Downloaded from Blogs on 29.11.2026 05:26

By PwC Deutschland | 14 November 2025

LG München: Urteil GEMA gegen Open AI

Die auf das Urheberrecht spezialisierte 42. Zivilkammer des Landgerichts München I hat mit Urteil vom 11. November 2025 den von der GEMA gegen zwei Unternehmen der Unternehmensgruppe Open AI geltend gemachten Ansprüchen auf Unterlassungs-, Auskunfts- und Schadensersatz im Wesentlichen stattgegeben (Az. 42 O 14139/24).

Downloaded from Blogs on 29.11.2026 05:26

Soweit die Klägerin darüber hinaus Ansprüche auf Grund einer**Verletzung des allgemeinen Persönlichkeitsrechts**wegen fehlerhafter Zuschreibung veränderter Liedtexte geltend gemacht hat, hat die Kammer die Klage abgewiesen.

Das Urteil betrifft die Liedtexte neun bekannter deutscher Urheberinnen und Urheber(darunter „Atemlos“ von Kristina Bach oder „Wie schön, dass du geboren bist“ von Rolf Zuckowski).

Die**Klägerin**ist eine Verwertungsgesellschaft und hat die Ansprüche als solche geltend gemacht. Zur Begründung hatte sie vorgetragen, die Liedtexte seien in den Sprachmodellen der Beklagten memorisiert und würden bei Nutzung des Chatbots auf einfache Anfragen der Nutzer als Antworten (Outputs) in weiten Teilen originalgetreu ausgegeben.

Die**Beklagten**sind Betreiber von Sprachmodellen und darauf basierender Chatbots. Sie hatten gegen die erhobenen Ansprüche eingewandt, ihre Sprachmodelle speicherten oder kopierten keine spezifischen Trainingsdaten, sondern reflektierten in ihren Parametern, was sie basierend auf dem gesamten Trainingsdatensatz erlernt hätten. Da die Outputs nur als Folge von Eingaben von Nutzern (Prompts) generiert werden würden, seien nicht die Beklagten, sondern der jeweilige Nutzer als Hersteller des Outputs für diese verantwortlich. Ohnehin seien eventuelle Rechtseingriffe von den Schranken des Urheberrechts, insbesondere der Schranke für das sogenannten Text- und Data-Mining gedeckt.

Nach der Entscheidung der erkennenden Kammerstehen der Klägerin die geltend gemachten Ansprüche sowohl aufgrund der gegebenen Vervielfältigung der Texte in den Sprachmodellen als auch durch ihre Wiedergabe in den Outputs zu.

Sowohl durch die Memorisierung in den Sprachmodellen als auch durch die Wiedergabe der Liedtexte in den Outputs des Chatbot lägen Eingriffe in die urheberrechtlichen Verwertungsrechte vor. Diese seien nicht durch Schrankenbestimmungen, insbesondere die Schranke für das Text und Data Mining gedeckt.

Im Einzelnen:

Nach Überzeugung der Kammerseien die streitgegenständlichen Liedtexte reproduzierbar**in den Sprachmodellen 4 und 4o der Beklagten enthalten**. Aus der informationstechnischen Forschung sei bekannt, dass Trainingsdaten in Sprachmodellen enthalten sein können und sich als Outputs extrahieren lassen. Dies werde als Memorisierung bezeichnet. Eine solche liege vor, wenn die Sprachmodelle beim Training dem Trainingsdatensatz nicht nur Informationen entnähmen, sondern sich in den nach dem Training spezifizierten Parametern eine vollständige Übernahme der Trainingsdaten finde. Eine solche Memorisierung sei durch einen Abgleich der Liedtexte, die in den Trainingsdaten enthalten waren, mit den Wiedergaben in den Outputs festgestellt. Angesichts der Komplexität und Länge der Liedtexte sei der Zufall als Ursache für die Wiedergabe der Liedtexte ausgeschlossen.

Durch die Memorisierung sei eine Verkörperung als Voraussetzung der urheberrechtlichen**Vervielfältigung**der streitgegenständlichen Liedtexte durch Daten in den spezifizierten

Downloaded from Blogs on 29.11.2026 05:26

Parametern des Modells gegeben. Die streitgegenständlichen Liedtexte seien reproduzierbar in den Modellen festgelegt. Gemäß Art. 2 InfoSoc-RL liege eine Vervielfältigung „auf jede Art und Weise und in jeder Form“ vor. Die Festlegung in bloßen Wahrscheinlichkeitswerten sei hierbei unerheblich. Neue Technologien wie Sprachmodelle wären vom Vervielfältigungsrecht nach Art. 2 InfoSoc-RL und § 16 UrhG erfasst. Nach der Rechtsprechung des Unionsgerichtshofes sei für die Vervielfältigung ausreichend eine mittelbare Wahrnehmbarkeit, die gegeben sei, wenn das Werk unter Einsatz technischer Hilfsmittel wahrgenommen werden könne.

Diese Vervielfältigung in den Modellen sei weder durchdie**Schrankenbestimmungen**des Text und Data Mining des § 44b UrhG noch durch § 57 UrhG als unwesentliches Beiwerk gedeckt.

Zwar unterfielen Sprachmodelle grundsätzlich dem Anwendungsbereich der**Text und Data Mining Schranken.** Die Vorschriften deckten erforderliche Vervielfältigungen beim Zusammenstellen des Datenkorpus für das Training, wie etwa die Vervielfältigung eines Werks durch seine Überführung in ein anderes (digitales) Format oder Speicherungen im Arbeitsspeicher. Hintergrund hierfür sei der Gedanke, dass diese Vervielfältigungen lediglich zu **nachfolgenden Analysezwecken erstellt würden und damit**die Verwertungsinteressen des Urhebers am Werk nicht beeinträchtigten. Da diese für das Text und Data Mining rein vorbereitenden Handlungen kein Verwertungsinteresse berührten, sehe das Gesetz keine Vergütungspflicht gegenüber dem Urheber vor**.**

Würden beim Training – wie hier - nicht nur Informationen aus Trainingsdaten extrahiert, sondern Werke vervielfältigt, stelle dies nach Auffassung der Kammer**kein Text und Data Mining**dar.**Die Prämisse des Text und Data Mining und der diesbezüglichen Schrankenbestimmungen, dass**durch die automatisierte Auswertung von bloßen Informationen selbst**keine Verwertungsinteressen berührt sind, greife in dieser Konstellation nicht. Im Gegenteil, durch die gegebenen Vervielfältigungen im Modell werde in das Verwertungsrecht der Rechteinhaber eingegriffen.**

Eine andere, mutmaßlich technik- und innovationsfreundliche Auslegung, die ebenfalls Vervielfältigungen im Modell von der Schranke als gedeckt ansehen wollte, verbiete sich angesichts des klaren Wortlauts der Bestimmung. Auch eine**analoge Anwendung komme nicht in Betracht.**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,**mangele es an einer vergleichbaren Interessenlage**. Die Schrankenregelung normiere mit der Zulässigkeit vorbereitender Vervielfältigungshandlungen beim Text und Data Mining einen Sachverhalt, bei dem die Verwertungsinteressen der Urheber nicht gefährdet seien, weil bloße Informationen extrahiert und das Werk als solches gerade nicht vervielfältigt werde.**Bei Vervielfältigungen im Modell werde 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. Das Risiko der Memorisierung stamme allein aus der Sphäre der Beklagten. Bei einer Analogie der Schranke würde ausschließlich der verletzte Rechteinhaber dieses Risiko tragen.

Downloaded from Blogs on 29.11.2026 05:26

Mangels Vorliegens eines Hauptwerks stellten die Vervielfältigungen der streitgegenständlichen Liedtexte**kein unzulässiges Beiwerk**nach§ 57 UrhG dar.Entgegen der Ansicht der Beklagten seien die Liedtexte nicht neben dem gesamten Trainingsdatensatz als nebensächlich und verzichtbar anzusehen. Hierfür wäre erforderlich, dass es sich bei dem gesamten Trainigsdatensatz ebenfalls um ein urheberrechtlich geschütztes Werk handele.

Der Eingriff der Beklagten in die Verwertungsrechte der Klägerin sei auch nicht durch eine**Einwilligung der Rechteinhaber gerechtfertigt,**da das Training von Modellen nicht als eine übliche und erwartbare Nutzungsart zu werten sei, mit der der Rechteinhaber rechnen müsse.

Auch durchWiedergabe der Liedtexte in den Outputs des Chatbotshätten die Beklagten nach der Entscheidung der Kammer unberechtigt die streitgegenständlichen Liedtexte vervielfältigt und öffentlich zugänglich gemacht. In den Outputs wären die originellen Elemente der Liedtexte stets wiedererkennbar.

Hierfür seien die Beklagten und nicht die Nutzer**verantwortlich**. Die Outputs seien durch einfach gehaltene Prompts generiert worden. Die Beklagten betrieben die Sprachmodelle, für die die Liedtexte als Trainingsdaten ausgewählt und mit denen sie trainiert worden sind. Sie seien für die Architektur der Modelle und die Memorisierung der Trainingsdaten verantwortlich. Damit hätten die von den Beklagten betriebenen Sprachmodelle die ausgegebenen Outputs maßgeblich beeinflusst, der konkrete Inhalt der Outputs werde von den Sprachmodellen generiert.

Der Eingriff in die Verwertungsrechte durch die Outputs sei ebenfalls nicht durch eine**Schrankenbestimmung gedeckt.**

Das Urteil ist nicht rechtskräftig.

Fundstelle

LG München I, **Pressemitteilung** vom11. November 2025.

Keywords

Gesetzgebung, Urheberrecht, Urheberrechtsverletzung