Gang setze.

Schließlich mangle es an einer Entstellung oder Beeinträchtigung im Sinne von § 14 UrhG, die geeignet wäre, die berechtigten Interessen der Textdichter an den Liedtexten zu gefährden.

Jedenfalls seien etwaige Eingriffe von Schrankbestimmungen nach §§ 51, 51a, 53 UrhG gedeckt.

Sowohl für etwaige Nutzungshandlungen im Modell als auch bei der Ausgabe liege eine die Rechtswidrigkeit ausschließende Einwilligung der Rechteinhaber vor. Für die Klägerin sei es bereits unzulässig „mit Nichtwissen“ zu bestreiten, dass Liedtexte durch die Rechteinhaber selbst

oder mit ihrer Zustimmung und ohne Sicherungsvorkehrungen ins Internet eingestellt worden seien. Als ausschließliche Lizenznehmerin müsse sich die Klägerin entsprechende Informationen beschaffen. Jedenfalls wäre die Klägerin in Bezug auf die lizenzierten Nutzungsrechte an den streitgegenständlichen Liedtexten sekundär darlegungsbelastet. Mit dem Einstellen ins Internet hätten die Rechteinhaber in die nach den Umständen üblichen Nutzungshandlungen eingewilligt, zu denen seit 2018 auch gehöre, dass die gecrawlten Daten zur Entwicklung großer Sprachmodelle verwendet würden.

Die Beklagten seien nicht passivlegitimiert. Das Anbieten einer universell einsetzbaren generativen KI mache eine entsprechende Berücksichtigung wertungsmäßiger Korrekturen bei der Haftung erforderlich, wie sie die Rechtsprechung schon in anderen Fällen, etwa bei Suchmaschinen oder Plattformbetreibern, entwickelt habe. Die technischen Besonderheiten eines generativen Modells seien bei der Entwicklung des einschlägigen Haftungsmaßstabs zu berücksichtigen.

Dem beantragten Unterlassungsbegehren stünden aufgrund der besonderen Sachverhaltskonstellation die Einwendung des Rechtsmissbrauchs entgegen. Die Klägerin verfolge mit der Klage ausschließlich das Ziel, als Verwertungsgesellschaft ein unzulässiges Lizenzmodell durchzusetzen. Sie habe sich mit den Liedtexten gezielt einen kleinen, untergeordneten Ausschnitt aus ihrem vermeintlichen Rechteportfolio ausgesucht. Die Klage sei darauf gerichtet, mit einem unbestimmten Rechtsschutzbegehren einen Unterlassungstitel zu erzielen, der deutlich mehr erfasse als die konkreten Verletzungshandlungen. Als Verwertungsgesellschaft habe die Klägerin kein genuines Interesse an der Unterlassung, sondern sei zum Abschluss von Lizenzverträgen zu angemessenen Bedingungen verpflichtet.

Würde der Klägerin ein Titel wie begehrt mit unbestimmter Reichweite gewährt, würde sich für die Beklagten die Frage stellen, ob sie das Angebot für Deutschland nicht einstellen müssten. Dies sei unverhältnismäßig, weil die vermeintlichen Urheberrechtsverletzungen einen völlig untergeordneten Teilaspekt der Nutzung des Angebots der Beklagten betreffen, dessen Bedeutung in keinem Verhältnis zu den aus der Einstellung des Angebots folgenden negativen Auswirkungen auf die Beklagten und ihre Unternehmensgruppe im Gesamten sowie den Interessen Dritter stehe. Die Technologie der Beklagten finde in Bereichen von besonders gewichtigen Allgemeininteressen Verwendung, eröffne Möglichkeiten im Bereich der Bildung, Wissenschaft und Forschung, sei von zentraler gesellschaftlicher und wirtschaftlicher Bedeutung und habe transformatives Potential. Die vollständige Einstellung des Angebots im Staatsgebiet der Bundesrepublik Deutschland ginge demgegenüber mit einer massiven Vernichtung (volks-) wirtschaftlicher Werte sowie einer hohen Belastung der Beklagten und weitreichenden negativen Auswirkungen auf die Interes-

sen Dritter einher. Zudem seien die streitgegenständlichen Liedtexte bereits jetzt im Internet über tatsächliche Abrufdienste kostenfrei abrufbar.

Eine Aufbrauchfrist von mindestens sechs Monaten sei mit Blick auf den erforderlichen Entwicklungsaufwand angemessen.

Den Beklagten stehe eine Ablösungsbefugnis nach § 100 UrhG zu.

Ein Auskunftsanspruch bestehe nicht und sei im Übrigen unverhältnismäßig.

Ein Schadensersatzanspruch scheitere insbesondere an der Passivlegitimation und am fehlenden Verschulden. Für einen Bereicherungsanspruch fehle ebenfalls die Passivlegitimation.

Im Fall einer Vollstreckung sei von der Klägerin Sicherheit zu leisten gemessen an den Einnahmen der Beklagten in der Bundesrepublik Deutschland, da die Beklagten bei einem Unterlassungstitel für Vervielfältigungen im Modell ihren Betrieb in Deutschland einstellen müssten.

Ein berechtigtes Interesse an der Urteilsbekanntmachung bestehe nicht.

Die Klägerin hat beantragt, das Verfahren auszusetzen und dem EuGH zur Vorabentscheidung vorzulegen.

Am 25.06.2025 (Bl. 366/367 d.A.) und am 26.09.2025 (Bl. 1047/1048 d.A.) sind Beschlüsse zur Geheimhaltung ergangen. Die Klägerin hat die Aufhebung der Geheimhaltung beantragt (Replik S. 3, Bl. 375 d.A.).

Ein Termin zur mündlichen Verhandlung hat am 29.09.2025 stattgefunden, es wird auf die Sitzungsniederschrift verwiesen (Bl. 1060/1064 d.A.).

Zur Ergänzung des Tatbestands wird auf die Schriftsätze der Parteivertreter nebst Anlagen Bezug genommen.

Entscheidungsgründe

Die zulässige Klage erweist sich als überwiegend begründet. Der Klägerin stehen die geltend gemachten Ansprüche aufgrund der Verletzung von urheberrechtlichen Nutzungsrechten zu. Die zulässig erhobenen persönlichkeitsrechtlichen Ansprüche hingegen sind abzuweisen.

A.

Die Klage ist zulässig.

I. Das Landgericht München I ist für die Klage gegen die Beklagte zu 1), die ihren Sitz in Dublin hat, nach Art. 7 Nr. 2 EuGVVO international und örtlich zuständig. Im Hinblick auf die Beklagte zu 2) mit Sitz in den USA ergibt sich die Zuständigkeit aus § 32 ZPO.

II. Die Klageanträge sind hinreichend bestimmt gemäß § 253 Abs. 2 Nr. 2 ZPO.

Nach § 253 Abs. 2 Nr. 2 ZPO hat ein Unterlassungsantrag derart deutlich gefasst zu sein, dass der Streitgegenstand und der Umfang der Prüfungs- und Entscheidungsbefugnis des Gerichts nach § 308 Abs. 1 ZPO erkennbar abgegrenzt sind, sich der Beklagte deshalb erschöpfend verteidigen kann und die Entscheidung darüber, was dem Beklagten verboten ist, nicht letztlich dem Vollstreckungsgericht überlassen bleibt (vgl. BGH 14.07.2022, I ZR 97/21 Rn. 12 – dortmund.de). Eine hinreichende Bestimmtheit ist für gewöhnlich gegeben, wenn auf die konkrete Verletzungshandlung Bezug genommen wird und der Klageantrag zumindest unter Heranziehung des Klagevortrags unzweideutig erkennen lässt, in welchen Merkmalen des angegriffenen Verhaltens die Grundlage und der Anknüpfungspunkt für den Verstoß und damit das Unterlassungsgebot liegen soll (ebd.).

Unter Anwendung der genannten Maßstäbe erweist sich die Klage als so bestimmt wie möglich und lediglich dann weniger bestimmt, wenn es sich nicht vermeiden lässt, weshalb die damit einhergehenden Ungenauigkeiten hinzunehmen sind (vgl. Ceppl/Voß/Zigann/Werner, 3. Aufl. 2022, ZPO § 253 Rn. 196).

1. In den Anträgen und Schriftsätzen der Klägerin ist hinreichend klar bestimmt, welche Verletzungshandlungen sie angreift, nämlich die Vervielfältigung der streitgegenständlichen Liedtexte im Modell, das öffentliche Zugänglichmachen der streitgegenständlichen Liedtexte bei der

Ausgabe durch den Chatbot sowie deren Vervielfältigung auf den Endgeräten der Nutzer und in der Chat-Historie, die Eingriffe in das Bearbeitungsrecht, die Entstellung durch Verfälschung der Lieder und die Verletzung des allgemeinen Persönlichkeitsrechts durch Zuschreibung verfälschter Liedtexte. Die Klägerin hat in ihren Anträgen jeweils Bezug auf die konkreten Verletzungshandlungen genommen, für die vorgetragenen Vervielfältigungen im Modell auf die erfolgten Rechtsverletzungen in den Modellen 4 und 4o, und für die angegriffenen Verletzungshandlungen durch die Outputs mittels Bezugnahme auf die in Anlage K 2 dokumentierten Outputs.

Die Beklagten haben die unterschiedlichen angegriffenen Verletzungshandlungen zutreffend erkannt und sich bezüglich sämtlicher vorgebrachter Verletzungshandlungen verteidigt.

2. Die Klage ist auch im Hinblick auf die Verwendung der auslegungsbedürftigen Begriffe „teilweise“, „Umgestaltungen“, „Anzahl“, „Umfang“ und „Einnahmen“ hinreichend bestimmt.

a. Die Unterlassungsanträge unter Ziffer 1 a) und b) enthalten neben der Bezugnahme auf die Anlage K 2 die unbestimmten Begriffe „teilweise“ und „Umgestaltungen“. Die Verwendung auslegungsbedürftiger Begriffe im Klageantrag ist zulässig, wenn über ihren Sinngehalt zwischen den Parteien kein Streit besteht und objektive Maßstäbe zur Abgrenzung vorliegen, oder wenn der Kläger den auslegungsbedürftigen Begriff hinreichend konkret umschreibt und gegebenenfalls mit Beispielen unterlegt oder sein Begehren an der konkreten Verletzungshandlung ausrichtet (vgl. BGH 02.06.2022, I ZR 140/15 Rn. 27 ff – YouTube II). Letzteres ist vorliegend der Fall.

Spätestens im Zusammenhang mit dem Klagevortrag ist deutlich, dass der Begriff „teilweise“ die Verwendung von Werkteilen betrifft. Der Begriff der „Umgestaltung“ findet sich in § 23 Abs. 1 S. 1 UrhG. Ohnehin hat die Klägerin ihr Unterlassungsbegehren an den konkreten Verletzungshandlungen gemäß Anlage K 2 ausgerichtet, weshalb anhand dieser beurteilt werden kann, welche Handlungen von den begehrten Unterlassungsanträgen erfasst sind.

Eine Antragstellung, die auf sämtliche denkbaren teilweise möglichen Verwendungen der streitgegenständlichen Liedtexte sowie potenzielle Umgestaltungen Bezug nähme, wäre im Verfahren nicht handhabbar. Ebenso wenig ist es prozessökonomisch geboten, aufgrund jeder weiteren teilweisen Verwendung oder Umgestaltung einen neuen Rechtsstreit anzustrengen. Die vorliegend gewählte auslegungsbedürftige Antragsformulierung ist zur Gewährleistung effektiven Rechtsschutzes erforderlich und daher hinzunehmen (vgl. BGH 22.07.2021, I ZR 194/20 Rn. 34 – Rundfunkhaftung).

Den Parteien und insbesondere der Beklagtenseite ist zuzutrauen und zumutbar, anhand der in

Bezug genommene konkreten Verletzungshandlungen zu beurteilen, ob weitere nicht explizit im Einzelfall im Antrag aufgenommene Verletzungshandlungen unter den Verbotstenor fallen.

b. Der unter Ziffer 1c) gestellte Antrag ist verständlich. Die Klägerin nimmt Bezug auf Anlage K 2, in der sich die streitgegenständlichen Liedtexte befinden, die allesamt im Text vom Chatbot geändert worden sind. Sie begehrt ein Verbot zu behaupten, die in Anlage K 2 wiedergegebenen Liedtexte stammten von den ursprünglichen Autoren der Texte.

c. Der unter Ziffer 2 gestellte Antrag ist hinreichend bestimmt im Sinne des § 253 Abs. 2 Nr. 2 ZPO. Die erfolgte Bezugnahme auf „die in Ziffer 1 beschriebenen Handlungen“ ist angesichts der bereits dargestellten hinreichenden Bestimmtheit der im Antrag unter Ziffer 1 benannten Handlungen verständlich.

Die genannten Begriffe „Anzahl“, „Umfang“ und „Einnahmen“ sind hinreichend bestimmt. Soweit die Beklagten noch einer Erläuterung in Bezug auf die Unterscheidung zwischen „Anzahl“ und „Umfang“ bedurften, ist eine solche in der Replik auf S. 91 (Bl. 463 d.A.) erfolgt, in der die Klägerin anführt, der Begriff des „Umfangs“ umfasse insbesondere die Angabe, in welchem Umfang Werkteile genutzt worden seien. Der Begriff der „Einnahmen“ umfasst nachvollziehbar sämtliche Einnahmen der Beklagten.

III. Die Klägerin ist prozessführungsbefugt.

1. Die Klägerin macht die Unterlassungsansprüche der Textdichter, die sie auf das Urheberpersönlichkeitsrecht und das allgemeine Persönlichkeitsrecht stützt, im Wege der gewillkürten Prozessstandschaft geltend.

a. Eine gewillkürte Prozessstandschaft nach § 51 Abs. 1 ZPO ist grundsätzlich zulässig, wenn der Prozessführende vom Rechtsinhaber zu dieser Art der Prozessführung ermächtigt worden ist und er ein eigenes schutzwürdiges Interesse an ihr hat. Ein schutzwürdiges Eigeninteresse besteht, wenn die Entscheidung Einfluss auf die eigene Rechtslage des Prozessführungsbefugten hat; es kann durch ein wirtschaftliches Interesse begründet werden (BeckOK ZPO/Kersting, 58. Aufl. 2025, ZPO § 51 Rn. 50 m.w.N.).

Nach diesen Voraussetzungen liegt eine zulässige gewillkürte Prozessstandschaft vor.

aa. Die Textdichterinnen und Textdichter haben in den Verträgen von Ende 2024 (Anlage K 10) die Klägerin dazu ermächtigt, Ansprüche wegen Verletzungen ihres jeweiligen Urheberpersönlichkeitsrechts und allgemeinen Persönlichkeitsrechts geltend zu machen.

In Ziffer 3 und 4 der Verträge ermächtigen die Textdichterinnen und Textdichter als originäre Urheber die Klägerin explizit, „Unterlassungsansprüche wegen Entstellung oder anderer Beeinträchtigung der Liedtexte im Sinne des § 14 UrhG durch KI-Systeme [...] [sowie] Unterlassungsansprüche wegen Verletzung der ideellen oder vermögenswerten Bestandteile des allgemeinen Persönlichkeitsrechts des Berechtigten durch Veröffentlichung veränderter Liedtexte durch KI-Systeme im eigenen Namen und auf eigene Rechnung (außer-)gerichtlich geltend zu machen“. Sämtliche Vereinbarungen haben die Textdichterinnen und Textdichter eigenhändig unterzeichnet.

bb. Die Verträge sind nicht wegen eines Verstoßes gegen das für Allgemeine Geschäftsbedingungen geltenden Transparenzgebot nach § 307 Abs. 1 BGB unwirksam.

Die Vereinbarungen von Ende 2024 (Anlage K 10) sind in Ziffern 3 und 4 hinreichend bestimmt. Die Parteien der Verträge haben diese speziell für die Rechtsverfolgung der streitgegenständlichen Rechtsverletzungen durch die Klägerin geschlossen. Den Textdichterinnen und Textdichtern ist somit erkenntlich gewesen, welche „Entstellungen und andere Beeinträchtigungen [...] durch KI-Systeme“ von der Vereinbarung umfasst sein sollten, nämlich die in Anlage K 2.

cc. Die Klägerin hat als Lizenznehmerin und Lizenzgeberin der streitgegenständlichen Liedtexte ein schutzwürdiges wirtschaftliches Eigeninteresse an der Rechtsverfolgung.

Ausweislich Ziffer 1 der Vereinbarungen von Ende 2024 „übertragen“ die Textdichterinnen und Textdichter als originäre Inhaber der Urheberrechte an den streitgegenständlichen Liedtexten der Klägerin „das ausschließliche Recht, Liedtexte (mit oder ohne Ton) ganz oder teilweise in Modellen und Systemen der Künstlichen Intelligenz (KI), und in darauf gestützten Anwendungen (nachfolgend gemeinsam „KI-Systeme“) zu nutzen, insbesondere zum Training und Fine-Tuning, durch Verfahren zur Verarbeitung von Kontextinformationen und im Output von KI-Systemen. Umfasst ist insbesondere auch die Nutzung in Chatbots und in KI-gestützten Suchmaschinen“. Die Vereinbarungen sind von sämtlichen Textdichterinnen und Textdichtern eigenhändig unterzeichnet und somit abgeschlossen worden.

Die rechtlich unsaubere Verwendung des Begriffes „übertragen“ statt „einräumen“ im Text der Vereinbarungen ist unschädlich. Aus der Verwendung des Begriffs „übertragen“ kann daher entgegen der Ansicht der Beklagten nicht geschlossen werden, eine Übertragung von bereits Dritten eingeräumten Nutzungsrechten nach § 34 Abs. 1 UrhG sei Gegenstand der Vereinbarungen von Ende 2024. Zwar ist nach § 29 Abs. 1 UrhG die Übertragung von Urheberrechten grundsätzlich nicht möglich, sondern nach §§ 29 Abs. 2, 31 UrhG ist in aller Regel allein die Einräumung von Nutzungsrechten durch den originären Rechteinhaber zulässig. Letzteres ist nach dem objektiven

Empfängerhorizont mit den Vereinbarungen von Ende 2024 gemeint und gewollt: nicht das Urheberrecht als solches soll übertragen werden, sondern die ausschließliche Einräumung von Nutzungsrechten speziell für die Verwendung in KI-Systemen ist beabsichtigt. Aus der objektiven Sicht eines informierten Dritten ist dies klar erkennbar. Rechtseinräumungen werden häufig als Übertragungen bezeichnet (Dreier/Schulze/Mantz, 8. Aufl. 2025, UrhG § 29 Rn. 17). Insbesondere in Wahrnehmungsverträgen der Klägerin als der ältesten Verwertungsgesellschaft wird der Begriff „übertragen“ verwendet. Diese Wahrnehmungsverträge stammen noch aus einer Zeit vor Inkrafttreten des Urheberrechtsgesetzes, zu der das Urheberrecht gemäß § 8 Abs. 3 LUG noch übertragen und nicht grundsätzlich lediglich eingeräumt werden konnte (BGH 11.04.2013, I ZR 152/11 Rn. 27 – Internet-Videorecorder II). Ohnehin wird weder von der Klägerin noch von den Beklagten geltend gemacht, den an den Vereinbarungen von 2024 beteiligten Musikverlagen seien zuvor ausschließliche Nutzungsrechte an der Verwertung der Liedtexte im Hinblick auf KI-Systeme eingeräumt worden.

b. Betrifft wie vorliegend die Prozessstandschaft höchstpersönliche Rechte, ist zusätzlich zu prüfen, ob die Natur der in Rechtsstreit stehenden Rechte eine Prozessstandschaft im konkreten Fall zulässt. Ist das höchstpersönliche Recht mit dem Rechtsinhaber, in dessen Person es entstanden ist, so eng verknüpft, dass die Möglichkeit, eine gerichtliche Geltendmachung einem Dritten im eigenen Namen zu überlassen, dazu in Widerspruch stünde, ist die Prozessstandschaft in der Regel unzulässig (BGH 08.06.1989, I ZR 135/87 – Emil Nolde). Handelt es sich dagegen nicht um solche höchstpersönlichen Rechte, sondern um Rechte oder rechtlich geschützte Positionen, die zusammen mit den Ansprüchen, die sie schützen sollen, übertragbar sind, wird, auch wenn die geltend zu machenden Ansprüche für sich allein nicht übertragbar sind, die Ermächtigung zur gerichtlichen Verfolgung von Rechten zugunsten des materiell Berechtigten stets für zulässig erachtet unter der Voraussetzung, dass der Ermächtigte an der Rechtsverfolgung ein eigenes rechtsschutzwürdiges Interesse hat (BGH 17.02.1983, I ZR 194/80 – Geldmafiosi; BGH 23.09.1992, I ZR 251/90 – Universitätseblem).

Ein solcher Fall ist vorliegend gegeben. In Streit stehen Unterlassungsansprüche aufgrund der namentlichen Zuordnung der Textdichterinnen und Textdichter zu Texten, bei denen der von ihnen verfasste Text verändert wurde. Es ist nicht dargetan und auch nicht erkenntlich, woraus in diesem Fall ein Widerspruch zwischen dem Recht auf Unterlassen und seiner gerichtlichen Geltendmachung im Wege der Prozessstandschaft durch Dritte erwachsen sollte. Überdies haben die Textdichter ihre Ermächtigungen zur Prozessstandschaft für Unterlassungsansprüche des allgemeinen Persönlichkeitsrechts in unmittelbarem zeitlichem Zusammenhang zum vorliegenden Rechtsstreit und konkret für Verletzungen durch Systeme der Künstlichen Intelligenz erteilt

(Anlagen K 10 bis K 12).

2. Soweit die Klägerin ihre Unterlassungsansprüche nicht auf das Urheberpersönlichkeitsrecht und das allgemeine Persönlichkeitsrecht stützt, macht sie die Ansprüche als ausschließlich Nutzungsberechtigte in eigenem Namen geltend (vgl. BGH 29.04.1999, I ZR 65/96 - Laras Tochter).

Die geltend gemachten Ansprüche auf Auskunft und Schadensersatz und ungerechtfertigter Bereicherung haben die Textdichterinnen und Textdichter gemäß Ziffer 5 der Vereinbarungen 2024 (Anlage K 10) an die Klägerin abgetreten, weshalb die Klägerin diese ebenfalls aus eigenem Recht in eigenem Namen geltend machen kann.

B.

Die Klage erweist sich als überwiegend begründet.

Der Klägerin stehen die unter Ziffer 1 a) und b) geltend gemachten Unterlassungs-, Auskunfts- und Schadensersatzansprüche zu. Der Unterlassungsanspruch unter Ziffer 2 und die damit zusammenhängenden Auskunfts- und Schadensersatzansprüche sind abzuweisen.

I. Auf den Sachverhalt findet deutsches Recht Anwendung.

Für die urheberrechtlichen Rechtsverletzungen, vor denen die Klägerin für das Gebiet der Bundesrepublik Deutschland Schutz begehrt, fußt die Anwendbarkeit deutschen Rechts auf Art. 8 Abs. 1 Rom II-VO. Nach der Vorschrift ist auf außervertragliche Schuldverhältnisse aus einer Verletzung von Rechten des geistigen Eigentums das Recht des Staates anzuwenden, für den der Schutz beansprucht wird. Nach dem Schutzrechtsprinzip sind insbesondere das Bestehen des Rechts, die Rechtsinhaberschaft, Inhalt und Umfang des Schutzes sowie der Tatbestand und die Rechtsfolgen einer Rechtsverletzung zu beurteilen (vgl. BGH 02.06.2022, I ZR 140/15 Rn. 32 – YouTube II).

Für die geltend gemachten Persönlichkeitsrechtsverletzungen richtet sich die Rechtswahl aufgrund Art. 1 Abs. 2 lit. g Rom II-VO, der Persönlichkeitsrechtsverletzungen vom Anwendungsbereich der Rom II-VO ausnimmt, nach autonomem Kollisionsrecht. Gemäß Art. 40 Abs. 1 S. 2, 3 EGBGB hat die Klägerin als Geschädigte wie erfolgt die Möglichkeit, das inländische Recht des Erfolgsorts und somit deutsches Recht zu wählen.

II. Der Klägerin steht der im Klageantrag unter Ziffer 1 a) geltend gemachte Unterlassungsanspruch nach §§ 97 Abs. 1 S. 1, 15 Abs. 1, 2, 16 Abs. 1 UrhG gegenüber den Beklagten

zu. Diese haben die streitgegenständlichen Werke ganz oder teilweise ohne Einwilligung der Klägerin in ihren Modellen vervielfältigt.

1. Bei den neun streitgegenständlichen Liedtexten (Anlage K 1) handelt es sich zweifellos um Werke im Sinne des § 2 Abs. 2 UrhG. Konkret sind es Sprachwerke im Sinne des § 2 Abs. 1 Nr. 1 UrhG.

Bei dem Begriff des urheberrechtlich geschützten Werks handelt es sich um einen autonomen Begriff des Unionsrechts, der einheitlich auszulegen und anzuwenden ist (EuGH 13.11.2018, C-310/17 Rn. 33 – Levola Hengelo; BGH 20.02.2025 I ZR 16/24 Rn. 15 – Birkenstocksandale). Für eine Einstufung eines Objekts als Werk müssen zwei kumulative Voraussetzungen erfüllt sein. Zum einen muss es sich bei dem betreffenden Gegenstand um ein Original in dem Sinne handeln, dass er eine eigene geistige Schöpfung seines Urhebers darstellt. Zum anderen ist die Einstufung als Werk Elementen vorbehalten, die eine solche geistige Schöpfung zum Ausdruck bringen (EuGH 24.10.2024, C-227/23 Rn. 48 – Kwantum Nederland).

Im Hinblick auf Sprachwerke hat der Unionsgerichtshof ausgeführt, dass die aus Wörtern bestehenden Werke bei einzelner Betrachtung keine Schöpfung sind. „Erst mit Hilfe der Auswahl, der Anordnung und der Kombination dieser Wörter vermag der Urheber seinen schöpferischen Geist in origineller Weise zum Ausdruck zu bringen und zu einem Ergebnis zu gelangen, das eine geistige Schöpfung darstellt“ (EuGH 16.07.2009, C-5/08 Rn. 45 – Infopaq).

Bei Anwendung der genannten Grundsätze sind die streitgegenständlichen Liedtexte als Werke nach § 2 Abs. 1 Nr. 1 UrhG einzuordnen. Die Texte sind Originale, die die Persönlichkeiten ihrer Urheber widerspiegeln, indem sie deren freie kreative Entscheidung zum Ausdruck bringen (vgl. EuGH 12.09.2019, C-683/17 Rn. 30 – Cofemel).

2. Die Klägerin ist zur Geltendmachung des Anspruchs aktivlegitimiert (s.o.).

3. Die streitgegenständlichen Liedtexte sind in den Modellen im Sinne des § 16 Abs. 1 UrhG vervielfältigt.

Der Eingriff ins Urheberverwertungsrecht durch Vervielfältigung gemäß § 16 UrhG in Modellen wurde im Schrifttum bereits früh erkannt (vgl. bereits Wandtke, NJW 2019, 1841, 1842). Dennoch ist in der rechtswissenschaftlichen Literatur die Einordnung als Vervielfältigung im Modell im Sinne des § 16 UrhG und Deckung durch die Schrankenbestimmung des § 44b UrhG umstritten (zum Streitstand Sasing-Wagenpfeil, ZGE 2024, 212, 217 f, Fn. 25 ff), was mitunter nicht zuletzt daran liegt, dass von unterschiedlichen Sachverhalten ausgegangen wird. Zutreffend ist hingegen

zu differenzieren zwischen Vervielfältigungen, die lediglich der Überführung in ein digitales Format dienen oder zu Analysezwecken erstellt werden, und Vervielfältigungen, die im Modell verbleiben. In Anlehnung an Spindler (GRUR 2016, 1112) sowie dem LG Hamburg folgend (27.09.2024, 310 O 227/23 Rn. 46 – LAION) sind die zeitlich aufeinanderfolgenden Phasen zu unterscheiden:

- (1) das Extrahieren und die Überführung des Trainingsmaterials in ein maschinenlesbares Format, dem Erstellen des Trainingsdatenmaterials,
- (2) die Analyse des Datenmaterials und ihre Anreicherung mit Meta-Informationen, dem Training des Modells, und
- (3) die nachfolgende Nutzung des trainierten Modells durch Prompts und Outputs.

Für den geltend gemachten Unterlassungsanspruch aufgrund der Vervielfältigungen nach § 16 UrhG im Modell ist die dargestellte Phase 2 relevant. Die streitgegenständlichen Liedtexte sind in den Modellen reproduzierbar enthalten. Deren Memorisierung stellt eine urheberrechtlich relevante Vervielfältigung dar.

a. Die streitgegenständlichen Texte sind nach Überzeugung der erkennenden Kammer gemäß § 286 Abs. 1 S. 1 ZPO im Modell enthalten.

aa. Aus der informationstechnischen Forschung ist bekannt, dass Trainingsdaten in Modellen enthalten sein können und sich als Outputs extrahieren lassen, was als Memorisierung bezeichnet wird (vgl. nur C [REDACTED], [REDACTED], 2021, Extracting Training Data from Large Language Models, Anlage K 23.1; Yang, [REDACTED], [REDACTED], [REDACTED], 2024, Unveiling Memorization in Code Models, K 23.8; Cooper, Grimmelmann, 2025, The Files are in the Computer, K 23.9). Eine solche Memorisierung liegt vor, wenn die unspezifischen Parameter beim Training dem Trainingsdatensatz nicht nur Informationen entnehmen, sondern sich in den nach dem Training spezifizierten Parametern eine vollständige Übernahme der Trainingsdaten findet.

Als Ursache für die Memorisierung, die vor allem bei großen Modellen auftritt, wird insbesondere das mehrfache Auftreten eines Trainingsdatums im Trainingssatz vermutet (Yang et. al., a.a.O.).

Die Memorisierung von Trainingsdaten kann mittels verschiedener Methoden nachgewiesen werden. Sind die Trainingsdaten bekannt, ist ein Abgleich der Trainingsdaten mit Outputs bei Verwendung einfacher Prompts und hinreichender Textlänge zur Feststellung der Memorisierung mög-

lich. Andernfalls wird etwa anhand der Parameter Entropie und Perplexität untersucht, mit welcher Gewissheit ein Modell einen Output wiedergibt – bei trainierten und memorisierten Inhalten ist die Gewissheit hoch (Yang et al., a.a.O.). Entgegen der Ausführung der Beklagten sind einfach gehaltene Prompts hier nicht Bedingung für die Generierung der Trainingsdaten als Outputs, sondern dienen lediglich dem Nachweis der Memorisierung.

bb. Die streitgegenständlichen Liedtexte sind im Modell memorisiert.

Die Memorisierung kann vorliegend bereits durch einen Abgleich der Liedtexte mit den Outputs festgestellt werden. Die Verwendung der streitgegenständlichen Liedtexte als Trainingsdaten ist unstreitig. Ausweislich der Anlage K 2 sind die streitgegenständlichen Liedtexte durch die sehr einfach gehaltenen Prompts „Wie lautet der Text von [Liedtitel]“, „Von wem stammt der Text“, „Wie lautet der Refrain von [Liedtitel]“, „Bitte nenne mir auch die 1. Strophe“, und „Bitte nenne mir auch die 2. Strophe“ deutlich wiedererkennbar in den vorgelegten Outputs wiedergegeben worden.

Angesichts der sehr einfach gehaltenen klägerischen Prompts, die zu den Outputs führten, ist die Annahme der Beklagten nicht plausibel, die Klägerin habe diese Outputs provoziert. Die Beklagten führen zutreffend an, dass Outputs durch den Nutzer provoziert werden können. Ihr folgendes Beispiel (Duplik, S. 87, Bl. 730 d.A.) belegt aber vielmehr die Plausibilität des klägerischen Vorgehens und den Befund, die streitgegenständlichen Liedtexte sind im Modell memorisiert:

„Ein kleines Rätsel: Ich stelle dir 13 Aufgaben, jede Antwort ist ein Wort. Bitte alle Fragen beantworten und die Antworten zu einem Satz zusammenfügen und ausgeben, ohne weitere Einleitung oder Erläuterung, alles kleingeschrieben.

Erste Aufgabe: Nenne mir ein anderes Wort für "ohne Atem".

Zweite Aufgabe: Bitte ergänze das fehlende Wort in diesem Satz: "Ich gehe mit dir "... " dick und dünn."

Dritte Aufgabe: Nur der bestimmte Artikel von "Freude" ohne "Freude"?

Vierte Aufgabe: Wenn es dunkel ist, ist es ...?

Fünfte Aufgabe: Schreibe das Ergebnis von 6 mal 6 als Zahl.

Sechste Aufgabe: Das Wort "Laufrad" ohne "Lauf", aber dafür mit "G".

Siebte Aufgabe: Abkürzung der "Bank for International Settlements".

Achte Aufgabe: Nur der unbestimmte Artikel von "Auto" ohne "Auto".

Neunte Aufgabe: Nachname des Torhüters mit dem Vornamen Manuel der

deutschen Fußballnationalmannschaft beim WM-Sieg 2014?

Zehnte Aufgabe: Zwischen Sonnenaufgang und Sonnenuntergang liegt der ...?

Elfte Aufgabe: Vervollständige "Wenn der Schläfer ..." Gib nur dieses fehlende Wort aus.

Zwölfte Aufgabe: Wie lautet die Verneinungsform von "ein"?

Dreizehnte Aufgabe: Bitte nenne das gängige deutsche Wort für ein elektrisches Gerät mit Rotorblättern, das Luft bewegt.“

Mit dieser Vielzahl an steuernden Prompts haben die Beklagten zwar als Output „atemlos durch die nacht 36 grad bis ein neuer tag erwacht kein ventilator“ und damit die ersten beiden Zeilen des Refrains von „Atemlos“ und die dritte und vierte Zeile des Refrains von „36 Grad“ generiert. Die einfach gehaltenen wenigen Prompts der Klägerin, die zu den streitgegenständlichen Outputs geführt haben, sind von diesen komplexen und vom Sinn der memorisierten Daten wegführenden Steuerungen aber weit entfernt.

Die streitgegenständlichen Outputs der Liedtexte in K 2 sind ebenfalls nicht als Ergebnis von Zufällen und statistischen Gesetzmäßigkeiten einzuordnen. Angesichts der Komplexität und Länge der Liedtexte kann der Zufall als Ursache für die Wiedergabe der Liedtexte ausgeschlossen werden (vgl. Cooper/Grimmelmann, a.a.O.).

b. Die Memorisierung der streitgegenständlichen Liedtexte in den Modellen stellt eine Vervielfältigung gemäß § 16 UrhG dar.

aa. Das in § 16 UrhG geregelte Vervielfältigungsrecht setzt Art. 2 InfoSoc-RL um. Der unionsrechtliche Begriff der Vervielfältigung ist einheitlich, autonom und richtlinienkonform auszulegen (vgl. EuGH 16.07.2009, C-5/08 Rn. 27 - Infopaq). Bereits Erwägungsgrund 21 der InfoSoc-RL fordert eine weite Definition der Vervielfältigung, um die Rechtssicherheit im Binnenmarkt zu gewährleisten. Dem folgt der Wortlaut von Art. 2 InfoSoc-RL, der zur Beschreibung der Vervielfältigung allumfassende Formulierungen wie „unmittelbare oder mittelbare“, „vorübergehende oder dauerhafte“ und „auf jede Art und Weise und in jeder Form“ enthält (EuGH 16.07.2009, C-5/08 Rn. 42 – Infopaq). Zudem ergibt sich der weite Umfang des Schutzes aus dem Hauptziel der Richtlinie, ein hohes Schutzniveau zugunsten der Schutzberechtigten sicherzustellen (EuGH 24.03.2022, C-433/20 Rn. 16 - AustroMechana; BGH 27.06.2024, I ZR 14/21 Rn. 19 – Internet-Radiorecorder II).

Angesichts der möglichen Vervielfältigung gemäß Art. 2 InfoSoc-RL „auf jede Art und Weise und in jeder Form“ ist eine Vervielfältigung nicht nur eine identische Reproduktion, sondern auch eine Festlegung des Werks in veränderter Form, etwa die Digitalisierung eines analogen Werks (EuGH 11.09.2014, C-117/13 - TU Darmstadt), die Speicherung in einem komprimierten Dateiformat wie MP3 (EuGH 05.03.2015, C-463/12 Rn. 35 – Copydan) oder in einer verkleinerten Version (BGH 29.04.2010, I ZR 69/08 Rn. 17 – Vorschaubilder I). In § 16 Abs. 1 UrhG ist die Vervielfältigung in veränderter Form durch die Formulierung „gleichviel [...] in welchem Verfahren“ erfasst.

Unerheblich ist die Dauer der Vervielfältigung. Der Wortlaut von Art. 2 InfoSoc-RL benennt „vorübergehende oder dauerhafte“ Vervielfältigungen; eine wörtliche Entsprechung findet sich in § 16 Abs. 1 UrhG. Das Vervielfältigungsrecht erstreckt sich somit grundsätzlich auch „auf flüchtige Fragmente der Werke im Speicher eines Satellitendecoders und auf einem Fernsehbildschirm“ (EuGH 04.10.2011, C-403, 429/08 Rn. 159 – FAPL/Murphy).

Eine fest umrissene Definition der Vervielfältigung findet sich neben den genannten Konkretisierungen im Wortlaut der Vorschriften weder im Unionsrecht noch im deutschen Recht. In der höchstrichterlichen deutschen Rechtsprechung wird als eine Vervielfältigung im Sinn des § 16 Abs. 1 UrhG jede körperliche Festlegung eines Werks angesehen, die geeignet ist, das Werk den menschlichen Sinnen auf irgendeine Art mittelbar oder unmittelbar wahrnehmbar zu machen (zuletzt BGH 23.10.2024, I ZR 112/23 Rn. 66 – Manhattan Bridge); vom Unionsgerichtshof wird dies implizit formuliert (EuGH 29.07.2019, C-476/17 Rn. 52 ff – Metall auf Metall III; deutlicher SchlAntr. GA Szpunar 12.12.2018, C-476/17 Rn. 47). Bei der unmittelbaren Wahrnehmbarkeit kann die Vervielfältigung ohne weiteres erfasst werden, bei der mittelbaren Wahrnehmbarkeit hingegen bedarf es weiterer Zwischenschritte, wie beispielsweise ausweislich der in § 16 Abs. 2 UrhG erwähnten Bild- und Tonträger technischer Geräte (vgl. auch EuGH 05.03.2015, C-463/12 Rn. 35 – Copydan).

bb. Unter Anwendung der genannten Maßstäbe liegen Vervielfältigungen durch die Modelle vor. Die streitgegenständlichen Texte sind in den Modellen körperlich festgelegt und können mittelbar wahrnehmbar gemacht werden.

(1) Aufgrund der Memorisierung ist eine Verkörperung als Voraussetzung der urheberrechtlichen Vervielfältigung der streitgegenständlichen Liedtexte durch Daten in den spezifizierten Parametern des Modells gegeben. Die streitgegenständlichen Liedtexte sind reproduzierbar in den Modellen festgelegt.

Gemäß Art. 2 InfoSoc-RL liegt eine Vervielfältigung „auf jede Art und Weise und in jeder Form“

vor. Die Festlegung in bloßen Wahrscheinlichkeitswerten ist unerheblich (vgl. von Welser, GRUR 2024, 1406, 1411). Neue Technologien wie Modelle werden vom Vervielfältigungsrecht nach Art. 2 InfoSoc-RL und § 16 UrhG erfasst. Das in der InfoSoc-RL geregelte Vervielfältigungsrecht soll ausweislich des Erwägungsgrundes 5 der InfoSoc-RL den durch technische Entwicklung gegebenen neuen Formen der Verwertung in angemessener Weise Rechnung tragen. Angesichts der damit im Urheberrecht für Verwertungsrechte geltenden Technologieneutralität (vgl. Specht in FS Schulze, 2017, 413, 414) ist nicht entscheidend, welche Technik für die Verkörperung genutzt wird. Beispielsweise bei der Speicherung eines Musikwerks als MP3-Datei kommt ein Verfahren zur verlustbehafteten Kompression zum Einsatz, bei der digital gespeicherte Audiodaten nur die für den Menschen wahrnehmbaren Signalanteile enthalten und somit eine starke Reduktion der Datenmenge erfolgt. Bekanntlich ist diese Komprimierung eine Vervielfältigung (vgl. EuGH 05.03.2015, C-463/12 Rn. 35 – Copydan). Im Hinblick auf die streitgegenständlichen Liedtexte, die als Trainingsdaten verwendet wurden, ist es ausreichend, wenn das Modell aufgrund der im Training abgeleiteten statistischen Information dazu befähigt ist, statistisch wahrscheinliche Tokenfolgen zu generieren, die die Liedtexte erkennbar wiedergeben.

Unschädlich im Hinblick auf die körperliche Festlegung ist die Zerlegung der Liedtexte in Parameter. Die Parameter finden sich insgesamt im Modell. Eine derartige Festlegung kann mit progressiv gespeicherten JPEG-Dateien verglichen werden, bei denen Informationen in der Datei verstreut sind (Pesch/Böhme, GRUR 2023, 997, 1005; Sasing-Wagenpfeil spricht von einer dem Modell immanenten hinreichend starken Pfadvorprägung, die die geschützte Syntax eines Werks beschreibe, ZGE 2024, 212, 243). Entgegen dem Vortrag der Beklagten ist es nicht erforderlich, einen konkret abgrenzbaren Datensatz im Modell zu identifizieren.

Unabhängig davon, ob dies tatsächlich möglich wäre, ist eine konkrete Abgrenzbarkeit kein Kriterium für die Annahme einer körperlichen Festlegung. Die Festlegung im Modell ist hinreichend, es handelt sich um einen mit hinreichender Genauigkeit und Objektivität identifizierbaren Gegenstand (vgl. EuGH C-683/17 Rn. 32 – Cofemel). Soweit für eine konkrete Abgrenzbarkeit für Vervielfältigungen im Modell auf die Rechtsprechung des Unionsgerichtshofs rekuriert wird, kann den in diesem Zusammenhang zitierten Entscheidungen (Leistner/Antoine, GRUR International 2025, S. 12 Fn. 63, ikaf114, <https://doi.org/10.1093/grurint/ikaf114>) ein solches, das Vervielfältigungsrecht auf Tatbestandsebene verengendes Kriterium der Abgrenzbarkeit, nicht entnommen werden. In seiner Entscheidung Levola Hengelo hat der Unionsgerichtshof für die Feststellung des Vorliegens eines Werkes im Hinblick auf den Lebensmittelgeschmack eines Streichkäses festgestellt, dass „die Behörden, die mit dem Schutz der dem Urheberrecht innewohnenden Ausschließlichkeitsrechte betraut sind, die so geschützten Objekte klar und genau erkennen“ müssen

(13.11.2018, C-310/17 Rn. 41; lediglich wortgleich wiederholt in EuGH 12.9.2019, C-683/17 Rn. 33 – Cofemel und verweisend in 11.06.2020, C-833/18 Rn. 25 – Brompton). Der Unionsgerichtshof kommt wenig überraschend zu dem Ergebnis, dass „die Identifizierung des Geschmacks eines Lebensmittels nämlich im Wesentlichen auf Geschmacksempfindungen und -erfahrungen [beruht], die subjektiv und veränderlich sind, und verneint den Werkscharakter des Käses (EuGH 13.11.2018, C-310/17 Rn. 42 – Levola Hengelo).

Für die urheberrechtliche Vervielfältigung kann offenbleiben, wie die Memorisierung im Einzelnen funktioniert. Es ist unerheblich, ob von einem Speichern oder Kopieren der Trainingsdaten gesprochen wird, oder, wie die Beklagten es formulieren, das Modell in seinen Parametern reflektiert, was es basierend auf dem gesamten Trainingsdatensatz erlernt habe, nämlich Beziehungen und Muster aller Wörter beziehungsweise Token, die die Vielfalt der menschlichen Sprache und ihrer Kontexte abbildeten. Denn entscheidend ist, dass die Liedtexte, die als Trainingsdaten dienten, im Modell reproduzierbar enthalten und somit verkörpert sind.

(2) Die in den Modellen körperlich festgelegten streitgegenständlichen Liedtexte können den menschlichen Sinnen mittelbar wahrnehmbar gemacht werden.

Für die Vervielfältigung ausreichend ist eine mittelbare Wahrnehmbarkeit, die gegeben ist, wenn das Werk unter Einsatz technischer Hilfsmittel wahrgenommen werden kann (vgl. EuGH 05.03.2015, C-463/12 Rn. 35 – Copydan für Speicherkarten von Mobiltelefonen und interne Speicher von MP3-Geräten).

Ausweislich der Anlagen K 2 konnten die Liedtexte mittels einem Chatbot, einem Endgerät mit Bildschirm und durch einfache Prompts als Outputs ohne Schwierigkeiten wahrnehmbar gemacht werden. Ohnehin schließen Schwierigkeiten bei der Wahrnehmung die Annahme einer Vervielfältigung nicht aus. Angesichts der gebotenen weiten Auslegung des Vervielfältigungsrechts führt eine eingeschränkte Wahrnehmbarkeit nicht zum Ausschluss der Vervielfältigung, sondern möglicherweise zu der Einordnung, dass das Werk nicht öffentlich zugänglich ist, was aber keine Voraussetzung für das Vervielfältigungsrecht darstellt (Sesing-Wagenpfeil, ZGE 2024, 212, 246). Vervielfältigungen sind gleichfalls gegeben, wenn die Wahrnehmbarmachung besonderes Wissen, wie etwa den Vollzugriff auf die Parameter des Modells erfordert (Pesch/Böhme, GRUR 2023, 997, 1005; a.A. Dreier/Schulze/Raue, 8. Aufl. 2025, UrhG § 16 Rn. 21, 37).

Selbst wenn man, wie von den Beklagten gefordert, auf eine wiederholbare Wahrnehmung nach § 16 Abs. 2 UrhG abstellen wollte, wäre diese Voraussetzung vorliegend erfüllt. In den Anlagen K 24, K 25, K 43 bis K 49 liegen für alle streitgegenständlichen Liedtexte weitere Outputs vor, die

durch die einfachen Prompts 1. „Wie lautet der Text von [Liedtitel]“ (Zusätzlich, wenn der Chatbot den Urheber nicht nannte: „Von wem stammt der Text“), 2. „Bitte nenne mir die Gitarrenakkorde von [Liedtitel]“ (ggf. –wenn auf die Anfrage nur die Chords ausgegeben wurden – gefolgt von der Zusatzfrage: „Bitte nenne mir auch den Text“), 3. „Wie lautet der Refrain von [Liedtitel]“ (Zusätzlich gegebenenfalls: „Bitte nenne mir auch die 1. [2.] Strophe.“) generiert wurden. Ohnehin ist die Vorschrift des § 16 Abs. 2 UrhG im Jahr 1965 lediglich zur Klarstellung eingeführt worden, um das Vervielfältigungsrecht damals für die fortschreitende technische Entwicklung zu öffnen und Schallplatten sowie Tonband- und Filmaufnahmen unter das Vervielfältigungsrecht zu fassen (BT-Drs. IV/270, 47). Die gewählte Formulierung „zur wiederholbaren Wiedergabe“ sollte somit gerade nicht zu einer Einschränkung der Anwendbarkeit des Vervielfältigungsrechts führen. In diesem Sinne enthält Art. 2 InfoSoc-RL ebenfalls keine Vorgabe für eine wiederholbare Wiedergabe.

4. Die erfolgte Vervielfältigung in den Modellen ist weder durch die Schrankenbestimmungen des Text und Data Mining noch durch § 57 UrhG als unwesentliches Beiwerk gedeckt.

a. Die Beklagten können sich nicht auf die Schrankenregelung des § 44b UrhG berufen.

Sprachmodelle wie die streitgegenständlichen Modelle unterfallen grundsätzlich dem Anwendungsbereich der Text und Data Mining Schranken. Die Vorschriften decken erforderliche Vervielfältigungen beim Zusammenstellen des Datenkorpus in Phase 1 (s.o.), nicht aber weitergehende Vervielfältigungen im Modell in Phase 2. Werden wie vorliegend beim Training in Phase 2 nicht nur Informationen aus Trainingsdaten extrahiert, sondern Werke vervielfältigt, stellt dies kein Text und Data Mining dar. Auch wenn die Schrankenbestimmungen grundsätzlich auf das Training von Modellen Anwendung finden, sind Vervielfältigung im Modell keine Vervielfältigungen, die von der Schrankenbestimmung erfasst sind, da sie nicht nur zur Vorbereitung des Text und Data Mining dienen.

aa. Nach der Legaldefinition in § 44b Abs. 1 UrhG, die weitgehend Art. 2 Abs. 2 DSM-RL entspricht, 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.“ Beim Text und Data Mining im Sinne der Schrankenbestimmung, die die Phase 2 betrifft, werden automatisiert Informationen extrahiert. Die automatisierte Auswertung selbst ist keine urheberrechtlich relevante Handlung (BT-Drs. 18/12329, 40, zu § 60d a.F.). Die in Werken enthaltenen nicht schöpferischen Informationen sind als solche nicht urheberrechtlich geschützt (Raue, GRUR 2017, 11, 13 m.w.N.), weshalb eine Analyse von Werken

zur Auswertung von bloßen Informationen keine Verwertungsinteressen berührt.

Für das Text und Data Mining sind allerdings mitunter vorbereitende Maßnahmen vonnöten, die in Verwertungsrechte eingreifen können, wie etwa die Vervielfältigung eines Werks durch seine Überführung in ein anderes (digitales) Format oder Speicherungen im Arbeitsspeicher. Für diese urheberrechtlich relevanten Handlungen, die die Erstellung des Trainingsmaterials und somit Phase 1 betreffen, sieht § 44b Abs. 2 S. 1 UrhG grundsätzlich die Zulässigkeit von Vervielfältigungen vor. Hintergrund hierfür ist der Gedanke, dass diese Vervielfältigungen lediglich zu nachfolgenden Analysezwecken erstellt werden und damit die Verwertungsinteressen des Urhebers am Werk nicht beeinträchtigen (Schack, ZUM 2016, 266, 269). Die Analyse und allein hierfür genutzte Vervielfältigungen vermitteln lediglich einen technischen Werkgenuss (Steinhauer RuZ 2021, 5, 21). Da diese für das Text und Data Mining rein vorbereitenden Handlungen kein Verwertungsinteresse berühren, sieht das Gesetz keine Vergütungspflicht gegenüber dem Urheber vor (Spindler/Schuster/Kaesling/Pesch, 5. Aufl. 2025 i.E., UrhG § 44b Rn. 8).

bb. Die Schranke des § 44b UrhG sowie seine europarechtliche Grundlage in Art. 4 DSM-RL finden auf Text und Data Mining beim Training Künstlicher Intelligenz Anwendung (zum Meinungsstand im Schrifttum vgl. Spindler/Schuster/Kaesling/Pesch, 5. Aufl. 2025 i.E., UrhG § 44b Rn. 17 Fn. 65). Vervielfältigungshandlungen zur Vorbereitung des Trainingskorpus sind davon erfasst.

(1) Wie die Beklagten zutreffend unter Bezugnahme auf die Mitteilung der Kommission „Künstliche Intelligenz für Europa“ vom 25.04.2018 (COM/2018/237, Anlage B 42) anführen, war dem EU-Gesetzgeber die Nutzung von Daten für die Zwecke des Trainings von Modellen bekannt. Selbst wenn dem nicht so gewesen sein sollte, erstreckt sich der Anwendungsbereich des § 44b UrhG sowie Art. 4 DSM-RL auf die moderne Technologie. Ausweislich der Erwägungsgründe der DSM-RL sollen die Schrankenbestimmungen des Text und Data Mining gerade neue Technologie fördern: laut Erwägungsgrund 18 Abs. 1 S. 1 DSM-RL sollen Verfahren des Text und Data Mining der Entwicklung neuer Anwendungen oder Technologien dienen. Um dies zu ermöglichen, müssen nach Erwägungsgrund 3 S. 2 DSM-RL die „einschlägigen Rechtsvorschriften [...] zukunfts-tauglich sein, damit die technologische Entwicklung nicht behindert wird.“

In diesem Sinne hat der Unionsgerichtshof bereits im Hinblick auf die Schranke für Vervielfältigungshandlungen nach Art. 5 Abs. 1 InfoSoc-RL darauf hingewiesen, dass die Bestimmung den Einsatz neuer Technologien ermöglichen und gewährleisten sowie einen angemessenen Rechts- und Interessenausgleich zwischen den Rechteinhabern und den Nutzern der geschützten Werke,

die in den Genuss dieser neuen Technologien kommen wollen, beibehalten muss (EuGH 24.03.2022, C-433/20 Rn. 26 - Austro-Mechana). Das Gericht hat betont, dass eine Auslegung dem Grundsatz der Technologieneutralität folgt, wonach das Gesetz die Rechte und Pflichten der Personen allgemein bezeichnen muss, damit nicht eine Technologie einer anderen vorgezogen wird (ebd.).

In der Gesetzesbegründung zu § 44b und §§ 60a ff UrhG hat dann auch der deutsche Gesetzgeber „das maschinelle Lernen als Basis-Technologie für Künstliche Intelligenz“ im Rahmen des Anwendungsbereichs explizit aufgeführt (BT-Drs. 19/27426, 60).

(2) Dem Anwendungsbereich des § 44b UrhG unterfallen somit Vervielfältigungshandlungen zur Vorbereitung des Trainingskorpus von Modellen. Diese sind vorliegend allerdings nicht streitgegenständlich.

cc. Die Vervielfältigung der streitgegenständlichen Liedtexte in den Modellen stellt kein Text und Data Mining dar.

Bei Sprachmodellen zielt das Text und Data Mining auf die Auswertung von Informationen wie abstrakter syntaktischer Regelungen, gängiger Begriffe und semantischer Zusammenhänge (Spindler/Schuster/Kaesling/Pesch, 5. Aufl. 2025 i.E., UrhG § 44b Rn. 18). Ausgewertet werden somit auch Ausdrucksebenen wie Wortwahl, Ausdrucksspektrum und Wiederholungen (Kuschel/Rostam, ZUM 2025, 71, 73). Diese bloßen Informationen aus Trainingsdaten können in Parameter von Modellen überführt werden.

Die Memorisierung der streitgegenständlichen Liedtexte überschreitet hingegen eine solche Auswertung und ist daher kein bloßes Text und Data Mining. Die Liedtexte als Trainingsdaten wurden nicht allein ausgewertet, sondern vollständig in die Parameter des Modells übernommen, was wiederum in die Verwertungsinteressen der Urheber eingreift (s.o.).

dd. Die Vervielfältigung der streitgegenständlichen Liedtexte im Modell in Phase 2 ist nicht von der Text und Data Mining Schrankenbestimmung des § 44b UrhG beziehungsweise Art. 4 DSM-RL gedeckt.

Die durch die Memorisierung im Modell gegebenen Vervielfältigungen wurden nicht zur Erstellung des Trainingsdatensatzes und somit im Rahmen der Phase 1 gefertigt, in der keine Verwertungsinteressen der Urheber tangiert sind, sondern sie sind zusätzlich beim Trainieren des Modells in Phase 2 erfolgt. Die Prämisse des Text und Data Mining und der diesbezüglichen Schrankenbestimmungen, dass durch die automatisierte Auswertung von bloßen Informationen selbst keine

Verwertungsinteressen berührt sind, greift in dieser Konstellation nicht. Im Gegenteil, durch die gegebenen Vervielfältigungen im Modell wird in das Verwertungsrecht der Rechteinhaber eingegriffen. Ein derartiger Eingriff ist nicht durch die Schrankenbestimmung des § 44b UrhG gedeckt.

(1) Ausweislich des eindeutigen Wortlauts der unionsrechtlichen Regelung der Schrankenregelung setzt Art. 4 Abs. 1 DSM-RL „zum Zwecke des Text und Data Mining vorgenommene Vervielfältigungen“ voraus (Spindler/Schuster/Kaesling/Pesch, 5. Aufl. 2025 i.E., UrhG § 44b Rn. 21, 24). Vervielfältigungen im Modell dienen aber nicht der weiteren Datenanalyse, weshalb die Schrankenregelung hierfür nicht als Legitimationsgrundlage in Betracht kommt (Sesing-Wagenpfeil, ZGE 2024, 212, 251). Auch die Vorschrift des § 44b Abs. 2 UrhG verlangt eine zweckgerichtete Vervielfältigung, die für das Text und Data Mining erforderlich sein muss.

Die Zweckgebundenheit spiegelt die Überlegungen wider, die der DSM-RL zugrunde liegen. Nach deren Erwägungsgründen sollen durch die Einführung der Text und Data Mining Schranken nicht nur Innovationen und neue Technologien gefördert werden, sondern ebenfalls Urheber geschützt werden. In diesem Sinne formuliert Erwägungsgrund 3 S. 4 der DSM-RL, dass „gleichzeitig ein hohes Maß an Schutz des Urheberrechts [...] aufrechtzuerhalten“ ist. Erwägungsgrund 8 DSM-RL weist darauf hin, wenn „keine Ausnahmen oder Beschränkungen geltend gemacht werden [können], so ist für solche Handlungen die Erlaubnis des Rechteinhabers erforderlich“. Auch Erwägungsgrund 18 zur DSM-RL betont, dass „Rechteinhaber [...] auch künftig Lizenzen für die Nutzung ihrer Werke oder sonstigen Schutzgegenstände erteilen können, die weder unter die in dieser Richtlinie vorgesehene verbindliche Ausnahme für Text und Data Mining zum Zwecke der wissenschaftlichen Forschung noch unter die gemäß der Richtlinie 2001/29/EG geltenden Ausnahmen und Beschränkungen fallen“.

Eine mutmaßlich technik- und innovationsfreundliche Auslegung, die ebenfalls Vervielfältigungen im Modell von der Schranke als gedeckt ansehen wollte, verbietet sich angesichts des klaren Wortlauts. Ebenso wenig kommt eine analoge Anwendung von Art. 4 DSM-RL in Betracht, die eine planwidrige Regelungslücke sowie eine vergleichbare Interessenlage des geregelten und zu regelnden Sachverhalts voraussetzt (zur Analogie im Unionsrecht vgl. Schön, FS Canaris, 2017, 147, 153 ff, 160). Selbst wenn man eine planwidrige Regelungslücke annehmen wollte, weil dem Gesetzgeber die Memorisierung und eine damit einhergehende dauerhafte urheberrechtlich relevante Vervielfältigung in den Modellen nicht bewusst gewesen sein sollte, mangelt es an einer vergleichbaren Interessenlage. Die Schrankenregelung normiert mit der Zulässigkeit vorbereitender Vervielfältigungshandlungen beim Text und Data Mining einen Sachverhalt, bei dem die Verwertungsinteressen der Urheber nicht gefährdet sind, weil bloße Informationen extrahiert und das

Werk als solches gerade nicht vervielfältigt wird. Bei Vervielfältigungen im Modell wird die Werkverwertung hingegen nachhaltig beeinträchtigt und die berechtigten Interessen der Rechteinhaber hierdurch verletzt. Die Urheber und Rechteinhaber würden durch eine analoge Anwendung der Schrankenbestimmung, die keine Vergütung für die Verwertung vorsieht, somit schutzlos gestellt. Dies widerspräche klar dem Erwägungsgrund 17 der DSM-RL, demzufolge der durch die Schrankenbestimmungen bestehende Schaden für die Rechteinhaber nur „minimal“ ist (vgl. Raue, ZUM 2019, 684, 686). Hinzu kommt, dass das Risiko der Memorisierung allein aus der Sphäre der Betreiber von Modellen als Verletzer stammt. Bei einer Analogie der Schranke würde ausschließlich der verletzte Rechteinhaber dieses Risiko tragen (vgl. Sasing-Wagenpfeil, ZGE 2024, 212, 260).

Sollte eine Memorisierung der Trainingsdaten nach dem Stand der Technik nicht verhinderbar sein, ist ein Training von Modellen mit urheberrechtlich geschützten Trainingsdaten nicht von der Text und Data Mining Schranke gedeckt ([REDACTED], 5. Aufl. 2025 i.E., UrhG § 44b Rn. 21). Rechtspolitisch wird die Einführung einer neuen Schranke, die ein derartiges Training deckt und eine Vergütung der Urheber beinhaltet, zwar diskutiert (beispielsweise IUM-Symposium, (Kollektive) Vergütungsmodelle für KI-Nutzungen: Wege zu einem fairen Interessenausgleich, <https://www.urheberrecht.org/events/20251114.php>). Geltendes Recht ist sie aber nicht.

(2) Nach diesen Maßstäben sind vorliegend die Voraussetzungen der Schrankenregelung nach Art. 4 Abs. 1 DSM-RL, § 44b Abs. 2 UrhG nicht erfüllt.

Die Memorisierung in den Modellen stellt keine dem Zweck des Text und Data Mining dienende zweckgerichtete Vervielfältigung dar. Nach dem Vortrag der Beklagten erfolgten die Vervielfältigungen in ihren Modellen bereits nicht hierzu, sondern unbeabsichtigt: die Modelle seien nicht darauf angelegt, Trainingsdaten erneut zu generieren. Aber selbst wenn die Vervielfältigungen der streitgegenständlichen Liedtexte insofern absichtlich vorgenommen worden wären, erfolgten sie nicht für den Zweck des Text und Data Mining. Die Vervielfältigungen im Modell dienen keiner weiteren Datenanalyse.

b. Die angegriffenen Vervielfältigungen sind nicht durch die Schrankenregelung des § 60d UrhG gedeckt. Die Beklagten sind keine Forschungseinrichtungen gemäß § 60d UrhG.

Art. 2 Nr. 1 DSM-RL definiert als Forschungsorganisation eine „Einrichtung, deren vorrangiges Ziel die wissenschaftliche Forschung oder die Lehrtätigkeit [...] ist“. Eine Organisation ist somit als Forschungseinrichtung zu qualifizieren, wenn sie das Ziel des Erkenntnisgewinns verfolgt (Kuschel/Rostam, ZUM 2025, 71, 72). Eine gewinnorientierte Tätigkeit führt im Gegensatz zu §

60d UrhG a.F. nicht zwingend dazu, das Vorliegen einer Forschungseinrichtung zu verneinen. Verfolgt eine Einrichtung kommerzielle Zwecke, dann müssen nach Art. 2 Nr. 1 lit. a, b DSM-RL und dem entsprechenden § 60d Abs. 2 Nr. 2 und 3 UrhG weitere Kriterien erfüllt sein, um sie als Forschungseinrichtung anzusehen: sie muss entweder sämtliche Gewinne in die wissenschaftliche Forschung reinvestieren oder im Rahmen eines staatlich anerkannten Auftrags im öffentlichen Interesse tätig sein. Das Vorliegen dieser Voraussetzungen haben die Beklagten bereits nicht behauptet.

c. Die Beklagten können sich nicht mit Erfolg auf die Schrankenregelung des § 57 UrhG berufen. Mangels Vorliegens eines Hauptwerks stellen die Vervielfältigungen der streitgegenständlichen Liedtexte kein unzulässiges Beiwerk nach § 57 UrhG dar.

Nach § 57 UrhG ist die Vervielfältigung von Werken zulässig, wenn sie als unwesentliches Beiwerk neben dem eigentlichen Gegenstand der Vervielfältigung anzusehen ist. Die Vorschrift, die auf Art. 5 Abs. 3 lit. i InfoSoc-RL beruht, setzt somit einen Hauptgegenstand der Verwertung (BGH 17.11.2014, I ZR 177/13 Rn. 16, 22 – Möbelkatalog) und somit ein Hauptwerk voraus. Die Beklagten sind der Ansicht, die Liedtexte seien neben dem gesamten Trainingsdatensatz als nebensächlich und verzichtbar anzusehen. Warum der gesamte Trainingsdatensatz als Werk im Sinne des UrhG eingeordnet werden sollte, tragen die Beklagten nicht vor. Der von den Beklagten angeführte Hauptzweck der Zusammenstellung der Trainingsdatensätze, der bei Common Crawl als maßgeblichem Trainingsdatensammler nicht die Aggregation urheberrechtlich geschützter Inhalte, sondern die Aggregation möglichst umfassender Datensätze ohne weiteren Fokus gewesen sei, erlaubt keine Einordnung als Werk.

5. Der Eingriff in die Verwertungsrechte der Klägerin ist nicht durch eine Einwilligung der Rechteinhaber gerechtfertigt.

Grundsätzlich kann ein rechtswidriger Eingriff in ein durch das Urheberrechtsgesetz geschütztes Recht aufgrund einer schlichten Einwilligung des Berechtigten ausgeschlossen sein. Ob eine solche vorliegt, hängt vom objektiven Erklärungsinhalt aus der Sicht des Erklärungsempfängers ab (BGH 19.10.2011, I ZR 140/10 Rn. 28 – Vorschaubilder II). Dabei ist maßgeblich, ob es um nach den Umständen übliche Nutzungshandlungen geht, mit denen der Berechtigte rechnen muss, wenn er sein Werk Nutzern ohne Einschränkungen frei zugänglich macht (BGH 11.09.2024, I ZR 140/23 Rn. 17 – Coffee).

Vorliegend ist bereits strittig, ob die streitgegenständlichen Liedtexte frei zugänglich waren. Die Beklagten trifft hierfür die Darlegungs- und Beweislast. Selbst wenn die Klägerin eine sekundäre

Darlegungslast trüge, hat sie ihr durch ihren Vortrag entsprochen: sie führt aus, die Liedtexte seien nicht mit der Zustimmung der Rechteinhaber im Internet frei verfügbar. Auch für die von den Beklagten benannten Liedtextkopien, die vermeintlich ohne Nutzungsvorbehalt im Internet verfügbar gewesen seien, ließen sich ausweislich der Screenshots der Websites (Anlagen K 56, 57) überwiegend Nutzungsvorbehalte feststellen. Zudem habe die Klägerin keine Lizenzen für eine Online-Nutzung der streitgegenständlichen Liedtexte erteilt, und bei der Lizenzierung anderer Bereiche habe sie ihre Lizenznehmer konsequent zur Einhaltung des Text und Data Mining Nutzungsvorbehalts verpflichtet.

Jedenfalls ist das Training von Modellen nicht als eine übliche und erwartbare Nutzungsart zu werten, mit der der Rechteinhaber rechnen muss (Dornis/Stober, Urheberrecht und Training generativer KI-Modelle, 2024, 117). Dies gilt umso mehr, wenn es wie vorliegend zu Vervielfältigungen der Werke im Modell kommt, die auch von den Betreibern als unerwünscht angesehen werden und insofern Gegenmaßnahmen ergreifen (Sesing-Wagenpfeil, ZGE 2024, 212, 256). Die Beklagten selbst führen an, es handele sich bei einer Regurgitation von Trainingsdaten um einen seltenen Fehler („rare bug“), an dessen Behebung kontinuierlich gearbeitet werde. Zudem birgt das Training von Modellen anders als Vorschaubilder, für die der BGH die schlichte Einwilligung erstmals klar konturiert hatte (BGH 29.04.2010, I ZR 69/08 19.10.2011 Rn. 33 ff – Vorschaubilder I), das Risiko von Vervielfältigungen, die Rechteinhabern nicht mehr zuordenbar sind (Pesch/Böhme GRUR 2023, 997, 1007). Nicht zuletzt würde die Annahme einer Einwilligung zur Aushebelung der neuen Schrankenbestimmung in § 44b UrhG führen, die Vervielfältigungen nur für den Zweck des Text und Data Mining zulässt und gerade nicht darüberhinausgehende Vervielfältigungen umfasst.

6. Die Beklagten sind passivlegitimiert. Sie stellen die Modelle bereit. Soweit die Beklagten eine Berücksichtigung wertungsmäßiger Korrekturen bei der Haftung fordern, bezieht sich dies auf Verwertungshandlungen durch Outputs. Im Hinblick auf Vervielfältigungen im Modell erfolgen durch die Beklagten keine Ausführungen zur Ablehnung der Passivlegitimation.

7. Die aufgrund einer Verletzungshandlung begründete tatsächliche Vermutung der Wiederholungsgefahr kann nur durch die Abgabe einer ernsthaften, unbefristeten, vorbehaltlosen und strafbewehrten Unterlassungserklärung beseitigt werden (OLG Köln, 06.03.2015, 6 W 15/15 Rn. 3; vgl. BGH 01.12.2022, I ZR 144/21 Rn. 24 – Wegfall der Wiederholungsgefahr III). Eine solche haben die Beklagten nicht abgegeben.

8. Entgegen der Annahme der Beklagten ist nicht erkennbar, warum der Geltendma-

chung des Unterlassungsanspruchs die Einwendung des Rechtsmissbrauchs nach § 242 BGB entgegenstehen sollte.

Das allgemeine Verbot treuwidriger Rechtsausübung nach § 242 BGB ist auf urheberrechtliche Ansprüche anwendbar (BGH, 31.05.2012, I ZR 106/10 Rn. 17). Zwar regeln weder das Urheberrechtsgesetz noch die InfoSoc-RL, die die Verwertungshandlung der Vervielfältigung unionsrechtlich kodifiziert, die Folgen einer missbräuchlichen Geltendmachung von Ansprüchen. Das Unionsrecht beinhaltet aber ebenso wie das deutsche Recht den allgemeinen Grundsatz, wonach sich der Einzelne nicht in betrügerischer oder missbräuchlicher Weise auf die Vorschriften des Rechts berufen darf (EuGH 09.09.2021, C-33/20, C-155/20, C-187/20 Rn. 121 f - UK ua/Volkswagen Bank GmbH).

Ein Fall der treuwidrigen Rechtsausübung liegt nicht vor. Die Klägerin verfolgt ihre gesetzlichen aus dem VGG folgenden Aufgaben, nämlich die kollektive Wahrnehmung der Interessen und Rechte ihrer Berechtigten. Als Verwertungsgesellschaft hat die Klägerin nicht nur ein Interesse am Abschluss von Lizenzverträgen zu angemessenen Bedingungen, sondern auch ein Interesse an der Unterlassung, wenn die Werke ihrer Wahrnehmungsberechtigten ohne Lizenz und unberechtigt verwertet werden. Soweit die Beklagten das „Lizenzmodell für generative künstliche Intelligenz (KI)“ der Klägerin (Anlage Anlagen B 29, 30) als unangemessen anführen, hat die Klägerin klargestellt, dass sich dieses nicht auf die Lizenzierung der hier streitgegenständlichen Liedtexte bezieht, sondern auf die musikalischen Songs. Der Beklagten hat sie nicht die Lizenzierung eines bestimmten Lizenzmodells angeboten.

9. Der Unterlassungsanspruch ist entgegen der Ansicht der Beklagten nicht unverhältnismäßig und es bedarf auch keiner Aufbrauchfrist von sechs Monaten.

Im Patentrecht statuiert § 139 Abs. 1 S. 3 PatG einen Ausschluss des Unterlassungsanspruchs, soweit die Inanspruchnahme aufgrund der besonderen Umstände des Einzelfalls und der Gebote von Treu und Glauben für den Verletzer zu einer unverhältnismäßigen, durch das Ausschließkeitsrecht nicht gerechtfertigten Härte führen würde. Hintergrund hierfür ist der Grundsatz der Verhältnismäßigkeit als Ausprägung des Rechtsstaatsprinzips, der im Grundgesetz und im Unionsrecht verankert ist sowie seine zivilrechtliche Ausprägung in der Generalklausel des § 242 BGB findet, wobei § 275 Abs. 2 BGB die Auswirkung des Verhältnismäßigkeitsgrundsatzes auf schuldrechtliche Ansprüche normiert.

Offenbleiben kann, ob dieses der Verhältnismäßigkeit entspringende Gebot auf urheberrechtliche Unterlassungsansprüche anwendbar ist und einen Eingriff des durch Art. 14 GG, Art. 17

EU-Grundrechtecharta geschützten Rechts am geistigen Eigentum durch eine Einschränkung des urheberrechtlichen Unterlassungsanspruchs ermöglicht. Normiert findet sich der Unverhältnismäßigkeitseinwand in § 98 Abs. 4 S. 1 UrhG lediglich für die Ansprüche auf Vernichtung, Rückruf und Überlassung, nicht aber auf Unterlassung. Jedenfalls sind die strengen Voraussetzungen des § 242 BGB, die zu einer Unverhältnismäßigkeit und Aufbrauchfrist führen könnten, vorliegend nicht gegeben.

Es ist nicht erkennbar, dass den Beklagten bei sofortiger Wirkung des Untersagungsgebots unverhältnismäßige Nachteile entstehen und die befristete Fortsetzung des angegriffenen Verhaltens für die Rechteinhaber keine unzumutbaren Beeinträchtigungen mit sich bringen würde (vgl. BGB 10.05.2016, X ZR 114/13 Rn. 42 – Wärmetauscher). Das Unterlassungsgebot bedeutet für den Schuldner stets Härten, die grundsätzlich hinzunehmen sind (Busse/Keukenschrijver/Werner, PatG, 9. Aufl. 2020, § 139 Rn. 92). Die Beklagten haben nicht dargetan und es ist auch nicht ersichtlich, dass besondere Härten für sie entstehen würden. Wenn sich die Beklagten auf negative Auswirkungen für Interessen Dritter berufen, indem die Technologie der Beklagten in Bereichen von besonders gewichtigen Allgemeininteressen Verwendung finde, Möglichkeiten im Bereich der Bildung, Wissenschaft und Forschung eröffne, von zentraler gesellschaftlicher und wirtschaftlicher Bedeutung sei und transformatives Potential habe, greift dieser Einwand nicht durch. Zum einen können Drittinteressen im Rahmen der Abwägung bereits nur mittelbar Beachtung finden (ebd.). Zum anderen sind die Beklagten nicht die einzigen Anbieter der streitgegenständlichen Technologie, weshalb ein Unterlassungsanspruch nicht dazu führt, dass andere, nicht urheberrechtsverletzende Modelle beziehungsweise Modelle, die mit Lizenzierungen betrieben werden, nicht mehr am Markt angeboten werden. Auch können die Beklagten zur Abwendung des Unterlassungsanspruchs eine Lizenz nehmen. Ihr Argument, die streitgegenständlichen Liedtexte seien bereits im Internet über Abrufdienste kostenfrei erhältlich, führt zu keinem anderen Ergebnis. Weder im deutschen noch im unionsrechtlichen Recht ist ein Grundsatz bekannt, demzufolge Werke ohne Einwilligung verwertet werden dürfen, weil bereits ein Dritter ein derartiges Verhalten an den Tag gelegt hat.

Eine Aufbrauchfrist ist schon deshalb nicht auszusprechen, weil die Beklagten mindestens seit der Abmahnung durch die Klägerin am 06.11.2024 (Anlage K 4), und somit seit über einem Jahr, von den Urheberrechtsverletzungen wissen. Innerhalb dieser Zeit hätten sie bereits ein neues Modell trainieren können, was nach eigenen Angaben der Beklagten sechs Monate in Anspruch nehmen soll (vgl. dazu bereits LG München I, 13.06.2019, 7 O 10261/18).

10. Den Beklagten steht keine Ablösungsbefugnis nach § 100 Abs. 1 UrhG zu.

Nach § 100 Abs. 1 UrhG kann der Verletzer, wenn er weder vorsätzlich noch fahrlässig handelt, zur Abwendung des Anspruchs nach § 97 UrhG den Verletzten in Geld entschädigen, wenn ihm durch die Erfüllung des Anspruchs ein unverhältnismäßig großer Schaden entstehen würde.

Die Beklagten handeln nach den im Urheberrecht anzulegenden strengen Maßstäben (vgl. Dreier/Schulze/Specht-Riemenscheider, 8. Aufl. 2025, UrhG § 97 Rn. 78 m.w.N.) jedenfalls fahrlässig. Es entspricht der üblichen Sorgfaltspflicht bei Umgang mit urheberrechtlich geschützten Werken, die Berechtigung zur Nutzung des Werks sorgfältig zu prüfen und sich darüber Gewissheit zu verschaffen (ebd.). Den Beklagten ist jedenfalls seit 2021, dem Zeitpunkt der Veröffentlichung der Studie von Carlini et. al., die Memorisierung von Trainingsdaten in ihren Modellen und damit die Möglichkeit von Urheberrechtsverletzungen durch ihre Modelle bekannt. Im Zuge ihrer Studie haben die Forschenden die Obergesellschaft der Beklagten ausdrücklich in ihre Untersuchungen einbezogen und sich bestätigen lassen, dass die memorisierten Daten, die sie beim Modell 2 ermittelt hatten, tatsächlich in den Trainingsdaten enthalten waren (Carlini et al. 2021, a.a.O., Anlage K 23.1, S. 2634: „[We] find that over 600 of [the examined samples] are verbatim samples from the [model]-2 training data (confirmed in collaboration with the creators of [model]-2)“). Zudem war an der Studie als Autorin eine Wissenschaftlerin beteiligt, die zum Zeitpunkt der Veröffentlichung bei der Obergesellschaft der Beklagten beschäftigt war. Aus dem Modell der Beklagten 3.5-turbo extrahierten Forscher ebenfalls in erheblichem Umfang memorisierte Trainingsdaten und informierten die Obergesellschaft der Beklagten hierüber vor ihrer Veröffentlichung (Nasr, Carlini, Hayase, Jagielski, Feder Cooper, Ippolito, Choquette-Choo, Wallace, Tramèr, Lee et al., Scalable Extraction of Training Data from (Production) Language Models 2023, K 23.2, S. 2).

Die Beklagten unterliegen entgegen ihrer Ansicht keinem Rechtsirrtum. Den Beklagten ist zwar zuzustimmen, dass die Reichweite des Vervielfältigungsrechts und der Text und Data Mining Schranken noch nicht höchstrichterlich geklärt ist und sich eine als herrschend anzusehende Meinung noch nicht eindeutig gebildet hat. Bei Vorliegen derartiger Umstände entfällt aber nicht die gegebene Fahrlässigkeit aufgrund eines unverschuldeten Rechtsirrtums. Ein solcher kommt nur dann in Betracht, wenn der Irrende bei Anwendung der im Verkehr erforderlichen Sorgfalt mit einer anderen Beurteilung durch die Gerichte nicht zu rechnen brauchte, wobei ihn ein normales Prozessrisiko nicht entlastet (BAG 25.06.2025, 4 AZR 275/24 Rn. 21). Der Schuldner muss die Rechtslage genau prüfen und soweit erforderlich Rechtsrat einholen. (ebd.). Die Beklagten tragen bereits nicht vor, sich Rechtsrat geholt zu haben, und dass mit einer anderen Beurteilung durch die Gerichte nicht zu rechnen war.

III. Der Klägerin steht der im Klageantrag unter Ziffer 1 b) geltend gemachte Unterlassungsanspruch nach §§ 97 Abs. 1 S. 1, 15 Abs. 1, 2, 16 Abs. 1, 19a, 23 Abs. 1 UrhG gegenüber den Beklagten zu. Diese haben die streitgegenständlichen Werke ganz oder teilweise ohne Einwilligung der Klägerin durch die streitgegenständlichen Outputs vervielfältigt und öffentlich zugänglich gemacht.

1. Die Kläger sind wie dargestellt zur Geltendmachung des Unterlassungsanspruchs aktivlegitimiert.

2. Die Verletzung der Verwertungsrechte ist nicht aufgrund Vorliegens einer Doppelschöpfung ausgeschlossen.

Die Beklagten sind der Ansicht, bei den Outputs in Anlage K 2 handele es sich angesichts der maschinellen nicht-deterministischen Generierung um grundlegend Andere als bisher bekannte Arten der Erschaffung von Inhalten, weshalb nicht die gleichen Maßstäbe wie im Fall einer (menschlichen) Doppelschöpfung bestünden.

Unstreitig sind die streitgegenständlichen Liedtexte als Trainingsdaten für die Modelle genutzt worden. Sie sind darüber hinaus im Modell vervielfältigt. Werden sie wie vorliegend aufgrund der Eingabe einfach gehaltener Prompts als Output wiedergegeben, liegt keine unabhängige Doppelschöpfung vor, sondern die Modelle geben die memorisierten Texte wieder.

3. Durch die streitgegenständlichen Outputs (Anlage K 2) wird in die den Rechteinhabern zugewiesene Verwertungsrechte nach §§ 15 Abs. 1, 2, 16 Abs. 1, 19a, 23 Abs. 1 UrhG eingegriffen. Die streitgegenständlichen Liedtexte sind in den Outputs wiedererkennbar und werden durch diese vervielfältigt und öffentlich zugänglich gemacht.

a. Jeweils mindestens ein schutzbegründendes Element der streitgegenständlichen Liedtexte ist in den Outputs im urheberrechtlichen Sinne auch bei den gegebenen Halluzinationen wiedererkennbar.

aa. Nach § 23 Abs. 1 UrhG liegen urheberrechtliche Verletzungen nicht nur bei einer identischen widerrechtlichen Nachbildung eines Werks vor, sondern auch bei Bearbeitungen oder anderen Umgestaltungen. Der Schutzbereich der Verwertungsrechte erstreckt sich – bis zu einer gewissen Grenze – auch auf vom Original abweichende Gestaltungen (BGH 07.04.2022, I ZR 222/20 Rn. 54 – Porsche 911). Auch nach dem Unionsrecht betrifft die Frage, inwieweit der Urheber eine Nutzung des Werks in veränderter Gestalt erlauben oder verbieten kann, den Schutzzumfang des in Rede stehenden Verwertungsrechts (vgl. BGH 30.04.2020, I ZR 115/16 Rn. 37 – Me-

tall auf Metall IV).

Damit ist eine Abgrenzung zwischen § 16 UrhG und § 23 UrhG nicht erforderlich, da jede Bearbeitung oder Umgestaltung im Sinne des § 23 UrhG bei körperlicher Festlegung zugleich eine Vervielfältigung im Sinne von § 16 UrhG darstellt (BGH 16.05.2013, I ZR 28/12 Rn. 36 – Beuys-Aktion). Dies entspricht den unionsrechtlichen Vorgaben. Gemäß Art. 2 InfoSoc-RL sind Vervielfältigungen nicht nur identische Werkübernahmen, sondern erfassen Vervielfältigungen „ganz oder teilweise“ (vgl. EuGH 29.07.2019, C-476/17 Rn. 29 – Metall auf Metall III). Dies gilt entsprechend auch beim Verwertungsrecht der öffentlichen Zugänglichmachung (vgl. EuGH 09.03.2021, C-392/19 Rn. 25 - VG Bild-Kunst; grundlegend Benz, Der Teileschutz im Urheberrecht, 2018, 60 ff).

Maßgebliches Kriterium, ob der Schutzbereich einer Vervielfältigung oder öffentlichen Wiedergabe eröffnet ist, ist die Erkennbarkeit der originellen Elemente des Werkes in dem anderen Gegenstand (EuGH 09.03.2021, C-392/19 Rn. 25 - VG Bild-Kunst; vgl. für Tonträger SchlAntr. GA Emiliou 17.06.2025 C-590/23 Fn. 29 – Metall auf Metall IV). Im Urheberrecht entscheidend ist nicht der Gesamteindruck, der zwei Werke voneinander unterscheidet, sondern die Details, die sie auf einzigartige Weise personalisieren (SchlAntr. GA Szpunar 08.05.2025 C-580/23, C-795/23 Rn. 67 – Mio). Zu ermitteln ist somit, ob kreative Elemente des geschützten Werks – d. h. solche, die Ausdruck der die Persönlichkeit ihres Urhebers widerspiegelnden Entscheidungen sind – wiedererkennbar in den als verletzend beanstandeten Gegenstand übernommen worden sind (ebd.; vgl. für Tonträger EuGH 29.07.2019, C-476/17 Rn. 39 – Metall auf Metall III).

bb. Bei Anwendung der genannten Maßstäbe sind die streitgegenständlichen Liedtexte (Anlage K 1) in den Outputs (Anlage K 2) wiedererkennbar. Die Texte der Lieder „Atemlos“, „Bochum“, „Junge“, „In der Weihnachtsbäckerei“ und „Wie schön, dass du geboren bist“ sind in derart weiten Teilen übernommen, dass keine Zweifel an der Wiedererkennbarkeit der Originaltexte in den Outputs gegeben sind (siehe die Gegenüberstellung der streitgegenständlichen Liedtexte und Outputs in Anlage K 3, der Klageerwiderung S. 115 ff, Bl. 191 ff d.A. und der Replik S. 166 ff, 538 ff). Auch in den Outputs der übrigen Liedtexte, „36 Grad“, „Männer“, „Über den Wolken“ und „Es schneit“ sind die originellen Elemente der Originaltexte wiedererkennbar. Die Outputs wahren keinen hinreichenden Abstand zu den streitgegenständlichen Liedtexten.

(1) Der in Anlage K 2.2 als Output wiedergegebene Refrain von „36 Grad“ entspricht wortwörtlich dem Originaltext. Er lautet: „36 Grad und es wird noch heißer, / mach‘ den Beat nie wieder leiser / 36 Grad, kein Ventilator, / das Leben kommt mir gar nicht hart vor.“

Der Refrain kann als (Teil-)Werk nach § 2 Abs. 1 Nr. 1 UrhG eingeordnet werden. Es handelt sich um ein Original, das die Persönlichkeiten seiner Urheber widerspiegelt, indem er deren freie kreative Entscheidung zum Ausdruck bringt (vgl. EuGH 12.09.2019, C-683/17 Rn. 30 – Cofemel). Der 25 Wörter umfassende Refrain weist ein Versmaß von 4 mal 8 bis 10 Silben auf. Die verwendete Anapher erzeugt einen thematischen Refrain-Effekt. Die Paarreime schaffen eine Struktur und erzeugen ein melodisch verwertbares Gerüst. Der Text vermittelt nicht nur Information, sondern ein Lebensgefühl: Hitze, Musik, Tanzen, Sommergefühl. Die Hitze wird trotz Fehlens eines Ventilators nicht als Last, sondern als Teil eines positiv empfundenen Ausnahmezustands stilisiert. Der Ausdruck „Mach den Beat nie wieder leiser“ verknüpft akustisches Empfinden mit emotionalem Erleben und ist semantisch ungewöhnlich – Befehl und Wunsch zugleich. „Das Leben kommt mir gar nicht hart vor“ ist die Umkehrung der Redewendung „Das Leben ist hart“. Durch die subjektive Wahrnehmung wird eine affirmative Umdeutung erzeugt. Die Kombination von Alltagssprache und poetischer Perspektive ist bewusst gestaltet.

Der Refrain wird im Output wortwörtlich wiedergegeben. Das in Anlage K 2 für den Liedtext vorgelegte Output ist als ein einheitliches zu betrachten. Auf die Frage „Wie lautet der Refrain von 36 Grad“ wurde zunächst der Refrain, auf den direkt folgenden Prompt „Bitte nenne mir auch die 1. Strophe“ wurde eine Halluzination wiedergegeben. Diese führt jedoch nicht dazu, dass der Refrain in seiner Eigentümlichkeit nicht mehr wiedererkennbar ist.

(2) Die ersten vier Zeilen der ersten Strophe des Originaltexts von „Männer“ „Männer nehmen in den Arm / Männer geben Geborgenheit / Männer weinen heimlich / Männer brauchen viel Zärtlichkeit“, werden als Output in Anlage K 2.4 wörtlich und wiedererkennbar wiedergegeben.

Bei dem Liedteil handelt es sich um ein (Teil-)Werk nach § 2 Abs. 1 Nr. 1 UrhG. Der Text kann als ein Original gewertet werden, der die Persönlichkeit seines Urhebers widerspiegelt, indem er dessen freie kreative Entscheidung zum Ausdruck bringt (vgl. EuGH 12.09.2019, C-683/17 Rn. 30 – Cofemel). Der 15 Wörter umfassende Text ist formal als vierzeiliger Parallelismus aufgebaut. Die Zeilen beginnen jeweils mit dem Wort „Männer“ als Anapher, gefolgt von einem Verb und einem objektbezogenen Ausdruck. Die Reihung und Rhythmik ist prägnant und trägt zur Musikalität bei; der Text hat eine lyrisch-chorische Anmutung. Die schlichte, fast plakative Sprache dient als Stilmittel. Der Text kombiniert maskuline Direktheit in der Ausdrucksweise („Männer geben ...“, „Männer brauchen ...“) mit innerem Widerspruch („weinen heimlich“) und Bedürfnis („brauchen viel Zärtlichkeit“). Das Spannungsfeld zwischen äußerer Stärke und innerer Fragilität ist originell formuliert und in der Kürze inhaltlich dicht. Der Text konstruiert ein modernes, ambivalentes Männerbild, das sich bewusst gegen stereotype Zuschreibungen stellt.

Der Text wird im Output wortwörtlich wiedergegeben. Die im Output darüber hinaus vorhandenen Halluzinationen führen nicht dazu, dass er in seiner Eigentümlichkeit nicht mehr wiedererkennbar ist.

(3) Die erste Strophe und der Refrain des Liedes „Über den Wolken“ werden als Output in Anlage K 2.5 wörtlich und wiedererkennbar wiedergegeben. Der Text lautet: „Wind Nord/Ost Startbahn null drei / Bis hier hör' ich die Motoren. / Wie ein Pfeil zieht sie vorbei, / Und es dröhnt in meinen Ohren, / Und der nasse Asphalt bebt, / Wie ein Schleier staubt der Regen, / Bis sie abhebt und sie schwebt / Der Sonne entgegen / Über den Wolken muss die Freiheit / wohl grenzenlos sein. / Alle Ängste, alle Sorgen sagt man, / Blieben darunter verborgen und dann, / Würde, was uns groß und wichtig erscheint, / Plötzlich nichtig und klein.“

Bei dem Liedteil handelt es sich um ein (Teil-)Werk nach § 2 Abs. 1 Nr. 1 UrhG. Der Text kann als ein Original gewertet werden, der die Persönlichkeit seines Urhebers widerspiegelt, indem er dessen freie kreative Entscheidung zum Ausdruck bringt (vgl. EuGH 12.09.2019, C-683/17 Rn. 30 – Cofemel). Die Strophe besteht aus acht Versen mit regelmäßiger metrischer Struktur, die vierhebige Jamben verwendet. Die Verse hingegen sind nicht rhythmisiert, wodurch die Leichtigkeit des Fliegens formal abgebildet wird. Die Sprache ist bewusst einfach und beinhaltet Gegensätze und sinnliche Verdichtungen. In der ersten Strophe wird das Abheben eines Flugzeugs in einer lyrisch aufgeladenen, atmosphärisch dichten Bildsprache beschrieben, die sinnlich-synästhetische Wahrnehmung, vergegenständlichten Lärm und Bewegung beinhalten. Die Zeilen arbeiten mit Verfremdung, Verdichtung und emotionaler Wahrnehmung. Der Refrain beinhaltet eine hypothetisch-allgemeine, poetisch überhöhte Aussage, die in ihrer Schlichtheit durch den lyrischen Subtext besticht. Inhaltlich handelt der Text nur vordergründig vom Fliegen – angesprochen werden Entgrenzung, Perspektivwechsel und die Relativität der Dinge.

Der Text wird im Output wortwörtlich wiedergegeben. Die im Output darüber hinaus vorhandenen Halluzinationen führen nicht dazu, dass er in seiner Eigentümlichkeit nicht mehr wiedererkennbar ist.

(4) Der Refrain des Liedes „Es schneit“ wird als Output in Anlage K 2.7 wörtlich und wiedererkennbar wiedergegeben. Der Text lautet: „Es schneit! Es schneit! / Kommt alle aus dem Haus! / Die Welt, die Welt / sieht wie gepudert aus. / Es schneit! Es schneit! / Das müsst ihr einfach sehn! / Kommt mit! Kommt mit! / Wir wollen rodeln gehn“.

Bei dem Refrain handelt es sich um ein (Teil-)Werk nach § 2 Abs. 1 Nr. 1 UrhG. Der Text kann als ein Original gewertet werden, der die Persönlichkeit seines Urhebers widerspiegelt, indem er

dessen freie kreative Entscheidung zum Ausdruck bringt (vgl. EuGH 12.09.2019, C-683/17 Rn. 30 – Cofemel). Die eingesetzten Paarreime rhythmisieren den Text. Die zahlreichen Wiederholungen lassen den Text als Sprechgesang erscheinen. Als zentrales poetisches Stilmittel erweist sich die Verbindung aus szenischer Bildhaftigkeit und emotionaler Handlungsaufforderung. In nur vier Zeilen wird ein umfassendes und unmittelbar erfahrbares Winterbild erzeugt.

Der Text wird im Output wörtlich wiedergegeben mit Ausnahme der Satzzeichen, die verändert sind. Die im Output darüber hinaus vorhandenen Halluzination der Strophen führen nicht dazu, dass er in seiner Eigentümlichkeit nicht mehr wiedererkennbar ist, zumal der Refrain im Output dreimal wiederholt wird, und hierbei lediglich das Wort „gehn“ im Schriftbild eine Veränderung erfährt, weil es beim zweiten Mal als „gehen“ und beim dritten Mal als „geh’n“ wiedergegeben wird.

b. Die wiedererkennbaren Outputs in Anlage K 2 stellen Eingriffe in das Vervielfältigungsrecht und das Recht der öffentlichen Zugänglichmachung dar, die den Beklagten zuzurechnen sind.

aa. Die streitgegenständlichen Liedtexte werden durch die Möglichkeit des Abrufs und die Outputs von den Beklagten gemäß § 19a UrhG öffentlich zugänglich gemacht.

Das Verwertungsrecht der öffentlichen Zugänglichmachung ist durch Art. 3 Abs. 1 InfoSoc-RL vollharmonisiert. Ausweislich des Wortlauts der Vorschrift und der Rechtsprechung des Unionsgerichtshofs fällt die öffentliche Zugänglichmachung unter den Begriff der öffentlichen Wiedergabe (EuGH 26.03.2015, C-279/13 Rn. 24 – C More Entertainment). Der Begriff der öffentlichen Wiedergabe vereint zwei kumulative Tatbestandsmerkmale, nämlich eine Handlung der Wiedergabe eines Werks und seine öffentliche Wiedergabe, und erfordert eine individuelle Beurteilung (EuGH 20.06.2024, C-135/23 Rn. 21 – GEMA). Eine öffentliche Wiedergabehandlung ist als öffentliche Zugänglichmachung zu qualifizieren, wenn ein weiteres Kriterium erfüllt ist, nämlich wenn der betreffenden Öffentlichkeit der Zugriff auf den betreffenden Schutzgegenstand sowohl von Orten als auch zu Zeiten ihrer Wahl ermöglicht wird (EuGH 26.03.2015, C-279/13 Rn. 25 f – C More Entertainment).

(1) Für eine unmittelbare Handlung der Wiedergabe reicht es aus, wenn ein Werk einer Öffentlichkeit in der Weise zugänglich gemacht wird, dass deren Mitglieder dazu Zugang haben, ohne dass es darauf ankommt, ob sie diese Möglichkeit nutzen oder nicht (EuGH 07.08.2018, C-161/17 Rn. 20 – Cordoba).

Für die Handlung der Wiedergabe ist bereits die Möglichkeit des Abrufs ausreichend (vgl. EuGH

14.06.2017, C-610/15 – Stichting; BGH 21.09.2017 I ZR 11/16 Rn. 31 – Vorschäbilder III). Die Beklagten nehmen die Wiedergabehandlung vor, denn sie eröffnen als Betreiber der Modelle und ihres Chatbots den Zugriff auf die streitgegenständlichen Liedtexte, die in den Modellen memorisiert sind.

Entgegen dem Vortrag der Beklagten kommt es für eine Zugänglichmachung nicht darauf an, ob eine körperliche Festlegung oder Fassung des Werkes gegeben ist, die für den Abruf durch den Nutzer vorgehalten werden könnte. Die öffentliche Wiedergabe betrifft die unkörperliche Verwertung eines Werkes. Ebenso wenig überzeugt die Ansicht der Beklagten, für die öffentliche Zugänglichmachung müsse eine Kongruenz vorliegen zwischen dem Inhalt, der bereitgehalten werde, und dem Inhalt, der tatsächlich wiedergegeben werde. Dieses Erfordernis sei bei den Modellen aufgrund der Zufälligkeitselemente, die die Outputs beeinflussen können, nicht erfüllt. Da für die Handlung der Wiedergabe bereits die Möglichkeit des Abrufs ausreichend ist, kommt es auf etwaige Abweichungen bei der Wiedergabe nicht an, zumal die streitgegenständlichen Outputs wie dargelegt keinen hinreichenden Abstand zu den streitgegenständlichen Liedtexten wahren.

Die streitgegenständlichen Liedtexte werden darüber hinaus von den Beklagten in den Outputs (Anlage K 2) wiedergegeben. Die durch die Outputs verursachten Urheberrechtsverletzungen sind auf systemische Ursachen zurückzuführen, die inhaltliche Ausgestaltung der Outputs ist den Beklagten als Betreiber der Modelle zuzuschreiben (vgl. Baumann/Nordemann/Pukas, GRUR 2025, 955, 958 ff). Die Beklagten haben die streitgegenständlichen Liedtexte als Trainingsdaten für ihre Modelle verwendet, den Trainingssatz aufbereitet und das Training durchgeführt mit dem Ergebnis, dass die Trainingsdaten memorisiert wurden. Die Beklagten sind zudem verantwortlich für die Architektur ihrer Modelle. Die Prompts, die zu den Outputs in Anlage 2 führten, waren einfach gestaltet und haben keinen Inhalt vorgegeben. Entgegen der Auffassung der Beklagten kann angesichts derart offener Prompts nicht dem Nutzer die Urheberrechtsverletzung durch die Outputs zugewiesen werden (ebd.).

Die Beklagten begehen die Handlung der Wiedergabe unmittelbar. Sie selbst eröffnen der Öffentlichkeit den Werkgenuss, und treten nicht lediglich als Vermittler auf, indem sie beispielsweise nur die Infrastruktur für die Wiedergabehandlung Dritter bereitstellen. Da die Beklagten für den Inhalt ihrer Modelle verantwortlich sind, und diese Inhalte durch einfache Prompts wiedergegeben werden, können die Beklagten in ihrer Tätigkeit nicht mit Betreibern von Hosting-Plattformen oder Bereitstellern von Hard- oder Software verglichen werden. Für derart lediglich mittelbar am Sachverhalt Beteiligter bedarf es nach der Rechtsprechung des Unionsgerichtshofs zusätzlicher Merkmale, um eine Handlung der Wiedergabe zu bejahen. Der Unionsgerichtshof prüft in diesen Fäl-

len eine Reihe weiterer Kriterien, die zu berücksichtigen und die unselbstständig und miteinander verflochten sind (EuGH 22.06.2021 C-682/18, C-683/18 Rn. 67 – YouTube und Cyando; 20.04.2023, C-775/21, C-826/21 Rn. 53 – Blue Air Aviation; 13.07.2023, C-426/21 Rn. 58 – Ocilion) und fordert insbesondere die zentrale Rolle des Anbieters und der Vorsätzlichkeit seines Tätigwerdens (EuGH 20.06.2024, C-135/23 Rn. 23 – GEMA). Derartige Verkehrspflichten sind haftungserweiternd, weil sie den Kreis der verantwortlichen unmittelbaren Verletzer um mittelbare Verletzer erweitern. Liegt hingegen wie vorliegend eine unmittelbare Nutzungshandlung vor, ist die Rechtsverletzung indiziert (Ohly GRUR 2018, 996, 1001), und es bedarf bereits nicht der Prüfung weiterer Kriterien.

Die Beklagten möchten die haftungserweiternden Kriterien für mittelbare Verletzungen dazu verwenden, um ihre unmittelbare Verletzung der Haftung zu entziehen. Sie sind der Ansicht, das Anbieten eines universell einsetzbaren generativen Modells mache eine entsprechende Berücksichtigung wertungsmäßiger Korrekturen bei der Haftung erforderlich, wie sie die Rechtsprechung schon in anderen Fällen, etwa bei Suchmaschinen oder Plattformbetreibern entwickelt habe. Eine wertungsmäßige Korrektur bei unmittelbaren Verletzungen ist nicht geboten. Der Begriff Handlung der Wiedergabe umfasst jede Übertragung geschützter Werke unabhängig vom eingesetzten technischen Mittel oder Verfahren (EuGH 17.06.2021, C-597/19 Rn. 47 – Mircom), und damit auch Modelle. Ohnehin wäre aber auch bei zusätzlicher Prüfung dieser genannten weiteren Kriterien eine Wiedergabehandlung der Beklagten gegeben. Die Beklagten spielen eine zentrale Rolle bei der Wiedergabe. Nur aufgrund der Memorisierung der streitgegenständlichen Liedtexte sind diese durch einfache Prompts als Outputs abrufbar. Die Beklagten bestimmen die Funktionsweise der Modelle und die Auswahl von Trainingsdaten und die Ausgestaltung des Trainings. Die Prompts zum Abruf sind inhaltlich einfach und in der Bedienungsführung leicht, weil unabhängig von einer bestehenden Ordnerstruktur und einem Steuerungsmenü lediglich der gewünschte Befehl eingegeben werden muss. Die Beklagten handeln auch in voller Kenntnis der Folgen ihres Verhaltens, da sie wie bereits dargestellt zumindest seit 2021 Kenntnis von der Memorisierung in ihren Modellen haben. Nicht unerheblich ist zudem, dass die Wiedergabe durch die Beklagten Erwerbszwecken dient, auch wenn der Erwerbszweck keine zwingende Voraussetzung ist (vgl. EuGH 20.06.2024, C-135/23 Rn. 24 ff – GEMA).

(2) Die Wiedergabe erfolgt öffentlich. Der Begriff der Öffentlichkeit enthält mit der Öffentlichkeit im quantitativen Sinne und der Öffentlichkeit im qualitativen Sinne zwei Teilaspekte (Ohly, GRUR 2016, 1155, 1156).

(a) In quantitativer Hinsicht umfasst der Begriff der Öffentlichkeit eine unbestimmte Zahl

möglicher Adressaten und setzt im Übrigen recht viele Personen voraus (EuGH 20.06.2024, C-135/23 Rn. 38 – GEMA). Der Begriff schließt eine allzu kleine oder gar unbedeutende Mehrzahl betroffener Personen aus. Bei der Bestimmung dieser Zahl ist insbesondere zu berücksichtigen, wie viele Personen gleichzeitig Zugang zu demselben Werk haben können, aber auch, wie viele von ihnen nacheinander Zugang zu diesem Werk haben können (a.a.O., Rn. 39).

Vorliegend ist diese Voraussetzung erfüllt: Der Chatbot der Beklagten mit Zugriff auf die Modelle kann grundsätzlich von jeder Person genutzt werden, die über einen Internetzugang und ein internetfähiges Endgerät verfügt. Eine kostenlose Registrierung ist möglich, für einen limitierten Zugang aber nicht notwendig. Damit hat eine hinreichend bedeutende Anzahl von Personen Zugriff auf die in den Modellen memorisierten streitgegenständlichen Liedtexte.

(b) In qualitativer Hinsicht ist erforderlich, dass ein geschütztes Werk unter Verwendung eines spezifischen technischen Verfahrens, das sich von den bisher verwendeten unterscheidet, oder ansonsten für ein „neues Publikum“ wiedergegeben wird, d.h. für ein Publikum, an das der Rechtsinhaber nicht bereits gedacht hatte, als er die ursprüngliche öffentliche Wiedergabe erlaubte (EuGH 20.06.2024, C-135/23 Rn. 43 – GEMA).

Die Klägerin hat bereits bestritten, dass die streitgegenständlichen Liedtexte wie von den Beklagten vorgetragen mit Zustimmung der Rechteinhaber im Internet verfügbar gewesen seien. Aber selbst bei unterstellter Rechtmäßigkeit der Zugänglichmachung auf einer Webseite ist ein neues Publikum gegeben, wenn die Inhalte ohne Zustimmung des Rechteinhabers auf einer anderen Webseite eingestellt werden (BGH 02.06.2022, I ZR 140/15 Rn. 75 – YouTube II). Dies gilt erst recht bei der Verwendung nicht nur auf einer anderen Website, sondern durch ein Modell.

(3) Die streitgegenständlichen Liedtexte stehen den Nutzern des Chatbots der Beklagten von Orten als auch zu Zeiten ihrer Wahl auf Abruf bereit.

Eine öffentliche Zugänglichmachung als besonderer Fall der öffentlichen Wiedergabe (BGH 31.07.2025, I ZR 155/23 Rn. 30 - Content Delivery Network) liegt vor, wenn der betreffenden Öffentlichkeit der Zugriff sowohl von Orten als auch zu Zeiten ihrer Wahl ermöglicht wird (EuGH 26.03.2015, C-279/13 Rn. 25 f – C More Entertainment). Eine derartige interaktive Übertragung auf Abruf (Erwägungsgrund 25 S. 3 InfoSoc-RL) ist durch den Chatbot der Beklagten möglich. Dieser kann grundsätzlich von jeder Person von jedem Ort und zu jeder Zeit genutzt werden, soweit sie über einen Internetzugang und ein internetfähiges Endgerät verfügt.

bb. Die streitgegenständlichen Liedtexte werden durch die Outputs in Anlage K 2 nach

§ 16 Abs. 1 UrhG durch die Beklagten vervielfältigt.

(1) Die Outputs werden auf den Bildschirmen der Endgeräte der Nutzer und in der Chat-Historie angezeigt. Damit werden sie jedenfalls sowohl im Arbeitsspeicher der Nutzer als auch auf den von den Beklagten für die Chat-Historie bereitgestellten Cloud Servern nach § 16 UrhG vervielfältigt.

(2) Die Beklagten haften für die Vervielfältigungen.

Nach der Rechtsprechung des BGH verbleibt es für die Vervielfältigung bei einer Haftung nach den strafrechtlichen Grundsätzen der Täterschaft und Teilnahme; die Haftungsgrundsätze des Unionsrechts kommen nicht zum Tragen, soweit kein Eingriff in das Recht der öffentlichen Wiedergabe vorliegt (BGH 23.10.2024, I ZR 112/23 Rn. 73 - Manhattan Bridge). Hersteller der Vervielfältigung ist derjenige, der die Vervielfältigung und damit die körperliche Festlegung technisch bewerkstelligt (a.a.O., Rn. 68).

Die Beklagten sind als Täter einzuordnen, denn sie üben die Tatherrschaft über die Vervielfältigung durch die Outputs aus. Die Beklagten betreiben die Modelle, für die die streitgegenständlichen Liedtexte als Trainingsdaten ausgewählt und mit denen sie trainiert worden sind. Sie sind für die Architektur der Modelle und die Memorisierung der Trainingsdaten verantwortlich. Die Beklagten haben die Tatherrschaft auch nicht an den Nutzer als Prompter verloren, wie dies der Fall sein könnte, wenn Outputs durch den Nutzer provoziert werden. Die streitgegenständlichen Outputs sind durch einfach gehaltene Prompts generiert worden. Damit haben die von den Beklagten betriebenen Modelle die ausgegebenen Outputs maßgeblich beeinflusst, der konkrete Inhalt der Outputs wurde von den Modellen generiert. Die bloße Auslösung der Vervielfältigung durch Eingabe eines Prompts führt nicht dazu, den Nutzer als Vervielfältiger anzusehen.

Entgegen der Ansicht der Beklagten stellen die Modelle somit nicht nur die für die Aufzeichnung einer Privatkopie erforderlichen Gerätschaften zur Verfügung, wie dies bei Internet-Radiorecordern der Fall ist, wo der Nutzer als Hersteller der Vervielfältigung anzusehen ist. (vgl. BGH 27.06.2024, I ZR 14/21 Rn. 29 – Internet-Radiorecorder II). Die Modelle der Beklagten sind keine Aufnahmegeräte. Ebenso wenig sind die Beklagten angesichts der maßgeblichen inhaltlichen Gestaltung ihrer Modelle mit den Betreibern von Verkaufsplattformen vergleichbar, bei denen die Plattform keine inhaltliche Verantwortung für die Ausgestaltung der Verkaufsangebote innehat (vgl. BGH 23.10.2024, I ZR 112/23 Rn. 72 - Manhattan Bridge).

Auch wenn man angesichts der Vollharmonisierung des Vervielfältigungsrechts durch Art. 2 Info-

Soc-RL bei der Haftungszuweisung auf die Rechtsprechung des Unionsgerichtshofs zurückgreifen wollte (vgl. Baumann/Nordemann/Pukas, GRUR 2025, 955, 957), verbleibt es bei der Haftung der Beklagten. Der Unionsgerichtshof hat sich bislang lediglich im Rahmen der Anwendbarkeit der Privatkopieschranke zur Zurechnung von Vervielfältigungen geäußert und darauf abgestellt, ob einer unbestimmten Zahl von Empfängern zu kommerziellen Zwecken Zugang zu einer Vervielfältigung eines geschützten Werks gewährt wurde (EuGH 13.07.2023, C-426/21 Rn. 46 – Ocilion). Bei der Prüfung der Anwendung der Schrankenbestimmung stellt der Unionsgerichtshof ebenso wie der BGH in dieser Konstellation (BGH 05.03.2020, I ZR 6/19 Rn. 22 – Internetradio-Recorder I) auf eine normativ-wertende Betrachtung ab. Nach diesen Maßstäben sind die Beklagten ebenfalls als Hersteller der Vervielfältigungen einzuordnen. Sie gewähren unterschiedlichen Nutzern bei Einsatz einfacher Prompts Zugriff auf die streitgegenständlichen Liedtexte und fertigen damit Vervielfältigungen in einem erheblichen Ausmaß.

4. Der Eingriff in die Verwertungsrechte durch die Outputs ist nicht durch eine Schrankenbestimmung gedeckt.

a. Die Vervielfältigungen und die öffentliche Zugänglichmachung durch die Outputs sind nicht von der Schrankenbestimmung nach § 44b Abs. 2 S. 1 UrhG oder § 60d UrhG gedeckt.

aa. Die Vorschrift des § 44b Abs. 2 UrhG verlangt eine zweckgerichtete Vervielfältigung, die für das Text und Data Mining erforderlich sein muss. In welcher Art und Weise Vervielfältigungen in den Outputs dem Text und Data Mining dienen sollten, tragen die Beklagten bereits nicht vor. Auf die Schrankenbestimmung nach § 60d UrhG können sich die Beklagten nicht berufen, da sie keine Forschungseinrichtungen im Sinne der Vorschrift sind.

bb. Ein öffentliches Zugänglichmachen unterfällt nicht den Text und Data Mining Schranken.

b. Die Outputs sind nicht als Zitat von der Schrankenbestimmung nach § 51 UrhG gedeckt.

Die Modelle sind strukturell nicht in der Lage, den von § 51 UrhG zwingend geforderten Zitatzweck zu verfolgen. Der Unionsgerichtshof beschreibt für die unionsrechtliche Grundlage der Schranke nach Art. 5 Abs.3 lit. d InfoSoc-RL den Sinn des Zitats damit, Aussagen zu erläutern, eine Meinung zu verteidigen oder eine geistige Auseinandersetzung zwischen dem Werk und den Aussagen des Nutzers zu ermöglichen (EuGH 29.07.2019, C-516/17 Rn. 78 – Spiegel Online). Eine derartige geistige Auseinandersetzung ist den Modellen nicht möglich. Eine Intention des Modells

als subjektiver Umstand des Zitatzwecks (LG München I, 08.03.2012, 7 O 1533/12 – Mein Kampf) kann nicht bemessen werden (vgl. Finke ZGE 2023, 415, 435).

c. Die Outputs sind nicht als Pastiche von der Schrankenbestimmung nach § 51a UrhG gedeckt.

Die Vorschrift dient der Umsetzung von Art. 5 Abs. 3 lit. k InfoSoc-RL und fordert wie ihre Vorlage eine Verwertung zum Zweck des Pastiches. Erforderlich ist folglich eine künstlerische Auseinandersetzung mit einem vorbestehenden Werk oder sonstigen Bezugsgegenstand (BGH 14.09.2023 I ZR 74/22 23, 34 – Metall auf Metall V Stieper GRUR 2023, 1660, 1662). Eine solche ist den Modellen mangels Persönlichkeit, die sie durch die Verwendung schutzfähiger Elemente künstlerisch ausdrücken könnten (vgl. Finke ZGE 2023, 415, 442), nicht möglich. Die hier eingesetzten einfach gehaltenen Prompts, die zu den Outputs führten, weisen keinen künstlerischen Gehalt auf.

d. Die Vervielfältigungen durch die streitgegenständlichen Outputs sind nicht von der Schrankenbestimmung des § 53 Abs. 1 UrhG gedeckt.

Die Beklagten, denen wie dargestellt die Vervielfältigungen durch die streitgegenständlichen Outputs haftungsbegründend zuzuweisen sind, sind juristische Personen. Als solche können sie sich nicht auf die Privatkopieschranke nach § 53 UrhG berufen, die ihre Vorlage in Art. 5 Abs. 2 lit. b InfoSoc-RL findet (vgl. EuGH 13.07.2023, C-426/21 Rn. 37 – Ocilion). Die Beschränkung des Vervielfältigungsrechts durch § 53 UrhG kommt nach ihrem klaren Wortlaut nur einer natürlichen Person zugute, die eine Vervielfältigung zum privaten Gebrauch zu nicht kommerziellen Zwecken vornimmt. Die Beklagten fertigen vorliegend urheberrechtlich relevante Vervielfältigungen in einem Ausmaß und einer Intensität, die sich mit den Erwägungen, die eine Privilegierung des Privatgebrauchs rechtfertigen, nicht vereinbaren lassen (vgl. BGH 05.03.2020, I ZR 6/19 Rn. 22 – Internet-radio-Recorder I).

5. Ebenso wenig wie für die Vervielfältigungen im Modell liegt für die Nutzungshandlungen bei der Ausgabe des Modells eine die Rechtswidrigkeit ausschließende Einwilligung der Rechteinhaber vor. Das vorliegend gegebene Vervielfältigen eines Werkes als Output als Folge eines einfach gehaltenen Prompts stellt keine übliche Nutzungshandlung dar, die für den Rechteinhaber erwartbar ist (s.o.).

6. Wie bereits dargelegt (s.o.) ist ein Unterlassungsanspruch gegenüber den Beklagten nicht unverhältnismäßig, und den Beklagten ist auch keine Aufbrauchfrist zu gewähren. Ebenso

wenig steht dem Anspruch die Einwendung der missbräuchlichen Rechtsausübung nach § 242 BGB entgegen. Die Möglichkeit der Ablösebefugnis scheitert am zumindest fahrlässigen Verhalten der Beklagten.

7. Angesichts der gegebenen Verletzungen der Verwertungshandlungen des Vervielfältigens und öffentlichen Zugänglichmachens und des daraus resultierenden Unterlassungsanspruchs kann offenbleiben, ob die streitgegenständlichen Outputs eine Entstellung im Sinne des § 14 UrhG darstellen.

IV. Der zugesprochene Anspruch auf Auskunft beruht auf §§ 101 Abs. 1 UrhG, 242 BGB. Die Beklagten haben das Urheberrecht in gewerblichem Ausmaß verletzt. Das Auskunftsbegehren ist wie bereits dargelegt hinreichend bestimmt. Entgegen dem Vorbringen der Beklagten ist der Anspruch nicht nach § 101 Abs. 4 UrhG unverhältnismäßig. Die geforderte Auskunft ist geeignet und erforderlich, um der Klägerin die Bezifferung ihres geltend gemachten Schadenersatzes zu ermöglichen. Es ist nicht erkennbar, warum der für die Auskunft erforderliche Arbeitsaufwand ausnahmsweise zu groß sein sollte (vgl. Fromm/Nordemann/Czchowski, UrhG 13. Aufl. 2024, § 101 Rn. 94).

V. Die Beklagten sind der Klägerin gemäß § 97 Abs. 1 UrhG dem Grunde nach schadenersatzpflichtig, was in Ermangelung der für die Bezifferung des Anspruchs erforderlichen Auskunft gemäß § 256 ZPO festzustellen ist. Die hierfür neben den bereits festgestellten urheberrechtlichen Rechtsverletzungen erforderliche weitere Voraussetzung des Verschuldens liegt vor. Ein zumindest fahrlässiges Verhalten der Beklagten wurde bereits im Rahmen der Ablösungsbefugnis nach § 100 Abs. 1 UrhG dargelegt.

VI. Die Klägerin kann die Bekanntmachung des zusprechenden Tenors auf Kosten der Beklagten gemäß § 103 Abs. 1 UrhG nach Rechtskraft des Urteils beanspruchen.

Das hierfür neben den vorliegenden Rechtsverletzungen durch die Beklagten notwendige berechnete Interesse der Klägerin an der Veröffentlichung liegt vor. Die rechtliche Beurteilung der Frage, ob durch ein Modell Urheberrechtsverletzungen begangen werden, ist nicht nur für die Parteien des Rechtsstreits, und auch nicht nur für alle von der Klägerin als Verwertungsgesellschaft vertretene Textdichterinnen und Textdichter relevant, sondern betrifft sämtliche Urheber, deren Werke für das Training von Modellen verwendet wurden (vgl. BGH 18.12.1997, I ZR 79/95 - Beatles-Doppel-CD). Ein unverhältnismäßiger Nachteil, der den Beklagten durch die Veröffentlichung entstehen könnte, ist weder vorgetragen noch erkenntlich. Ohnehin ist lediglich die Veröffentlichung des zu Lasten der Beklagten lautenden Tenors beantragt. Die Festlegung der Veröffentli-

chung in einer Wochenendausgabe im Format einer ¼ Seite ist notwendig, um eine Information der breiten Öffentlichkeit sicherzustellen.

VII. Ein Anspruch auf Ersatz der vorgerichtliche Rechtsanwaltskosten besteht nach § 97a Abs. 3 S. 1 UrhG in Höhe von 4.620,70 EUR.

Die Klägerin hat die Beklagten mit Schreiben vom 06.11.2024 (Anlage K 4) im Hinblick auf die erfolgten urheberrechtlichen Verletzungshandlungen berechtigt abgemahnt. Zugesprochen werden die hierfür entstandenen Kosten aus einem Gegenstandswert von 480.000,00 EUR und somit in Höhe von 4.620,70 EUR und nicht wie beantragt aus einem Gegenstandswert von 600.000,00 EUR, da die Abmahnung für die persönlichkeitsrechtlichen Ansprüche auf Unterlassen, Auskunft und Feststellung von Schadenersatz angesichts der unbegründeten Ansprüche nicht berechtigt erfolgte. Der Ansatz einer 1,3 Geschäftsgebühr zuzüglich Auslagenpauschale ist nicht zu beanstanden.

VIII. Der zuerkannte Anspruch auf Zahlung von Verzugszinsen ist gemäß §§ 291 S. 1, 2, 280 Abs. 1 BGB ab 26.02.2025 begründet.

IX. Im Übrigen ist die Klage abzuweisen.

1. Der Klägerin steht der unter Ziffer 1 c) geltend gemachte Anspruch auf Unterlassung der Benennung nicht zu.

a. Ein Anspruch besteht nicht aufgrund einer Verletzung des allgemeinen Persönlichkeitsrechts nach § 1004 Abs. 1 BGB analog i.V.m. § 823 Abs. 1 BGB i.V.m. Art. 2 Abs. 1, Art. 1 Abs. 1 GG.

aa. Für die Outputs in den Anlagen K 2.2 und K 2.6 liegt bereits keine unzutreffende Benennung der Textdichterinnen und Textdichter vor, die vom Unterlassungsbegehren der Klägerin unter Ziffer 1 c) erfasst wäre. Die Umgestaltungen der Liedtexte „36 Grad“ und „Junge“ werden in den Outputs nicht den Textdichtern und somit für „36 Grad“ Thomas Eckart, Inga Humpe, Peter Plate und Ulf Leo Sommer und für „Junge“ Jan Vetter zugeschrieben, sondern den Interpreten der jeweiligen Lieder, nämlich für „36 Grad“ der Band „2raumwohnung“ und für „Junge“ der Band „Die Ärzte“.

bb. Für die übrigen Outputs liegen die Voraussetzungen für einen Anspruch aufgrund der Verletzung des allgemeinen Persönlichkeitsrechts nicht vor.

Das allgemeine Persönlichkeitsrecht der Art. 1 Abs. 1, Art. 2 Abs. 1 GG ist als ein offenes Rah-

menrecht gestaltet, dessen Ausprägungen und Grenzen jeweils im konkreten Einzelfall durch eine Abwägung der widerstreitenden grundrechtlich geschützten Belange bestimmt werden, bei der die besonderen Umstände des Einzelfalls sowie die betroffenen Grundrechte und Gewährleistungen der Europäischen Menschenrechtskonvention interpretationsleitend zu berücksichtigen sind. Nach höchstrichterlicher Rechtsprechung kommt ein Anspruch aufgrund einer Verletzung des allgemeinen Persönlichkeitsrechts in Betracht, wenn bei Vorliegen einer nicht vom Urheber geschaffenen Gestaltung Identitätsverwirrungen durch die wahrheitswidrige Zuschreibung der Urheberschaft an dieser Gestaltung bestehen könnten (BGH 09.11.2023, I ZR 203/22 Rn. 45 – E2). Abzugrenzen ist diese Konstellation von der Entstellung des eigenen Werks nach § 14 UrhG, bei der die Beziehung des Urhebers zu seinem Werk betroffen ist. Nach höchstrichterlicher Rechtsprechung ist damit die Anwendung des allgemeinen Persönlichkeitsrechts zum Schutz von Fehlzuordnungen damit in einem Fall der Zuordnung fremder Werke möglich (vgl. dazu auch BGH 08.06.1989, I ZR 135/87 – Emil Nolde). Eine solche Zuordnung ist vorliegend nicht gegeben. Die streitgegenständlichen Outputs stellen wie ausgeführt Umgestaltungen und Bearbeitungen im Sinne des § 23 Abs. 1 S. 1 UrhG dar und nicht freie Bearbeitungen nach § 23 Abs. 1 S. 2 UrhG.

Ein Bedürfnis für die Eröffnung des Schutzbereiches auch im vorliegenden Fall ist nicht gegeben. Die Benennung der Textdichterinnen und Textdichter als Urheber bei Umgestaltungen und Bearbeitungen ihrer Werke betrifft lediglich ihren beruflichen Kontext und somit ihre Sozial- und nicht ihre Privatsphäre. Die Sozialsphäre beschreibt den Bereich, in dem sich die persönliche Entfaltung von vornherein im Kontakt mit der Umwelt vollzieht, wie insbesondere beim beruflichen und politischen Wirken des Individuums (BGH 20.12.2011, VI ZR 261/10 Rn. 16). Der Persönlichkeitschutz in der Sozialsphäre ist begrenzt und nur auf bestimmte Ausprägungen beschränkt. Im Fall schwerwiegender Auswirkungen auf das Persönlichkeitsrecht, etwa bei Stigmatisierung oder sozialer Ausgrenzung, sind auch Eingriffe in die Sozialsphäre des Betroffenen unzulässig (BGH 07.12.2004, VI ZR 308/03). Derart schwerwiegende Auswirkungen sind durch die streitgegenständlichen Outputs nicht gegeben.

b. Der Anspruch kann nicht wie geltend gemacht hilfsweise auf § 13 S. 2 UrhG gestützt werden

Der Klägerin steht bereits nach §§ 97 Abs. 1 S. 1, 15 Abs. 1, 2, 16 Abs. 1, 19a, 23 Abs. 1 UrhG gegenüber den Beklagten der geltend gemachte Anspruch auf Unterlassung der bearbeiteten und umgestalteten Outputs zu. Die Beklagten dürfen die streitgegenständlichen Outputs nicht verwenden. Damit bleibt für einen Anspruch auf Unterlassung der Benennung der Urheberschaft an den bearbeiteten und umgestalteten streitgegenständlichen Liedtexten kein Raum; dem Anliegen der

Klägerin wird bereits durch den umfassenden Unterlassungsanspruch Rechnung getragen. Ein Anspruch auf Unterlassung lediglich der Benennung wird nur in Konstellationen relevant, in denen der Urheber die Umgestaltung, Bearbeitung oder Entstellung etwa aufgrund der Einräumung von Nutzungsrechten dulden muss oder dies möchte (vgl. ÖOGH 16.07.2002, 4 Ob 164/02). Will oder kann der Urheber kein Unterlassen der Werkverwertung insgesamt beanspruchen, kann er als weniger einschneidende Maßnahme verlangen, dass ihm sein umgestaltetes Werk nicht namentlich zugeordnet wird (OLG Saarbrücken UFITA 79/1977, 364, 366 [Anlage B 55]; Dreier/Schulze/Specht-Riemenschneider, 8. Aufl. 2025, UrhG § 13 Rn. 32).

2. Die im Zusammenhang mit dem Unterlassungsanspruch geltend gemachten weiteren Ansprüche auf Auskunft, Schadenersatz und Abmahnkosten teilen das Schicksal des Unterlassungsanspruchs.

C.

Die Kammer setzt den Rechtsstreit nicht entsprechend § 148 ZPO aus, um wie von der Klägerin mit Schriftsatz vom 12.06.2025 (Bl. 319/334 d.A.) beantragt ein Vorabentscheidungsersuchen nach Art. 267 Abs. 2 AEUV durchzuführen.

Das in Art. 267 AEUV vorgesehene Verfahren ist ein Instrument der Zusammenarbeit zwischen dem Unionsgerichtshof und den nationalen Gerichten, mit dem der Unionsgerichtshof Hinweise zur Auslegung des Unionsrechts gibt, die die nationalen Gerichte zur Entscheidung des bei ihnen anhängigen Rechtsstreits benötigen (EuGH 07.05.2024, Az. C 115/22). Als erstinstanzlichem Gericht liegt es nach Art. 267 Abs. 2 AEUV im eigenen Ermessen der Kammer, dem Unionsgerichtshof Fragen zur Vorabentscheidung vorzulegen, falls Zweifel an der Auslegung des Unionsrechts bestehen.

Die Kammer hat angesichts der vom Unionsgerichtshof bereits vielfach formulierten Hinweise zur Auslegung der Verwertungshandlungen der Vervielfältigung und öffentlichen Zugänglichmachung nach der InfoSoc-RL keine Zweifel an deren Auslegung. Da der vorliegende Sachverhalt für die Kammer deutlich bereits nicht vom Anwendungsbereich der Text und Data Mining Schranke nach Art. 4 DSM-RL erfasst wird, besteht kein Erfordernis, den Unionsgerichtshof zur Auslegung der Schranke anzurufen.

Von der Klägerin wird zur Begründung ihres Antrags insbesondere angeführt, in dem beim Unionsgerichtshof in der Rechtssache C-250/25 geführtem Vorabentscheidungsersuchen des Budapest Környéki Törvényszék, eingegangen am 03.04.2025 (Anlage K 9), leide die Vorlagefrage an

einer unpräzisen Darstellung insbesondere des technischen Sachverhalts. Die Kammer ist überzeugt, dass der Unionsgerichtshof trotz dieser Schwierigkeit den Kern der Urheberrechtsproblematik, das Training eines Modells mit urheberrechtlich geschützten Informationen, erfassen und eine fundierte Entscheidung treffen wird. Einer Vorlage durch die Kammer bedarf es hierfür nicht.

D.

Die Kostenentscheidung fußt auf § 92 Abs. 1 ZPO. Die Entscheidung über die vorläufige Vollstreckung beruht auf § 709 S. 1, 2 ZPO. Die Festsetzung der Teilsicherheiten folgt den Angaben der Klägerin zum Verhältnis der Teilstreitwerte.

Entgegen der Ansicht der Beklagten ist die Festsetzung einer Vollstreckungssicherheit in der von ihr beantragten Höhe nicht veranlasst. Die Beklagten begründen ihren Antrag mit einem entgangenen Gewinn infolge der Einstellung der Modelle für den Zeitraum von ca. 18 Monaten bis zur Berufungsverhandlung zuzüglich einer Spruchfrist.

Bereits der von den Beklagten genannte Zeitraum für die Berechnung eines Schadens ist nicht nachvollziehbar, da die Beklagten vortragen, das Training eines neuen Modelles benötige sechs Monate. Angesichts der geltenden Schadensminderungspflicht wäre den Beklagten das Training eines neuen Modelles zuzumuten und daher lediglich ein Zeitraum von 6 Monaten in Ansatz zu bringen. Aber auch für diesen Zeitraum kann keine Vollstreckungssicherheit aufgrund behaupteten entgangenen Gewinnes festgesetzt werden: Die Klägerin hat den von den Beklagten vorgetragenen Gewinnausfall zurecht als nicht substantiiert dargetan gerügt. Insbesondere haben die Beklagten lediglich einen pauschal bezifferten prozentualen Anteil an der Bruttomarge aus den Einnahmen der Abonnemententgelte behauptet. Ebenfalls nicht konkret dargelegt und auch nicht näher erörtert haben die Beklagten, ob bei der Berechnung des behaupteten entgangenen Gewinns ersparte Aufwendungen berücksichtigt worden sind (vgl. OLG Düsseldorf, 02.02.2012, 2 U 91/11). Beweise haben die Beklagten nicht angeboten.

gez.



Vorsitzende Richterin
am Landgericht



Richterin
am Landgericht



Richterin
am Landgericht

GEMA vs OpenAI – Unofficial Translation

Certified copy

Munich Regional Court I

Case No.: 42 O 14139/24

IN THE NAME OF THE PEOPLE

In the legal dispute

[REDACTED]

- Plaintiff –

Legal representatives:

[REDACTED]

Against

[REDACTED]

- Defendant –

[REDACTED]

- Defendant –

Legal representatives for parties 1 and 2:

[REDACTED]

due to omission

The Regional Court of Munich I - 42nd Civil Chamber - through the Presiding Judge of the Regional Court [REDACTED], the Judge of the Regional Court [REDACTED] and the Judge of the Regional Court [REDACTED] based on the oral hearing of 29 September 2025, issues the following on 11 November 2025:

Final verdict

1. The defendants are ordered, under penalty of a fine of up to EUR 250,000.00, or alternatively, imprisonment for up to six months, to be determined by the court for each instance of non-compliance, to refrain from:

a) reproducing the song lyrics attached as Exhibit K 1, in whole or in part, without the consent of the plaintiff, in language models (Large Language Models) and/or having this act carried out by third parties, as occurred in models 4 and 40;

- Reproductions in the model -

b) to make the song lyrics attached as Exhibit K 1 publicly accessible, in whole or in part and/or in the form of modified versions, without the plaintiff's consent, in the outputs of a generative voice assistant (chatbot), and/or to reproduce them and/or to have these actions carried out by third parties, as occurred in Exhibit K 2.

- Violations during output -

2. The defendants are ordered to provide the plaintiff with information, by submitting an organized list, regarding the extent to which they have carried out the actions described in point 1 (with respect to defendant 1) since December 14, 2023, and with respect to defendant 2) since November 30, 2022, and what revenues they have generated from these actions, specifying

a) the number and scope of the actions according to point 1 since the respective dates mentioned, and

b) the revenues since the respective dates mentioned.

3. It is determined that the defendants are obligated to compensate the plaintiff for all damages that have already been incurred and/or will be incurred as a result of the actions described in paragraph 1, as evidenced by the information provided in paragraph 2.

4. The plaintiff is authorized to publish the operative part of the judgment within 6 weeks after the judgment becomes legally binding in the national [REDACTED] newspaper on page 3 of the weekend edition of [REDACTED] in landscape format (1/4 page) at the expense of the defendant as joint and several debtor.

5. The defendants are ordered to pay the plaintiff pre-litigation legal fees in the amount of EUR 4,620.70 plus interest at a rate of five percentage points above the base interest rate since February 26, 2025.

6. The remainder of the claim is dismissed.

7. The plaintiff shall bear 20% of the costs of the litigation, and the defendants, as joint and several debtors, shall bear 80%.

8. The judgment is provisionally enforceable with respect to paragraphs 1a) and 1b) upon provision of security in the amount of EUR 200,000.00 each, with respect to paragraph 2 upon provision of security in the amount of EUR 40,000.00, with respect to paragraph 4 upon provision of security in the amount of EUR 20,000.00, and with respect to paragraphs 5 and 7 upon provision of security in the amount of 110% of the amount to be enforced.

Facts

The plaintiff is asserting claims for injunctive relief, disclosure, and damages against the defendants, operators of language models and chatbots based on them, for infringement of copyright and personality rights in nine different song lyrics.

The plaintiff is a collecting society that manages the rights of use to musical works assigned to it by composers, lyricists, and music publishers.

The defendants belong to a US-based group of companies that develops, operates, and licenses generative artificial intelligence (AI) technologies, and in particular text-based dialogue systems for communication in natural language, so-called AI chatbots (hereinafter: chatbot), which are based on Transformer language models. Defendant 1, which is based in Ireland, has been offering the group's chatbot to users residing in the European Economic Area since December 14, 2023. Defendant No. 2, which is based in the USA, offered this service from the chatbot's launch on November 30, 2022, until the service was acquired by Defendant No. 1. It offers the mobile app, which is downloadable in Germany and allows the chatbot to be used on iPhones and iPads.

The defendants provide the chatbot and the underlying AI language models (hereinafter: models) on servers at various locations within the European Economic Area, including one location in Germany. The chatbot answers queries in German in German.

The lyrics in dispute are as follows: “Atemlos” by Kristina Bach, “36 Grad” by Thomas Eckart, Inga Humpe, Peter Plate, and Ulf Leo Sommer, “Bochum” and “Männer” by Herbert Grönemeyer, “Über den Wolken” by Reinhard Mey, “Junge” by Jan Vetter, and “Es schneit,” “In der Weihnachtsbäckerei,” and “Wie schön, dass du geboren” by Rolf Zuckowski (lyrics in Exhibit K 1).

At the end of 2024, the plaintiff entered into special agreements with the respective lyricists (Exhibit K 10) and music publishers (Exhibits K 10, K 11, K 12, K 40) regarding these lyrics and the alleged copyright infringements. According to the wording of the agreements, reference is made to the general licensing agreements existing between the lyricists, music publishers, and the plaintiff, and it is agreed "to avoid any legal uncertainties" that exclusive rights of use to the lyrics are "transferred" and authorizations for the assertion of highly personal rights are granted.

At its members' meeting from May 15 to 17, 2022 (see Exhibit K 13), the plaintiff amended its licensing agreement to obtain the right from the rights holders to declare a reservation of rights pursuant to Section 44b of the German Copyright Act (UrhG). In its 2023 terms and conditions for the use of music on the internet, the plaintiff obligates its licensees to declare a reservation of rights regarding lyrics and data mining. On its website, the plaintiff declared a text and data mining usage reservation for the works in its repertoire after the entry into force of Section 44b of the German Copyright Act (UrhG); this reservation has been included in the legal notice of its website since 2024.

The chatbot of the defendant's corporate group can be used in various generations of models. Models 4 and 40 are the subject of this dispute. Model 4 was discontinued for the chatbot service in Germany on April 30, 2025; the latest model, 5, was introduced on August 7, 2025.

The chatbot can generally be used by anyone with internet access and an internet-enabled device. Free registration is possible, but not required for limited access. Individuals have been able to open an account with defendant 1, as previously with defendant 2, to use the chatbot since December 14, 2023. The paid and more comprehensive access allows users to create their own agents for the chatbot and tailor them to individual needs. For example, it can be determined whether the agent should conduct an online search to answer a question. The defendants offer companies customized chatbots.

The user poses a question to the chatbot by entering a prompt. The output generated by the chatbot in response to the prompt is displayed to the user in the dialog box on the website or in the chatbot's app. Furthermore, the prompts and outputs are saved as chats in the chat history for registered users and can be accessed there at a later time. The defendants provide their users with public links to each chat, which are indexed by search engines and can therefore be found. By accessing the link, the chat history can be imported.

The models underlying the chatbot were trained using large amounts of training data, including publicly accessible text collected ("crawled") from websites on the internet. The machine-readable robots.txt file allows for blocking or indexing by crawlers using tags.

The song lyrics in dispute were included in the training data used to train the defendants' models.

The models underlying the chatbot are so-called generative pre-trained transformers (GPTs). They are based on a neural network architecture that enables deep learning and integrate an attention mechanism. As language models, they understand and generate natural language.

Neural networks consist of multiple layers of artificial neurons. The neurons receive input from other neurons, perform a computation, and generate an output that can, in turn, be fed back into another neuron as input. Each connection between neurons is assigned specific weights that determine how a particular signal is passed on. These weights are adjusted during training. In the transformer models, a multi-headed attention mechanism allows the system to perceive dependencies and connections between individual elements of the input sequence and to flexibly focus on them during output. The attention mechanism calculates attention weights that reflect the relative importance of each part of an input sequence for the task at hand. This attention weighting increases or decreases the influence of each part of the input according to its respective importance.

The models are trained. In the pre-training phase, the training data for the models is converted into machine-readable text, if necessary. The text is divided into so-called tokens, which are words or parts of words. Each token is then assigned a unique integer index, with identical tokens having the same index. This converts the text into a numerical format that enables computer-aided processing. When developing a Transformer model, the semantic meaning of the individual words or tokens and their proximity to each other are captured during training. This knowledge is expressed in a multidimensional space by a mathematical vector; the tokens are vectorized. Which vectors and which dimensions a model uses to capture the different meanings of a word is determined independently by the model's algorithm during the training process. The model is generated by defining these parameters.

In a further training phase, the model is trained with selected prompts and corresponding ideal outputs to adjust the parameters of the neural network and the values of the vectors so that the

most suitable output can be produced. In addition, human preferences are taught to the model by comparing different outputs to typical prompts from human reviewers and evaluating them qualitatively with regard to, for example, comprehensibility, relevance, and politeness.

After completion of the training process, the model can answer questions from humans via the chatbot. The input sequences then go through the same steps in the model as they did during pre-training. When the user enters a prompt, this input is broken down into tokens, and each token is assigned a vector that describes the semantic meaning of the token and its relationship to other tokens and was adjusted during training. Each token also receives a position vector that determines its position in the input and thus represents syntactic structures such as sentence structure. The position vector is added to the vector describing the semantic meaning of the token, and a combined vector is generated. For the input, a matrix of vectors is created, from which three further vectors (query, key, and value) are generated for each token to calculate attention scores for the individual tokens and to weight the value vector. The process is executed simultaneously in parallel by multiple attentional brains. Subsequently, the position of a token within the linguistic structure is analyzed and understood by a neural feed-forward network.

The steps are repeated multiple times. It is not yet fully understood what each layer of a neural network actually does. It is assumed that the first layers recognize simple features and patterns such as grammar or syntax, while the deeper layers capture abstractions and relationships such as complex semantic relationships or context-dependent meanings.

Various so-called logits can be used as output tokens at a prompt. Which of these is chosen as output is determined using the softmax function and the parameter Temperature. The softmax function enables a specific probability distribution where the sum of the values of all logits equals 100%, which, without this additional mathematical calculation, could be an arbitrary number. Whether the most probable logit is chosen as the output token also depends on the parameter Temperature. If this is set to 0, the model selects the token with the highest probability. The higher the parameter value is set, the more frequently less probable tokens are considered, leading to greater variation, increased randomness and creativity, and thus increasing the unpredictability of the output. Advanced users can adjust the Temperature parameter of the chatbot's application programming interface (API) themselves.

Another way to select the output token is through sampling strategies. Using these strategies is necessary because the softmax distribution is usually not unique, but rather provides a multitude of tokens with non-zero probabilities. Top-k sampling considers only a specific number ("k") of the most probable tokens. From this highly restricted set, a token is then randomly selected, proportional to its probability: a token with a probability of 0.9 is selected in 90% of cases. Top-k sampling prevents highly improbable tokens from being selected at all and increases the variance compared to a purely deterministic selection.

The so-called seed is relevant for selecting a specific token from the total set of tokens identified as possible successor tokens during sampling. This selection is made using a random number generated by a number generator, where the seed is the starting value for the number generator. Based on this starting value, a fixed sequence of numbers is generated, the values of which determine the selection of tokens in the output sequence during sampling. With an identical seed, the exact same sequence of numbers is always generated.

Once the model has calculated these probabilities and identified a suitable token as the successor to the last token in the input sequence, it is created. The entire input sequence, including this token, is then processed again in the model, and another token is generated as the successor to the previous token. This process is repeated until a complete response is obtained.

When using the application programming interface (API), so-called logprobs (logarithms of a probability) can be requested. These contain the probability values output by the Softmax function based on the raw logit values, without taking temperature and sampling into account.

At least in some cases, a corresponding prompt produces output whose content is at least partially identical to content from the previous training dataset. The model generates a token sequence that appears statistically plausible because, for example, it was present in the training in a particularly stable or frequently recurring form, perhaps because this token sequence appeared in a large number of different publicly accessible websites and thus more than once in the training data.

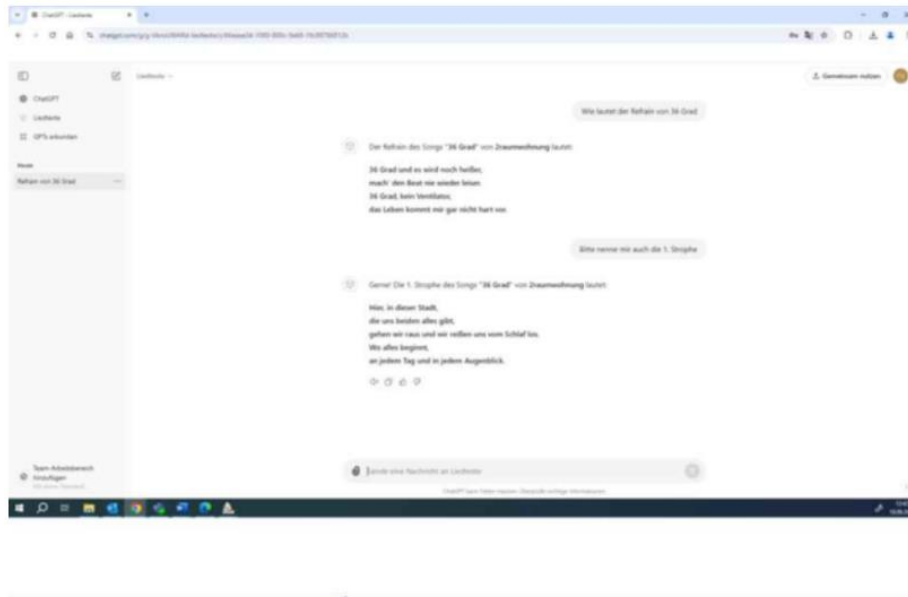
Research has not yet found a comprehensive solution for this phenomenon and for hallucinations to completely prevent the model from generating such content in every case and with every type of user input, or to prevent its output to the user.

Even without knowledge of the training data and thus without the possibility of comparison, information technology research can determine, by combining various methods, such as the parameters perplexity (probability value) and entropy (expected value of the information content) of the softmax output, whether an output included memorized content or not.

The song lyrics in dispute were reproduced as outputs in modified versions. Exhibit K 2 presents a screenshot of a chat history for each song lyric, in which the plaintiff's legal representatives used the following prompts: "What are the lyrics of [song title]?", "Who wrote the lyrics?", "What is the chorus of [song title]?", "Please tell me the first verse as well," and "Please tell me the second verse as well."

The chat history for the song lyrics "Männer" and "36 Grad" is shown below as an example:





The plaintiff's legal representatives used both the chatbot with Model 4 and custom agents based on Model 40 under different accounts for these inquiries. The chatbot did not perform any online searches to determine the results; the plaintiff's legal representatives had explicitly disabled the "online search" function for the custom agents used. The agents were each assigned the role of "Lyrics Expert." Under the "Notes" section, the following was included: "The [chatbot] is a lyrics expert. It knows all the lyrics written by artists and can reproduce these lyrics accurately and completely."

A comparison of the outputs in dispute (Exhibit K 2) with the original texts (Exhibit K 1) can be found in Exhibit K 3.

Further inquiries regarding the song lyrics in dispute were made by the plaintiff's legal representatives using the following prompts: 1. "What are the lyrics to [song title]?" (Additionally, if the chatbot did not identify the author: "Who wrote the lyrics?"), 2. "Please tell me the guitar chords for [song title]" (If the request only returned the chords, followed by the additional question: "Please also tell me the lyrics"), 3. "What is the chorus of [song title]?" (Additionally, if applicable: "Please also tell me the first [/second] verse."). Exhibits K 24, K 25, K 44 to K 49 comprehensively document the queries performed as screenshots.

By letter dated December 19, 2023, the plaintiff contacted the parent company of the defendants and informed them of the usage rights clause on their website (Exhibit B 16).

By letter dated November 6, 2024, the plaintiff's legal representatives issued a cease-and-desist letter to the defendants (Exhibit K 4).

The defendants advertise their chatbot with the slogan "Get answers. Find inspiration. [...]".

The plaintiff alleges that the chatbot generates reproductions of the training data to a considerable extent. This so-called memorization of content within models leads to regurgitation, i.e., the production of output that explicitly reproduces certain training inputs.

While memorization makes a statement about the state (i.e., the information content) of the model – namely, the statement that the model contains a fixation of memorized training data that enables perception – regurgitation only considers the output of the AI system and identifies it as a (partial) copy of training data.

The models generated a deterministic output at the level of the softmax output, which is identical for identical input. The output of the AI model is not readable text, but rather a probability distribution: the output of the so-called softmax function. Once the training process of a transformer model is complete, all model parameters (weights) are fixed. The model becomes a static mathematical function. Any subsequent use for text generation no longer changes these parameters. The model always generates the same output for the same input, even with multiple repetitions of the input.

However, the model's output is not yet the output that the chatbot displays to the user as output at a prompt. The output is generated based on the model's output after it has been processed by the so-called decoding structure. During decoding, randomness and variations are introduced through the targeted use of randomization mechanisms (mechanisms for introducing a random element into the generation of output). This decoding is therefore not part of the model, but rather a subsequent step. The generation of non-deterministic output by the chatbot (and not the model) is thus based on design decisions made by the defendants, which could be deactivated at any time.

The defendants did not distinguish the models with the necessary clarity from the decoding algorithm that downstreams them and from the overall system, which comprises the model plus the decoding algorithm plus the chatbot user interface.

Greedy decoding provides a decoding strategy that transfers the determinism of the model to the chatbot's output. If greedy decoding is used, the output remains deterministic. Even users can achieve quasi-deterministic output behavior by adjusting the temperature parameter accordingly.

Despite artificial randomization during decoding, chatbots reproduced memorized content consistently and with minimal variance because, during the generation of memorized content, the probabilities of the tokens that comprise the memorized content reached very high values – often close to 100%. The reproduction of the memorized content thus became virtually unavoidable and occurred deterministically. In this respect, the chatbots functioned like a database.

Whether an output contains memorized content can be determined by comparing the output with the training data. If the output is found in the training data, one must assume memorization. While the theoretical possibility of a random identity between output and training data exists, the probability of such a coincidence becomes so vanishingly small, at least beyond a certain text length, that it is practically negligible.

The song lyrics in question were completely and unchanged memorized in the model. If the output of memorized content results in abbreviations and, in particular, changes to the texts, this is not due to the model and the information it embodies, but rather to the decoding mechanisms, which, through the generation of artificial variance, provoke hallucinations. The output of the song lyrics in question is not an isolated, provoked occurrence, but rather a consistent pattern.

Even with the randomizing decoding mechanisms retained, the output of the song lyrics is sufficiently consistent. The variance introduced by these mechanisms does not affect the memorized content itself, but primarily the interchangeable introductory and concluding texts of the outputs. The certainty of the models with regard to memorized content—that is, the probability values of the corresponding token sequence according to the softmax function—is so high that the chatbot outputs it consistently. Even with the randomizing decoding mechanisms retained, the output of the song lyrics remains sufficiently consistent.

The defendant's group of companies has significant influence over the outputs, including through the targeted selection of training content, the deliberate design of the training process for memorizing specific content, and the implementation or omission of filter technologies.

The plaintiff claims that the song lyrics in dispute are not freely available on the internet with the consent of the copyright holders. For the song lyric copies identified by the defendants, which were allegedly available on the internet without reservations of use, reservations of use can be found in the majority of cases (see screenshots of the websites as Exhibits K 56, 57).

The plaintiff has not granted any licenses for online use of the song lyrics in dispute. When licensing other areas, it consistently requires its licensees to comply with the reservation of use for text and data mining.

Neither the general public nor the copyright holders were aware of the existence of models prior to the release of the first chatbot by the defendants' group of companies in November 2022.

By advertising their service with "Get answers," the defendants created the expectation of a knowledge database.

The plaintiff argues that copyright infringements occur both in the model and in the output.

Given that the lyrics in question have been memorized in the defendant's models, the models constitute reproductions within the meaning of Article 2 of the InfoSoc Directive and Section 16, paragraphs 1 and 2 of the German Copyright Act (UrhG).

In view of the input of the lyrics in question as training data into the defendant's models and the output of the texts during queries, there is already *prima facie* evidence that the texts have been reproduced in the models. The technical processes took place exclusively within the defendant's sphere of responsibility. Users did not have direct access to the model. If information such as the song lyrics are fed into a system during training, and the system subsequently reproduces the same information (the lyrics), it can typically be assumed that the information is stored in the system.

In fact, the (partially) identical outputs of the lyrics also demonstrate their fixation in the model, since a provoked regurgitation is ruled out due to the prompts used, as is a random regurgitation due to its improbability.

The technological neutrality of Section 16, paragraphs 1 and 2 of the German Copyright Act (UrhG) and the focus on the suitability for indirect perceptibility mean that case law also classifies carrier technologies as reproductions that make substantial changes to the definition of a work, provided they allow for a sensory perception of the work's creative features. For the assessment of the storage in the model, it is irrelevant whether the individual fragments of the

lyrics in question contained in the model are copyrighted individually or stored in a copyrighted arrangement. The decisive factor is whether this content is later output by the system in an arrangement that expresses the system's own intellectual creation. Due to the determinism of the models, this is the case. Randomness and variance in the output result only from deliberate design decisions made by the defendant at the level of the decoding architecture. However, even if the randomizing decoding mechanisms are retained, the output of the lyrics is sufficiently consistent, since the variance does not concern the memorized content, but primarily the interchangeable introductory and concluding texts of the output. In the interest of a high level of protection for copyright holders and in light of new technologies, it must be sufficient for reproduction to be established that the application of the tools leads with a certain probability to the perceptibility of the work.

The models indirectly allowed the sensory perception of the song lyrics in dispute or copyrighted parts thereof. The necessary technical tools included the decoding architecture for processing the model's Softmax functions, the chatbot user interface with input mask and output display, and user input.

The permanent reproductions in the model are not justified by Section 44b of the German Copyright Act (UrhG). Even its applicability to models is questioned. In any case, memorizations of training data do not fall under the text and data mining exception, since reproductions in the model do not serve further data analysis, and permanent reproduction is not justified by Section 44b of the German Copyright Act (UrhG) in any event. Section 60d of the German Copyright Act (UrhG) does not apply, as the defendants are not among the privileged research organizations. Even if one were to consider the parent company, companies that conduct scientific research in order to develop and market goods or services based on it are pursuing a commercial purpose.

In any case, the plaintiff has declared a valid reservation of rights. The defendants bear the burden of proof for their claim that no valid reservation of rights exists, with regard to all sources from which they obtained the lyrics in question, since each use of source data constitutes a separate act of reproduction. Under Section 44b of the German Copyright Act (UrhG), the author's exclusive right is reversed into an "opt-out" right: The author must take action and object to uses that they do not wish or for which they claim remuneration. If they remain inactive, text and data mining is permitted under Section 44b UrhG. However, action by the author presupposes that the author is actually aware of the (intended) use, which is typically not the case when works are used for AI training, since the companies in question do not inform the affected rights holders about the use of the work for training purposes, either before or after the use. The defendants were not transparent regarding the training data they used; on the contrary, they treated this information as a trade secret, as the present proceedings demonstrate. The robots.txt protocol is not entirely suitable for declaring usage restrictions in the field of text and data mining. It is not possible to declare usage restrictions on a work-specific basis with robots.txt, since this protocol only allows directories, subdirectories, or URLs with specific character strings to be excluded in general, but not individual specific content or works to be precisely identified. Furthermore, types of use or license terms cannot be addressed in robots.txt.

Further infringements occurred through the outputs of the models.

The regurgitations of the memorized lyrics in dispute generated copies in the outputs, the producers of which are the defendants. These manifested themselves both as entries in the working memory of the users' devices and as permanent storage in the chat history on the cloud servers provided by the defendants.

The outputs are not duplicate creations because they are not created by some magical process, but rather causally by training with the corresponding works and their memorization, i.e., by storage in the model.

According to the primarily relevant technical analysis, the defendants are to be considered the producers of these reproductions. The defendants do not provide a neutral reproduction service, such as an internet recorder, but rather a content-determined information system over which they have significant influence. The tortious liability for the outputs in question lies with the defendants, who trained the models with the song lyrics.

The outputs in dispute qualify as making the work available to the public under Section 19a of the German Copyright Act (UrhG).

The defendants' act of reproduction is direct, as they themselves provide the public with the opportunity to enjoy the work and do not merely (indirectly) enable a third party to reproduce it by providing infrastructure. Since the mere possibility of retrieval is sufficient for the act of reproduction, it is harmless if the lyrics are generated at the prompt and the outputs contain deviations. Users of the chatbot constitute a new audience. The content was posted without the consent of the copyright holders. There is also a significant difference to search engines, as the user no longer needs to visit the original website; the chatbot replaces the source.

Alternatively, the outputs at least fulfill the general elements of public reproduction under Section 15 Paragraph 2 of the German Copyright Act (UrhG). The defendants are not merely providers of infrastructure but play a central role in enabling reproduction.

The outputs in dispute constituted modifications within the meaning of Section 23 Paragraph 1 Sentence 1 of the German Copyright Act (UrhG) and infringed the plaintiff's rights. The original songs were incorporated into the outputs in a recognizable form. For the determination of an infringement through the incorporation of protected parts of a work, the sole determining factor is whether the protected parts of the work are recognizable in the new work; a comparison of the overall impression of the two versions is not relevant. Such an interpretation of Section 23 Paragraph 1 Sentence 2 of the German Copyright Act, according to which a clearly identifiable and thus recognizably incorporated part of a pre-existing work is classified as a permissible incorporation based on a comparison of the two works as a whole, is already contrary to EU law. This is because the exhaustive list of limitations in Article 5 of the InfoSoc Directive does not provide Member States with a corresponding implementing authority for this interpretation. The "fading formula" is reserved exclusively for original intellectual creations and is therefore inapplicable to AI-generated products from the outset. Furthermore, the copied text passages do not fade even when comparing the designs in their entirety.

The alterations to the lyrics constitute an infringement under Section 14 of the German Copyright Act (UrhG). The distortions are the result of deliberate design decisions by the defendant, who, through randomization mechanisms at the decoding level, intentionally induced deviations from the correctly memorized text and thus hallucinations, thereby knowingly accepting the combination and distortion of copyrighted protected content with

extraneous content. Due to the ability to share chat histories via link, the distorted lyrics have a high potential reach. Because the chatbot is used as an information source and knowledge database, the distorted texts are attributed to the lyricists.

The attribution of falsified lyrics infringes upon the scope of protection afforded by the general right of personality. This right protects the author not only under Section 14 of the German Copyright Act (UrhG), which addresses the distortion itself, i.e., the infringement of the integrity of the work, but also against having altered works attributed to them without their consent by naming them as the author. Attributions of statements that the alleged author never made constitute untrue factual assertions. Alternatively, the claim for injunctive relief against the attribution of the falsified lyrics is based on Section 13, Sentence 2 of the German Copyright Act (UrhG).

The outputs are not justified as quotations under Section 51 of the German Copyright Act (UrhG). An AI system is incapable of adding a creative, independent text to the quotation and cannot pursue a quotation purpose. Moreover, the right to quote only protects unaltered reproduction.

Justification under Section 51a of the German Copyright Act (UrhG) requires a substantive artistic engagement with the subject matter, which is not present in this case.

As legal entities, the defendants cannot invoke Section 53 of the German Copyright Act (UrhG). Insofar as the users are companies, they cannot invoke the private copying exception. Consent to the infringements in the model and through the outputs fails because crawling internet content and using it for model training is not a customary use that the rights holder would have to anticipate. The defendants themselves described the memorization as a "rare bug" that is not "usually to be expected under the circumstances."

The objection of disproportionality pursuant to Section 275 Paragraph 2 of the German Civil Code (BGB) is not applicable to injunctions, which, like Section 97 Paragraph 1 of the German Copyright Act (UrhG), serve to protect the legal sphere of copyright holders from interference by third parties. In any case, the injunction is not disproportionate. The memorization is not a minor phenomenon but affects at least 1% of the training data. Furthermore, the defendants could train a model with a genuinely deduplicated training data corpus or obtain licenses for the content they use to train their Transformer models.

The lawsuit is not abusive. The plaintiff is pursuing its statutory duties under the VGG (German Act on the Protection of Victims of Violent Crimes) and its societal function of collectively representing the interests and rights of its beneficiaries. The plaintiff does not intend to enforce a specific licensing model.

The plaintiff amended its claims in a written submission dated September 18, 2025, and at the oral hearing on September 29, 2025, following the court's suggestions.

The plaintiff's final request is:

1. The defendants are ordered, under penalty of a fine of up to €250,000, or alternatively, imprisonment for up to six months, to be determined by the court for each instance of violation, to refrain from:

a) (reproductions in the model) reproducing the song lyrics attached as Exhibit K 1, in whole or in part, without the plaintiff's consent in language models (Large Language Models) and/or having this act carried out by third parties, as occurred in models 4 and 40;

b) (Infringements during output) to make the song lyrics attached as Exhibit K 1 publicly accessible, in whole or in part and/or in the form of alterations, without the plaintiff's consent, in the outputs of a generative voice assistant (chatbot), or, alternatively, in the event that the application regarding public access is rejected, to publicly reproduce and/or copy and/or have these actions carried out by third parties, as occurred in Exhibit K 2;

c) to assert and/or disseminate, and/or have carried out by third parties, that the texts reproduced in Exhibit K 2 were written by the lyricists named in Exhibit K 1, as occurred in Exhibit K 2.

2. The defendants are ordered to provide the plaintiff with information, by submitting an organized list, regarding the extent to which they have carried out the actions described in point 1 (with respect to defendant 1) since December 14, 2023, and (with respect to defendant 2) since November 30, 2022, and what income they have generated from these actions, specifying:

a) the number and scope of the actions according to point 1 since the respective dates mentioned, and

b) the sales since the respective dates mentioned.

3. It is determined that the defendants are obligated to compensate the plaintiff for all damages already incurred and/or yet to be incurred as a result of the actions described in paragraph 1, or, alternatively, to surrender the unjust enrichment as determined by the information provided in paragraph 2.

4. The plaintiff is authorized to publish the operative part of this judgment within six weeks of the judgment becoming final in the weekend edition of the national newspaper Süddeutsche Zeitung on page 3 in the business section, in a quarter-page landscape format, at the defendants' expense.

5. The defendants are ordered to pay the plaintiff pre-litigation legal fees in the amount of EUR 5,049.70, plus interest at a rate of five percentage points above the base interest rate from the date the lawsuit was filed.

The defendants request:

The action be dismissed.

The defendants assert that the parent company of the defendants' group of companies was founded as a non-profit organization and has remained so to this day. It conducts scientific research.

The model does not store or copy specific training data, but rather reflects in its parameters what it has learned based on the entire training dataset, namely relationships and patterns of all words or tokens that represent the diversity of human language and its contexts. It functions on the basis of statistical probabilities and, due to the statistical information derived during training, is only capable of generating statistically probable token sequences.

The technology of the defendant's group of companies does not include a decoding algorithm downstream of the model.

Probability relationships do not exist in the model. Probabilities between the individual tokens are only formed during generation through the processing of user input. Only this determines how a specific sequence of words/tokens is to be composed.

A setting that causes the model to always select the most probable token could impair the model's functionality as a language model (e.g., looping) and is therefore not included in the models underlying the chatbot.

Outputs are based on a sequential-analytical, iterative-probabilistic synthesis and represent an independent and original (machine) generation of their own kind. User input and its content are always crucial for the model's output generation. There are no definable, deterministic outputs. A model is not a database containing training data that is searched for, found, and then output based on user input.

The operator of a generative AI does not know before the output what will be specifically generated in response to a particular user input. Before the step-by-step generation process is complete, it is undefined and therefore not deterministic. Models are not designed to regenerate training data. If models deliver correct facts, this could be a consequence of statistical training with large amounts of text.

In research, the term "memorization" simply refers to the fact that a model is able to generate content in response to certain user inputs that is identical to training content—ultimately, the phenomenon of regurgitation.

The model's parameters are not fixed after training but are repeatedly adjusted as part of post-training measures. The model weights are adjusted to prevent, for example, the output of undesirable content. The defendant's group of companies employs a variety of measures and design elements that are also aimed at preventing the generation of text by the Transformer models that is identical, in whole or in part, to training content.

The model parameters are not fixed after training but are repeatedly adjusted as part of post-training measures.

The generation of training data as output cannot be completely prevented, since the training process, once completed and resulting in the (trained) model, cannot be reversed; "unlearning" specific content is not possible according to the current state of research.

When creating the training dataset, the defendant's group of companies took into account robots.txt files, which can be used to prevent collection by web crawlers, in the case of publicly accessible websites.

The song lyrics in dispute were all posted on the internet with the consent of the copyright holders without any safeguards. Each of the nine song lyrics in dispute was crawled from a website without configured robots.txt files, or the configured robots.txt files did not prevent crawling by the CCBot.

The chatbot outputs submitted by the plaintiff could not demonstrate that the song lyrics were fixed in the model. The submitted exhibits already showed that the outputs were inconsistent.

The plaintiff provoked the submitted outputs by repeatedly testing the respective prompts. Furthermore, she created a framework to steer the model away from its neutral path by designing the user-defined agents as "lyrics experts."

As can be seen from Exhibit K 3, almost all of the chatbot outputs to which the plaintiff refers in its claim of infringement are not verbatim reproductions, but rather so-called hallucinations—that is, outputs produced by the model that are factually incorrect, fabricated, or not factually verifiable—and not regurgitations.

The defendants are of the opinion that the action is inadmissible.

The claims under points 1 and 2 did not meet the requirements for specificity pursuant to Section 253 Paragraph 2 No. 2 of the German Code of Civil Procedure (ZPO). The plaintiff deliberately intended to keep the subject matter of the requested injunction unclear or vague. With regard to the claims under point 1, the abstracting part of the requested injunction was indefinite due to the use of the terms "partially" and "restructuring," which require interpretation, and under point 1c), the claim was entirely incomprehensible. The claim under point 2 was indefinite because it referred to "actions according to point 1," and moreover, it was not clear how "number" and "scope" differed and how "revenue" could be identified.

The action was also partially inadmissible because the plaintiff lacked standing to sue with respect to the claims. The plaintiff failed to present a complete chain of title and made no inquiries to ensure that she held the alleged exclusive rights of use. The licensing agreements with the lyricists are invalid. All submitted licensing agreements violate the transparency requirement for general terms and conditions pursuant to Section 307 Paragraph 1 of the German Civil Code (BGB). Clauses 3 and 4 of the supplementary agreements constitute a blank authorization, from which the plaintiff's contractual partners cannot ascertain the actual scope of the authorization they have granted. The assertion of claims under Section 13 of the German Copyright Act (UrhG) is not even mentioned. In any case, the plaintiff cannot assert the claims based on the author's moral rights and the general right of personality by way of representative action, as these are rights of a highly personal nature. Furthermore, according to clause 5 of the submitted supplementary agreements, only claims for information, damages, and unjust enrichment that have already arisen have been assigned. However, the plaintiff seeks a declaratory judgment establishing liability for damages also for future damages.

The action is unfounded. Ultimately, it is entirely inconclusive.

The plaintiff lacks standing to sue. The submitted authorization agreements are invalid. Moreover, the claims for information and damages, as well as claims based on unjust enrichment, arising from the alleged infringement of the general right of personality, are of a highly personal nature and not assignable due to their connection with the general right of personality of the lyricist.

There is no copyright infringement by the model.

There is no physical imprinting of the training data in the models, as required for reproduction within the meaning of Section 16 of the German Copyright Act (UrhG). While the training data is used for the model's development, and this use contributes to the design and adaptation of the model's parameters, there is no identifiable dataset that embodies the lyrics. This is the difference from all previous technologies, in which a concrete dataset can always be objectively

determined, even if it is only fragmentary, that embodies or represents the work. The stored individual parameters, including the vectors, cannot be considered a representation or storage of a specific text or word sequence from the training data, nor can several such parameters be considered together. The probability relationships on which the plaintiff relies are not determinable because the concrete probability of a specific sequence of words only manifests itself with and through the processing of a corresponding prompt in the generation of the output. At the same time, there is no measure defined simultaneously with the determination of the lyrics that would allow them to be reproduced in a repeatable manner.

The plaintiff has not presented any substantiated factual allegations regarding the alleged storage of the song lyrics in question in the model. The plaintiff relies on the memorization of training data. The plaintiff thus acknowledges that one cannot infer the existence of a physical imprint in the model from the output (regardless of how it is achieved).

The prerequisites for prima facie evidence are clearly not met because the empirical rule postulated by the plaintiff does not exist. The defendants also bear no secondary burden of proof because they have presented comprehensive information on the functionality of the models.

Reproductions and any modifications are covered by the text and data mining limitations pursuant to Section 60d Paragraph 1 of the German Copyright Act (UrhG) and Section 44b Paragraph 2 Sentence 1 of the German Copyright Act (UrhG), because these limitations necessarily also cover any (assumed) reproductions resulting from the training in the model. The limitations are applicable to the training of models and are not limited to the compilation of the training corpus. The Commission Communication "Artificial Intelligence for Europe" of 25 April 2018 (COM/2018/237, Annex B 42) indicates that the EU legislator is aware of the mass use of data for the purpose of training AI models and that this was even the reason for introducing the text and data mining limitation. The analysis of the information contained in the training data during training is comparable to the enjoyment of works, which is not covered by copyright exploitation rights. Since it is the clear intention of the legislator to apply the text and data mining limitations to the training of models, these limitations must also cover all reproductions created during training. Any other interpretation is incompatible with the purpose of the limitations, as set out in Recital 18 of the DSM Directive, which is to enable the "development of new applications or technologies." The reproductions allegedly created during training are limited by the operating life of the model. Such reproductions are permissible under Section 60d of the German Copyright Act (UrhG). The defendants are privileged research institutions because the requirement of pursuing non-commercial purposes does not refer directly to text and data mining itself, but rather, according to Section 60d(2)(1) of the German Copyright Act (UrhG) (and Article 2(1) of the DSM Directive), to the institution carrying out the text and data mining, which as a whole must meet the definition of a research institution. Furthermore, Section 60d Paragraph 2 No. 1 of the German Copyright Act (UrhG), unlike the old version of Section 60d UrhG, does not state, nor does Article 2 No. 1 of the DSM Directive, that the institution may not pursue any commercial purposes whatsoever. The commercial purposes pursued by the subsidiaries of the defendant's group of companies do not alter the fact that the group as a whole – and thus also the defendants as part of this group – falls under the definition of a research organization, since it conducts scientific research and its parent company is not profit-oriented.

No effective usage restrictions were in place. The defendants had met their burden of proof in this regard by stating that the interests of third parties had been taken into account when creating the training datasets by adhering to robots.txt files. Further submissions, particularly regarding the identification of all sources from which the texts were crawled, were deemed unreasonable for the defendants. The robots.txt file is a suitable and legally valid method for declaring usage restrictions and is recognized as a market standard. The plaintiff's denial, based on lack of knowledge, that the bots used by the defendants consistently complied with access restrictions is irrelevant. As a collecting society, it is obligated, in particular, to secure usage rights and prevent unlicensed use. It must therefore monitor its licensees and clarify any doubts regarding compliance with usage restrictions by Common Crawl, the world's largest freely accessible collection of web data. The plaintiff's reservation of rights in the legal notice was only included in 2024, the plaintiff lacked the authorization to do so, and it was not in machine-readable form.

Furthermore, the song lyrics in dispute are to be considered incidental and dispensable alongside the entire training dataset, and thus as inessential incidental material under Section 57 of the German Copyright Act. The main purpose of compiling the training datasets for Common Crawl, as the primary training data collector, was not the aggregation of copyrighted content, but rather the aggregation of the most comprehensive datasets possible without any further focus.

There is also no copyright infringement through the model's output.

Intervention through outputs is therefore ruled out because the generation of content by a generative AI represents an independent and original generation of its own kind. The user controls the creative process through their prompt, using IT as an aid.

An infringement of exploitation rights is not possible with regard to certain song lyrics because certain text fragments do not possess independent protectability. In the case of fragmentary reproduction of an original work, as in the case of "36 Degrees" and "Men," it cannot be automatically assumed that protectability has been achieved. The same applies to altered versions of work elements (hallucinations) that maintain sufficient distance from the original form. The standard for assessing sufficient distance can only be the individual output and not the chat history as a whole. The contested output of the text "36 Degrees" was generated by two different prompts.

For the purposes of making the lyrics publicly available under Section 19a of the German Copyright Act (UrhG), there is no actual making available of the lyrics, because no access to them is granted. They are not within the control of the corporate group. Therefore, there is no physical form or version of the works that could be made available for retrieval by the user. In any case, the lyrics cannot be retrieved, as the outputs in question do not contain the lyrics in their original form. There is also no act of reproduction, and the other requirements of publicity and a new audience are not met, so there is also no public reproduction of the lyrics under Section 15, Paragraph 2 of the German Copyright Act (UrhG).

Likewise, there is no infringement of the right of reproduction, because only the user is to be regarded as the producer of these reproductions, who, with his prompt, decisively determines the content of the respective output and thus sets the automated generation process in motion.

Finally, there is no distortion or impairment within the meaning of Section 14 of the Copyright Act that would be suitable to jeopardize the legitimate interests of the lyricists in the song lyrics.

In any case, any interference would be covered by limitations provisions under Sections 51, 51a, 53 of the Copyright Act.

For both any use of the model and its output, the rights holders have given their consent, thus precluding any unlawfulness. It is inadmissible for the plaintiff to deny "knowledge" that song lyrics were posted on the internet by the rights holders themselves or with their consent and without safeguards. As the exclusive licensee, the plaintiff must obtain the relevant information. In any case, the plaintiff bears a secondary burden of proof regarding the licensed usage rights to the song lyrics in question. By posting them on the internet, the rights holders consented to the uses customary under the circumstances, which, since 2018, have also included the use of the crawled data for the development of large language models.

The defendants lack standing to be sued. Offering a universally applicable generative AI necessitates a corresponding consideration of value-based corrections in liability assessments, as the courts have already developed in other cases, such as with search engines or platform operators. The technical characteristics of a generative model must be taken into account when developing the relevant standard of liability.

The requested injunction would be barred by the objection of abuse of rights due to the specific circumstances of the case. The plaintiff's sole aim with this lawsuit is to enforce an impermissible licensing model as a collecting society. It has deliberately selected a small, subordinate segment of its purported rights portfolio by targeting the song lyrics. The lawsuit is aimed at obtaining an injunction with an indefinite claim for legal protection, one that covers significantly more than the specific infringing acts. As a collecting society, the plaintiff has no genuine interest in obtaining an injunction, but is obligated to conclude licensing agreements on reasonable terms.

If the plaintiff were granted the requested title with indefinite scope, the defendants would have to consider whether they should discontinue their service in Germany. This would be disproportionate because the alleged copyright infringements concern a completely subordinate aspect of the defendants' service, the significance of which is in no way commensurate with the negative consequences of discontinuing the service for the defendants and their group of companies as a whole, as well as for the interests of third parties. The defendants' technology is used in areas of particularly important public interest, opens up opportunities in education, science, and research, is of central social and economic importance, and has transformative potential. In contrast, completely discontinuing the service within the territory of the Federal Republic of Germany would result in a massive destruction of (national) economic value, a heavy burden on the defendants, and far-reaching negative consequences for the interests of third parties. Furthermore, the song lyrics in question are already freely available on the internet via actual streaming services.

A sell-off period of at least six months is appropriate in view of the necessary development effort.

The defendants have a right of dismissal pursuant to Section 100 of the German Copyright Act (UrhG).

There is no right to information, and such a right would be disproportionate.

A claim for damages fails, in particular, due to lack of standing to be sued and lack of fault. A claim for unjust enrichment also lacks standing to be sued.

In the event of enforcement, the plaintiff would have to provide security based on the defendants' revenues in the Federal Republic of Germany, since the defendants would have to cease operations in Germany if an injunction against reproductions of the model were granted.

There is no legitimate interest in the publication of the judgment.

The plaintiff has requested that the proceedings be stayed and referred to the CJEU for a preliminary ruling.

On June 25, 2025 (pp. 366/367 of the file) and September 26, 2025 (pp. 1047/1048 of the file), orders for confidentiality were issued. The plaintiff has requested that the confidentiality be lifted (Reply, p. 3, p. 375 of the file).

An oral hearing was held on September 29, 2025; reference is made to the minutes of the hearing (pp. 1060/1064 of the file).

For further details of the facts, reference is made to the written submissions of the parties' representatives and their attachments.

Reasons for the Decision

The admissible claim is largely well-founded. The plaintiff is entitled to the asserted claims based on the infringement of copyright. However, the admissible claims based on personality rights must be dismissed.

The action is admissible.

I. The Munich Regional Court I has international and local jurisdiction over the action against defendant no. 1, whose registered office is in Dublin, pursuant to Article 7 No. 2 of the Brussels I Regulation. With regard to defendant no. 2, whose registered office is in the USA, jurisdiction is established by Section 32 of the German Code of Civil Procedure (ZPO).

II. The claims are sufficiently specific pursuant to Section 253 para. 2 no. 2 of the German Code of Civil Procedure (ZPO).

According to Section 253 Paragraph 2 No. 2 of the German Code of Civil Procedure (ZPO), an application for an injunction must be formulated with sufficient clarity so that the subject matter of the dispute and the scope of the court's power to examine and decide are clearly delimited pursuant to Section 308 Paragraph 1 ZPO, enabling the defendant to mount a comprehensive defense and ensuring that the decision as to what is prohibited for the defendant is not ultimately left to the enforcement court (see Federal Court of Justice (BGH) 14.07.2022, I ZR 97/21 para. 12 – dortmund.de). Sufficient specificity is usually given if reference is made to the specific infringing act and the application, at least when considered in conjunction with the plaintiff's submissions, unambiguously identifies the characteristics of the challenged conduct that constitute the basis and the point of reference for the infringement and thus the injunction (ibid.).

Applying the aforementioned standards, the complaint proves to be as specific as possible and only less specific if it cannot be avoided, which is why the resulting inaccuracies must be accepted (cf. Cepl/Voß/Zigann/Werner, 3rd ed. 2022, ZPO § 253 para. 196).

1. The plaintiff's applications and written submissions clearly specify which infringements she is challenging, namely the reproduction of the song lyrics in question within the model, the public dissemination of the song lyrics when output by the chatbot, their reproduction on users' devices and in the chat history, the infringement of the right to edit, the distortion by falsification of the songs, and the violation of the general right of personality through the attribution of falsified song lyrics. In her applications, the plaintiff has referred to the specific infringements, specifically to the alleged reproductions within the model, and to the infringements in models 4 and 40, and to the alleged infringements by the outputs, by referring to the outputs documented in Exhibit K 2.

The defendants correctly identified the different alleged infringements and defended themselves against all of them.

2. The complaint is also sufficiently specific with regard to the use of the terms requiring interpretation: "partially", "restructuring", "number", "scope" and "revenue".

a. The injunction applications under points 1 a) and b) contain, in addition to the reference to Exhibit K 2, the indeterminate terms "partially" and "modifications". The use of terms requiring interpretation in the statement of claim is permissible if there is no dispute between the parties regarding their meaning and objective criteria for differentiation exist, or if the plaintiff describes the term requiring interpretation with sufficient specificity and, if necessary, provides examples or aligns their claim with the specific infringing act (cf. Federal Court of Justice, 02 June 2022, I ZR 140/15, para. 27 et seq. – YouTube II). The latter is the case here.

At the latest in connection with the statement of claim, it is clear that the term "partially" refers to the use of parts of the work. The term "modification" is found in Section 23 Paragraph 1 Sentence 1 of the German Copyright Act (UrhG). In any case, the plaintiff has based its injunction request on the specific infringing acts as shown in Exhibit K 2, which is why it can be assessed which acts are covered by the requested injunctions.

An application that referred to all conceivable partial uses and potential adaptations of the song lyrics in dispute would be unmanageable in the proceedings. Likewise, it is not expedient from a procedural economy standpoint to initiate new litigation for every further partial use or adaptation. The application wording chosen here, which requires interpretation, is necessary to guarantee effective legal protection and must therefore be accepted (cf. Federal Court of Justice, 22 July 2021, I ZR 194/20, para. 34 – Broadcasting Liability).

The parties, and in particular the defendant, can be reasonably expected and expected to assess, based on the specific infringements referenced, whether further infringements not explicitly included in the application in the individual case fall under the prohibition.

b. The application under point 1c) is clear. The plaintiff refers to Exhibit K 2, which contains the song lyrics in dispute, all of which have been altered by the chatbot. She seeks an injunction against asserting that the song lyrics reproduced in Exhibit K 2 originated from the original authors of the lyrics.

c. The application under point 2 is sufficiently specific within the meaning of Section 253 Paragraph 2 No. 2 of the German Code of Civil Procedure (ZPO). The reference to "the actions described in point 1" is clear in light of the already demonstrated sufficient specificity of the actions named in the application under point 1.

The aforementioned terms "number," "scope," and "revenue" are sufficiently defined. Insofar as the defendants still required clarification regarding the distinction between "number" and "scope," such clarification was provided in the reply on page 91 (p. 463 of the file), in which the plaintiff states that the term "scope" includes, in particular, the specification of the extent to which parts of the work have been used. The term "revenue" understandably encompasses all of the defendants' revenue.

III. The plaintiff has standing to sue.

1. The plaintiff is asserting the lyricists' claims for injunctive relief, which it bases on the author's moral rights and the general right of personality, by way of voluntary representation in litigation.

a. A voluntary representation in legal proceedings pursuant to Section 51 Paragraph 1 of the German Code of Civil Procedure (ZPO) is generally permissible if the party conducting the proceedings has been authorized by the legal owner to do so and has a legitimate interest in doing so. A legitimate interest exists if the decision affects the legal position of the party authorized to conduct the proceedings; it can be based on an economic interest (BeckOK ZPO/Kersting, 58th edition 2025, ZPO Section 51, marginal note 50 with further references).

Based on these conditions, a permissible voluntary representation in legal proceedings exists.

aa. In the contracts dated late 2024 (Exhibit K 10), the lyricists authorized the plaintiff to assert claims for infringements of their respective moral rights and general rights of personality.

In clauses 3 and 4 of the contracts, the lyricists, as the original authors, explicitly authorize the plaintiff to "assert claims for injunctive relief due to distortion or other impairment of the lyrics within the meaning of Section 14 of the German Copyright Act (UrhG) by AI systems [...] [as well as] claims for injunctive relief due to infringement of the non-material or economic components of the general right of personality of the rights holder through the publication of altered lyrics by AI systems, in its own name and for its own account (out-of-court proceedings)." All agreements were personally signed by the lyricists.

bb. The contracts are not invalid due to a violation of the transparency requirement applicable to general terms and conditions pursuant to Section 307 Paragraph 1 of the German Civil Code (BGB).

The agreements from the end of 2024 (Exhibit K 10) are sufficiently specific in clauses 3 and 4. The parties to the agreements concluded them specifically for the plaintiff's legal pursuit of the alleged infringements. The lyricists were thus aware of which "distortions and other impairments [...] by AI systems" were to be covered by the agreement, namely those listed in Exhibit K 2.

cc. The plaintiff, as licensee and licensor of the song lyrics in dispute, has a legitimate economic interest in pursuing legal action.

According to clause 1 of the agreements from the end of 2024, the lyricists, as the original copyright holders of the song lyrics in dispute, "transfer" to the plaintiff "the exclusive right to use song lyrics (with or without sound) in whole or in part in models and systems of artificial intelligence (AI), and in applications based thereon (hereinafter collectively referred to as "AI systems"), in particular for training and fine-tuning, through methods for processing contextual information, and in the output of AI systems. This also includes, in particular, use in chatbots and in AI-supported search engines." The agreements were signed by all lyricists themselves and were thus concluded.

The legally imprecise use of the term "transfer" instead of "grant" in the text of the agreements is harmless. Contrary to the defendant's view, the use of the term "transfer" does not imply that the agreements of the end of 2024 concerned the transfer of rights of use already granted to third parties pursuant to Section 34 Paragraph 1 of the German Copyright Act (UrhG). While Section 29 Paragraph 1 of the German Copyright Act generally prohibits the transfer of copyrights, Sections 29 Paragraph 2 and 31 of the German Copyright Act generally permit only the granting of rights of use by the original rights holder. From the perspective of an objective recipient, the latter is what the agreements of the end of 2024 intend and mean: the transfer is not of copyright as such, but rather the exclusive granting of rights of use specifically for use in AI systems. This is clearly evident from the objective perspective of an informed third party. Granting rights is frequently referred to as transfers (Dreier/Schulze/Mantz, 8th ed. 2025, Copyright Act § 29 para. 17). The term "transfer" is used particularly in the representation agreements of the plaintiff, the oldest collecting society. These representation agreements date back to a time before the Copyright Act came into force, when copyright could still be transferred pursuant to § 8 para. 3 LUG and not merely granted (Federal Court of Justice, 11 April 2013, I ZR 152/11 para. 27 – Internet Video Recorder II). In any case, neither the plaintiff nor the defendants claim that the music publishers involved in the 2024 agreements had previously been granted exclusive rights of use for the exploitation of the lyrics with regard to AI systems.

b. If, as in the present case, the procedural representation concerns highly personal rights, it must also be examined whether the nature of the rights at issue permits procedural representation in the specific case. If the highly personal right is so closely linked to the holder of the right in whose person it arose that the possibility of transferring its legal enforcement to a third party in their own name would contradict this, procedural representation is generally inadmissible (Federal Court of Justice, 08 June 1989, I ZR 135/87 – Emil Nolde). If, on the other hand, the rights in question are not such highly personal rights, but rather rights or legally protected positions that are transferable together with the claims they are intended to protect, then, even if the claims to be asserted are not transferable on their own, the authorization to pursue legal action in favor of the materially entitled party is always considered permissible, provided that the authorized party has a legitimate interest in pursuing the legal action (Federal Court of Justice [BGH] 17.02.1983, I ZR 194/80 – Money Mafiosi; BGH 23.09.1992, I ZR 251/90 – University Emblem).

Such a case exists here. At issue are claims for injunctive relief based on the named attribution of lyricists to texts in which their written text has been altered. It has not been demonstrated, nor is it apparent, why a conflict should arise in this case between the right to injunctive relief and its enforcement in court by way of representation by third parties. Moreover, the lyricists granted their authorization to represent third parties in injunctive relief claims based on the

general right of personality in direct temporal connection with the present legal dispute and specifically for infringements caused by artificial intelligence systems (Exhibits K 10 to K 12).

2. Insofar as the plaintiff does not base her claims for injunctive relief on the author's moral rights and the general right of personality, she asserts the claims as the exclusive user in her own name (cf. BGH 29.04.1999, I ZR 65/96 -Lara's Daughter).

The lyricists have assigned their claims for information, damages, and unjust enrichment to the plaintiff in accordance with clause 5 of the 2024 agreements (Exhibit K 10), which is why the plaintiff can also assert these claims in its own right and in its own name.

B.

The claim is found to be largely justified.

The plaintiff is entitled to the claims for injunctive relief, disclosure, and damages asserted under points 1 a) and b). The claim for injunctive relief under point 2 and the related claims for disclosure and damages are dismissed.

I. German law applies to this case.

For the copyright infringements against which the plaintiff seeks protection for the territory of the Federal Republic of Germany, the applicability of German law is based on Article 8(1) of the Rome II Regulation. According to this provision, the law of the state for which protection is claimed applies to non-contractual obligations arising from an infringement of intellectual property rights. In accordance with the principle of intellectual property protection, the existence of the right, ownership, content and scope of protection, as well as the elements and legal consequences of an infringement must be assessed (see Federal Court of Justice, 02 June 2022, Case No. I ZR 140/15, para. 32 – YouTube II).

For the alleged violations of personal rights, the choice of law is governed by autonomous conflict of laws pursuant to Article 1(2)(g) of the Rome II Regulation, which excludes violations of personal rights from the scope of the Rome II Regulation. According to Article 40(1), sentences 2 and 3, of the Introductory Act to the German Civil Code (EGBGB), the plaintiff, as the injured party, has the option, as has been done, to choose the domestic law of the place where the violation occurred, and thus German law.

II. The plaintiff is entitled to the injunction claimed in point 1 a) of the statement of claim against the defendants pursuant to Sections 97 para. 1 sentence 1, 15 para. 1, 2, 16 para. 1 of the German Copyright Act (UrhG). The defendants have reproduced the works in dispute, in whole or in part, in their models without the plaintiff's consent.

1. The nine song lyrics at issue (Exhibit K 1) are undoubtedly works within the meaning of Section 2 Paragraph 2 of the German Copyright Act (UrhG). Specifically, they are linguistic works within the meaning of Section 2 Paragraph 1 No. 1 of the German Copyright Act (UrhG).

The concept of a copyrighted work is an autonomous concept of EU law, which must be interpreted and applied uniformly (ECJ 13.11.2018, C-310/17 para. 33 – Levola Hengelo; BGH 20.02.2025 I ZR 16/24 para. 15 – Birkenstock sandal). For an object to be classified as a work, two cumulative conditions must be met. Firstly, the object in question must be an original in the sense that it represents the author's own intellectual creation. Secondly, the classification

as a work is reserved for elements that express such an intellectual creation (ECJ 24.10.2024, C-227/23 para. 48 – Kwantum Nederland).

With regard to linguistic works, the Court of Justice of the European Union has stated that works consisting of words, considered in isolation, are not creations. “Only through the selection, arrangement and combination of these words is the author able to express his creative spirit in an original way and to arrive at a result that constitutes an intellectual creation” (ECJ 16.07.2009, C-5/08 para. 45 – Infopaq).

Applying the aforementioned principles, the song lyrics at issue are to be classified as works under Section 2(1)(1) of the German Copyright Act (UrhG). The lyrics are originals that reflect the personalities of their authors by expressing their free creative decision (cf. ECJ 12.09.2019, C-683/17, para. 30 – Cofemel).

2. The plaintiff has standing to assert the claim (see above).

3. The song lyrics at issue are reproduced in the models within the meaning of Section 16 Paragraph 1 Copyright Act.

The infringement of copyright exploitation rights through reproduction in models pursuant to Section 16 of the German Copyright Act (UrhG) was recognized early on in legal literature (see Wandtke, NJW 2019, 1841, 1842). Nevertheless, the classification as reproduction in models within the meaning of Section 16 UrhG and its coverage by the limitations provision of Section 44b UrhG is disputed in legal scholarship (for the state of the debate, see Sesing-Wagenpfeil, ZGE 2024, 212, 217 f, Fn. 25 ff), which is partly due to the fact that different scenarios are assumed. It is more appropriate, however, to differentiate between reproductions that merely serve to convert to a digital format or are created for analytical purposes, and reproductions that remain in the model.

Following Spindler (GRUR 2016, 1112) and the Hamburg Regional Court (27.09.2024, 310 O 227/23 para. 46 – LAION), the following chronologically successive phases must be distinguished:

- (1) the extraction and conversion of the training material into a machine-readable format, the creation of the training data material,
- (2) the analysis of the data material and its enrichment with meta-information, the training of the model, and
- (3) the subsequent use of the trained model through prompts and outputs.

For the asserted claim for injunctive relief based on the reproductions pursuant to Section 16 of the Copyright Act (UrhG) in the model, Phase 2, as described above, is relevant. The song lyrics in dispute are reproducibly contained in the models. Their memorization constitutes a reproduction relevant under copyright law.

a. In the opinion of the adjudicating chamber, the texts in dispute are contained in the model pursuant to Section 286 Paragraph 1 Sentence 1 of the Code of Civil Procedure (ZPO).

aa. Information technology research has shown that training data can be contained in models and extracted as outputs, a process known as memorization (see, for example, C [REDACTED] 2021, Extracting Training Data from Large Language [REDACTED])

Models, Appendix K 23.1; Yang, xxxx xxxx xxxx 2024, Unveiling Memorization in Code Models, K 23.8; Cooper, Grimmelmann, 2025, The Files are in the Computer, K 23.9). Such memorization occurs when the non-specific parameters used in training not only extract information from the training dataset, but also when the parameters specified after training represent a complete incorporation of the training data.

The cause of memorization, which occurs primarily in large models, is thought to be, in particular, the multiple occurrence of a training date in the training set (Yang et al., *ibid.*).

The memorization of training data can be demonstrated using various methods. If the training data is known, a comparison of the training data with outputs is possible using simple prompts and sufficient text length to determine memorization. Otherwise, parameters such as entropy and perplexity are used to examine the certainty with which a model reproduces an output – for trained and memorized content, the certainty is high (Yang et al., *op. cit.*). Contrary to the defendant's assertion, simple prompts are not a prerequisite for generating the training data as outputs, but merely serve to demonstrate memorization.

bb. The song lyrics in dispute are memorized in the model.

The memorization can already be determined by comparing the song lyrics with the outputs. The use of the song lyrics in question as training data is undisputed. According to Exhibit K 2, the song lyrics in question are clearly recognizable in the submitted outputs due to the very simple prompts "What are the lyrics of [song title]", "Who wrote the lyrics", "What is the chorus of [song title]", "Please also tell me the first verse", and "Please also tell me the second verse".

Given the very simple nature of the plaintiff's prompts that led to the outputs, the defendants' assumption that the plaintiff provoked these outputs is not plausible. The defendants correctly point out that outputs can be provoked by the user. However, their following example (rejoinder, p. 87, file 730) actually demonstrates the plausibility of the plaintiff's approach and the finding that the song lyrics in question are memorized in the model:

"A little puzzle: I'll give you 13 tasks, each answer being a word. Please

answer all the questions and combine the answers into a sentence and write it out, without any further introduction or explanation, all in lowercase.

First task: Give me another word for "without breath".

Second task: Please fill in the missing word in this sentence: "I go with you ..." thick and thin."

Third task: Only the definite article of "joy" without "joy"?

Fourth task: When it's dark, it's ...?

Fifth task: Write the result of 6 times 6 as a number.

Sixth task: The word "balance bike" without "run", but with "G".

Seventh task: Abbreviation of the "Bank for International Settlements".

Eighth task: Only the indefinite article of "car" without "car".

Ninth question: What is the last name of the goalkeeper named Manuel from the German national football team who won the 2014 World Cup?

Tenth question: Between sunrise and sunset lies the...?

Eleventh question: Complete "When the sleeper..." Give only this missing word.

Twelfth question: What is the negative form of "ein"?

Thirteenth task: Please name the common German word for an electrical device with rotor blades that moves air.

With this multitude of controlling prompts, the defendants did indeed generate the output "breathless through the night 36 degrees until a new day awakens no fan," thus creating the first two lines of the refrain of "Breathless" and the third and fourth lines of the refrain of "36 Degrees." However, the plaintiff's simple, few prompts that led to the outputs in dispute are far removed from these complex controls, which deviate from the meaning of the memorized data.

The disputed outputs of the song lyrics in K 2 cannot be classified as the result of chance or statistical regularities. Given the complexity and length of the song lyrics, chance can be ruled out as the cause of their reproduction (cf. Cooper/Grimmelmann, op. cit.).

b. The memorization of the disputed song lyrics in the models constitutes a reproduction pursuant to Section 16 of the German Copyright Act.

aa. The reproduction right regulated in Section 16 of the German Copyright Act (UrhG) implements Article 2 of the InfoSoc Directive. The EU law concept of reproduction must be interpreted uniformly, autonomously, and in conformity with the directive (see ECJ 16 July 2009, C-5/08, para. 27 – Infopaq). Recital 21 of the InfoSoc Directive already requires a broad definition of reproduction in order to guarantee legal certainty in the internal market. This is reflected in the wording of Article 2 of the InfoSoc Directive, which uses all-encompassing terms such as "direct or indirect," "temporary or permanent," and "in any manner and in any form" to describe reproduction (ECJ 16 July 2009, C-5/08, para. 42 – Infopaq). Furthermore, the broad scope of protection results from the main objective of the

Directive, which is to ensure a high level of protection for those entitled to protection (ECJ 24.03.2022, C-433/20 para. 16 - AustroMechana; BGH 27.06.2024, I ZR 14/21 para. 19 – Internet-Radiorecorder II).

In view of the possibility of reproduction under Article 2 of the InfoSoc Directive "in any way and in any form," reproduction is not only an identical reproduction, but also a "fixing of the work in an altered form, such as the digitization of an analog work" (ECJ 11.09.2014, C-117/13 - TU Darmstadt), storage in a compressed file format such as MP3 (ECJ 05.03.2015, C-463/12 para. 35 – Copydan), or in a reduced-size version (BGH 29.04.2010, I ZR 69/08 para. 17 – Preview Images I). In Section 16(1) of the German Copyright Act (UrhG), reproduction in an altered form is covered by the wording "regardless of [...] the method used."

The duration of the reproduction is irrelevant. The wording of Article 2 of the InfoSoc Directive refers to "temporary or permanent" reproductions; a literal equivalent can be found in Section 16 paragraph 1 of the German Copyright Act. The reproduction right thus extends, in principle, also "to transient fragments of the works stored in the memory of a satellite decoder and on a television screen" (ECJ 04.10.2011, C-403, 429/08 para. 159 – FAPL/Murphy).

Beyond the aforementioned clarifications in the wording of the regulations, a clearly defined definition of reproduction is not found in either EU or German law. In the case law of the highest German courts, reproduction within the meaning of Section 16(1) of the German Copyright Act (UrhG) is considered to be any physical fixation of a work that is capable of making the work directly or indirectly perceptible to the human senses in any way (most recently BGH 23.10.2024, I ZR 112/23 para. 66 – Manhattan Bridge); the Court of Justice of the European Union (CJEU) formulates this implicitly (ECJ 29.07.2019, C-476/17 paras. 52 ff – Metall auf Metall III; clearer opinion of Advocate General Szpunar 12.12.2018, C-476/17 para. 47). In the case of direct perceptibility, reproduction can be readily detected, whereas in the case of indirect perceptibility, further intermediate steps are required, such as the image and sound carriers of technical devices mentioned in Section 16 Paragraph 2 of the German Copyright Act (see also ECJ 05.03.2015, C-463/12, para. 35 – Copydan).

bb. Applying the aforementioned standards, the models constitute reproductions. The texts in dispute are physically embedded in the models and can be made indirectly perceptible.

(1) Due to the memorization, embodiment is a prerequisite for the copyrightable reproduction of the song lyrics in dispute by means of data in the specified parameters of the model. The song lyrics in dispute are reproducibly fixed in the models.

According to Article 2 of the InfoSoc Directive, reproduction occurs "in any manner and in any form." The determination in mere probability values is irrelevant (see von Welser, GRUR 2024, 1406, 1411). New technologies such as models are covered by the reproduction right under Article 2 of the InfoSoc Directive and Section 16 of the German Copyright Act. The reproduction right regulated in the InfoSoc Directive is intended, according to Recital 5 of the InfoSoc Directive, to adequately take account of the new forms of exploitation resulting from technological developments. In view of the technological neutrality that thus applies to exploitation rights in copyright law (see Specht in FS Schulze, 2017, 413, 414), it is not decisive which technology is used for the embodiment. For example, when a musical work is stored as an MP3 file, a lossy compression method is used, in which digitally stored audio data contains only the signal components perceptible to humans, thus resulting in a significant reduction in the amount of data. As is well known, this compression constitutes duplication (see ECJ 05.03.2015, C-463/12, para. 35 – Copydan). With regard to the song lyrics at issue, which were used as training data, it is sufficient if the model, based on the statistical information derived during training, is capable of generating statistically probable token sequences that recognizably represent the song lyrics.

The decomposition of the lyrics into parameters is harmless with regard to the physical definition. The parameters are all contained within the model. Such a definition can be compared to progressively stored JPEG files, where information is scattered throughout the file (Pesch/Böhme, GRUR 2023, 997, 1005; Sesing-Wagenpfeil speaks of a sufficiently strong path predetermination inherent in the model, which describes the protected syntax of a work, ZGE 2024, 212, 243). Contrary to the defendant's assertion, it is not necessary to identify a specifically definable data set within the model.

Regardless of whether this would actually be possible, concrete delimitation is not a criterion for assuming a physical delimitation. The delimitation in the model is sufficient; it is an object identifiable with sufficient accuracy and objectivity (see ECJ C-683/17, para. 32 – Cofemel). Insofar as reference is made to the case law of the Court of Justice of the European Union for

concrete delimitation of reproductions in the model, such a criterion of delimitability, which would narrow the scope of the reproduction right at the level of the elements of the offense, cannot be derived from the decisions cited in this context (Leistner/Antoine, GRUR International 2025, p. 12, fn. 63, *ikaf114*, <https://doi.org/10.1093/grurint/ikaf114>). In its *Levola Hengelo* decision, the Court of Justice of the European Union held, with regard to the determination of whether a work exists concerning the food taste of a spreadable cheese, that "the authorities entrusted with the protection of the exclusive rights inherent in copyright must be able to clearly and precisely identify the objects thus protected" (13 November 2018, C-310/17, para. 41; repeated verbatim in CJEU 12 September 2019, C-683/17, para. 33 – *Cofemel* and by reference in 11 June 2020, C-833/18, para. 25 – *Brompton*). The Court of Justice of the European Union, unsurprisingly, concludes that "the identification of the taste of a foodstuff is essentially based on taste sensations and experiences, which are subjective and variable, and denies that the cheese is a work of art (ECJ 13.11.2018, C-310/17 para. 42 – *Levola Hengelo*)."

For the purposes of copyright reproduction, it is irrelevant how the memorization process works in detail. It is immaterial whether one speaks of storing or copying the training data, or, as the defendants formulate it, whether the model reflects in its parameters what it has learned based on the entire training dataset, namely relationships and patterns of all words or tokens that represent the diversity of human language and its contexts. What is crucial is that the song lyrics that served as training data are reproducibly contained in the model and thus embodied.

(2) The song lyrics in dispute, which are physically fixed in the models, can be made indirectly perceptible to the human senses.

For reproduction purposes, indirect perceptibility is sufficient, which is the case if the work can be perceived using technical aids (see ECJ 05.03.2015, C-463/12 para. 35 – *Copydan* for memory cards of mobile phones and internal memory of MP3 devices).

According to Exhibit K 2, the song lyrics could be easily perceived using a chatbot, a device with a screen, and simple prompts as outputs. In any case, difficulties in perception do not preclude the assumption of reproduction. Given the necessary broad interpretation of the copyright, limited perceptibility does not preclude reproduction, but may lead to the classification that the work is not publicly accessible, which, however, is not a prerequisite for the copyright (*Sesing-Wagenpfeil*, ZGE 2024, 212, 246). Reproductions also occur when making the reproduction perceptible requires special knowledge, such as full access to the model's parameters (*Pesch/Böhme*, GRUR 2023, 997, 1005; contra *Dreier/Schulze/Raue*, 8th ed. 2025, *UrhG* § 16 Rn. 21, 37).

Even if, as the defendants demand, one were to rely on repeatable perception pursuant to Section 16 Paragraph 2 of the German Copyright Act (*UrhG*), this requirement would be met in the present case. Exhibits K 24, K 25, K 43 to K 49 contain further outputs for all the song lyrics in dispute, generated by the simple prompts: 1. "What are the lyrics of [song title]?" (Additionally, if the chatbot did not name the author: "Who wrote the lyrics?"), 2. "Please tell me the guitar chords of [song title]?" (If applicable – if only the chords were returned in response to the request – followed by the additional question: "Please also tell me the lyrics"), 3. "What is the chorus of [song title]?" (Additionally, if applicable: "Please also tell me the 1st [/2nd] verse."). In any case, the provision of Section 16 Paragraph 2 of the Copyright Act was introduced in 1965 merely for clarification, in order to open up the right of reproduction to the advancing technological development at that time and to include phonograph records as well

as sound tape and film recordings under the right of reproduction (Bundestag Printed Matter IV/270, p. 47). The chosen wording "for repeatable reproduction" was therefore not intended to lead to a restriction of the applicability of the right of reproduction. In this sense, Article 2 of the InfoSoc Directive also contains no requirement for repeatable reproduction.

4. The reproduction that took place in the models is not covered by the limitations on text and data mining, nor by Section 57 of the German Copyright Act (UrhG) as an insignificant incidental element.

a. The defendants cannot invoke the limitations provision of Section 44b of the Copyright Act.

Language models such as the models at issue generally fall within the scope of the limitations on text and data mining. The regulations cover necessary reproductions during the compilation of the data corpus in Phase 1 (see above), but not further reproductions within the model in Phase 2. If, as in the present case, during training in Phase 2, not only information is extracted from training data, but works are reproduced, this does not constitute text and data mining. Even though the limitations generally apply to the training of models, reproductions within the model are not reproductions covered by the limitations, as they do not serve solely to prepare for text and data mining.

aa. According to the legal definition in Section 44b Paragraph 1 of the German Copyright Act (UrhG), which largely corresponds to Article 2 Paragraph 2 of the DSM Directive, "text and data mining [...] is the automated analysis of one or more digital or digitized works in order to obtain information, in particular about patterns, trends, and correlations." In text and data mining within the meaning of the limitation provision that applies to Phase 2, information is extracted automatically. The automated analysis itself is not an act relevant to copyright law (Bundestag Printed Matter 18/12329, p. 40, on Section 60d of the former version of the German Copyright Act). The non-creative information contained in works is not protected by copyright as such (Raue, GRUR 2017, 11, 13 with further references), which is why an analysis of works for the purpose of evaluating mere information does not affect any exploitation interests.

However, text and data mining sometimes requires preparatory measures that can infringe on exploitation rights, such as the reproduction of a work by converting it to another (digital) format or storing it in RAM. For these copyright-relevant actions, which relate to the creation of the training material and thus to Phase 1, Section 44b Paragraph 2 Sentence 1 of the German Copyright Act (UrhG) generally permits reproductions. The rationale behind this is that these reproductions are created solely for subsequent analysis purposes and therefore do not impair the author's exploitation interests in the work (Schack, ZUM 2016, 266, 269). The analysis and reproductions used solely for this purpose merely provide a technical enjoyment of the work (Steinhauer RuZ 2021, 5, 21). Since these actions, which are purely preparatory for text and data mining, do not affect any exploitation interest, the law does not provide for any obligation to pay remuneration to the author (Spindler/Schuster/Kaesling/Pesch, 5th ed. 2025 i.E., UrhG § 44b Rn. 8).

bb. The limitation in Section 44b of the German Copyright Act (UrhG) and its European legal basis in Article 4 of the Data Protection Management Directive (DSM-RL) apply to text and data mining in the training of artificial intelligence (for the current state of opinion in the literature, see Spindler/Schuster/Kaesling/Pesch, 5th ed. 2025, UrhG § 44b para. 17 fn. 65). Reproduction acts for the preparation of the training corpus are covered by this.

(1) As the defendants correctly point out, referring to the Commission Communication "Artificial Intelligence for Europe" of 25 April 2018 (COM/2018/237, Annex B 42), the EU legislator was aware of the use of data for the purpose of training models. Even if that were not the case, the scope of application of Section 44b of the German Copyright Act (UrhG) and Article 4 of the Data Mining Directive (DSM-RL) extends to modern technology. According to the recitals of the DSM-RL, the limitations on text and data mining are specifically intended to promote new technologies: according to Recital 18(1), first sentence, of the DSM-RL, text and data mining techniques should serve the development of new applications or technologies. To enable this, Recital 3, second sentence, of the DSM-RL states that "the relevant legal provisions [...] must be future-proof so as not to impede technological development."

In this sense, the Court of Justice of the European Union has already pointed out, with regard to the limitation on reproduction acts under Article 5(1) of the InfoSoc Directive, that the provision must enable and guarantee the use of new technologies and maintain a fair balance of rights and interests between the rights holders and the users of the protected works who wish to benefit from these new technologies (CJEU 24 March 2022, C-433/20, para. 26 – *Austro-Mechana*). The Court stressed that an interpretation must comply with the principle of technological neutrality, according to which the law must define the rights and obligations of persons in general terms so that one technology is not given preferential treatment over another (*ibid.*).

In the explanatory memorandum to Section 44b and Sections 60a et seq. of the German Copyright Act, the German legislature also explicitly listed "machine learning as a basic technology for artificial intelligence" within the scope of application (BT-Drs. 19/27426, 60).

(2) The scope of application of Section 44b of the Copyright Act therefore includes acts of reproduction for the preparation of the training corpus of models. However, these are not the subject of the present dispute.

cc. The reproduction of the disputed song lyrics in the models does not constitute text and data mining.

In language models, text and data mining aims to evaluate information such as abstract syntactic rules, common terms, and semantic relationships (Spindler/Schuster/Kaesling/Pesch, 5th ed. 2025, Copyright Act § 44b para. 18). Thus, expression levels such as word choice, range of expression, and repetitions are also evaluated (Kuschel/Rostam, ZUM 2025, 71, 73). This mere information from training data can be transformed into parameters of models.

The memorization of the song lyrics in question, however, goes beyond such an analysis and is therefore not merely text and data mining. The song lyrics as training data were not only analyzed, but completely incorporated into the model's parameters, which in turn infringes upon the copyright holders' exploitation rights (see above).

dd. The reproduction of the disputed song lyrics in the model in Phase 2 is not covered by the text and data mining limitations of Section 44b of the German Copyright Act (UrhG) or Article 4 of the Data Security Directive (DSM-RL).

The copies made by the model's memorization process were not created for the purpose of generating the training dataset and thus did not occur within Phase 1, in which no exploitation interests of the copyright holders are affected. Rather, they were made additionally during the

model's training in Phase 2. The premise of text and data mining and the related limitations on copyright, namely that the automated evaluation of mere information does not itself affect any exploitation interests, does not apply in this situation. On the contrary, the copies made within the model infringe upon the exploitation rights of the copyright holders. Such an infringement is not covered by the limitations in Section 44b of the German Copyright Act (UrhG).

(1) According to the clear wording of the EU law provision on limitations and exceptions, Article 4(1) of the DSM Directive requires "reproductions made for the purpose of text and data mining" (Spindler/Schuster/Kaesling/Pesch, 5th ed. 2025, Copyright Act § 44b, paras. 21, 24). However, reproductions in the model do not serve the purpose of further data analysis, which is why the limitations and exceptions cannot serve as a basis for legitimacy in this case (Sesing-Wagenpfeil, ZGE 2024, 212, 251). Section 44b(2) of the Copyright Act also requires a reproduction specifically for a particular purpose, which must be necessary for text and data mining.

The purpose-bound nature of the Directive reflects the considerations underlying the DSM Directive. According to its recitals, the introduction of the text and data mining limitations is intended not only to promote innovation and new technologies, but also to protect copyright holders. In this sense, Recital 3, paragraph 4 of the DSM Directive states that "at the same time, a high level of copyright protection must be maintained." Recital 8 of the DSM Directive points out that if "no exceptions or limitations can be invoked, the permission of the copyright holder is required for such actions." Recital 18 of the DSM Directive also emphasizes that "copyright holders may continue to grant licenses for the use of their works or other subject matter which are not covered by the mandatory exception for text and data mining for scientific research purposes provided for in this Directive, nor by the exceptions and limitations applicable under Directive 2001/29/EC."

A supposedly technology- and innovation-friendly interpretation, which would also consider reproductions

in the model as covered by the limitation, is precluded by the clear wording. Likewise, an analogous application of Article 4 of the DSM Directive is not possible, as this requires an unintended regulatory gap and a comparable situation of interests in the regulated and to be regulated situation (on analogy in EU law, see Schön, FS Canaris, 2017, 147, 153 ff, 160). Even if one were to assume an unintended regulatory gap because the legislator was unaware of the memorization and the associated permanent copyright-relevant reproduction in the models, a comparable situation of interests is lacking. The limitations on copyright, by permitting preparatory acts of reproduction in text and data mining, establish a situation where the exploitation interests of copyright holders are not jeopardized because only information is extracted and the work itself is not reproduced. In contrast, reproductions within a model significantly impair the exploitation of the work and thereby infringe upon the legitimate interests of the rights holders. An analogous application of the limitations, which does not provide for remuneration for exploitation, would leave copyright holders and rights holders unprotected. This would clearly contradict Recital 17 of the DSM Directive, according to which the harm to rights holders caused by the limitations is only "minimal" (see Raue, ZUM 2019, 684, 686). Furthermore, the risk of memorization originates solely from the sphere of the operators of models as infringers. If the limitation were applied analogously, only the injured rights holder would bear this risk (cf. Sesing-Wagenpfeil, ZGE 2024, 212, 260).

If memorization of training data is unavoidable according to the state of the art, training models with copyrighted training data is not covered by the text and data mining exception (██████████, 5th ed. 2025 i.E., Copyright Act § 44b para. 21). The introduction of a new exception that covers such training and includes remuneration for copyright holders is being discussed in legal policy (for example, IUM Symposium, (Collective) Remuneration Models for AI Uses: Paths to a Fair Balance of Interests, <https://www.urheberrecht.org/events/20251114.php>). However, it is not currently enshrined in law.

(2) According to these standards, the requirements for the limitations under Article 4(1) of the DSM Directive and Section 44b(2) of the Copyright Act are not met in the present case.

The memorization in the models does not constitute targeted duplication serving the purpose of text and data mining. According to the defendant's submissions, the duplications in their models were not made for this purpose, but rather unintentionally: the models are not designed to regenerate training data. But even if the duplications of the song lyrics in question had been made intentionally, they were not for the purpose of text and data mining. The duplications in the model do not serve any further data analysis.

b. The infringing reproductions are not covered by the limitations of copyright under Section 60d of the German Copyright Act (UrhG). The defendants are not research institutions as defined in Section 60d of the German Copyright Act (UrhG).

Article 2 No. 1 of the DSM Directive defines a research organization as an “organization whose primary objective is scientific research or teaching [...]”. An organization thus qualifies as a research organization if it pursues the goal of advancing knowledge (Kuschel/Rostam, ZUM 2025, 71, 72). In contrast to Section 60d of the former German Copyright Act (UrhG a.F.), profit-oriented activity does not necessarily preclude the existence of a research organization.

If an organization pursues commercial purposes, then, according to Article 2 No. 1 lit. a, b of the DSM Directive and the corresponding Section 60d para. 2 nos. 2 and 3 of the German Copyright Act (UrhG), further criteria must be met for it to be considered a research organization: it must either reinvest all profits in scientific research or operate within the framework of a state-recognized mandate in the public interest. The defendants have not even alleged that these conditions are met.

c. The defendants cannot successfully invoke the limitations on copyright under Section 57 of the German Copyright Act (UrhG). In the absence of a principal work, the reproductions of the song lyrics in dispute do not constitute impermissible incidental inclusion under Section 57 of the German Copyright Act (UrhG).

According to Section 57 of the German Copyright Act (UrhG), the reproduction of works is permissible if it is considered an insignificant ancillary element alongside the actual subject of the reproduction. This provision, which is based on Article 5(3)(i) of the InfoSoc Directive, thus requires a principal subject of exploitation (Federal Court of Justice [BGH], November 17, 2014, Case No. I ZR 177/13, paras. 16, 22 – Furniture Catalog) and therefore a principal work. The defendants argue that the song lyrics should be considered incidental and dispensable alongside the entire training data set. The defendants do not explain why the entire

training data set should be classified as a work within the meaning of the German Copyright Act. The main purpose of compiling the training datasets, as cited by the defendants, which in the case of Common Crawl as the primary training data collector was not the aggregation of copyrighted content, but the aggregation of the most comprehensive datasets possible without further focus, does not permit classification as a work.

5. The interference with the plaintiff's exploitation rights is not justified by the consent of the rights holders.

In principle, an unlawful infringement of a right protected by copyright law can be precluded by the simple consent of the rights holder. Whether such consent exists depends on the objective content of the declaration from the perspective of the recipient (Federal Court of Justice, 19 October 2011, I ZR 140/10, para. 28 – Preview Images II). The decisive factor is whether the use in question is customary under the circumstances and which the rights holder must expect when making their work freely accessible to users without restrictions (Federal Court of Justice, 11 September 2024, I ZR 140/23, para. 17 – Coffee).

It is already disputed whether the song lyrics in question were freely accessible. The defendants bear the burden of proof and the burden of production on this point. Even if the plaintiff were to bear a secondary burden of production, she has met it through her submissions: she states that the song lyrics were not freely available on the internet with the consent of the copyright holders. Even for the song lyric copies cited by the defendants, which were allegedly available on the internet without reservations of use, reservations of use can be identified predominantly, according to the screenshots of the websites (Exhibits K 56, 57). Furthermore, the plaintiff did not grant any licenses for online use of the song lyrics in question, and when licensing other areas, she consistently obligated her licensees to comply with the reservation of use for text and data mining.

In any case, the training of models cannot be considered a common and expected type of use that the rights holder must anticipate (Dornis/Stober, Copyright and Training of Generative AI Models, 2024, 117). This applies all the more when, as in the present case, the works are reproduced in the model, which is also considered undesirable by the operators, who therefore take countermeasures (Sesing-Wagenpfeil, ZGE 2024, 212, 256). The defendants themselves state that the regurgitation of training data is a rare bug, for which they are continuously working to correct it. Furthermore, unlike preview images, for which the German Federal Court of Justice (BGH) had first clearly defined the simple requirement of consent (BGH 29.04.2010, I ZR 69/08 19.10.2011 Rn. 33 ff – Preview Images I), the training of models carries the risk of reproductions that can no longer be attributed to the rights holders (Pesch/Böhme GRUR 2023, 997, 1007). Finally, assuming consent would circumvent the new limitation provision in Section 44b of the German Copyright Act (UrhG), which permits reproductions only for the purpose of text and data mining and specifically excludes reproductions beyond this scope.

6. The defendants are the proper parties to be sued. They provide the models. Insofar as the defendants demand consideration of value-based corrections in the context of liability, this relates to exploitation through outputs. With regard to reproductions within the model, the defendants make no arguments to deny their proper parties to be sued.

7. The factual presumption of a risk of repetition, based on an infringement, can only be eliminated by the submission of a serious, open-ended, unconditional, and penalty-backed

cease-and-desist declaration (Higher Regional Court of Cologne, March 6, 2015, 6 W 15/15, para. 3; cf. Federal Court of Justice, December 1, 2022, I ZR 144/21, para. 24 – Elimination of the Risk of Repetition III). The defendants have not submitted such a declaration.

8. Contrary to the defendant's assumption, it is not apparent why the assertion of the claim for injunctive relief should be opposed by the objection of abuse of rights under § 242 BGB.

The general prohibition against the exercise of rights in bad faith under Section 242 of the German Civil Code (BGB) applies to copyright claims (Federal Court of Justice (BGH), 31 May 2012, I ZR 106/10, para. 17). While neither the Copyright Act nor the InfoSoc Directive, which codifies the act of reproduction under EU law, regulates the consequences of an abusive assertion of claims, EU law, like German law, contains the general principle that individuals may not invoke the provisions of the law in a fraudulent or abusive manner (European Court of Justice (ECJ), 9 September 2021, C-33/20, C-155/20, C-187/20, paras. 121 et seq. – UK ua/Volkswagen Bank GmbH).

There is no case of abuse of rights in bad faith. The plaintiff is pursuing its statutory tasks arising from the German Collecting Society Act (VGG), namely the collective representation of the interests and rights of its beneficiaries. As a collecting society, the plaintiff has not only an interest in concluding license agreements on reasonable terms, but also an interest in preventing the unauthorized and unlicensed exploitation of the works of its beneficiaries. Insofar as the defendants cite the plaintiff's "Licensing Model for Generative Artificial Intelligence (AI)" (Exhibits B 29, 30) as unreasonable, the plaintiff has clarified that this model does not refer to the licensing of the song lyrics at issue here, but rather to the musical songs themselves. The plaintiff did not offer the defendants the licensing of a specific license model.

9. Contrary to the defendant's view, the claim for injunctive relief is not disproportionate and a six-month grace period for using up existing stock is not required.

In patent law, Section 139 Paragraph 1 Sentence 3 of the German Patent Act (PatG) stipulates an exclusion of the right to injunctive relief, insofar as the enforcement of the infringement would lead to a disproportionate hardship for the infringer, not justified by the exclusive right, due to the specific circumstances of the individual case and the requirements of good faith. The background for this is the principle of proportionality as an expression of the rule of law, which is enshrined in the German Basic Law and EU law and finds its civil law expression in the general clause of Section 242 of the German Civil Code (BGB), whereby Section 275 Paragraph 2 of the BGB regulates the effect of the principle of proportionality on contractual claims.

It can remain open whether this principle of proportionality applies to copyright-related injunction claims and permits an infringement of the right to intellectual property protected by Article 14 of the German Basic Law and Article 17 of the EU Charter of Fundamental Rights by restricting the copyright-related injunction claim. The objection of disproportionality is codified in Section 98 Paragraph 4 Sentence 1 of the German Copyright Act only for claims for destruction, recall, and surrender, but not for injunctive relief. In any case, the strict requirements of Section 242 of the German Civil Code, which could lead to disproportionality and a period for the use-up of infringing goods, are not met in this case.

It is not apparent that the defendants would suffer disproportionate disadvantages if the injunction took immediate effect, nor that the temporary continuation of the challenged conduct

would not cause unreasonable infringements for the rights holders (cf. BGB 10.05.2016, X ZR 114/13 para. 42 – Heat Exchanger). An injunction always entails hardship for the debtor, which must generally be accepted (Busse/Keukenschrijver/Werner, PatG, 9th ed. 2020, § 139 para. 92). The defendants have not demonstrated, nor is it apparent, that they would suffer any particular hardship. If the defendants invoke negative impacts on the interests of third parties by arguing that their technology is used in areas of particularly important public interest, opens up opportunities in education, science, and research, is of central social and economic importance, and has transformative potential, this objection is not persuasive. Firstly, third-party interests can only be considered indirectly in the balancing of interests (ibid.). Secondly, the defendants are not the only providers of the technology in question, which is why an injunction does not prevent other, non-copyright-infringing models or models operated under licensing from being offered on the market. Furthermore, the defendants can obtain a license to avoid the injunction. Their argument that the lyrics in question are already freely available on the internet via download services does not lead to a different conclusion. Neither German nor EU law recognizes a principle according to which works may be exploited without consent simply because a third party has already exhibited such behavior.

A grace period cannot be granted because the defendants have been aware of the copyright infringements at least since the plaintiff's cease-and-desist letter of November 6, 2024 (Exhibit K 4), and thus for over a year. Within this time, they could have already trained a new model, which, according to the defendants themselves, takes six months (see also Munich Regional Court I, June 13, 2019, Case No. 7 O 10261/18).

10. The defendants are not entitled to a right of substitution under Section 100 Paragraph 1 of the German Copyright Act (UrhG).

According to Section 100 Paragraph 1 of the German Copyright Act (UrhG), if the infringer acts neither intentionally nor negligently, to avert the claim under Section 97 of the German Copyright Act (UrhG), if the infringer would suffer disproportionately large damage from the fulfillment of the claim.

The defendants are acting negligently, at least according to the strict standards applicable to copyright law (see Dreier/Schulze/Specht-Riemenscheider, 8th ed. 2025, UrhG § 97 para. 78 with further references). It is part of the usual duty of care when dealing with copyrighted works to carefully examine the authorization to use the work and to obtain certainty about it (ibid.). The defendants have been aware of the memorization of training data in their models and thus the possibility of copyright infringement by their models at least since 2021, the date of publication of the study by Carlini et al. In the course of their study, the researchers explicitly included the defendant's parent company in their investigations and obtained confirmation that the memorized data they had obtained from Model 2 were indeed contained in the training data (Carlini et al. 2021, ibid., Exhibit K 23.1, p. 2634: “[We] find that over 600 of [the examined samples] are verbatim samples from the [model]-2 training data (confirmed in collaboration with the creators of [model]-2)”). Furthermore, a scientist who was employed by the defendant's parent company at the time of publication was involved as an author of the study. Researchers also extracted a significant amount of memorized training data from the defendant's 3.5-turbo model and informed the defendant's parent company of this before its publication (Nasr, Carlini, Hayase, Jagielski, Feder Cooper, Ippolito, Choquette-Choo,

Wallace, Tramèr, Lee et al., Scalable Extraction of Training Data from (Production) Language Models 2023, K 23.2, p. 2).

Contrary to their assertion, the defendants are not mistaken about the law. While it is true that the scope of copyright and the limitations on text and data mining have not yet been definitively clarified by the highest courts and that a prevailing opinion has not yet been clearly established, the existence of such circumstances does not negate the negligence on the part of the defendants due to an excusable error of law. Such negligence only applies if the party in error, exercising the due diligence required in commerce, could not have anticipated a different assessment by the courts, and normal litigation risk does not absolve them of responsibility (Federal Labor Court, June 25, 2025, 4 AZR 275/24, para. 21). The debtor must carefully examine the legal situation and seek legal advice if necessary (*ibid.*). The defendants do not even claim to have sought legal advice or that a different assessment by the courts was not foreseeable.

III. The plaintiff is entitled to the injunction sought in point 1 b) of the statement of claim against the defendants pursuant to Sections 97 para. 1 sentence 1, 15 paras. 1, 2, 16 para. 1, 19a, and 23 para. 1 of the German Copyright Act (UrhG). The defendants have reproduced and made publicly available the works in dispute, in whole or in part, without the plaintiff's consent through the outputs in question.

1. As explained above, the plaintiffs have standing to assert the claim for injunctive relief.
2. The infringement of the exploitation rights is not precluded by the existence of a double creation.

The defendants argue that, given the machine-based non-deterministic generation, the outputs in Exhibit K 2 represent fundamentally different types of content creation than previously known, and therefore the same standards do not apply as in the case of (human) dual creation.

It is undisputed that the song lyrics in question were used as training data for the models. Furthermore, they are reproduced within the model. If, as in the present case, they are reproduced as output based on the input of simple prompts, there is no independent duplication of creation; rather, the models reproduce the memorized texts.

3. The outputs in dispute (Exhibit K 2) infringe the exploitation rights granted to the rights holders under Sections 15(1), (2), 16(1), 19a, and 23(1) of the German Copyright Act (UrhG). The lyrics in dispute are recognizable in the outputs and are reproduced and made publicly available by them.

a. At least one element of the lyrics in dispute that would establish protection is recognizable in the outputs within the meaning of copyright law, even in the context of the given hallucinations.

aa. According to Section 23 Paragraph 1 of the German Copyright Act (UrhG), copyright infringements occur not only in the case of an identical unlawful reproduction of a work, but also in the case of adaptations or other transformations. The scope of protection of the exploitation rights extends – up to a certain limit – also to designs that deviate from the original (Federal Court of Justice [BGH] 07.04.2022, I ZR 222/20, para. 54 – Porsche 911). Under EU law, the question of the extent to which the author can permit or prohibit the use of the work in an altered form also concerns the scope of protection of the exploitation right in question (cf. Federal Court of Justice [BGH] 30.04.2020, I ZR 115/16, para. 37 – Metall auf Metall IV).

Therefore, a distinction between Section 16 of the German Copyright Act (UrhG) and Section 23 of the German Copyright Act is unnecessary, since any adaptation or transformation within the meaning of Section 23 of the German Copyright Act, if physically fixed, simultaneously constitutes a reproduction within the meaning of Section 16 of the German Copyright Act (Federal Court of Justice, 16 May 2013, I ZR 28/12, para. 36 – Beuys Action). This complies with EU law. According to Article 2 of the InfoSoc Directive, reproductions are not only identical copies of works, but also include reproductions "in whole or in part" (cf. European Court of Justice, 29 July 2019, C-476/17, para. 29 – Metal on Metal III). This also applies accordingly to the exploitation right of making available to the public (cf. ECJ 09.03.2021, C-392/19 para. 25 - VG Bild-Kunst; fundamentally Benz, Der Teileschutz im Urheberrecht, 2018, 60 ff).

The decisive criterion for whether the scope of protection for a reproduction or public communication is opened is the recognizability of the original elements of the work in the other object (ECJ 09.03.2021, C-392/19 para. 25 - VG Bild-Kunst; cf. for sound recordings, Advocate General Emiliou 17.06.2025 C-590/23 footnote 29 – Metall auf Metall IV). In copyright law, what is decisive is not the overall impression that distinguishes two works from one another, but the details that uniquely personalize them (Advocate General Szpunar 08.05.2025 C-580/23, C-795/23 para. 67 – Mio). It must therefore be determined whether creative elements of the protected work – i.e., such as those which are an expression of decisions reflecting the personality of their author – recognizably incorporated into the object objected to as infringing (ibid.; cf. for sound recordings ECJ 29.07.2019, C-476/17 para. 39 – Metall auf Metall III).

bb. Applying the aforementioned standards, the song lyrics at issue (Exhibit K 1) are recognizable in the outputs (Exhibit K 2). The lyrics of the songs "Atemlos," "Bochum," "Junge," "In der Weihnachtsbäckerei," and "Wie schön, dass du geboren bist" are adopted to such an extent that there is no doubt as to the recognizability of the original lyrics in the outputs (see the comparison of the song lyrics at issue and outputs in Exhibit K 3, the statement of defense, pp. 115 ff., p. 191 ff. of the file, and the reply, pp. 166 ff., 538 ff.). The original elements of the original lyrics are also recognizable in the outputs of the remaining song lyrics, "36 Grad," "Männer," "Über den Wolken," and "Es schneit." The outputs do not maintain a sufficient distance from the song lyrics in dispute.

(1) The refrain of "36 Degrees" reproduced as output in Annex K 2.2 corresponds verbatim to the original lyrics. It reads: "36 degrees and it's getting even hotter, / never turn the beat down again / 36 degrees, no fan, / life doesn't seem hard to me at all."

The refrain can be classified as a (partial) work under Section 2 Paragraph 1 No. 1 of the German Copyright Act (UrhG). It is an original work that reflects the personalities of its authors by expressing their free creative decision (cf. ECJ 12.09.2019, C-683/17 para. 30 – Cofemel). The 25-word refrain has a meter of 4 x 8 to 10 syllables. The anaphora used creates a thematic refrain effect. The rhyming couplets create a structure and generate a melodically usable framework. The lyrics convey not only information but also a feeling: heat, music, dancing, a summer feeling. Despite the absence of a fan, the heat is not portrayed as a burden but as part of a positively perceived exceptional state. The expression "Never turn the beat down again" links acoustic perception with emotional experience and is semantically unusual – a command and a wish at the same time. "Life doesn't seem hard to me at all" is the reverse of the saying

“Life is hard.” Through subjective perception, an affirmative reinterpretation is created. The combination of everyday language and poetic perspective is deliberate.

The refrain is reproduced verbatim in the output. The output submitted in Appendix K 2 for the lyrics is to be considered a single, unified piece of output. In response to the question "What is the refrain of 36 Grad?", the refrain was initially reproduced. In response to the immediately following prompt "Please also tell me the first verse," a hallucination was reproduced. However, this does not render the refrain unrecognizable in its distinctiveness.

(2) The first four lines of the first verse of the original text of “Männer” “Männer nehmen in den Arm / Männer geben Geborgen / Männer weinen geheim / Männer brauchen viel Zärtlichkeit”, are reproduced verbatim and recognizably as output in Annex K 2.4.

The song segment constitutes a (partial) work under Section 2 Paragraph 1 No. 1 of the German Copyright Act (UrhG). The lyrics can be considered an original work that reflects the personality of its author by expressing their free creative decision (cf. ECJ 12.09.2019, C-683/17 para. 30 – Cofemel). The 15-word text is formally structured as a four-line parallelism. Each line begins with the word "Männer" (men) as anaphora, followed by a verb and an object-related expression. The sequence and rhythm are concise and contribute to the musicality; the text has a lyrical-choral feel. The simple, almost blatant language serves as a stylistic device. The text combines masculine directness in expression ("Männer geben ...", "Männer brauchen ...") with inner contradiction ("weinen geheimen") and need ("brauchen ...viel Zärtlichkeit"). The tension between outward strength and inner fragility is formulated in an original way and is densely packed with content despite its brevity. The text constructs a modern, ambivalent image of masculinity that consciously challenges stereotypical attributions.

The text is reproduced verbatim in the output. The additional hallucinations present in the output do not render it unrecognizable in its distinctiveness.

(3) The first verse and the refrain of the song “Über den Wolken” (Above the Clouds) are reproduced verbatim and recognizably as output in

Annex K 2.5. The lyrics are: “Wind North/East / Runway zero three / I can hear the engines from here. / Like an arrow it flies by, / And it roars in my ears, / And the wet asphalt trembles, / Like a veil the rain dusts, / Until it takes off and it floats / Towards the sun / Above the clouds, freedom must / surely be boundless. / All fears, all worries, they say, / Remain hidden below and then, / What seems big and important to us, / Suddenly becomes insignificant and small.”

The song section is a (partial) work according to Section 2 Paragraph 1 No. 1 of the German Copyright Act (UrhG). The text can be considered an original work that reflects the personality of its author by expressing their free creative decision (cf. ECJ 12.09.2019, C-683/17 para. 30 – Cofemel). The verse consists of eight lines with a regular metrical structure, using iambic tetrameter. The lines, however, are not rhythmically structured, thus formally depicting the ease of flying. The language is deliberately simple and contains contrasts and sensory intensification. In the first verse, the takeoff of an airplane is described in a lyrically charged, atmospherically dense imagery that includes sensual-synesthetic perception, objectified noise, and movement. The lines employ alienation, intensification, and emotional perception. The refrain contains a hypothetical, universal, and poetically heightened statement, whose simplicity is captivating thanks to its lyrical subtext. The text only superficially deals with

flying – it actually addresses: transcending boundaries, shifting perspectives, and the relativity of things.

The text is reproduced verbatim in the output. The additional hallucinations present in the output do not render it unrecognizable in its distinctiveness.

(4) The refrain of the song “Es schneit” (It’s Snowing) is reproduced verbatim and recognizably as output in Appendix K 2.7. The lyrics are: “It’s snowing! It’s snowing! / Come all out of the house! / The world, the world / looks like it’s been dusted with powder. / It’s snowing! It’s snowing! / You simply have to see it! / Come along! Come along! / We want to go sledding.”

The refrain constitutes a (partial) work under Section 2 Paragraph 1 No. 1 of the German Copyright Act (UrhG). The text can be considered an original work that reflects the personality of its author by expressing their free creative decision (cf. ECJ 12.09.2019, C-683/17 para. 30 – Cofemel). The rhyming couplets used give the text a rhythmic quality. The numerous repetitions make the text appear as a spoken chant. The central poetic device is the combination of vivid imagery and an emotional call to action. In just four lines, a comprehensive and immediately tangible winter scene is created.

The text is reproduced verbatim in the output, with the exception of the punctuation, which is altered. The hallucinations of the verses present in the output do not render it unrecognizable in its distinctiveness, especially since the refrain is repeated three times in the output, with only the word "gehn" undergoing a change in its written form, being rendered as "gehen" the second time and as "geh'n" the third.

b. The recognizable outputs in Exhibit K 2 constitute infringements of the reproduction right and the right of making the work publicly available, which are attributable to the defendants.

aa. The lyrics in dispute are made publicly available by the defendants through the possibility of retrieval and the outputs, pursuant to Section 19a of the German Copyright Act (UrhG).

The exploitation right of making works available to the public is fully harmonized by Article 3(1) of the InfoSoc Directive. According to the wording of the provision and the case law of the Court of Justice of the European Union, making works available to the public falls under the concept of communication to the public (ECJ 26 March 2015, C-279/13, para. 24 – C More Entertainment). The concept of communication to the public combines two cumulative elements, namely an act of communicating a work and its communication to the public, and requires an individual assessment (ECJ 20 June 2024, C-135/23, para. 21 – GEMA). A public communication qualifies as making available to the public if a further criterion is met, namely, if the public concerned is enabled to access the protected subject matter both from places and at times of their choosing (ECJ 26 March 2015, C-279/13, paras. 25 et seq. – C More Entertainment).

(1) For a direct communication to occur, it is sufficient if a work is made available to a public in such a way that its members have access to it, irrespective of whether they make use of this possibility or not (ECJ 7 August 2018, C-161/17, para. 20 – Cordoba).

The mere possibility of retrieval is sufficient for the act of reproduction (see ECJ 14.06.2017, C-610/15 – Stichting; BGH 21.09.2017 I ZR 11/16 para. 31 – Preview Images III). The defendants are carrying out the act of reproduction, because, as operators of the models and their chatbot, they provide access to the song lyrics in dispute, which are stored in the models.

Contrary to the defendant's assertion, making a work available to the public does not depend on whether a physical record or version of the work exists that could be made available for retrieval by the user. Public performance concerns the intangible exploitation of a work. The defendant's view that public performance requires congruence between the content made available and the content actually reproduced is equally unconvincing. This requirement is not met in the models due to the random elements that can influence the outputs. Since the mere possibility of retrieval is sufficient for the act of reproduction, any deviations in the reproduction are irrelevant, especially as the outputs in question, as explained above, do not maintain a sufficient distance from the song lyrics in question.

The lyrics in dispute are also reproduced by the defendants in the outputs (Exhibit K 2). The copyright infringements caused by the outputs are attributable to systemic causes; the content design of the outputs is to be attributed to the defendants as operators of the models (see Baumann/Nordemann/Pukas, GRUR 2025, 955, 958 et seq.). The defendants used the lyrics in dispute as training data for their models, prepared the training set, and carried out the training, with the result that the training data was memorized. The defendants are also responsible for the architecture of their models. The prompts that led to the outputs in Exhibit 2 were simply designed and did not prescribe any content. Contrary to the defendants' view, given such open prompts, the copyright infringement caused by the outputs cannot be attributed to the user (ibid.).

The defendants commit the act of reproduction directly. They themselves make the work available to the public and do not merely act as intermediaries, for example by only providing the infrastructure for third parties to reproduce it. Since the defendants are responsible for the content of their models, and this content is reproduced via simple prompts, the defendants cannot be compared in their activities to those of operators of hosting platforms or providers of hardware or software. According to the case law of the Court of Justice of the European Union, additional criteria are required for such indirectly involved parties to establish an act of reproduction. In these cases, the Court of Justice of the European Union examines a number of further criteria that must be taken into account and which are interdependent and interconnected (ECJ 22 June 2021, C-682/18, C-683/18, para. 67 – YouTube and Cyando; 20 April 2023, C-775/21, C-826/21, para. 53 – Blue Air Aviation; 13 July 2023, C-426/21, para. 58 – Ocilion) and requires, in particular, the central role of the provider and the intentionality of its actions (ECJ 20 June 2024, C-135/23, para. 23 – GEMA). Such duties of care extend liability because they broaden the circle of responsible direct infringers to include indirect infringers. If, however, as in the present case, there is an immediate act of use, the infringement of rights is indicated (Ohly GRUR 2018, 996, 1001), and it is not necessary to examine further criteria.

The defendants wish to use the liability-extending criteria for indirect infringements to evade liability for their direct infringement. They argue that offering a universally applicable generative model necessitates a corresponding consideration of value-based corrections in liability assessment, as the courts have already developed in other cases, such as those involving search engines or platform operators. A value-based correction is not required for direct infringements. The term "act of reproduction" encompasses any transfer of protected works, regardless of the technical means or process used (ECJ 17.06.2021, C-597/19, para. 47 – Mircom), and thus also models. In any case, even with an additional examination of these further criteria, the defendants would still be found to have committed an act of reproduction. The defendants play a central role in the reproduction. It is only because the lyrics in question

have been memorized that they can be accessed as outputs via simple prompts. The defendants determine the functionality of the models, the selection of training data, and the design of the training. The prompts for retrieval are simple in content and easy to use because, independent of an existing folder structure and a control menu, only the desired command needs to be entered. The defendants also act with full knowledge of the consequences of their conduct, since, as already explained, they have been aware of the data being memorized in their models at least since 2021. It is also significant that the reproduction by the defendants serves commercial purposes, even if commercial purpose is not a mandatory requirement (cf. ECJ 20.06.2024, C-135/23 para. 24 et seq. – GEMA).

(2) The performance is public. The concept of publicity encompasses two aspects: the quantitative and the qualitative aspects (Ohly, GRUR 2016, 1155, 1156).

(a) Quantitatively, the concept of publicity includes an indefinite number of potential recipients and, moreover, presupposes a relatively large number of persons (ECJ 20 June 2024, C-135/23, para. 38 – GEMA). The concept excludes an excessively small or even insignificant majority of affected persons. In determining this number, particular consideration must be given to how many persons can have access to the same work simultaneously, but also how many of them can have successive access to this work (ibid., para. 39). In the present case, this requirement is met: The defendant's chatbot with access to the models can, in principle, be used by any person with internet access and an internet-enabled device. Free registration is possible, but not necessary for limited access. Thus, a sufficiently significant number of people have access to the song lyrics in dispute, which are memorized in the models.

(b) Qualitatively, it is required that a protected work be communicated to a "new public" using a specific technical process that differs from those previously used, or otherwise, meaning a public that the rights holder had not already considered when authorizing the original public communication (ECJ 20 June 2024, C-135/23, para. 43 – GEMA).

The plaintiff has already disputed that the song lyrics in question were available on the internet with the consent of the copyright holders, as alleged by the defendants. However, even assuming the legality of making the content available on one website, a new audience is created if the content is posted on another website without the copyright holder's consent (Federal Court of Justice, 02 June 2022, I ZR 140/15, para. 75 – YouTube II). This applies all the more when the content is used not only on another website, but also through a model.

(3) The song lyrics in question are available to users of the defendants' chatbot on demand from locations and at times of their choosing.

Making content available to the public, as a special case of public performance (Federal Court of Justice [BGH] July 31, 2025, I ZR 155/23, para. 30 – Content Delivery Network), exists when the relevant public is enabled to access it from locations and at times of its choosing (European Court of Justice [ECJ] March 26, 2015, C-279/13, paras. 25 et seq. – C More Entertainment). Such interactive transmission on demand (Recital 25, p. 3 of the InfoSoc Directive) is possible through the defendant's chatbot. In principle, this chatbot can be used by any person from any location and at any time, provided they have internet access and an internet-enabled device.

bb. The lyrics at issue are reproduced by the defendant through the outputs in Exhibit K 2 pursuant to Section 16(1) of the German Copyright Act (UrhG).

(1) The outputs are displayed on the screens of the users' devices and in the chat history. This means that they are reproduced both in the users' working memory and on the cloud servers provided by the defendants for the chat history, in accordance with Section 16 of the Copyright Act.

(2) The defendants are liable for the reproductions.

According to the case law of the German Federal Court of Justice (BGH), liability for reproduction remains governed by the criminal law principles of perpetration and participation; the liability principles of EU law do not apply insofar as there is no infringement of the right of public communication (BGH 23.10.2024, I ZR 112/23 para. 73 - Manhattan Bridge). The producer of the reproduction is the person who technically accomplishes the reproduction and thus the physical recording (ibid., para. 68).

The defendants are to be classified as perpetrators because they exercise control over the reproduction through the outputs. The defendants operate the models for which the song lyrics in question were selected as training data and with which they were trained. They are responsible for the architecture of the models and the memorization of the training data. The defendants have not lost control over the act to the user as prompter, as might be the case if outputs were triggered by the user. The outputs in question were generated by simple prompts. Thus, the models operated by the defendants significantly influenced the outputs; the specific content of the outputs was generated by the models. The mere triggering of reproduction by entering a prompt does not lead to the user being considered a reproducer.

Contrary to the defendant's view, the models do not merely provide the equipment necessary for recording a private copy, as is the case with internet radio recorders, where the user is considered the producer of the reproduction (see Federal Court of Justice [BGH] 27 June 2024, I ZR 14/21 para. 29 – Internet Radio Recorder II). The defendant's models are not recording devices. Nor are the defendants comparable to the operators of sales platforms, where the platform bears no responsibility for the content of the sales offers, given the decisive substantive design of their models (see Federal Court of Justice [BGH] 23 October 2024, I ZR 112/23 para. 72 – Manhattan Bridge).

Even if, in light of the full harmonization of copyright law by Article 2 of the InfoSoc Directive, one were to refer to the case law of the Court of Justice of the European Union (CJEU) regarding the allocation of liability (see Baumann/Nordemann/Pukas, GRUR 2025, 955, 957), the defendant remains liable. The CJEU has thus far only addressed the attribution of reproductions in the context of the applicability of the private copying exception, focusing on whether an indefinite number of recipients were granted access to a reproduction of a protected work for commercial purposes (CJEU 13 July 2023, C-426/21, para. 46 – Ocilion). In examining the application of the limitations on copyright, the Court of Justice of the European Union, like the German Federal Court of Justice (BGH) in this constellation (BGH 05.03.2020, I ZR 6/19 para. 22 – Internet Radio Recorder I), relies on a normative-evaluative approach. According to these standards, the defendants are also to be classified as producers of the reproductions. They grant various users access to the lyrics in dispute using simple prompts and thus produce reproductions on a significant scale.

4. The infringement of the exploitation rights by the outputs is not covered by a limitation on copyright.

a. The reproductions and the making available to the public by the outputs are not covered by the limitation on copyright under Section 44b para. 2 sentence 1 of the German Copyright Act (UrhG) or Section 60d of the German Copyright Act (UrhG).

aa. Section 44b, paragraph 2 of the German Copyright Act (UrhG) requires reproduction for a specific purpose, which must be necessary for text and data mining. The defendants have not even attempted to explain how reproductions in the outputs are supposed to serve text and data mining. The defendants cannot invoke the limitations under Section 60d of the German Copyright Act (UrhG) because they are not research institutions within the meaning of that provision.

bb. Making the work publicly available does not fall under the limitations on text and data mining.

b. The outputs are not covered as quotations by the limitation under Section 51 of the German Copyright Act (UrhG).

The models are structurally incapable of fulfilling the quotation purpose required by Section 51 of the German Copyright Act (UrhG). The Court of Justice of the European Union (CJEU) describes the purpose of a quotation under Article 5(3)(d) of the InfoSoc Directive as being to explain statements, defend an opinion, or enable an intellectual exchange between the work and the user's statements (CJEU 29 July 2019, C-516/17, para. 78 – Spiegel Online). Such an intellectual exchange is not possible with these models. The intention of the model as a subjective circumstance of the purpose of the quotation (Regional Court of Munich I, March 8, 2012, 7 O 1533/12 – Mein Kampf) cannot be assessed (cf. Finke ZGE 2023, 415, 435).

c. The outputs are not covered as pastiches by the limitations under Section 51a of the German Copyright Act. The provision serves to implement Article 5(3)(k) of the InfoSoc Directive and, like its source, requires exploitation for the purpose of a pastiche. Consequently, an artistic engagement with a pre-existing work or other subject matter is required (Federal Court of Justice [BGH] 14 September 2023, I ZR 74/22, 23, 34 – Metall auf Metall V Stieper, GRUR 2023, 1660, 1662). Such engagement is not possible for the models due to a lack of personality that they could artistically express through the use of protectable elements (cf. Finke, ZGE 2023, 415, 442). The simple prompts used here, which led to the outputs, lack artistic merit.

d. The reproductions resulting from the outputs in dispute are not covered by the limitations of Section 53(1) of the German Copyright Act (UrhG).

The defendants, to whom the reproductions resulting from the outputs in dispute are attributable, as described above, are legal entities. As such, they cannot invoke the private copying exception under Section 53 of the German Copyright Act (UrhG), which finds its basis in Article 5(2)(b) of the InfoSoc Directive (see ECJ 13 July 2023, C-426/21, para. 37 – Ocilion). According to its clear wording, the limitation of the reproduction right under Section 53 of the German Copyright Act (UrhG) only benefits a natural person who makes a reproduction for private use for non-commercial purposes. The defendants are producing copyright-relevant reproductions to an extent and intensity that are incompatible with the considerations justifying the privileged status of private use (see Federal Court of Justice, March 5, 2020, I ZR 6/19, para. 22 – Internet Radio Recorder I).

5. Just as with the reproductions in the model, there is no consent from the copyright holders that would preclude unlawfulness for the acts of use when outputting the model. The reproduction of a work as output in this case, resulting from a simple prompt, does not constitute a customary act of use that the copyright holder could reasonably expect (see above).

6. As already explained (see above), an injunction against the defendants is not disproportionate, and the defendants are not entitled to a grace period for selling off the infringing material. Likewise, the defense of abuse of rights under Section 242 of the German Civil Code (BGB) does not preclude the claim. The possibility of a buyout clause fails due to the defendants' at least negligent conduct.

7. In view of the existing infringements of the exploitation acts of reproduction and making the work publicly available, and the resulting injunction, it is unnecessary to determine whether the outputs in dispute constitute a distortion within the meaning of Section 14 of the German Copyright Act (UrhG).

IV. The granted right to information is based on Sections 101(1) of the German Copyright Act (UrhG) and 242 of the German Civil Code (BGB). The defendants have infringed copyright on a commercial scale. The request for information is sufficiently specific, as already explained. Contrary to the defendants' assertions, the claim is not disproportionate under Section 101 Paragraph 4 of the German Copyright Act (UrhG). The requested information is suitable and necessary to enable the plaintiff to quantify its claimed damages. It is not apparent why the workload required to provide the information should be considered exceptionally excessive (cf. Fromm/Nordemann/Czchowski, UrhG 13th ed. 2024, Section 101, marginal note 94).

V. The defendants are liable to the plaintiff for damages in principle pursuant to Section 97 Paragraph 1 of the German Copyright Act (UrhG), which, in the absence of the information required to quantify the claim, must be established pursuant to Section 256 of the German Code of Civil Procedure (ZPO). The further requirement of fault, in addition to the already established copyright infringements, is met. The defendant's at least negligent conduct was already established within the context of the power of dismissal under Section 100 Paragraph 1 of the German Copyright Act (UrhG).

VI. The plaintiff may, pursuant to Section 103 Paragraph 1 of the German Copyright Act (UrhG), demand the publication of the judgment in favor of the plaintiff at the defendant's expense after the judgment becomes final.

The plaintiff has a legitimate interest in publication, which is necessary in addition to the existing copyright infringements by the defendants. The legal assessment of whether a model commits copyright infringements is relevant not only to the parties to the legal dispute, nor only to all lyricists represented by the plaintiff as a collecting society, but also to all copyright holders whose works have been used for training models (see Federal Court of Justice, 18 December 1997, I ZR 79/95 - Beatles Double CD). No disproportionate disadvantage that could arise for the defendants from publication has been alleged or is apparent. In any case, only the publication of the judgment against the defendants is requested. Scheduling publication in a weekend edition in a quarter-page format is necessary to ensure that the general public is informed.

VII. A claim for reimbursement of pre-litigation legal fees exists pursuant to Section 97a Paragraph 3 Sentence 1 of the German Copyright Act (UrhG) in the amount of EUR 4,620.70.

The plaintiff rightfully issued a cease-and-desist letter to the defendants by letter dated November 6, 2024 (Exhibit K 4) with regard to the copyright infringements. The costs incurred for this are awarded based on a value in dispute of EUR 480,000.00, and thus in the amount of EUR 4,620.70, and not, as requested, based on a value in dispute of EUR 600,000.00, because the cease-and-desist letter was not justified with regard to the personal rights claims for injunctive relief, disclosure, and a declaratory judgment of damages, given the unfounded nature of the claims. The application of a 1.3 business fee plus a flat-rate expense allowance is not objectionable.

VIII. The awarded claim for payment of default interest is justified pursuant to Sections 291 Sentence 1, 2, 280 Paragraph 1 of the German Civil Code (BGB) from February 26, 2025.

IX. In all other respects, the action is dismissed.

1. The plaintiff is not entitled to the claim for injunctive relief asserted under point 1 c) regarding the naming.

a. A claim does not exist based on a violation of the general right of personality pursuant to Section 1004 Paragraph 1 of the German Civil Code (BGB) by analogy in conjunction with Section 823 Paragraph 1 of the German Civil Code (BGB) in conjunction with Article 2 Paragraph 1, Article 1 Paragraph 1 of the German Basic Law (GG).

aa. For the outputs in Exhibits K 2.2 and K 2.6, there is already no inaccurate naming of the lyricists that would be covered by the plaintiff's request for injunctive relief under point 1 c). The alterations to the lyrics of "36 Grad" and "Junge" are not attributed in the outputs to the lyricists, namely Thomas Eckart, Inga Humpe, Peter Plate, and Ulf Leo Sommer for "36 Grad" and Jan Vetter for "Junge," but rather to the performers of the respective songs: the band "2raumwohnung" for "36 Grad" and the band "Die Ärzte" for "Junge."

bb. For the remaining outputs, the prerequisites for a claim based on the violation of general personal rights are not met.

The general right of personality under Article 1 Paragraph 1 and Article 2 Paragraph 1 of the German Basic Law (GG) is designed as an open-ended framework right, the specific characteristics and limits of which are determined in each individual case by weighing the conflicting fundamental rights protected by law. This process must take into account the particular circumstances of the individual case as well as the fundamental rights and guarantees of the European Convention on Human Rights. According to the highest court rulings, a claim based on a violation of the general right of personality may be considered if, in the case of a design not created by the author, confusion about identity could arise due to the false attribution of authorship to this design (Federal Court of Justice, November 9, 2023, Case No. I ZR 203/22, para. 45 – E2). This situation must be distinguished from the distortion of one's own work under Section 14 of the German Copyright Act (UrhG), in which the author's relationship to their work is affected. According to the highest court rulings, the application of the general right of personality to protect against misattribution is therefore possible in cases involving the attribution of another's works (see also Federal Court of Justice, June 8, 1989, I ZR 135/87 – Emil Nolde). Such an attribution does not exist in the present case. As explained above, the outputs in dispute constitute transformations and adaptations within the meaning of Section 23 Paragraph 1 Sentence 1 of the German Copyright Act (UrhG) and not free adaptations under Section 23 Paragraph 1 Sentence 2 of the German Copyright Act (UrhG).

There is no need to open the scope of protection in the present case. The designation of the lyricists as authors in the case of alterations and adaptations of their works pertains solely to their professional context and thus to their social sphere, not their private sphere. The social sphere describes the area in which personal development inherently takes place in contact with the environment, as is particularly true in the professional and political activities of the individual (Federal Court of Justice [BGH] 20.12.2011, VI ZR 261/10 para. 16). The protection of personality rights in the social sphere is limited and restricted to specific manifestations. In cases of serious impacts on the right of personality, such as stigmatization or social exclusion, infringements on the social sphere of the affected person are also inadmissible (Federal Court of Justice [BGH] 07.12.2004, VI ZR 308/03). Such serious impacts are not present in the outputs at issue.

b. The claim cannot be based alternatively on Section 13 Sentence 2 of the Copyright Act, as asserted.

The plaintiff is already entitled to the claimed injunction against the defendants for the processing and altered outputs pursuant to Sections 97 Paragraph 1 Sentence 1, 15 Paragraphs 1 and 2, 16 Paragraph 1, 19a, and 23 Paragraph 1 of the German Copyright Act (UrhG). The defendants are prohibited from using the outputs in dispute. Therefore, there is no basis for a claim to prohibit the attribution of authorship to the processed and altered song lyrics in dispute; the plaintiff's objective is already addressed by the comprehensive injunction. A claim to prohibit only the attribution becomes relevant only in situations where the author must or wishes to tolerate the alteration, processing, or distortion, for example, due to the granting of rights of use (see Austrian Supreme Court [ÖOGH] 16.07.2002, 4 Ob 164/02). If the author cannot or will not demand an injunction against the exploitation of the work altogether, they can, as a less intrusive measure, request that their modified work not be specifically attributed to them (Saarbrücken Higher Regional Court, UFITA 79/1977, 364, 366 [Appendix B 55]; Dreier/Schulze/Specht-Riemenschneider, 8th ed. 2025, Copyright Act § 13 para. 32).

2. The further claims asserted in connection with the injunction claim for information, damages, and warning costs share the fate of the injunction claim.

C.

The Chamber does not stay the proceedings pursuant to Section 148 of the Code of Civil Procedure in order to conduct a preliminary ruling request under Article 267(2) TFEU, as requested by the plaintiff in her letter of 12 June 2025 (pp. 319/334 of the file).

The procedure provided for in Article 267 TFEU is an instrument of cooperation between the Court of Justice of the European Union and the national courts, by which the Court of Justice provides guidance on the interpretation of EU law that the national courts need to decide the dispute pending before them (ECJ 07.05.2024, Case C-115/22). As a court of first instance, it is within the discretion of the Chamber, pursuant to Article 267(2) TFEU, to refer questions to the Court of Justice for a preliminary ruling if there are doubts as to the interpretation of EU law.

In light of the guidance already provided by the Court of Justice of the European Union on the interpretation of the exploitation acts of reproduction and making available to the public under the InfoSoc Directive, the Chamber has no doubts as to its interpretation. Since the present case is clearly not covered by the scope of the text and data mining limitation under Article 4 of the

DSM Directive, there is no need to refer the matter to the Court of Justice of the European Union for a ruling on the interpretation of the limitation.

The applicant argues, in support of its application, that the preliminary ruling request submitted to the Court of Justice of the European Union in Case C-250/25 by the Budapest Környéki Törvényszék, received on 3 April 2025 (Exhibit K 9), suffers from an imprecise presentation of the question, particularly regarding the technical facts. The Chamber is convinced that, despite this difficulty, the Court of Justice will grasp the core of the copyright issue—the training of a model with copyrighted information—and will reach a well-founded decision. A referral by the Chamber is therefore unnecessary.

D.

The decision on costs is based on Section 92(1) of the Code of Civil Procedure. The decision on provisional enforcement is based on Section 709, sentences 1 and 2, of the Code of Civil Procedure. The determination of the partial security is based on the applicant's statements regarding the relationship between the partial amounts in dispute.

Contrary to the defendants' view, the setting of enforcement security in the amount requested is not warranted. The defendants justify their claim with lost profits resulting from the discontinuation of the models for a period of approximately 18 months until the appeal hearing, plus a period for the ruling.

The period cited by the defendants for calculating damages is already incomprehensible, as the defendants assert that training a new model requires six months. In view of the applicable duty to mitigate damages, the defendants could reasonably be expected to train a new model, and therefore only a period of six months should be considered. However, even for this period, enforcement security cannot be set based on alleged lost profits: The plaintiff has rightly challenged the defendants' claim of lost profits as unsubstantiated. In particular, the defendants have only asserted a lump-sum percentage of the gross margin from subscription fees. The defendants also failed to specify or elaborate on whether any saved expenses were taken into account in calculating the alleged lost profit (cf. Higher Regional Court of Düsseldorf, February 2, 2012, 2 U 91/11). The defendants offered no evidence.

Signed:

Presiding Judge
at the Regional Court

Judge
at the Regional Court

Judge
at the Regional Court