„Weggeforscht“ – der Podcast der Forschungsstelle Recht

Alle Informationen am Ende der Ausgabe

infobrief recht DFN

5 / 2026
Mai 2026

Mit digitalen Wasserzeichen zu klareren Informationsökosystemen?

Wie die EU KI-generierte Inhalte erkennbar machen will

Vertrau niemals deinem Auftragsverarbeiter

BGH zur Haftung des Verantwortlichen wegen des Verbleibs personenbezogener Daten beim Auftragsverarbeiter nach Beendigung des Auftragsverarbeitungsverhältnisses

Mehr öffentliche Daten für die Forschung?

Der Entwurf für ein Forschungsdatengesetz

Kurzbeitrag: Minest du das im Ernst?

Neue Rechtsprechung führt zu Unklarheiten bezüglich der Zulässigkeit von KI-Training mit urheberrechtlich geschützten Werken

Mit digitalen Wasserzeichen zu klareren Informationsökosystemen?

Wie die EU KI-generierte Inhalte erkennbar machen will

Von Dr. Paul Friedl

Katzenvideos, Forschungsdaten, Fotobeweise für Kriegsverbrechen: Seit Kurzem lässt sich all das täuschend echt von Künstlicher Intelligenz (KI) generieren. Jetzt verabschiedet die EU einen „Verhaltenskodex für die Kennzeichnung und Kennzeichnung von KI-generierten Inhalten" .[1]Können die Vorgaben zu KI-Wasserzeichen und anderen technischen Lösungen Abhilfe schaffen?

I. „Slop" und Co.: KI-generierte Gesellschaftsherausforderungen

Jedes Jahr im Dezember kürt das bekannte englische Wörterbuch Merriam Webster ein Wort des Jahres. 2020 war das „pandemic", passend zum grassierenden Corona-Virus, 2022 „gaslighting", ein Modebegriff zur Bezeichnung psychischer Manipulationsstrategien und 2024 „polarization", wenig überraschend im US-amerikanischen Wahljahr. Letztes Jahr kürte die Jury nun „slop" zum Gewinner. Damit werden nach der hauseigenen Definition „geringwertige digitale Inhalte, die für gewöhnlich in großen Mengen durch Künstliche Intelligenz produziert werden," bezeichnet.[2] Darunter fallen etwa Videos von alten Frauen, die angeblich ihren hundertzwanzigsten Geburtstag feiern oder Bilder eines aus Shrimp bestehenden Jesus.[3] Auf Deutsch lässt sich „slop" wörtlich mit „Pampe" oder „Mansch" übersetzen. Durchaus passend also, denn es steht zu befürchten, dass von generativer KI produzierter „slop" unsere ohnehin schon angeschlagenen Informationsökosysteme vertrübt, im schlimmsten Fall sogar verseucht. Was echt und was falsch ist, lässt sich schließlich immer schlechter unterscheiden.

II. Abhilfe durch Nachverfolgbarkeit und Transparenz?

Schon länger wird debattiert, diesen Problemen mit einer Transparenzoffensive zu begegnen: KI-generierte Inhalte müssten als solche gelabelt und Nutzer:innen diese Information zugänglich präsentiert werden. Dadurch solle sichergestellt werden, dass KI-generierte Inhalte nachverfolgt und potenziell auch moderiert werden können. Endnutzer:innen würden befähigt, Inhalte richtig einzuschätzen und darauf aufbauend ihr Verhalten anzupassen.

Um solche Nachvollziehbarkeit bzw. Transparenz zu erreichen, stehen verschiedene technologische Mittel zur Verfügung. Die technische Identifikation eines Inhalts als KI-generiert kann beispielsweise über Metadaten erfolgen, in denen der künstliche Ursprung einer Datei, klassischerweise einer Audio-, Video- oder Bilddatei, an die Hauptdatei angehängt und im Datentransfer an jede:n Empfänger:in (bspw. eine Social Media-Plattform oder ein:e Social Media-Nutzer:in) mitübermittelt werden. Derartige Metadaten lassen sich aber technisch leicht entfernen und können die Nachverfolgbarkeit des künstlichen Ursprungs eines Inhalts deshalb nur begrenzt sicherstellen. Deswegen gewinnen

1 So die eigenartige, KI-generierte deutsche Übersetzung des im englischen Original als „Code of Practice on Marking and Labelling of AI-generated content" betitelten Dokuments, https://digital-strategy.ec.europa.eu/en/library/commission-publishes-second-draft-code-practice-marking-and-labelling-ai-generated-content, alle Links dieses Beitrages wurden zuletzt abgerufen am: 18.03.2026.

2 Merriam Webster, „slop", Bedeutung 1.a, https://www.merriam-webster.com/dictionary/slop.

3 Beide Beispiele sind besonders weit verbreitet und deswegen bekannt, siehe https://www.theguardian.com/technology/2025/dec/27/from-shrimp-jesus-to-erotic-tractors-how-viral-ai-slop-took-over-the-internet.

seit einigen Jahren auch sogenannte digitale Wasserzeichen an Beliebtheit. Bei diesen wird die Struktur der generierten Datei noch im Erstellungsvorgang selbst so verändert, dass mit spezifisch dafür entwickelten Programmen der künstliche Ursprung der Datei ausgelesen werden kann, ohne dass dies für das menschliche Auge oder Ohr erkennbar wäre. Technisch wird dies beispielsweise durch „unsichtbare" Veränderungen der Pixelstruktur eines Bildes oder Videos erreicht. Auch Wasserzeichen bieten allerdings keinen hundertprozentigen Schutz; auch sie können - teils ohne größeren technischen Aufwand - überschrieben oder zerstört werden und sind insbesondere bei künstlich generierten Texten schwer zu bewerkstelligen und fehleranfällig.

Neben diesen Mitteln, die die *technische* Erkennbarkeit von KI-Inhalten gewährleisten können, bedarf es aber auch Techniken, die die Erkennbarkeit von KI-Inhalten *gegenüber Menschen* bewirken können. Hier kommen in erster Linie sichtbare Wasserzeichen, beispielsweise in Gestalt kleiner Logos, in Betracht, die auf das Medium „aufgedruckt" werden. KI-Systeme wie OpenAIs *Sora* oder Googles *Gemini* verwenden derartige sichtbare Wasserzeichen schon heute in der Bild- und Videogeneration.

III. Das Regelungskonzept der KI-VO

Auch die 2024 verabschiedete KI-Verordnung (KI-VO)[4] der EU sieht Regelungen vor, die die Nachvollziehbarkeit und Erkennbarkeit von KI-generierten Inhalten garantieren oder zumindest erhöhen sollen. Konkret sieht Art. 50 KI-VO Transparenzpflichten für zwei Arten von Akteuren vor: Nach Art. 50 Abs. 2 KI-VO müssen *Anbieter*[5] von KI-Systemen, die "synthetische Audio-, Bild-, Video- oder Textinhalte erzeugen", sicherstellen, dass die „Ausgaben des KI-Systems in einem maschinenlesbaren Format gekennzeichnet und als künstlich erzeugt oder manipuliert erkennbar sind". Diese Mittel müssen weiter möglichst „wirksam, interoperabel, belastbar und zuverlässig" sein. Praktisch bedeutet dies, dass GenAI-Anbieter wie Google, Meta oder OpenAI dafür sorgen müssen, dass die mit ihren Systemen generierten Inhalte (wirksam, interoperabel, etc.) als solche erkannt werden können. Daneben sieht Art. 50 Abs. 4 KI-VO vor, dass *Betreiber*[6] eines KI-Systems offenlegen müssen, dass Inhalte „künstlich erzeugt oder manipuliert wurden", wenn es sich entweder um a) „Deepfakes"[7] oder b) „Text, der … die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse informieren" soll, handelt, wobei in beiden Fällen gewisse Ausnahmen bestehen. Praktisch bedeutet dies, dass etwa OpenAI sicherstellen muss, dass mit ChatGPT erzeugte Deepfakes als solche erkennbar sind. Gleiches gilt nach der zweiten Alternative für Online-Magazine, die KI-generierte Artikel veröffentlichen wollen, sofern diese nicht "einem Verfahren der menschlichen Überprüfung oder redaktionellen Kontrolle unterzogen wurden". Wie genau diese Pflichten aber erfüllt werden sollen, lässt sich aus dem doch sehr abstrakten Normtext nicht ableiten.

IV. Ein bisschen konkreter: Der neue Praxisleitfaden

Zur Konkretisierung der abstrakten Pflichten sieht die KI-Verordnung deshalb auch die Ausarbeitung eines Praxisleitfadens (engl.: *Code of Practice*) vor. Solche Praxisleitfäden können von der Europäischen Kommission als rechtsverbindliche Konkretisierungen der genannten gesetzlichen Pflichten anerkannt werden. Kommt es zu dieser Anerkennung, reicht die Einhaltung der Vorgaben des Praxisleitfadens auch zur Einhaltung des Gesetzes aus. Aus diesem Grund kommt den Praxisleitfäden erhebliche Bedeutung zu, nicht zuletzt, weil derzeit alle beteiligten Akteure von einer Annahme durch die Kommission auszugehen scheinen.

---

4 Verordnung (EU) 2024/1689.

5 „Anbieter" sind in Art. 3 Nr. 4 KI-Verordnung definiert als „natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, die ein KI-System oder ein KI-Modell mit allgemeinem Verwendungszweck entwickelt oder entwickeln lässt und es unter ihrem eigenen Namen oder ihrer Handelsmarke in Verkehr bringt oder das KI-System unter ihrem eigenen Namen oder ihrer Handelsmarke in Betrieb nimmt, sei es entgeltlich oder unentgeltlich".

6 „Betreiber" sind in Art. 3 Nr. 4 KI-Verordnung definiert als „natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, die ein KI-System in eigener Verantwortung verwendet, es sei denn, das KI-System wird im Rahmen einer persönlichen und nicht beruflichen Tätigkeit verwendet".

7 „Deepfakes" sind Art. 3 Nr. 60 KI-VO definiert als ein „durch KI erzeugter oder manipulierter Bild-, Ton- oder Videoinhalt, der wirklichen Personen, Gegenständen, Orten, Einrichtungen oder Ereignissen ähnelt und einer Person fälschlicherweise als echt oder wahrheitsgemäß erscheinen würde". Im Hochschulkontext sind an die Transparenzpflichten für Betreiber nach Art. 50 Abs. 4 KI-VO beispielsweise bei der Erstellung von Lehrmaterial zu denken. Siehe dazu Schöbel, Der AI Act und die Wissenschaft, DFN-Infobrief Recht 2/2025, S. 4.

Die Ausarbeitung des Praxisleitfadens zur Kennzeichnung von KI-generierten Inhalten (engl: *Code of Practice on Marking and Labelling of AI-generated content*) stieß das Büro für Künstliche Intelligenz der EU im September 2025 an: KI-Entwickler, Nutzende, Forschende und Vertreter der Zivilgesellschaft (u. a.) wurden in einem Konsultationsverfahren dazu aufgerufen, Input zu verschiedenen offenen Fragen bezüglich Regelungsbereich und -inhalt der genannten Pflichten zu liefern. Ausgehend von diesen Einreichungen haben seitdem von der Kommission ausgewählte wissenschaftliche Expert:innen in Zusammenarbeit mit den eingeladenen Stakeholdern bisher zwei Entwürfe des Praxisleitfadens veröffentlicht.

Im ersten Entwurf dieses Praxisleitfadens wurden bereits erste große Weichenstellungen getroffen. Was die Pflichten für Anbieter von generativen KI-Systemen angeht, einigte man sich darauf, dass diese nicht auf eine einzige Kennzeichnungstechnologie setzen dürften, sondern vielmehr einen „multilayered approach", d. h. eine Kombination an Methoden, verfolgen sollten, da keine Technologie 100 % sicher sei. Zu diesen Technologien gehörten zumindest 1) für das menschliche Auge unsichtbare, für dazu spezialisierte Systeme aber erkennbare „Wasserzeichen", 2) Metadaten-Tags und 3) sogenannte Fingerprinting. Bei Letzterem wird für jedes erstellte Medium (bspw. ein Bild) ein digitaler Kurzschlüssel erstellt und gespeichert, mit dem in der Folge nachgewiesen werden kann, dass das betroffene Medium mit dem jeweiligen KI-System erstellt wurde. Außerdem sollten Hersteller ein Interface bereitstellen, mit dem Dritte kostenlos überprüfen lassen könnten, ob ein bestimmtes Medium aus dem betroffenen KI-System stammt und generell zur Entwicklung zuverlässigerer KI-Detektionssysteme beitragen. Betreiber von KI-Systemen wiederum müssten eine einheitliche Taxonomie zur Identifizierung von Deepfakes und KI-generierten Texten anwenden und derartige Inhalte mit einem leicht erkennbaren und barrierefreien Icon versehen. Zudem müssten Feedbackkanäle bereitgestellt werden, über die Dritte nicht- oder falsch-gelabelte Deepfakes oder KI-Texte melden könnten; Betreiber müssten Fehler in der Folge auch kategorisieren.

Im Anfang März veröffentlichten zweiten Entwurf des Praxisleitfadens wurde dann bestimmt, dass Hersteller von generativen KI-Systemen mindestens zwei Identifizierungs-Schichten verwenden müssten (Metadaten und unsichtbare Wasserzeichen). Außerdem sollten Identifizierungsmarker in einem von der EU noch zu entwickelnden, öffentlichen Verzeichnis bereitgestellt und hier gespeichert werden, um interoperable Identifizierungstechniken zu ermöglichen. Was die Pflichten für Betreiber betrifft, wurde klargestellt, dass die Verwendung eines von der EU noch zu entwickelnden Kennzeichnungs-Icons freiwillig bleiben würde. Zusätzlich wurde für die Entwicklung dieses Icons die Einsetzung einer EU-Taskforce in den Raum gestellt.

V. Potenzial des Leitfadens und Ausblick

Bei dem Leitfaden handelt es sich um einen ambitionierten Vorschlag, der dem „Problem" KI-generierter Inhalte mit einem umfangreichen und ausgefeilten Lösungsarsenal beikommen will und dabei auch auf die zentralen technologischen Herausforderungen eingeht, die zur Herstellung von Nachverfolgbarkeit und Transparenz bewältigt werden müssen. Die Technologieoffenheit, die dem Leitfaden an wichtigen Stellen eingeschrieben ist, und auch die vorgeschlagenen Kooperationsmechanismen zwischen KI-Herstellern, KI-Betreibern und Aufsichtseinrichtungen überzeugen.

Der Praxisleitfaden bietet also eine gute Lösung. Als Außenstehender fragt man sich allerdings auch: Für welches Problem eigentlich? Wie eingangs schon erwähnt wurde, spielen KI-generierte Inhalte in vielen Bereichen eine Rolle. Einige der problematischsten, etwa der Einsatz für Desinformationskampagnen oder Betrugsmaschen, zeichnen sich durch Aspekte aus, die in dem Leitfaden wenig bis gar nicht adressiert werden. Insbesondere müssen hier nämlich Akteure eingefangen werden, die die Identifikation ihrer Inhalte als KI-generiert aktiv verhindern wollen, und aus diesem Grund auch auf (außereuropäische) Anbieter zurückgreifen, die sich nicht an die Regeln der EU halten. Auch im Übrigen scheint der bisherige Entwurf zu wenig Augenmerk auf die Verbreitung KI-generierter Inhalte und die Akteure dieser Verbreitung, primär Social-Media-Plattformen, zu legen. Entscheidend ist in erster Linie nicht, dass Deepfakes oder andere (schädliche) KI-Inhalte nicht generiert werden, sondern, dass sie nicht oder nur mit adäquatem Kontext verbreitet werden. Um das zu erreichen, müssen Content-Verbreiter womöglich nicht nur befähigt, sondern auch verpflichtet werden, risikobehaftete Inhalte zu markieren oder zu verhindern. Auch wenn solche Pflichten potenziell auch aus dem Digital Services Act abgeleitet werden können, wäre eine stärkere Einbindung dieser Akteure auch in das Regelungsgefüge des Art. 50 KI-VO und den hier besprochenen Praxisleitfaden sicherlich wünschenswert. Bleibt zu hoffen, dass dies in der finalen, für den Juni angekündigten Version des Praxisleitfadens berücksichtigt wird. Denn aller

Voraussicht nach treten die Transparenzpflichten des Art. 50 KI-VO schon im August dieses Jahres in Kraft.

Das kann durchaus auch für Hochschulen oder dort Beschäftigte relevant werden. Treten diese nämlich als Betreiber eines KI-Systems auf – so etwa die Universität, die ihren Studierenden eine eigens getrimmte generative KI bereitstellt, oder die Lehrperson, die eine solche zur Erstellung von Vorlesungsmaterialien nutzt –, sind sie es, die das Einhalten der Kennzeichnungspflichten bezüglich Deepfakes und Texten von öffentlichem Interesse einhalten müssen.

Vertrau niemals deinem Auftragsverarbeiter

BGH zur Haftung des Verantwortlichen wegen des Verbleibs personenbezogener Daten beim Auftragsverarbeiter nach Beendigung des Auftragsverarbeitungsverhältnisses

von Johannes Müller-Westphal

Der Bundesgerichtshof (BGH) musste sich in seinem Urteil vom 11. November 2025 (Az. VI ZR 396/24) mit der Frage beschäftigen, in welchem Umfang ein Verantwortlicher auch für einen Datenverlust haften muss, der nicht beim Verantwortlichen selbst, sondern einem Auftragsverarbeiter nach Beendigung des Auftragsverarbeitungsverhältnisses aufgetreten ist. Hierzu konkretisierte der BGH die Pflichten des Verantwortlichen zum Schutz personenbezogener Daten nach Beendigung von Auftragsverarbeitungsverhältnissen.

I. Datenverarbeitungen durch Auftragsverarbeiter

Die Datenschutz-Grundverordnung (DSGVO)kennt zwei rechtliche Einordnungen von Akteuren, die an einer Datenverarbeitung beteiligt sein können: den Verantwortlichen und den Auftragsverarbeiter. Verantwortlicher ist gemäß Art. 4 Nr. 7 DSGVO eine Person oder Einrichtung, die allein oder gemeinsam über die Zwecke und Mittel der Verarbeitung personenbezogener Daten entscheidet.[1] Der Verantwortliche ist der primäre Adressat der Pflichten aus der DSGVO. Er muss gemäß Art. 24 DSGVO sicherstellen, dass eine Verarbeitung personenbezogener Daten im Einklang mit den Vorschriften der DSGVO erfolgt. In vielen Fällen führt der Verantwortliche die Datenverarbeitung auch selbst (bzw. durch seine Angestellten) durch. Dies ist allerdings nicht zwingend erforderlich. Die Definition in Art. 4 Nr. 7 DSGVO stellt lediglich darauf ab, dass der Verantwortliche die maßgebliche Entscheidung über die Datenverarbeitung trifft. Nicht erforderlich ist hingegen, dass der Verantwortliche die Datenverarbeitung auch selbst durchführt.

Ein solches Auseinanderfallen von der Entscheidung über eine Datenverarbeitung und deren Ausführung ist praktisch möglich, indem der Verantwortliche einen Auftragsverarbeiter einsetzt.[2] Ein Auftragsverarbeiter ist gemäß Art. 4 Nr. 8 DSGVO eine Person oder Einrichtung, die personenbezogene Daten im Auftrag des Verantwortlichen verarbeitet. Für Auftragsverarbeitungsverhältnisse trifft die DSGVO eigene Regelungen. Gemäß Art. 28 Abs. 1 DSGVO darf der Verantwortliche nur mit solchen Auftragsverarbeitern arbeiten, die hinreichend Garantien dafür bieten, dass die Verarbeitung im Einklang mit den Regelungen der DSGVO erfolgt. Die Datenverarbeitung muss auf Grundlage eines Vertrags erfolgen, der die Verarbeitungsmodalitäten festlegt. Dieser Vertrag muss gemäß Art. 28 Abs. 3 lit. a DSGVO unter anderem vorsehen, dass der Auftragsverarbeiter die personenbezogenen Daten nur auf (dokumentierte) Weisung des Verantwortlichen verarbeitet. Ebenso muss der Auftragsverarbeiter verpflichtet werden, nach Erbringung der Verarbeitungsleistungen nach Wahl des Verantwortlichen entweder alle personenbezogenen Daten zu löschen oder an den Verantwortlichen zurückzugeben und etwaige noch vorhandene Kopien zu löschen, sofern der Auftragsverarbeiter nicht durch gesetzliche Bestimmungen zur Speicherung der Daten verpflichtet ist.

1 Zur gemeinsamen Verantwortlichkeit Tech, Die Heiligen Drei der gemeinsamen Verantwortlichkeit, DFN-Infobrief Recht 12/2025.

2 Allgemein zur Einschaltung von Auftragsverarbeitern Geiselmann, Für die IT-Sicherheit haben wir jetzt jemanden, DFN-Infobrief Recht 6/2025.

II. Haftung für Datenschutzverstöße im Rahmen des Auftragsverarbeitungsverhältnisses

Sofern einer betroffenen Person, deren Daten verarbeitet werden, aufgrund einer rechtswidrigen Datenverarbeitung ein Schaden entsteht, kann sie gemäß Art. 82 DSGVO Schadensersatz verlangen. Sofern die Verarbeitung im Rahmen eines Auftragsverarbeitungsverhältnisses erfolgte, kommen als potenzielle Gegner des Schadensersatzanspruchs sowohl der Verantwortliche als auch der Auftragsverarbeiter in Betracht. Beide werden in Art. 82 Abs. 1 DSGVO genannt.

Die Haftung des Auftragsverarbeiters wird genauer in Art. 82 Abs. 2 S. 1 DSGVO geregelt. Dieser sieht vor, dass der Auftragsverarbeiter für einen entstandenen Schaden lediglich haftet, wenn er spezifische, für Auftragsverarbeiter geltende Pflichten, verletzt hat oder wenn er gegen die Weisung des Verantwortlichen gehandelt hat.

Die Haftung des Verantwortlichen ist hingegen weitreichender. Gemäß Art. 82 Abs. 2 S. 1 DSGVO haftet jeder Verantwortliche, der an einer Verarbeitung beteiligt ist, für den Schaden, der durch eine rechtswidrige Verarbeitung verursacht wurde. Hiernach ist es nicht erforderlich, dass der Verantwortliche selbst die Pflichten der DSGVO verletzt. Für die Haftung genügt es grundsätzlich, dass der Verantwortliche an der Verarbeitung beteiligt war. Damit kann der Verantwortliche auch haften, wenn der Rechtsverstoß nicht durch ihn, sondern den Auftragsverarbeiter begangen wurde. Jedoch sieht Art. 82 Abs. 3 DSGVO die Möglichkeit vor, dass sich ein Verantwortlicher (oder Auftragsverarbeiter) von einer Haftung befreien kann, indem er nachweist, dass er in keinerlei Hinsicht für den Umstand, durch den der Schaden eingetreten ist, verantwortlich ist. Nicht eindeutig ist, ob Art. 82 Abs. 3 DSGVO auch auf Datenverarbeitungen im Rahmen eines Auftragsverarbeitungsverhältnisses anwendbar ist. So ist es in der juristischen Literatur umstritten, ob sich ein Verantwortlicher auch auf Art. 82 Abs. 3 DSGVO berufen kann, sofern nicht der Verantwortliche, sondern allein der Auftragsverarbeiter rechtswidrig gehandelt und hierdurch einen Schaden begründet hat. Einige Stimmen vertreten die Auffassung, dass Art. 82 Abs. 3 DSGVO auch greift, wenn der Verantwortliche nachweist, dass der Datenschutzverstoß durch den Auftragsverarbeiter begangen wurde, der dabei eigenmächtig oder gegen die Weisungen des Verantwortlichen handelte.[3] In einer solchen Konstellation könne die geschädigte Person allein gegen den Auftragsverarbeiter und nicht auch gegen den Verantwortlichen einen Schadensersatzanspruch geltend machen. Die wohl überwiegende Gegenauffassung nimmt hingegen an, dass Art. 82 Abs. 3 DSGVO in dieser Konstellation nicht greift und dass sich ein Verantwortlicher nicht von einer Haftung befreien kann, indem er nachweist, dass der Auftragsverarbeiter gegen die Weisung des Verantwortlichen rechtswidrig gehandelt hat.[4] Nach dieser Auffassung haftet der Verantwortliche also unbeschränkt für Schäden, die durch ein datenschutzwidriges Verhalten des Auftragsverarbeiters hervorgerufen wurden. Dies beschränkt jedoch nicht die Möglichkeit des Verantwortlichen, im Anschluss den Auftragsverarbeiter in Anspruch zu nehmen und von ihm Ausgleich für die Schadensersatzzahlung zu verlangen, die der Verantwortliche an die betroffene Person geleistet hat. Hierbei trägt allerdings der Verantwortliche das Risiko, dass der Auftragsverarbeiter zahlungsunfähig ist, mit der Folge, dass der Verantwortliche keinen Ausgleich für die geleisteten Schadensersatzzahlungen erhält.

Soweit ersichtlich, bestehen bisher kaum Gerichtsurteile, die sich mit diesen Rechtsfragen auseinandersetzen. Besondere Bedeutung kommt daher dem höchstinstanzlichen Urteil des BGH zu, das sich mit einer Haftung des Verantwortlichen im Rahmen eines Auftragsverarbeitungsverhältnisses auseinandersetzt.

III. Sachverhalt: Datenverlust beim Auftragsverarbeiter

Die beklagte Partei in dem Prozess betrieb einen Musikstreamingdienst. Hierbei setzte sie einen Auftragsverarbeiter ein. Entsprechend den Vorgaben aus der DSGVO vereinbarten sie, dass der Auftragsverarbeiter nach Ende des Auftragsverhältnisses entweder eine vollständige Kopie aller personenbezogenen Daten an den Betreiber des Streamingdienstes zurückzugeben und alle anderen Kopien zu löschen habe oder alle Daten sowie deren Kopien löschen müsse. Zwischen diesen beiden Optionen

3 Paal/Pauly/Frenzel, DSGVO, Art. 82 Rn. 12; Taeger/Gabel/Moos/Schefzig, DSGVO, Art. 82 Rn. 66.

4 Kühling/Buchner/Bergt, DSGVO, Art. 82 Rn. 55; Simitis/Hornung/Spiecker gen. Döhmann/Boehm, DSGVO, Art. 82 Rn. 26; Dickmann, r +s 2018, 345 (347); Geissler/Ströbel, NJW 2019, 3414 (3415); Kremer/Conrady, ZD 2021, 128 (129); Paal, MMR 2020, 14 (18); Wybitul/Neu/Strach, ZD 2018, 202 (204).

sollte der Betreiber des Musikstreamingdienstes wählen können. Kurz vor Ende des Auftragsverarbeitungsverhältnisses am 1.12.2019 informierte der Auftragsverarbeiter den Betreiber des Streamingdienstes darüber, dass er dessen Webseite und die dort befindlichen Daten löschen werde. Etwa drei Jahre später wurde jedoch bekannt, dass unbekannte Hacker Daten von Nutzern des Dienstes im Darknet anboten. Diese Daten waren vom Auftragsverarbeiter entgegen der Vereinbarung nach Auftragsende nicht gelöscht worden. Stattdessen führten Mitarbeitende des Auftragsverarbeiters sie in eine Testumgebung ein. Im Anschluss müssen sie entweder von Hackern erbeutet oder von Mitarbeitenden des Auftragsverarbeiters unbefugt weitergegeben worden sein. Nachdem der Verlust der Daten bekannt wurde, informierte der Musikstreamingdienst hierüber die hiervon betroffenen Personen.

Zu den betroffenen Personen, deren Daten im Darknet angeboten wurden, zählte auch der Kläger. Er war einer der Nutzer des Musikstreamingdienstes. Der Datensatz, der im Darknet angeboten wurde, enthielt seinen Vor- und Nachnamen, sein Geschlecht, seine E-Mail-Adresse, seine Sprache sowie das Datum seiner Registrierung. Er verlangte von dem Betreiber des Musikstreamingdienstes Ersatz für die ihm entstandenen immateriellen Schäden. Seitdem er Kenntnis von dem Datenverlust erlangt habe, mache er sich Sorgen über den Verbleib der Daten und über einen möglichen Missbrauch. Dieser Missbrauch könne in Form von Identitätsdiebstahl, Phishing, unzulässigen Werbeanrufen und Werbemails auftreten.

Bevor die Sache vor dem BGH verhandelt wurde, haben sich das Landgericht Dresden und das Oberlandesgericht (OLG) Dresden mit dem Fall befasst. Das OLG Dresden lehnte zuletzt das Bestehen eines Schadensersatzanspruchs mit der Begründung ab, dass der betroffenen Person kein ersatzfähiger immaterieller Schaden entstanden sei.

IV. Das Urteil des BGH

Zunächst prüfte der BGH, ob der für einen Entschädigungsanspruch nach Art. 82 DSGVO erforderliche Verstoß gegen die Vorschriften der DSGVO vorliegt. Ein solcher Verstoß könnte sowohl durch den Auftragsverarbeiter als auch durch den Verantwortlichen (hier den Betreiber des Streamingdienstes) begangen worden sein. Hätte der BGH lediglich eine Pflichtverletzung des Auftragsverarbeiters festgestellt, hätte er sich anschließend ausführlich mit der (in der juristischen Literatur umstrittenen) Frage beschäftigen müssen, ob auch der Verantwortliche hierfür unbeschränkt haftet oder ob er sich gemäß Art. 82 Abs. 3 DSGVO von einer Haftung befreien kann.

Eine solche Erörterung war jedoch nicht erforderlich, da der BGH annahm, dass der Betreiber des Streamingdienstes selbst Pflichten aus der DSGVO verletzt habe. Diese Pflichtverletzung sah der BGH darin, dass der Betreiber nicht sichergestellt habe, dass der Auftragsverarbeiter tatsächlich die personenbezogenen Daten nach Beendigung des Auftragsverhältnisses löscht. Das Gericht wies darauf hin, dass Art. 28 Abs. 3 DSGVO vorsieht, dass der Verantwortliche und der Auftragsverarbeiter vertraglich vereinbaren müssen, nach Abschluss des Auftragsverhältnisses alle Daten zu löschen oder die Kopien zurückzugeben und vorhandene Kopien zu löschen. Zudem müsse der Vertrag gemäß Art. 28 Abs.3 S. 2 lit. h DSGVO regeln, dass der Auftragsverarbeiter dem Verantwortlichen die Einhaltung seiner Pflichten nachweist. Der BGH betonte jedoch, dass sich der Verantwortliche nicht damit begnügen darf, dem Auftragsverarbeiter diese vertraglichen Regeln aufzulegen. Stattdessen müsse der Verantwortliche das seinerseits Erforderliche beitragen, um sicherzustellen, dass der Auftragsverarbeiter seine vertraglichen Pflichten erfüllt, er also tatsächlich die personenbezogenen Daten löscht.

Eine solche Pflicht zur tatsächlichen Sicherstellung der Datenlöschung leitete der BGH aus dem Grundsatz der Datenminimierung gemäß Art. 5 Abs. 1 lit. c DSGVO und dem Grundsatz der Speicherbegrenzung gemäß Art. 5 Abs. 1 lit. e DSGVO her. Die Einhaltung dieser Grundsätze müsse der Verantwortliche gewährleisten. Ebenso leitete der BGH die Pflicht aus Art. 32 DSGVO ab. Dieser verlangt, dass der Verantwortliche angemessene technische und organisatorische Maßnahmen ergreifen muss, um ein angemessenes Schutzniveau zu gewährleisten. Hierbei müsse das Risiko berücksichtigt werden, dass unbefugte Personen Zugriff auf die Daten erlangen. Ein solches Risiko bestehe nach Auffassung des BGH aber nicht lediglich im Fall eines Cyberangriffs durch außenstehende Dritte, sondern auch bei einem Zugriff eines Auftragsverarbeiters, dessen Zugangsrechte mit Auftragsende erloschen sind. Daher habe der Verantwortliche sicherzustellen, dass der Auftragsverarbeiter nach Ende des Auftragsverarbeitungsverhältnisses tatsächlich die Daten löscht. Der BGH nahm demnach an, dass der Verantwortliche den Auftragsverarbeiter dazu drängen muss, die Daten zu löschen. Unterlässt er dies, entlaste es den Verantwortlichen auch nicht, dass er eine Löschpflicht vertraglich vereinbart hat und der

Auftragsverarbeiter hiergegen verstoßen hat. Folglich ging der BGH davon aus, dass der Verantwortliche nicht lediglich auf ein vertragskonformes Verhalten des Auftragsverarbeiters vertrauen darf, sondern ihn zu einem solchen Verhalten drängen müsse.

Diese Pflicht habe der Betreiber des Musikstreamingdienstes verletzt, indem es ihm genügte, dass der Auftragsverarbeiter ankündigte, er werde die Webseite und alle Daten auf ihr löschen. Nach dem abgeschlossenen Vertrag hätte der Auftragsverarbeiter die fristgerechte Einhaltung der Lösch- bzw. Rückgabepflicht schriftlich bestätigen müssen. Eine solche Bestätigung unterblieb aber, ohne dass der Betreiber des Streamingdienstes hierauf reagierte. Demnach nahm der BGH eine eigene Pflichtverletzung des Streamingdienst-Anbieters als Verantwortlichen an. Aufgrund dieser eigenen Pflichtverletzung war eine Haftungsbefreiung gemäß Art. 82 Abs. 3 DSGVO zweifelsohne nicht möglich. Daher musste der BGH auch keine klare Position zu der Frage einnehmen, ob sich der Verantwortliche von einer Haftung befreien kann, wenn ausschließlich der Auftragsverarbeiter Pflichten aus der DSGVO verletzt hat. Der BGH scheint jedoch davon auszugehen, dass eine Haftungsbefreiung des Verantwortlichen möglich sein kann, wenn allein der Auftragsverarbeiter seine Pflichten verletzt und hierbei gegen die Vereinbarung mit dem Verantwortlichen verstößt und der Verantwortliche hiermit nicht rechnen musste. Eine ausführliche Auseinandersetzung mit dieser Frage erfolgte jedoch nicht.

Anders als noch das OLG Dresden nahm der BGH neben dem Bestehen einer Pflichtverletzung zudem das Entstehen eines immateriellen Schadens an, der gemäß Art. 82 Abs. 1 DSGVO ersatzfähig ist. Hierzu erfolgten ebenso vertiefte Ausführungen des Gerichts, die an dieser Stelle aber nicht weiter behandelt werden sollen, da die Thematik des Ersatzes immaterieller Schäden infolge von Datenschutzverstößen bereits wiederholt Gegenstand verschiedener Infobriefe war.[5] In dem vorliegenden Fall nahm der BGH an, dass ein immaterieller Schaden zum einen in Form der missbräuchlichen Datenverwendung infolge des Kontrollverlustes und zum anderen in Form der hiermit verbundenen Befürchtungen bestehen würde.

V. Relevanz für wissenschaftliche Einrichtungen

Auch wissenschaftliche Einrichtungen setzen regelmäßig unterschiedliche Auftragsverarbeiter ein, die im Auftrag der Einrichtung Datenverarbeitungen vornehmen. Für solche Auftragsverarbeitungsverhältnisse lassen sich dem BGH-Urteil praxisnahe Hinweise entnehmen, die auch von wissenschaftlichen Einrichtungen beachtet werden müssen. Das Urteil trifft die klare Aussage, dass es nicht bereits genügt, dass ein Verantwortlicher den Auftragsverarbeiter vertraglich zu datenschutzkonformem Verhalten verpflichtet. Stattdessen muss er auch darauf hinwirken, dass sich der Auftragsverarbeiter entsprechend rechtskonform verhält. Konkret muss eine wissenschaftliche Einrichtung bei Ende eines Auftragsverhältnisses den Auftragsverarbeiter dazu drängen, die vorhandenen personenbezogenen Daten zu löschen. Unterlässt die wissenschaftliche Einrichtung dies, begeht sie eine eigenständige Pflichtverletzung.

5 Geiselmann, Kurzbeitrag: Keiner will den Schaden – aber jeder Ersatz, DFN-Infobrief Recht 11, 2024; Tech, Wer den Schaden hat, braucht für den Ärger nicht zu sorgen, DFN-Infobrief Recht 8/2024; Müller, Ich glaub, es hackt, DFN-Infobrief Recht 04/2024; Voget, Kurzbeitrag: Nicht (un)erheblich?!, DFN-Infobrief Recht 07/2023; Müller, Schaden oder kein Schaden, das ist hier die Frage, DFN-Infobrief Recht 4/2023; Müller, Morgen Kinder werden wir klagen, DFN-Infobrief Recht 12/2022.

Mehr öffentliche Daten für die Forschung?

Der Entwurf für ein Forschungsdatengesetz

Von Philipp Schöbel

Das Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR) hat einen Referentenentwurf für ein Gesetz zur verbesserten Nutzung von Daten für die Forschung veröffentlicht.[1] Der Entwurf sieht Änderungen an bestehenden Gesetzen (wie dem Bundesstatistikgesetz und dem Hochschulstatistikgesetz) vor und enthält außerdem einen Vorschlag für ein gänzlich neues „Gesetz zum Zugang zu und zur Nutzung von Daten für die Forschung (Forschungsdatengesetz – FDG)".

I. Bisherige Entwicklung

Die Idee eines Forschungsdatengesetzes ist nicht gänzlich neu. Bereits im Koalitionsvertrag 2021 wurde ein Forschungsdatengesetz als Idee aufgeführt, um den Zugang zu Forschungsdaten für öffentliche und private Forschung umfassend zu verbessern sowie zu vereinfachen.[2] Im Februar 2024 legte das damalige Bundesministerium für Bildung und Forschung (BMBF) die Eckpunkte eines zukünftigen Forschungsdatengesetzes vor.[3] Auch im aktuellen Koalitionsvertrag findet sich das Forschungsdatengesetz wieder.[4] Im Dezember 2025 hat nun das BMFTR einen Referentenentwurf für ein Gesetz zur verbesserten Nutzung von Daten für die Forschung veröffentlicht, das auch einen Entwurf für ein Forschungsdatengesetz (FDG-E) enthält.

II. Ziele des FDGs

„Die datengetriebene Forschung ist das Fundament, auf dem unsere Innovationen und unsere wissenschaftliche Wettbewerbsfähigkeit stehen."[5] Mit diesem bedeutungsschweren Satz beginnt die Begründung des Gesetzesentwurfs. Das FDG-E dient der Verbesserung des Zugangs zu und der Nutzung von Daten zu Forschungszwecken (§ 1 Abs. 1 FDG-E). Das FDG-E verfolgt im Wesentlichen drei Ziele.[6] Vorhandene Lücken in der Datennutzung für die Forschung sollen mittels einer bereichsübergreifenden Rechtsgrundlage zum Zugang zu Daten für Forschungszwecke geschlossen werden (1). Zudem soll die Datenzusammenführung von Datenbeständen der öffentlichen Hand und der Forschungsdatenzentren mittels einer bereichsübergreifenden Rechtsgrundlage ermöglicht werden (2). Schließlich will man die Datenschutzaufsicht bei länderübergreifenden Forschungsvorhaben einheitlicher und anwendungsfreundlicher ausgestalten (3).

---

1 BMFTR, Referentenentwurf, Entwurf eines Gesetzes zur verbesserten Nutzung von Daten für die Forschung, 22.12.2025, abrufbar unter: https://www.bmftr.bund.de/SharedDocs/Downloads/DE/gesetze/forschungsdatengesetz/referentenentwurf/referentenentwurf_bessere_nutzung_von_daten_forschung.pdf?__blob=publicationFile&v=3 (alle Links dieses Beitrags wurden zuletzt am 30.04.2026 abgerufen).

2 Koalitionsvertrag 2021 - 2025 zwischen der Sozialdemokratischen Partei Deutschlands (SPD), Bündnis 90/Die Grünen und den Freien Demokraten (FDP), Mehr Fortschritt wagen - Bündnis für Freiheit, Gerechtigkeit und Nachhaltigkeit, 07.12.2021, S. 18.

3 BMBF, Eckpunkte BMBF Forschungsdatengesetz, 28.02.2024, abrufbar unter: https://www.bmftr.bund.de/SharedDocs/Downloads/DE/gesetze/forschungsdatengesetz/sonstige/Eckpunktepapier.pdf?__blob=publicationFile&v=3.

4 Koalitionsvertrag für die 21. Legislaturperiode zwischen CDU, CSU und SPD, „Verantwortung für Deutschland", 05.05.2025, S. 79.

5 BMFTR, Referentenentwurf, Entwurf eines Gesetzes zur verbesserten Nutzung von Daten für die Forschung, 22.12.2025, S. 1 u. 33.

6 Zu diesen drei Zielen: BMFTR, Referentenentwurf, Entwurf eines Gesetzes zur verbesserten Nutzung von Daten für die Forschung, 22.12.2025, S. 1 u. 35.

Teilweise bestehen zwar schon nach derzeitiger Rechtslage einzelne Rechtsgrundlagen für einen Datenzugang. Diese sind aber in vielen Einzelgesetzen verteilt und meist auf sehr konkrete Forschungsvorhaben beschränkt.[7] Das FDG-E soll Abhilfe leisten, indem einheitliche Rechtsgrundlagen für den Zugang zu und die Zusammenführung von Daten geschaffen werden.[8] Zudem sollen für Fälle, in denen datenschutzrechtliche Nutzungsmöglichkeiten der betroffenen Daten noch nicht bestehen, entsprechende Rechtsgrundlagen geschaffen werden.[9]

Das FDG-E soll aber nicht uneingeschränkt gelten. Die Anwendung des FDG-E ist ausgeschlossen, soweit andere Gesetze die Verarbeitung zu Forschungszwecken ausdrücklich ausschließen (§ 1 Abs. 2 FDG-E). Der Gesetzgeber hat hier Regelungen wie die des § 21 Abs. 2 S. 2 i. V. m Abs. 1 S. 2 Bundeskriminalgesetzes (BKAG) vor Augen.[10] Das BKAG erlaubt dem BKA unter engen Voraussetzungen die Weiterverwendung personenbezogener Daten zur wissenschaftlichen Forschung (§ 21 Abs. 1 S. 1 BKAG). Dies gilt aber nicht für personenbezogene Daten, die durch einen verdeckten Einsatz technischer Mittel in oder aus Wohnungen oder einen verdeckten Eingriff in informationstechnische Systeme erlangt wurden (§ 21 Abs. 1 S. 2 i. V. m. § 13 Abs. 3 BKAG). Das FDG-E soll also diese speziellen Verbote nicht unterlaufen.

III. Datenkategorien nach dem FDG-E

Die Kategorien von Daten, zu denen ein Zugang nach dem FDG-E beansprucht werden kann, sind im Gesetz abschließend aufgeführt. Dazu gehören Daten in Registern, die zum Zwecke der Registerführung dort eingetragen sind (§ 7 Abs. 1 Nr. 1 FDG-E). Weiterhin gehören dazu Daten, die vom Statistischen Bundesamt oder den statistischen Ämtern der Länder zu statistischen Zwecken verarbeitet werden (§ 7 Abs. 1 Nr. 2 FDG-E) oder solche Daten, die sie aus allgemein zugänglichen Quellen gewinnen (§ 7 Abs. 4 Nr. 3 FDG-E). Ebenfalls erfasst sind Daten oberster Bundesbehörden, die diese zur Erfüllung statistischer Berichtspflichten nach EU-Recht erhoben haben (§ 7 Abs. 4 Nr. 4 FDG-E). Dies umfasst auch Daten, die zu diesem Zweck im Auftrag oberster Bundesbehörden erhoben wurden. Eine weitere Kategorie von Daten sind solche, die Forschungsdatenzentren für Forschungszwecke bereithalten (§ 7 Abs. 4 Nr. 5 FDG-E). Die Forschungsdatenzentren werden in Anlage 2 des Gesetzes abschließend aufgeführt. Weitere Daten sind solche aus Statistiken, die von der Deutschen Bundesbank oder der Bundesagentur für Arbeit erstellt wurden (§ 7 Abs. 4 Nr. 6 FDG-E). Eine weitere Kategorie umfasst Daten, von Bundeseinrichtungen oder Landeseinrichtungen mit Forschungsaufgaben, soweit diese zu Forschungszwecken bereitgehalten werden (§ 7 Abs. 4 Nr. 7 FDG-E). Daten der Träger der Deutschen Rentenversicherung stellen auch eine eigene Datenkategorie dar (§ 7 Abs. 4 Nr. 8 FDG-E). Die vorletzte Datenkategorie umfasst Daten nach dem Verwaltungsdatenverwendungsgesetz (§ 7 Abs. 4 Nr. 9 FDG-E). Schließlich sind auch Daten erfasst, die dem Deutschen Zentrum für Mikrodaten (DZM) freiwillig zur Nutzung für Forschungszwecke zur Verfügung gestellt werden und deren Verfügbarkeit für Forschungszwecke nach dem FDG-E das DZM der Öffentlichkeit anzeigt. (§ 7 Abs. 4 Nr. 10 FDG-E).

IV. Deutsches Zentrum für Mikrodaten

Das DZM hat die Aufgabe, die im öffentlichen Interesse liegende Forschung zu fördern (§ 3 Abs. 3 FDG-E). Das DZM wird beim Statistischen Bundesamt als von dessen Verwaltungsbereich zu trennende, eigene und unabhängige Organisationseinheit eingerichtet (§ 3 Abs. 1 S. 1 FDG-E). Die Unabhängigkeit des DZM soll dem Abschottungs- und Trennungsgebot Rechnung tragen.[11] Nach diesem Gebot muss der Statistikbetrieb bei den statistischen Ämtern von einer Nichtstatistik-Einrichtung getrennt sein. Damit soll sichergestellt werden, dass Datensätze von personenbezogenen Angaben getrennt werden. Das Bundesverfassungsgericht hat bereits im Jahr 1983 in seinem Volkszählungsurteil dargelegt, dass die strikte Geheimhaltung der zu statistischen Zwecken erhobenen Einzelangaben für den Schutz des Rechts auf informationelle Selbstbestimmung

7 Vgl. BMFTR, Referentenentwurf, Entwurf eines Gesetzes zur verbesserten Nutzung von Daten für die Forschung, 22.12.2025, S. 1.

8 BMFTR, Referentenentwurf, Entwurf eines Gesetzes zur verbesserten Nutzung von Daten für die Forschung, 22.12.2025, S. 1 u. 48.

9 BMFTR, Referentenentwurf, Entwurf eines Gesetzes zur verbesserten Nutzung von Daten für die Forschung, 22.12.2025, S. 1 u. 48.

10 BMFTR, Referentenentwurf, Entwurf eines Gesetzes zur verbesserten Nutzung von Daten für die Forschung, 22.12.2025, S. 1 u. 48.

11 Siehe hierzu und zu den nachfolgenden beiden Sätzen: BMFTR, Referentenentwurf, Entwurf eines Gesetzes zur verbesserten Nutzung von Daten für die Forschung, 22.12.2025, S. 52.

unverzichtbar ist, solange ein Personenbezug noch besteht oder herstellbar ist.[12] Auch hat es schon damals festgestellt, dass das Vertrauen der Bürger:innen in die Abschottung notwendig ist, damit sie gegenüber dem Staat bereit sind, Angaben zu statistischen Zwecken zu machen.[13]

Das DZM soll aus einer Forschungsstelle und einer Vertrauensstelle bestehen (§ 3 Abs. 1 S. 1 FDG-E). Die Forschungsstelle soll insbesondere Anträge auf Datenzugang und Datenzusammenführung zu Forschungszwecken prüfen sowie die Verarbeitung von Forschungsdaten ermöglichen (vgl. § 3 Abs. 3 Nr. 2 FDG-E). Die Vertrauensstelle wirkt bei der Pseudonymisierung und der Erstellung von Pseudonymen mit (§ 5 Abs. 1 S. 1 FDG-E). Diese beiden Stellen sollen räumlich, organisatorisch und technisch voneinander getrennt geführt werden (§ 3 Abs. 2 FDG-E).

Die Förderung der Forschung durch das DZM erfolgt durch eine Reihe von Unteraufgaben. Das DZM soll für Anträge auf Zugang zu Forschungsdaten entsprechende elektronische Strukturen schaffen und bereitstellen (§ 3 Abs. 3 Nr. 1 FDG-E). Zudem soll es Anträge auf Datenzugang und Datenzusammenführung prüfen sowie die damit zusammenhängende Verarbeitung von Daten ermöglichen (§ 3 Abs. 3 Nr. 2 FDG-E). Weiterhin soll es eine kontrollierte, besonders gesicherte elektronische Verarbeitungsumgebung für den Datenzugang und die Datenzusammenführung bereitstellen (§ 3 Abs. 3 Nr. 3 FDG-E). Eine weitere Aufgabe des DZM ist es, Forschende im Rahmen der Datennutzung zu beraten und zu betreuen (§ 3 Abs. 3 Nr. 4 FDG-E). Dies umfasst auch eine Beratung bezüglich der Sicherstellung der Geheimhaltung der bereitgestellten Daten. Außerdem soll das DZM die Öffentlichkeit über seine Arbeit, über die gestellten Anträge, die Antragstellerinnen und Antragsteller sowie über publizierte Forschungsergebnisse in einem öffentlichen Antragsregister informieren (§ 3 Abs. 3 Nr. 5 FDG-E). Eine weitere Aufgabe ist die Unterstützung der Bundesregierung bei der Steigerung der Verfügbarkeit von Daten zu Forschungszwecken auf nationaler und europäischer Ebene (§ 3 Abs. 3 Nr. 6 FDG-E). Darüber hinaus soll es praktische Empfehlungen zur Datennutzung zu Forschungszwecken und zum Aufbau und zur Nutzung von sicheren Verarbeitungsinfrastrukturen geben (§ 3 Abs. 3 Nr. 7 FDG-E). Schließlich soll es auch im Rahmen seiner Zuständigkeit zur Verfügbarkeit qualitativ hochwertiger Daten beitragen (§ 3 Abs. 3 Nr. 8 FDG-E).

12 BVerfGE 65, 1, 50.

13 Ebd.

V. Datenanbietende und datenhaltende Stellen

Zwei weitere wichtige Akteure nach dem FDG-E sind die sogenannten datenanbietenden und datenhaltenden Stellen. Eine „datenhaltende Stelle“ ist jede Stelle, die über Daten im Sinne des FDG-E verfügt (§ 2 Nr. 13 S. 1 HS. 1 FDG-E). Hierbei ist es unerheblich, ob sie die Daten selbst erhebt oder ihr diese von anderer Stelle zur Verfügung gestellt werden (§ 2 Nr. 13 S. 1 HS. 2 FDG-E). Datenanbietende Stelle ist hingegen jede Einrichtung, die bei staatlichen Stellen oder bei wissenschaftlichen Institutionen angesiedelt ist, um wissenschaftliche Erkenntnisse zu ermöglichen, und die in der Regel eine der im Gesetz vorgesehenen Aufgaben hat (§ 2 Nr. 14 FDG-E). Zu diesen Aufgaben gehören Datenverarbeitung, Bereitstellung von Metadaten, Minimierung des Re-Identifikationsrisikos sowie das Aufbauen von datenspezifischen Fachkenntnissen. Datenhaltende Stellen können zugleich auch datenanbietende Stellen sein (§ 2 Nr. 15 FDG-E).

Die datenanbietenden und datenhaltenden Stellen sind verpflichtet, auf Ersuchen die angeforderten Daten an das DZM zu übermitteln (§ 6 S. 1 FDG-E). Sie können die Übermittlung unter strengen Voraussetzungen auch verweigern. Dies ist der Fall, soweit gesetzliche Verbote oder vertragliche Nutzungsbeschränkungen oder Bestimmungen über die Verwendungszweckbeschränkung und Geheimhaltungspflicht einer Übermittlung entgegenstehen (§ 6 S. 3 FDG-E).

VI. Zugang zu Daten zu Forschungszwecken

Das DZM hat Forschenden auf Antrag Zugang zu den im Antrag genannten Daten zu gewähren, soweit vier Voraussetzungen erfüllt sind (§ 7 Abs. 1 FDG-E). Die Forschenden müssen zunächst einer akkreditierten Einrichtung angehören (§ 7 Abs. 1 Nr. 1 FDG-E). Das DZM akkreditiert auf Antrag Einrichtungen nach § 2 Nr. 8-11 FDG-E (§ 8 Abs. 1 FDG-E). Zu diesen Einrichtungen gehören Hochschulen der Länder und des Bundes, private Hochschulen, Forschungseinrichtungen sowie Unternehmen, die Forschung betreiben. Die Einrichtungen müssen wissenschaftliche Standards und datenschutzrechtliche Anforderungen einhalten (§ 8 Abs. 2 Nr. 2, 3 FDG-E) sowie Datensicherheit durch

die Erfüllung von technischen und infrastrukturbezogenen Anforderungen gewährleisten (§ 8 Abs. 2 Nr. 4 FDG-E). Die zweite Voraussetzung für einen Antrag auf Zugang zu Forschungsdaten ist, dass die Daten für das konkrete Forschungsvorhaben und dessen Zweck erforderlich sind (§ 7 Abs. 1 Nr. 2 FDG-E). Darüber hinaus muss das Forschungsvorhaben im öffentlichen Interesse liegen (§ 7 Abs. 1 Nr. 3 FDG-E). Schließlich müssen schutzwürdige Interessen nicht beeinträchtigt werden oder das Interesse an der Forschung muss das Geheimhaltungsinteresse überwiegen (§ 7 Abs. 1 Nr. 4 S. 1 FDG-E). Schutzwürdige Interessen sind der Schutz personenbezogener Daten, der Schutz der öffentlichen Sicherheit und Ordnung sowie der Schutz von Geschäftsgeheimnissen (§ 7 Abs. 1 Nr. 4 S. 3 FDG-E).

Das DZM kann einen Antrag auch in vorgeschriebenen Fällen ablehnen. Dies kann es etwa, wenn der begründete Verdacht besteht, die Daten könnten für einen anderen Zweck als zu dem Zweck der Durchführung des dem Antrag zugrunde liegenden Forschungsvorhabens verarbeitet werden (§ 7 Abs. 2 Nr. 1 FDG-E). Ein weiterer Ablehnungsgrund wäre, die übermäßige Antragstellung durch Forschende, wenn dies dazu führt, dass die Kapazitäten des DZM unverhältnismäßig gebunden sind und seine Arbeitsfähigkeit gefährdet wird (§ 7 Abs. 2 Nr. 2 FDG-E). Schließlich kann das DZM auch Anträge ablehnen, wenn die datenhaltende oder datenanbietende Stelle die Übermittlung der Daten aus berechtigten Gründen verweigert hat (§ 7 Abs. 2 Nr. 3 FDG-E). Das DZM kann außerdem unter besonderen Umständen Forschende vom Datenzugang für bis zu zwei Jahre ausschließen (§ 7 Abs. 7 S. 2 FDG-E). Dies kann sie, wenn eine Datenschutzbehörde festgestellt hat, dass die Forschenden die bereitgestellten Daten entgegen den datenschutzrechtlichen Vorschriften oder entgegen den Auflagen des DZM verarbeitet haben (§ 7 Abs. 7 S. 1 FDG-E). Zusätzlich muss die Datenschutzbehörde auch eine Abhilfemaßnahme nach Art. 58 Abs. 2 lit. b - j DSGVO erlassen haben, also etwa eine Verwarnung gegenüber den Forschenden ausgesprochen haben (Art. 58 Abs. 2 lit. b DSGVO).

Die Forschenden können wählen, ob sie den Zugriff auf die Forschungsdaten über einen Fernzugriff oder in den dafür vorgesehenen Räumlichkeiten der Forschungsstelle wahrnehmen (§ 7 Abs. 6 FDG-E). Der Fernzugriff erfolgt über eine von der Forschungsstelle bereitgestellte, kontrollierte und besonders gesicherte Verarbeitungsumgebung. Besonderheiten gelten für personenbezogene Daten. So können diese inklusive der besonderen Kategorien personenbezogener Daten (Art. 9 Abs. 1 DSGVO) in der besonders gesicherten Verarbeitungsumgebung der Forschungsstelle des DZM in pseudonymisierter Form verarbeitet werden (§ 7 Abs. 8 S. 1 FDG-E). Hinsichtlich personenbezogener Daten sind Forschende zur Geheimhaltung verpflichtet und sie können diese Daten auch nur eingeschränkt nutzen. Sie dürfen diese etwa nur für die Zwecke verwenden, für die sie ihnen zugänglich gemacht werden (§ 12 Abs. 1 S. 1 Nr. 1 FDG-E). Außerdem dürfen sie die Daten nicht zur Herstellung eines Personenbezugs verwenden (§ 12 Abs. 1 S. 1 Nr. 2 FDG-E). Sie dürfen sie auch nicht an Dritte weitergeben oder in sonstiger Weise offenlegen (§ 12 Abs. 1 S. 1 Nr. 3 FDG-E). Dies gilt auch für Daten, die sich auf verstorbene Personen beziehen (§ 12 Abs. 1 S. 2 FDG-E).

VII. Zusammenführung von Daten zu Forschungszwecken

Ein weiterer gewichtiger Baustein neben dem Datenzugang ist die Zusammenführung von Daten zu Forschungszwecken. Der Gesetzesentwurf geht davon aus, dass bestimmte Erkenntnisse nur über die Zusammenführung von Daten gewonnen werden können.[14] Auf Antrag hat das DZM die Zusammenführung von Daten im Sinne des FDG-E (§ 7 Abs. 4 FDG-E) zu ermöglichen und den Zugang zu diesen Daten zu gewähren (§ 9 Abs. 1 S. 1 FDG-E). Dafür müssen vier Voraussetzungen erfüllt sein. Die Forschenden müssen einer akkreditierten Einrichtung angehören (§ 9 Abs. 1 S. 1 Nr. 1 FDG-E). Zudem muss die Datenzusammenführung für die Durchführung der konkreten Forschungsvorhaben erforderlich sein (§ 9 Abs. 1 S. 1 Nr. 2 FDG-E). Weiterhin muss das Forschungsvorhaben im öffentlichen Interesse liegen (§ 9 Abs. 1 S. 1 Nr. 3 FDG-E). Schließlich müssen schutzwürdige Interessen nicht beeinträchtigt werden oder das Interesse an der Forschung muss das Geheimhaltungsinteresse überwiegen (§ 9 Abs. 1 S. 1 Nr. 4 FDG-E).

VIII. Anonymisierungs- und Publikationspflicht von Forschungsergebnissen

Ein weiterer wichtiger Teil des Gesetzes sind die Pflichten zur Anonymisierung und Publikation von Forschungsergebnissen.

14 BMFTR, Referentenentwurf, Entwurf eines Gesetzes zur verbesserten Nutzung von Daten für die Forschung, 22.12.2025, S. 65.

Die Forschungsergebnisse dürfen nur in anonymisierter Form veröffentlicht werden (§ 10 Abs. 1 FDG-E). Vor einer Veröffentlichung müssen die Forschenden ihre Forschungsergebnisse dem DZM übermitteln (§ 10 Abs. 2 S. 1 FDG-E). Das DZM prüft dann, ob die Forschungsergebnisse einen Rückschluss auf Angaben über betroffene Personen oder auf nicht personenbezogene Daten mit Unternehmens-, Betriebs- oder Arbeitsstättenbezug zulassen (§ 10 Abs. 2 S. 2 FDG-E). Sofern eine Rückschlussmöglichkeit besteht, müssen die Forschenden diese beseitigen (§ 10 Abs. 2 S. 3 FDG-E).

Soweit für ein Forschungsvorhaben die vom DZM bereitgestellten Daten genutzt werden, müssen die Forschungsergebnisse innerhalb von 36 Monaten nach Abschluss des Forschungsvorhabens in anonymisierter Form der Allgemeinheit zur Verfügung gestellt werden (§ 10 Abs. 3 S. 1 FDG-E). Die Hauptergebnisse sind unentgeltlich zur Verfügung zu stellen, soweit nicht erhebliche Rechte und Interessen Dritter entgegenstehen. Die Forschenden müssen das DZM über die erfolgte Publikation unverzüglich informieren (§ 10 Abs. 3 S. 2 FDG-E).

IX. Ausblick

Bisher ist eine Vielzahl von Stellungnahmen zu dem Gesetzesentwurf eingegangen.[15] Das Gesetz wird grundsätzlich begrüßt, aber teilweise wird auch Nachbesserungsbedarf gesehen.[16] Als Nächstes wird die Bundesregierung einen Regierungsentwurf des Gesetzes ausarbeiten, der dann wahrscheinlich in den Bundestag eingebracht werden wird. Es bleibt abzuwarten, welche Regelungen noch angepasst werden. Bis zum Inkrafttreten des Forschungsdatengesetzes wird noch etwas Zeit vergehen. Es wird sich zeigen, ob das künftige FDG die Datensilos aufschließen und für die deutsche Forschung zugänglich machen kann.

15 Alle Stellungnahmen sind abrufbar unter: https://www.bmftr.bund.de/DE/Ministerium/Gesetze/Gesetze-Einzeln/Forschungsdatengesetz.html.

16 Forschung & Lehre, Organisationen aus der Wissenschaft begrüßen Referentenentwurf, 04.03.2026, abrufbar unter: https://www.forschung-und-lehre.de/politik/wissenschaftsorganisationen-begruessen-referentenentwurf-7566.

DFN Infobrief-Recht-Aktuell

• Strafrecht/Datenschutzrecht: Urteil des Europäischen Gerichtshofs (EuGH) zur Erhebung biometrischer Daten im Rahmen eines strafrechtlichen Ermittlungsverfahrens

Der EuGH hat am 19. März 2026 entschieden, dass die Erhebung biometrischer Daten für erkennungsdienstliche Maßnahmen nur möglich ist, wenn sie unbedingt erforderlich ist und geeignete Garantien für die Rechte und Freiheiten der betroffenen Person bestehen. Ihre Verarbeitung ist daher nur mit einer klaren Begründung zulässig, da biometrische Daten zu den sensiblen personenbezogenen Daten im Sinne des Unionsrechts gehören.

Hier erhalten Sie den Link zur Entscheidung (alle Links dieser Seite wurden zuletzt am 30.04.2026 abgerufen):
https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:62024CJ0371

• Datenschutzrecht: Auslaufen der Übergangsregelung zur freiwilligen Chatkontrolle

Die freiwillige Chatkontrolle wird nach der Abstimmung des EU-Parlaments nicht verlängert. Gemäß der Übergangsregelung der Verordnung (EU) 2021/1232 konnten Anbieter digitaler Kommunikationsdienste bis zum 3. April 2026 freiwillig private Kommunikation automatisiert zur Erkennung von Darstellungen sexuellen Kindesmissbrauchs sowie verdächtiger Inhalte durchsuchen. Diese bis zum 3. April 2026 bestehende Ausnahmeregelung läuft damit aus. Die Konferenz der unabhängigen Datenschutzbehörden des Bundes und der Länder (DSK) hatte bereits am 17. Oktober 2023 eine Entschließung veröffentlicht, in der sie sich gegen die Chatkontrolle aussprach.

Hier erhalten Sie den Link zur Entschließung der DSK:
https://www.datenschutzkonferenz-online.de/media/en/20231017DSKEntschliessungChatkontrolle.pdf

• Medienrecht: EU-Kommission veröffentlicht Leitlinien zum Schutz von Medieninhalten auf Plattformen

Am 5. Februar 2026 hat die EU-Kommission detaillierte Leitlinien zur Umsetzung von Art. 18 Abs. 1 des European Media Freedom Act (EMFA) veröffentlicht. Art. 18 EMFA sieht besondere Verfahrensgarantien zugunsten von Mediendiensteanbietern gegenüber sehr großen Online-Plattformen (VLOPs) vor. Nach diesen Schutzvorkehrungen sind VLOPs verpflichtet, Mediendiensteanbieter vor der Entfernung journalistischer Inhalte zu benachrichtigen und die Gründe hierfür darzulegen.

Hier erhalten Sie den Link zur zur Pressemitteilung der EU-Kommission:
https://eur-lex.europa.eu/eli/C/2026/901/oj/eng

• Arbeitsrecht/Datenschutzrecht: Urteil des Hessischen Landesarbeitsgerichts zu Schadensersatzansprüchen nach einem Hackerangriff

Mit Urteil vom 10. Februar 2026 hat das Hessische Landesarbeitsgericht entschieden, dass ein Schadensersatzanspruch gemäß Art. 82 Datenschutz-Grundverordnung im konkreten Fall nicht besteht. Nach einem Hackerangriff waren Dateinamen im Darknet zum Zwecke des Verkaufs veröffentlicht worden. Der Kläger konnte nicht hinreichend darlegen, welche konkreten personenbezogenen Daten betroffen waren. Zudem fehlte es an einem ersatzfähigen Schaden. Die Berufung war daher erfolglos.

Hier erhalten Sie den Link zur Entscheidung:
https://www.rv.hessenrecht.hessen.de/bshe/document/LARE260000341

Kurzbeitrag: Minest du das im Ernst?

Neue Rechtsprechung führt zu Unklarheiten bezüglich der Zulässigkeit von KI-Training mit urheberrechtlich geschützten Werken

Von Sofie Leonhard

Das Training von großen KI-Modellen findet regelmäßig mit Datensätzen statt, die urheberrechtlich geschützte Werke beinhalten. Es ist aber bislang unklar, ob und in welchem Umfang das KI-Training mit diesen Werken rechtlich zulässig ist. Dadurch ergeben sich auch für Forschungseinrichtungen erhebliche Unsicherheiten. Nachfolgend soll die aktuelle Rechtslage anhand der jüngsten Rechtsprechung dargestellt werden. Dabei wird deutlich: Es bedarf einer höchstrichterlichen Entscheidung, um die Frage final zu beantworten.

I. Hintergrund

1. Urheberrechtliche Schranken

Das Urheberrecht weist Schöpfer:innen eines Werkes ausschließliche Rechte zu.[1] Wenn Dritte das Werk verwerten wollen, also es verbreiten, öffentlich wiedergeben oder vervielfältigen möchten, bedarf es grundsätzlich deren Einwilligung oder einer vertraglichen Einräumung von Nutzungsrechten. Allerdings finden sich im Gesetz Ausnahmen von diesem Grundsatz: Die in §§ 44a ff. Urheberrechtsgesetz (UrhG) geregelten Schranken ermöglichen es, ein fremdes Werk unter bestimmten Voraussetzungen auch ohne Zustimmung der Rechteinhaber:innen zu nutzen.[2]

2. Die Text und Data Mining-Schranken

Für das KI-Training relevant sind dabei insbesondere die Schranken zum Text und Data Mining nach § 44b UrhG sowie die für Zwecke der wissenschaftlichen Forschung spezielle Text und Data Mining Schranken gemäß § 60d UrhG. Sie erlauben die Vervielfältigung urheberrechtlich geschützter Werke zum Zwecke des Text und Data Mining. § 60d UrhG enthält dabei besondere Privilegierungen für nicht kommerzielle Forschungseinrichtungen.[3] Unter Text und Data Mining (kurz: TDM) versteht man die automatisierte Analyse digitaler Daten, um daraus Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen. Beispielsweise werden Texte auf grammatikalische und semantische Regeln oder die Häufigkeit bestimmter Wörter untersucht. Häufig handelt es sich dabei um die Auswertung großer oder sogar „gigantischer" Datenmengen (sog. Big Data). Die TDM-Schranken finden aber auch auf die Analyse einzelner Werke Anwendung.[4]

II. (Wann) ist KI-Training Text und Data Mining?

Mit der Frage, wann das KI-Training unter die Text- und Data Mining Schranken fällt, befassten sich zuletzt das Landgericht (LG) München I in der Rechtssache GEMA/OpenAI und das

1 Vgl. zu den Voraussetzungen eines urheberrechtlich geschützten Werkes ausführlich Müller, Die Menge macht s, DFN-Infobrief Recht 11/2024.

2 Vgl. zu den Voraussetzungen eines urheberrechtlich geschützten Werkes ausführlich Müller, Die Menge macht s, DFN-Infobrief Recht 11/2024.

3 Vgl. zu den § 44b, § 60d UrhG und zum Text- and Data-Mining allgemein ausführlich Gielen, Die neue urheberrechtliche Schranke zum Text- und Data-Mining, DFN-Infobrief Recht 12/2019.

4 BT-Drs. 19/27426, S. 88.

Oberlandesgericht (OLG) Hamburg in der Rechtssache LAION. Ausgangspunkt des Rechtsstreits GEMA/OpenAI war die Feststellung, dass ChatGPT auf einfache Prompts der Nutzer:innen ganze Liedtexte fast Wort für Wort wiedergeben konnte. Tatsächlich neigen KI-Modelle dazu, sich bei ihrem Training nicht nur abstrakte Muster, sondern auch konkrete Fakten oder ganze Texte anzueignen (sog. Memorization, dt. Memorisierung). Die Verwertungsgesellschaft GEMA war der Ansicht, dass durch solche Outputs die Schöpfer:innen dieser Texte in ihrem Urheberrecht verletzt würden.[5]

Etwas anders gelagert war der Sachverhalt vor dem OLG Hamburg: Ein Fotograf wehrte sich gegen die Verwendung seiner Bilder in einem Datensatz, der online zum Training von KI-Modellen zur Verfügung gestellt wurde. Auch er war der Ansicht, diese Nutzung seiner Werke verletze sein Urheberrecht.

1. Übereinstimmungen

In ihren Urteilen gingen beide Gerichte davon aus, dass zunächst zwischen drei Phasen im „Leben" eines KI-Modells unterschieden werden muss: der Erstellung des Trainingsmaterials, dem eigentlichen Training des KI-Modells und der darauffolgenden Nutzung des (auf dem fertig trainierten Modell basierenden) KI-Systems.

In der ersten Phase, der Erstellung des Trainingsmaterials, werden die benötigten Daten kompiliert bzw. zusammengestellt. Die Daten kommen mithilfe von Web-Crawlern aus dem Internet oder sind in bereits bestehenden Datensätzen enthalten. Die/der Entwickler:in bereitet die Daten so auf, dass sie für das KI-Training genutzt werden können. Dabei erfolgen Vervielfältigungen von urheberrechtlich geschützten Daten, etwa durch deren Speicherung im Arbeitsspeicher. Da dies jedoch allein ihrer nachfolgenden Analyse und damit dem Text und Data Mining dient, waren sich die Gerichte einig, dass diese Phase von § 44b UrhG erfasst ist. Die Erstellung der Trainingsdaten stellt also keine Urheberrechtsverletzung dar.[6]

Auch gingen beide Gerichte davon aus, dass die dritte Phase, in der das auf dem fertig trainierten Modell aufbauende KI-System genutzt wird, also durch Prompts der Nutzer:innen Output generiert wird, *nicht* von den TDM-Schranken erfasst ist. In dieser Phase findet keine weitere Analyse der Trainingsdaten statt, mithin handelt es sich nicht um Text und Data Mining. Wird also, wie im Fall GEMA/OpenAI, durch den Output eines Large Language Models (LLMs) ein urheberrechtlich geschützter Text wiedergegeben, so stellt dies eine Urheberrechtsverletzung dar.[7]

2. Abweichungen

Die Unstimmigkeiten betreffen also die zweite Phase, in der das eigentliche Training stattfindet. Die zuvor aufbereiteten Daten werden nun im Modell ausgewertet. Dabei werden die enthaltenen Informationen auf abstrakte Zusammenhänge und Muster untersucht. So „lernt" das Modell, anhand dieser Korrelationen Outputs zu generieren (Generalization, dt. Generalisierung).[8] Ob auch diese Phase des KI-Trainings als Text und Data Mining i. S. d. § 44b UrhG einzuordnen ist, bewerteten die Gerichte unterschiedlich.

Das LG München I ging davon aus, dass das Training des KI-Modells urheberrechtlich unzulässig gewesen sei. Das Gericht begründete diese Entscheidung damit, dass zwar die Analyse abstrakter Daten beim Training von KI-Modellen, wie beispielsweise die Auswertung syntaktischer Regeln, Text und Data Mining i. S. d. § 44b UrhG darstelle. Bei einer vollständigen Memorisierung urheberrechtlich geschützter Werke im Modell, wie sie nachweislich im Trainingsprozess von ChatGPT erfolgt sei, handle es sich jedoch nicht um eine bloße Auswertung

5 Vgl. zum Vgl. zum Urteil des LG München I in der Rs. GEMA/OpenAI ausführlich Müller-Westphal, Was raus kommt, muss auch drin sein, DFN-Infobrief Recht 4/2026.

6 Ebenfalls entschied das LG Frankfurt a. M. in einem Urteil vom 17.12.2024 (Az: 2-06 O 401/25), eine Urheberrechtsverletzung liege auch dann vor, wenn ein urheberrechtlich geschützter Text mittels KI verändert und anschließend ohne Zustimmung des Urhebers/ der Urheberin veröffentlicht wird.

7 Etwas anderes kann allenfalls gelten, wenn zur Generierung eines Outputs neue, d. h. davor nicht in seinen Trainingsdaten vorhandene Informationen aus externen Quellen abgerufen werden (sog. Retrieval Augmented Generation) und diese dann im Rahmen des Groundings (Überprüfung der neuen Informationen mit verlässlichen Daten) analysiert werden. Auch dann ist jedoch wohl nur der Analysevorgang, nicht aber der Output selbst Text und Data Mining i. S. d. § 44b UrhG.

8 Vgl. zum Training von KI-Modellen ausführlicher Müller, Das kann sich doch niemand merken, DFN-Infobrief Recht 3/2025.

abstrakter Informationen.[9] Eine solche Nutzung urheberrechtlich geschützter Werke falle nicht unter § 44b UrhG und sei daher unzulässig. Dieses Ergebnis würde wohl dazu führen, dass das KI-Training immer dann unzulässig wäre, wenn eine Memorisierung urheberrechtlich geschützter Trainingsdaten durch das Modell nicht verhindert werde.

Zu einem anderen Ergebnis kam das OLG Hamburg in seinem Urteil vom 10. Dezember 2025.[10] Gegenstand des Rechtsstreits war zwar das Erstellen von KI-Trainingsdaten, also Phase 1, weshalb sich das Gericht nicht abschließend zur Zulässigkeit des Trainingsvorgangs äußern musste. Im Urteil finden sich dennoch Ausführungen, die auf eine von der Ansicht des LG München I abweichende Meinung schließen lassen. So ging das OLG Hamburg davon aus, dass das Training von generativen KI-Modellen grundsätzlich von den TDM-Schranken erfasst sei. Zur Begründung verweist es dabei zunächst auf den Willen des Gesetzgebers: Insbesondere der europäische Gesetzgeber habe durch Art. 4 der Richtlinie zum Urheberrecht im digitalen Binnenmarkt (DSM-RL), für dessen Umsetzung § 44b UrhG in Deutschland erlassen wurde, die Entwicklung von KI fördern wollen.[11] Das Gericht führte weiter aus, dass Urheber:innen die Möglichkeit hätten, rechtlich gegen den verletzenden Output vorzugehen oder durch einen Vorbehalt gem. § 44b UrhG die Nutzung ihres Werkes zumindest für das Training kommerzieller KI-Systeme von vornherein zu unterbinden. Daraus lässt sich schließen, dass das Gericht KI-Training selbst dann als Text und Data Mining einordnet, wenn es dabei zur Memorisierung urheberrechtlich geschützter Trainingsdaten kommt. Nach dieser Auffassung wäre das Training von KI-Modellen urheberrechtlich zulässig, sofern die weiteren Voraussetzungen des § 44b bzw. des § 60d UrhG vorliegen.[12]

III. Ausblick und Bedeutung für Hochschulen und Forschungseinrichtungen

Die unterschiedlichen Ergebnisse zeigen: Die weitere Entwicklung in dieser Rechtsfrage ist noch unklar, und es bedarf dringend einer höchstrichterlichen Klärung. Die Entwicklung der Rechtsprechung muss also im Auge behalten werden: Die Parteien im Rechtsstreit GEMA/OpenAI haben angekündigt, Revision einzulegen. Mit der Rechtssache GEMA/Suno[13] und der Rechtssache Like Company vs. Google Ireland[14] sind vor dem LG München I bzw. dem Europäischen Gerichtshof (EuGH) weitere Verfahren anhängig, in denen die Zulässigkeit des Trainings von KI-Modellen eine der entscheidenden Streitfragen darstellt. Hochschulen und Forschungseinrichtungen, die KI-Modelle entwickeln, sollten vorsichtig bleiben: Bis abschließend geklärt ist, ob das Training von KI-Modellen Text und Data Mining i. S. d. § 44b UrhG darstellt, ist eine Memorisierung urheberrechtlich geschützter Daten beim Trainingsprozess eines KI-Modells weiterhin zu vermeiden.

9 Vgl. LG München I, Endurteil vom 11.11.2025 - 42 O 14139/24 Rn. 250 ff.

10 OLG Hamburg, Urteil v. 10.12.2025, Az: 5 U 104/24 – Rs. LAION; vgl. ausführlicher zum erstinstanzlichen Urteil des LG Hamburg: Pesch, LTZ 1/2025, 65LTZ 1/2025, 65

11 Vgl. ErwGr. 18 Satz 1 und 4 DSM-RL.

12 Zu den zusätzlichen Voraussetzungen der § 44b und § 60d UrhG vgl. ausführlich Müller, Die Menge macht̀ s, DFN-Infobrief Recht 11/2024.

13 LG München I, Az: 42 O 763/25.

14 EuGH, Az: C-250/25.

Impressum

Herausgeber

Verein zur Förderung eines Deutschen Forschungsnetzes e. V.
DFN-Verein
Alexanderplatz 1, D-10178 Berlin
E-Mail: dfn-verein@dfn.de

Texte:

Forschungsstelle Recht im DFN
Ein Projekt des DFN-Vereins an der Humboldt-Universität Berlin.

Humboldt-Universität zu Berlin
Lehrstuhl für Bürgerliches Recht und Recht der Digitalisierung

Prof. Dr. Katharina de la Durantaye, LL. M. (Yale)
Unter den Linden 11, 10117 Berlin

Tel. (030) 838-66754

E-Mail: recht@dfn.de

Podcast der Forschungsstelle Recht im DFN

„Weggeforscht“, der Podcast der Forschungsstelle Recht im DFN, informiert knapp und verständlich über relevante juristische Entwicklungen und Fragestellungen im digitalen Umfeld. Neben einem kurzen Newsblock wird in jeder Folge ein aktuelles Thema erörtert.

Er erscheint regelmäßig ein- bis zweimal im Monat auf allen gängigen Podcast-Plattformen.

Link: https://anchor.fm/fsr-dfn