



Urheberrecht und KI-Training

IRIS

Eine Publikation
der Europäischen Audiovisuellen Informationsstelle

IRIS-2

Urheberrecht und KI-Training

Europäische Audiovisuelle Informationsstelle, Straßburg, 2026

ISSN 2079-1062

Verlagsleitung – Pauline Durand-Vialle, Geschäftsführende Direktorin

Redaktionelle Betreuung – Maja Cappello, Leiterin der Abteilung für juristische Informationen
Europäische Audiovisuelle Informationsstelle

Redaktionelles Team – Maja Cappello, Sophie Valais und Diego de la Varga
Europäische Audiovisuelle Informationsstelle

Verfasser

Diego de la Vega

Übersetzung

Stefan Pooth

Korrektur

Udo Lücke

Verlagsassistentz – Alexandra Ross

Presse und PR – Alison Hindhaugh, alison.hindhaugh@coe.int
Europäische Audiovisuelle Informationsstelle

Herausgeber

Europäische Audiovisuelle Informationsstelle
76, allée de la Robertsau, 67000 Straßburg, Frankreich
Tel.: +33 (0)3 90 21 60 00
iris.obs@coe.int
www.obs.coe.int

Titellayout – ALTRAN, Frankreich

Bitte zitieren Sie diese Publikation wie folgt:

De la Vega, D., *Urheberrecht und KI-Training*, IRIS, Europäische Audiovisuelle Informationsstelle, Straßburg, Juli 2026

© Europäische Audiovisuelle Informationsstelle (Europarat), Straßburg, 2026

Die in diesem Bericht enthaltenen Aussagen geben die Meinung der Verfasser wieder und stellen nicht unbedingt die Meinung der Europäischen Audiovisuellen Informationsstelle, ihrer Mitglieder oder des Europarats dar.

In diesem Dokument/Bericht verwenden wir zur besseren Lesbarkeit und Verständlichkeit geschlechtsspezifische Begriffe. Wo immer möglich, streben wir eine geschlechtsneutrale Formulierung an. Bitte beachten Sie, dass alle Bezeichnungen geschlechtsneutral zu verstehen sind und alle Geschlechter gleichermaßen einschließen.

Urheberrecht und KI-Training

Diego de la Vega

Vorwort

Mehr als drei Jahrhunderte ist es her, dass das *Statute of Anne* (Annes Gesetz) mit dem Versuch, einen Ausgleich zwischen den Rechten von Autoren und Herausgebern zu schaffen, den Grundstein für das moderne Urheberrecht legte. Der vollständige Titel dieses Gesetzes erklärt sowohl dessen Ziel („Förderung des Lernens“, also der Bildung) als auch das Mittel zur Einlösung dieses hohen Anspruchs („Übertragung der Rechte an Exemplaren gedruckter Bücher auf die Autoren oder Käufer dieser Exemplare während der darin genannten Zeiträume“). Im ersten Absatz wird dann ein Problem geschildert, das als neu dargestellt wird, in Wirklichkeit aber uralt ist: „[...] *Drucker, Buchhändler und andere Personen haben sich in letzter Zeit häufig die Freiheit genommen, [...] Bücher und andere Schriften ohne die Zustimmung der Autoren oder Eigentümer zu drucken, nachzudrucken und zu veröffentlichen [...], zu deren sehr großem Nachteil und allzu oft zum Ruin für sie und ihre Familien*“.

Während es im *Statute of Anne* von 1710 um gedruckte Bücher und unseriöse Verleger ging, richtete sich die von Pierre-Augustin Caron de Beaumarchais und anderen französischen Bühnenaufregern propagierte *Loi du 13 janvier 1791 relative aux spectacles* (Gesetz vom 13. Januar 1791 über Schauveranstaltungen) gegen die unbefugte öffentliche Aufführung von Theaterstücken. Und jede neue Ära brachte dann eigene technische Entwicklungen und Probleme, aber letztlich auch – praktische und rechtliche – Lösungen mit sich.

Heute befinden wir uns in der Ära der künstlichen Intelligenz (KI), einer revolutionären technologischen Entwicklung, die das Urheberrecht vor diverse neue, knifflige Probleme stellt. Probleme, die ironischerweise an den Wortlaut des *Statute of Anne* erinnern. Oder zumindest an ein bestimmtes Wort in dessen Titel:

Lernen.

Natürlich ging es im *Statute of Anne* darum, dass *Menschen* lernen sollten, und man war tief besorgt über rücksichtslose Profiteure, die *Bücher und andere Schriften* ohne *Zustimmung* verwerteten. Das mit dem Gesetz geschaffene Copyright förderte also das *Lernen*, indem es Autoren oder Eigentümern, die sonst ihres Eigentums beraubt worden wären, einen Lebensunterhalt sicherte.

Doch irgendwie haben die oben kursiv geschriebenen Wörter in der KI-Ära eine andere Bedeutung bekommen. Das *Lernen* erfolgt nun maschinell, und die Menschen dahinter bestehen darauf, dass dieses Lernen *ohne die Zustimmung der Autoren oder Eigentümer* stattfinden muss – sicherlich ohne viel Rücksicht auf den *sehr großen Nachteil*, den das für diese mit sich bringt.

Lernen. Zustimmung. Nachteil.

Der vorliegende Bericht erörtert diese drei Konzepte anhand einer Untersuchung der rechtlichen Fragen, die sich beim Training von KI-Systemen ergeben. Er geht hierzu der Frage nach, wie sich wichtige urheberrechtliche Grundsätze auf die „Inputphasen“ der KI anwenden lassen, und beleuchtet das aktuelle Zusammenspiel zwischen Urheberrecht und generativer und agentischer KI. Dies soll eine ausgewogene, fundierte Grundlage schaffen, die hilft, ein Thema zu verstehen, das die Urheberrechtspolitik prägt – nicht nur in Europa, sondern weltweit. Eine erste (und vielleicht naheliegende) Feststellung ist, dass der Umgang mit dem Thema Urheberrecht und KI nach wie vor stark davon abhängt, auf welche Seite des Atlantiks wir schauen.

Viel Vergnügen beim Lesen!

Maja Cappello

IRIS-Koordinatorin

Leiterin der Abteilung für juristische Informationen

Europäische Audiovisuelle Informationsstelle

Inhaltsverzeichnis

1. Inhaltsgenerierung mit KI im Kontext. Eine urheberrechtliche Perspektive	1
1.1. KI und Inhaltsgenerierung in Europa: aktueller Stand	1
1.2. KI-Funktionalitäten und der Prozess der Inhaltsgenerierung	4
1.3. Ein technologischer Ansatz für die Inhaltsgenerierung mit KI	6
1.4. Zusammenspiel zwischen Urheberrecht und KI	8

2. Die Rolle von Ausnahmen für Text- und Data-Mining beim KI-Training ..	11
2.1. KI-Regulierung: allgemeiner europäischer Rechtsrahmen für KI	11
2.2. Urheberrechtsrahmen	12
2.2.1. Europäischer Urheberrechtsrahmen und geltende Vorschriften für KI-Training	13
2.2.2. Text- und Data-Mining und Transparenzpflichten	17
2.2.3. Ausgewählte Beispiele für Ansätze zum Urheberrechtsschutz bei KI	20

3. Urheberrecht bei KI-Training und KI-Prompting	24
3.1. KI-Trainingsaktivitäten	24
3.2. KI-Prompting	26

4. Plattformen und Nutzer: die Nutzungsbedingungen	27
---	-----------

5. Wichtigste Erkenntnisse	30
---	-----------

Abbildungen

Abbildung 1.	Funktionsweise von agentischer KI	7
Abbildung 2.	Überblick über die grundlegenden Technologien von KI und GenAI	8
Abbildung 3.	Der technische Prozess des Trainierens von KI-Systemen	25

Tabellen

Tabelle 1.	Beim Training von KI-Systemen involvierte Urheberrechte	14
Tabelle 2.	Für den KI-Trainingsprozess relevante TDM-Ausnahmen	17
Tabelle 3.	Alternativen im Umgang mit dem KI-Training	21
Tabelle 4.	Europäische und nationale Ansätze im Umgang mit KI-Training	22
Tabelle 5.	Bestimmungen zum KI-Training in den Nutzungsbedingungen einiger der wichtigsten KI-Plattformen	27

1. Inhaltsgenerierung mit KI im Kontext. Eine urheberrechtliche Perspektive

1.1. KI und Inhaltsgenerierung in Europa: aktueller Stand

Seit OpenAI am 30. November 2022 die öffentliche Version von ChatGPT freigegeben hat,¹ hat der Einsatz von Tools mit künstlicher Intelligenz (KI) für private wie auch für berufliche Zwecke rasant zugenommen. Im Januar 2023, nur zwei Monate nach der Freigabe, hatte die Plattform bereits 100 Millionen Nutzerinnen und Nutzer.² Seitdem sind viele weitere KI-Plattformen mit einer enormen Funktionsvielfalt hinzugekommen.

Die Funktionsweise dieser Plattformen lässt sich, insbesondere im Zusammenhang mit Systemen der generativen KI (GenAI) und verwandten Technologien wie agentischer KI, in zwei Hauptphasen unterteilen. Die erste Phase umfasst das Training und den eigentlichen Einsatz der KI-Systeme, die zweite Phase der Erhalt des gewünschten Ergebnisses als Output, das ein konkretes Problem löst.

Um die urheberrechtlichen Implikationen dieser Technologien zu analysieren, befasst sich dieser Bericht mit der Inputphase generativer Prozesse, wie beispielsweise dem Training von KI-Systemen und der Eingabe von Anweisungen (Prompts) durch die Nutzer, um einen bestimmten Output zu erhalten. Das Training von KI-Systemen betrifft hauptsächlich die Anbieter dieser Systeme und das Prompten hauptsächlich die Nutzer, doch beides hat im aktuellen europäischen Rahmen erhebliche urheberrechtliche Implikationen.

In diesem Zusammenhang lassen sich KI-Anwendungen in Abhängigkeit davon, welche Ergebnisse die Nutzer anstreben, in verschiedene KI-Kategorien einordnen. So ging es bei der KI-Nutzung zuletzt vor allem um generative KI, doch wie Untersuchungen zeigen, verlagert sich der Marktfokus nun rasch auf andere Innovationen wie agentische KI, bei der KI-Agenten für die Nutzer autonom Aufgaben erledigen können. Andererseits steckt die reine GenAI zurzeit in einer Phase der Desillusionierung.³

¹ OpenAI-website, [ChatGPT ist da](#), 30. November 2022.

² Hu K., [ChatGPT sets record for fastest-growing user base - analyst note](#) (auf Englisch) [ChatGPT stellt Rekord für die am schnellsten wachsende Nutzerbasis auf – Analystenbericht], *Reuters*, 2. Februar 2023.

³ Khandabattu H., [Latest hype cycle for artificial intelligence goes beyond GenAI](#) (auf Englisch) [Neuester Hype-Zyklus rund um künstliche Intelligenz geht über GenAI hinaus], *Gartner*, 8. Juli 2025. Im Modell des Hype-Zyklus nach Gartner, der für alle aufkommenden Technologien gilt, befindet sich GenAI derzeit in der „Talsole der Desillusionierung“, in der „das Interesse nachlässt, weil Experimente und Umsetzungen die Erwartungen nicht

Dennoch ist gerade im Bereich GenAI schon heute eine umfangreiche Palette an Tools auf einem etablierten Markt in Europa verfügbar, wo diese Dienste in der Regel von großen Technologieunternehmen angeboten werden.⁴ Rund 20 % der europäischen Unternehmen und 55 % der Großunternehmen⁵ setzten im Jahr 2025 bereits KI (hauptsächlich GenAI) ein, was einem Anstieg von 6,5 % gegenüber 2024 entspricht. Bemerkenswert ist auch, dass etwa 10 % der europäischen Unternehmen, beispielsweise Produktionsfirmen, KI schon zur Erzeugung von Bildern, Videos und Audioinhalten nutzen.⁶ Allerdings zeichnen sich Unterschiede zwischen den einzelnen Regionen ab. Bei der geografischen Verteilung zeigen Statistiken, dass KI in Nordeuropa stärker genutzt wird, während die Länder im Süden hinterherhinken. So nutzen im Jahr 2026 beispielsweise 56,3 % der Schweden GenAI, während es bei den Italienern nur 19,9 % sind.⁷

Ereignisse wie die Einführung von ChatGPT im Jahr 2022 sind jedoch nur ein Meilenstein für die KI. Sicherlich ein wichtiger, aber eben nur einer von vielen auf einem Weg, der bereits in den 1950er Jahren begann.⁸ Die ständige Weiterentwicklung der KI führt dazu, dass den Nutzern von KI-Systemen laufend neue Funktionen angeboten werden, die ein sehr komplexes Marktumfeld widerspiegeln. Bei dieser Technologie, bei der das Interesse an ihrer Entwicklung in „Sommer“ und „Wintern“ gemessen wird,⁹ sind Autoren der Ansicht, dass die KI derzeit einen anhaltenden Sommer durchläuft, der gern als „Sommer der KI“ bezeichnet wird. Dies liegt an dem wachsenden, wenn auch etwas unbeständigen Interesse von Unternehmen und Nutzern an der Entwicklung der KI – im Gegensatz zu früher, als das Interesse an KI im Kontext der allgemeinen technologischen Entwicklung zweitrangig war.¹⁰

erfüllen. Anbieter der Technologie geben auf oder gehen unter. Investitionen fließen nur weiter, wenn die überlebenden Anbieter ihre Produkte so verbessern, dass die Early Adopter zufrieden sind“.

⁴ United Nations Conference on Trade and Development (UNCTA), [Technology and Innovation Report 2025](#) (auf Englisch) [Technologie- und Innovationsbericht 2025], Genf, April 2025.

⁵ Als Großunternehmen gelten in Europa Unternehmen mit 250 oder mehr Beschäftigten. Eurostat-website, [Statistics Explained](#) (auf Englisch) [Statistik erklärt].

⁶ Eurostat-Website, [Use of artificial intelligence in enterprises](#) (auf Englisch) [Einsatz künstlicher Intelligenz in Unternehmen], Dezember 2025.

⁷ Eurostat, [The use of artificial intelligence \(AI\) technologies in the European Union](#) (auf Englisch) [Der Einsatz von Technologien der künstlichen Intelligenz (KI) in der Europäischen Union], Key Results, Ausgabe 2026.

⁸ Die Grundlagen der KI finden sich in dem Werk „*Computing Machinery and Intelligence* / Können Maschinen denken?“ von Alan Turing. Turing, eine zentrale Figur der Digitaltechnologie, entwickelte auch den sogenannten Turing-Test, mit dem die Fähigkeit von Maschinen, menschliche Fragen zu beantworten, geprüft wurde – eine Vorwegnahme des Chatbot-Modells, das heute von vielen beliebten KI-Plattformen genutzt wird. Der Begriff „künstliche Intelligenz“ tauchte schon 1956 auf, und bereits in den 1960er Jahren wurden entscheidende Experimente durchgeführt. Siehe Blagoj D., Chrisa T., Uroš K., „[Historical Evolution of Artificial Intelligence](#)“ (auf Englisch) [Die historische Entwicklung der künstlichen Intelligenz], EUR 30221 EN, Publications Office of the European Union, Luxembourg, 2020.

⁹ Francesconi, E., [The winter, the summer and the summer dream of artificial intelligence in law](#). Address to the 18th International Conference on Artificial Intelligence and Law (auf Englisch) [Der Winter, der Sommer und der Sommertraum von künstlicher Intelligenz im Recht. Rede anlässlich der 18. Internationalen Konferenz über künstliche Intelligenz und Recht.], Artificial Intelligence and Law, Springer Nature, 3. Februar 2022.

¹⁰ Siehe Susskind, R. [The summer of AI](#) (auf Englisch) [Der Sommer der KI] in *How to think about AI: a guide for the perplexed* (auf Englisch) [Wie über KI denken: Ein Leitfaden für Verwirrte], Oxford Academic, 20. März 2025. Der Autor vertritt die Ansicht, dass Ereignisse wie die Entwicklung des World Wide Web in der Entwicklung der KI zu einem gewissen Stillstand geführt haben.

Gleichzeitig versuchen Gesetzgeber weltweit, einen geeigneten Rechtsrahmen zu schaffen, indem sie sich mit den verschiedenen rechtlichen Implikationen der KI auseinandersetzen. So wurden in den letzten Jahren beispielsweise das Rahmenübereinkommen des Europarats über künstliche Intelligenz¹¹ – der erste rechtsverbindliche internationale Vertrag zum Thema KI überhaupt – sowie einschlägige europäische Rechtsvorschriften wie die EU-Verordnung über künstliche Intelligenz (KI-Verordnung)¹² eingeführt. Parallel dazu sammelt die Europäische Kommission (EK) derzeit bis Juni 2026 Informationen, um die Wirksamkeit des bestehenden Rechtsrahmens zu bewerten und festzustellen, ob ergänzende Maßnahmen erforderlich sind.¹³ Genauer gesagt heißt es in der im Mai 2026 veröffentlichten Aufforderung zur Einreichung von Stellungnahmen zum Urheberrecht: „Die Vorbereitungsarbeiten für die gezielte Gesetzgebungsinitiative werden parallel zur Überprüfung der CDSM-Richtlinie laufen. Aus diesem Grund wird eine einzige Aufforderung zur Einreichung von Stellungnahmen veröffentlicht: Ziel ist es, die für die Überprüfung der CDSM-Richtlinie erforderlichen Informationen zu sammeln und Rückmeldungen zu den Herausforderungen, die insbesondere mit der Anwendung und Durchsetzung des Urheberrechts und verwandter Schutzrechte im Zusammenhang mit technologischen Entwicklungen verbunden sind, sowie zu möglichen Lösungen für diese Herausforderungen einzuholen.“¹⁴

Allgemeine Rechtsvorschriften sind zwar bereits in Kraft, doch über die Position des Gerichtshofs der Europäischen Union (GHEU) zum Thema GenAI und Urheberrecht, die die auch bereits geltenden europäischen Rechtsvorschriften ergänzen und auslegen soll, ist bisher nichts bekannt. Im Bereich des Urheberrechts gilt seine Position als wichtig, weil es keine spezifischen urheberrechtlichen Vorschriften für GenAI gibt. Hier fallen die urheberrechtlichen Regelungen immer noch in den Geltungsbereich des allgemeinen Urheberrechtsrahmens.

In diesem Zusammenhang wird davon ausgegangen, dass das erste Urteil zu dieser Materie auf EU-Ebene die Entscheidung des GHEU in der Rechtssache *Like Company/Google Ireland Limited*¹⁵ sein wird. Im Zusammenhang mit dem Schutz von

¹¹ Das [Rahmenübereinkommen](#) des Europarats über künstliche Intelligenz wurde am 5. September 2024 zur Unterzeichnung aufgelegt und soll Rechtslücken schließen, die sich aus Sicht von Menschenrechten, Demokratie und Rechtsstaatlichkeit aus dem raschen technologischen Fortschritt ergeben könnten. Es ist technologieneutral und regelt nicht die Technologie selbst, damit es nicht gleich wieder von den fortlaufenden KI-Entwicklungen überholt wird.

¹² [Verordnung \(EU\) 2024/1689](#) des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz – KI-Verordnung).

¹³ [Richtlinie \(EU\) 2019/790](#) des Europäischen Parlaments und des Rates vom 17. April 2019 über das Urheberrecht und die verwandten Schutzrechte im digitalen Binnenmarkt und zur Änderung der Richtlinien 96/9/EG und 2001/29/EG.

¹⁴ Europäische Kommission, [Aufruf zur Einreichung von Stellungnahmen](#), Bericht über die Überprüfung der Richtlinie über das Urheberrecht im digitalen Binnenmarkt / Gezielte Initiative für ein besseres urheberrechtliches Umfeld für europäische Kreativität und Innovation, Abschnitt A, 13. Mai 2026.

¹⁵ Europäische Union, [Rechtssache C-250/25](#), *Like Company*, Vorabentscheidungsersuchen. 3. April 2025 – *Like Company g. Google Ireland Limited*, *EUR-Lex*, 16 Juni 2025. Im Vorabentscheidungsersuchen des *Budapest Könyvtári Társulás* (Ungarn) geht es unter anderem um Orientierung zu Themen wie der Einstufung der Antworten eines auf LLM basierenden KI-Chatbots als öffentliche Wiedergabe, der Einstufung des

Presseveröffentlichungen im Hinblick auf die Online-Nutzung,¹⁶ wird dieses Urteil vermutlich einige entscheidende Aspekte der bestehenden europäischen Urheberrechtsvorschriften in Bezug auf KI klären. Zwar steht ein übergreifendes Urteil zur Harmonisierung der bestehenden Vorschriften noch aus, doch können andere Urteile, zuletzt etwa in der Rechtssache *Meta Platforms Ireland*¹⁷, indirekt eine gewisse Orientierung zu entscheidenden Aspekten von KI und Urheberrecht bieten – beispielsweise zur Rolle der ausschließlichen Urheberrechte im Rahmen der CDSM-Richtlinie, die das Recht auf Vergütung für Presseverlage nicht ersetzen können. Dies dürfte auch den verschiedenen innerstaatlichen Gerichten als Orientierung dienen, die bereits Urteile in Fällen sprechen, in denen es um KI und Urheberrecht geht.

1.2. KI-Funktionalitäten und der Prozess der Inhaltsgenerierung

Wie die KI-Verordnung einräumt, bezeichnet KI eine Reihe von Technologien, die sich rasant entwickeln. Der Verordnung zufolge können diese Technologien zu vielfältigem Nutzen für Wirtschaft, Umwelt und Gesellschaft über das gesamte Spektrum industrieller und gesellschaftlicher Tätigkeiten hinweg beitragen. Ihr Training erfordert die Analyse großer Datensätze, die dem Urheberrecht unterliegen können.

Insbesondere können diese Technologien zur Erstellung von Inhalten, zur Verbesserung von Vorhersageprozessen sowie zur Optimierung aller Arten von Abläufen des täglichen Lebens und der Ressourcenallokation beitragen sowie digitale Lösungen für Einzelpersonen und Organisationen personalisieren.¹⁸ Diese Funktionalitäten werden entweder eigenständig oder als Bestandteil eines bestimmten digitalen Produkts angeboten.¹⁹

Trainingsprozesses eines auf LLM basierenden Chatbots als Vervielfältigung, der allgemeinen Rolle von Ausnahmen für Text- und Data-Mining im Zusammenhang mit KI sowie bestimmten Prompting-Techniken.

¹⁶ Art. 15 CDSM-Richtlinie.

¹⁷ Gerichtshof der Europäischen Union, *Meta Platforms Ireland Ltd v. Autorità per le Garanzie nelle Comunicazioni* [GC], Rechtssache C-797/23, 12. Mai 2026. Hier stellt die Große Kammer in Randnummer 62 fest, dass „die Mitgliedstaaten Artikel 15 der Richtlinie 2019/790 nicht umsetzen (können), indem sie die darin vorgesehenen ausschließlichen Rechte vorbeugender Art durch einen bloßen Anspruch auf Entschädigung ersetzen, der es den Presseverlagen lediglich ermöglichen würde, eine Vergütung für die Online-Nutzung der Presseveröffentlichungen durch Anbieter von Diensten der Informationsgesellschaft zu erhalten, nicht aber, diese Nutzung zu verbieten.“

¹⁸ Erwägungsgrund 6 KI-Verordnung. Zu diesen Inhalten können beispielsweise auch Deepfakes im Zusammenhang mit Persönlichkeits- und Bildrechten gehören, definiert als 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. Weitere Auswirkungen können beispielsweise die Meinungsfreiheit betreffen. Siehe die [Guidance Note on the implications of Generative AI for freedom of expression](#) des Europarats (auf Englisch) [Leitfaden zu den Auswirkungen von generativer KI auf die Meinungsfreiheit], Dezember 2025.

¹⁹ So sind beispielsweise die [GenAI- und Agentic-AI-Systeme von Adobe](#) „in die Kreativ- und Marketinganwendungen von Adobe eingebettet“, [Microsofts Copilot](#) ist in das „Microsoft-365-Ökosystem“ integriert, während einige der gängigsten KI-Tools eigenständige Chatbots wie [ChatGPT](#) oder [Claude](#) sind.

Bei GenAI lassen sich in Abhängigkeit von dem Problem, das damit gelöst wird, unterschiedliche Funktionen unterscheiden. Zu den gängigsten Funktionen gehören virtuelle GenAI-Assistenten, die sich zur Generierung qualitativ hochwertiger Outputs auf große Sprachmodelle (*Large Language Models* – LLMs) stützen. Danach folgen andere KI-Systemmodelle wie die multimodale GenAI, die in der Lage ist, eine Mischung aus Inputs wie Bildern, Videos oder Audio zu verarbeiten und komplexe Outputs wie Text-zu-Bild- oder Skript-zu-Video-Inhalte zu generieren.²⁰

Die neuere agentische KI lässt sich ebenfalls als Form der GenAI einordnen, bei der Systeme nicht nur Inhalte generieren, sondern auch ohne ganz expliziten Input autonome Entscheidungen treffen und damit verbundene konkrete Aktionen durchführen. KI-Agenten können komplexe Ziele erreichen und Arbeitsabläufe mit einem unterschiedlichen Grad an menschlicher Intervention planen und umsetzen. Der agentischen KI dürfte in den nächsten drei bis sechs Jahren eine bedeutende Entwicklung bevorstehen.

Sehr schnelle Fortschritte machen dürften aus dem weiten Kreis der sich entwickelnden GenAI-Technologien im Übrigen auch KI-Programmierassistenten, die mit großer Autonomie und einfachen Prompts Software-Code erstellen können (z. B. für die App- oder Videospielentwicklung).²¹

Wie sich die Auswirkungen der KI messen lassen, ist in jeder Branche verschieden. In manchen können Wirkung und Rolle des Urheberrechts entscheidend sein, in anderen dagegen eher gering.²² Wo GenAI eine wichtige Rolle spielen kann, etwa im audiovisuellen Sektor,²³ insbesondere bei rein kreativen Aufgaben,²⁴ oder bei der Entwicklung von Videospielen,²⁵ sind die Auswirkungen der KI meist heute schon groß. Besonders stark von KI betroffen sind in der Regel wissensintensive Branchen, in denen kognitive Aufgaben anfallen. Dort wird davon ausgegangen, dass GenAI und die Erstellung synthetischer Inhalte die Auswirkungen bereits bestehender, nicht auf GenAI basierender Anwendungen noch verstärken werden.²⁶

²⁰ Zimmermann, A. et al., [Emerging Tech Impact Radar: Generative AI](#) (auf Englisch) [Radar für die Auswirkungen neuer Technologien: Generative KI], *Gartner*, 14. Februar 2025.

²¹ Ebd.

²² Eurostat, [The use of artificial intelligence \(AI\) technologies in the European Union](#), Key Results, 2026 Edition (auf Englisch) [Der Einsatz von Technologien der künstlichen Intelligenz (KI) in der Europäischen Union. Wichtigste Ergebnisse. Ausgabe 2026]. So sind beispielsweise Branchen wie das Baugewerbe weniger von KI betroffen als andere wie die Kreativwirtschaft, in der KI – und insbesondere GenAI – eine Vielzahl von Aufgaben übernehmen kann.

²³ Für einen konkreten Überblick über KI im audiovisuellen Sektor siehe Cappello M. (Hrsg.), [KI im audiovisuellen Sektor: ein Kompass für die aktuelle Rechtslandschaft](#), IRIS, Europäische Audiovisuelle Informationsstelle, Straßburg, Oktober 2024.

²⁴ Siehe beispielsweise die Richtlinien von Netflix zum Einsatz generativer KI bei der Produktion von Inhalten („[Using Generative AI in Content Production](#)“), die sich an Partnerfirmen richten, die audiovisuelle Inhalte für Netflix produzieren. Beim Prompting empfehlen die Richtlinien, keine Inputs (z. B. Prompts oder Bilder) zu verwenden, die ohne entsprechende Genehmigung auf urheberrechtlich geschütztes Material oder auf Abbildungen von Personen des öffentlichen Lebens oder von Verstorbenen Bezug nehmen.

²⁵ Heinisch O., Jadot C., [Some Implications of the EU AI Act on Video Game Developers](#) (auf Englisch) [Einige Auswirkungen der KI-Verordnung der EU auf Entwickler von Videospielen], *Sheppard*, 28. Februar 2025.

²⁶ OECD, [Macroeconomic productivity gains from Artificial Intelligence in G7 economies](#) (auf Englisch) [Makroökonomische Produktivitätsgewinne durch künstliche Intelligenz in den G7-Volkswirtschaften], *Artificial Intelligence Papers*, Juni 2025, Nr. 41.

1.3. Ein technologischer Ansatz für die Inhaltsgenerierung mit KI

GenAI ist eine Unterart von KI. Gemessen an all den Aufgaben, die GenAI erledigen kann, lässt sie sich technisch als eine Art von KI definieren, die auf Prompts und der Verwendung natürlicher Sprache basiert und auf Deep Learning (DL) beruht – einer Unterart des maschinellen Lernens (ML), die in der Lage ist, mithilfe von umfangreichen Datensätzen und geeignetem Training komplexe Aufgaben zu bewältigen, um neue Inhalte zu erstellen.²⁷

Mit Blick auf die rechtlichen Aspekte hat das Europäische Parlament kürzlich eine Entschließung zum Thema Urheberrecht und generative künstliche Intelligenz (den sogenannten Voss-Bericht) verabschiedet. Darin formuliert es, „dass es sich bei generativer KI – im Gegensatz zu anderen KI-Systemen, die in erster Linie für Einordnungen oder Prognosen konzipiert sind – um eine Art der KI handelt, die neue Inhalte wie Texte, Bilder, Musik, Videos und Code generiert, und zwar auf der Grundlage von Training unter Verwendung sehr großer Datensätze, aus denen sie Muster und Strukturen erlernt“. GenAI-Systeme nutzen dazu Prognosen auf der Grundlage statistischer Modelle, können menschliche Kreativität nachahmen und stützen sich auf bereits bestehende Inhalte.²⁸

Die KI-Verordnung verweist lediglich allgemein darauf, dass sie eine flexible Erzeugung von Inhalten ermöglicht, etwa in Form von Text- Audio-, Bild- oder Videoinhalten, die leicht ein breites Spektrum unterschiedlicher Aufgaben umfassen können.²⁹

Andere Klassifizierungen der KI-Technologie, wie etwa agentische KI, sind vorerst eine rein wirtschaftliche und technische Frage. Einen rechtlichen Definitionsansatz³⁰ gibt es für sie in Europa nicht. Branchenintern werden KI-Agenten als KI-Systeme definiert, die über die Verarbeitung natürlicher Sprache hinaus ein breites Spektrum an Funktionen abdecken können, darunter Entscheidungsfindung, Problemlösung, Interaktion mit externen Umgebungen und Ausführung von Handlungen.³¹

Außerhalb Europas hat Australien KI-Agenten gerade als „softwarebasiertes System“ definiert, „das Inputs aus seiner Umgebung wahrnimmt, einen internen Zustand aufrechterhält oder aktualisiert und mithilfe von APIs oder anderen Tools Aktionen auswählt und ausführt, um im Rahmen zugewiesener Berechtigungen und Einschränkungen definierte Ziele zu erreichen. Es kann planen, bei entsprechender Berechtigung aus Rückmeldungen lernen und sein Verhalten selbst überwachen, um seine Leistung im Laufe

²⁷ Siehe Microsoft, [Generative KI im Vergleich zu anderen KI-Typen](#).

²⁸ [Urheberrecht und generative künstliche Intelligenz – Chancen und Herausforderungen](#). Entschließung des Europäischen Parlaments vom 10. März 2026 zu dem Urheberrecht und generativer künstlicher Intelligenz – Chancen und Herausforderungen (2025/2058(INI)) (Voss-Bericht).

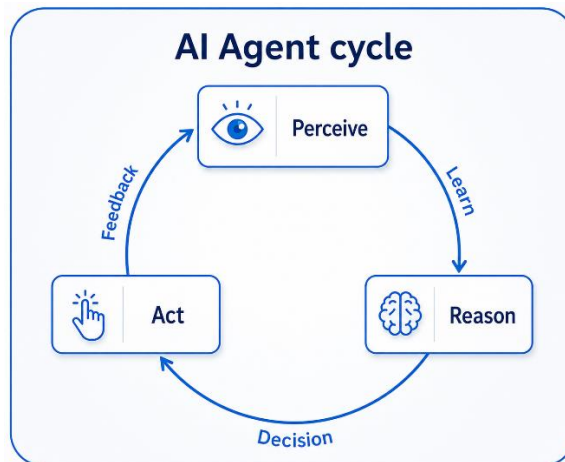
²⁹ Erwägungsgrund 99 KI-Verordnung.

³⁰ Aus informativer Sicht definiert der Europäische Datenschutzbeauftragte agentische KI als „Konzept der künstlichen Intelligenz (KI), das Systeme beschreibt, die mit begrenzten menschlichen Interaktionen (insbesondere ohne Schritt-für-Schritt-Anweisungen) autonom agieren, um nicht isolierte Aufgaben zu erledigen, sondern Ziele zu erreichen“. Goldstein A., [Agentic AI](#) (auf Englisch) [Agentische KI], European Data Protection Supervisor.

³¹ Gutowska A., [AI agents defined](#) (auf Englisch) [KI-Agenten definiert], *Think*, IBM.

der Zeit zu verbessern, wobei es innerhalb definierter Governance-, Aufsichts- und Kontrollmechanismen operiert“.³² KI-Agenten können letztlich auch Inhalte generieren.

Abbildung 1. Funktionsweise von agentischer KI



Quelle: Australische Regierung

Kurz gesagt ähneln die zugrunde liegenden technischen Konzepte denen der allgemeinen KI, bei der maschinelles Lernen eine grundlegende Rolle spielt. Maschinelles Lernen basiert auf dem Lernen aus bereits vorhandenen Datensätzen mithilfe von Algorithmen und dem Ableiten neuer Daten. Ein Teil davon ist das Deep Learning, der Branche zufolge eine „komplexe Reihe verschachtelter Gleichungen, die einen Input auf einen Output abbildet“.³³ Bei Foundation-Modellen wiederum handelt es sich um spezifische Deep-Learning-Modelle, die mit großen Mengen verfügbarer Daten trainiert werden,³⁴ wodurch LLMs in der Lage sind, dialogorientiert mit Menschen zu interagieren.³⁵

Gesetzliche Regelungen zur KI – ob aus urheberrechtlicher Sicht oder aus einer breiteren Perspektive – gehen davon aus, dass die technischen Ebenen der KI stets auf den oben genannten Prinzipien basieren. Allerdings könnten neu aufkommende Technologien wie die Quanten-KI³⁶ diese Annahme künftig in Frage stellen, weil das Training von KI-

³² Australian Government (Digital Transformation Agency), [Agentic AI addendum to the AI technical standard for Australian Government](#) (auf Englisch) [Nachtrag zu agentischer KI im technischen Standard KI für die australische Regierung], [digital.gov.au](#).

³³ Bergmann, D. [What is deep learning?](#) (auf Englisch) [Was ist Deep Learning?], *Think*, IBM

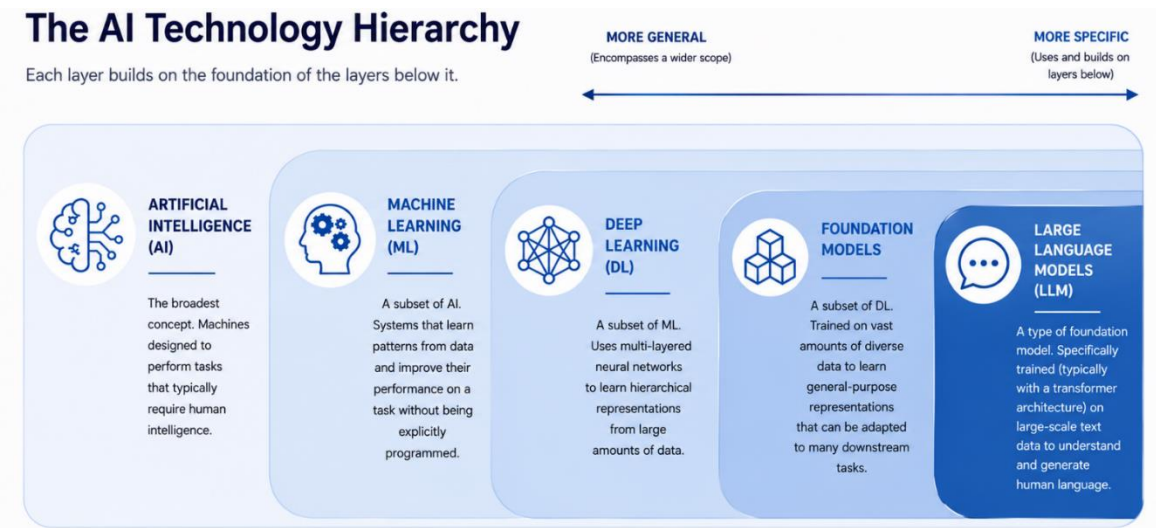
³⁴ Siehe Stanford Teaching Commons-Website, [Defining AI and chatbots](#) (auf Englisch) [Definition von KI und Chatbots], *Stanford University*. Siehe auch Bergmann, D., [What is deep learning?](#), op. cit., und Cappello M. (Hrsg.), [KI im audiovisuellen Sektor: ein Kompass für die aktuelle Rechtslandschaft](#), op. cit.

³⁵ Jurafsky D., Martin J. H., [Speech and Language Processing](#) (auf Englisch) [Sprachverarbeitung] (3. Aufl. Entwurf), *Stanford University*, Kapitel 7.

³⁶ Laut Amazon Web Services ist Quanten-KI die Nutzung von Quantentechnologien zum Betrieb von KI-Systemen – im Gegensatz zur klassischen KI, die auf herkömmlichem Computing basiert. Diese Technologie würde hypothetisch die bestehenden technischen Standards ablösen und die Funktionsweise von KI-Systemen und -Plattformen verbessern. Siehe die Amazon Web Services (aws)-Website, [Was ist Quanten-KI?](#).

Systemen ganz anders aussehen kann, wenn Quantencomputing und KI miteinander verschmelzen.³⁷

Abbildung 2. Überblick über die grundlegenden Technologien von KI und GenAI



Quelle: Europäische Audiovisuelle Informationsstelle

1.4. Zusammenspiel zwischen Urheberrecht und KI

Da GenAI Inhalte erstellen kann, ergeben sich schon im Rahmen des eigentlichen Generierungsprozesses urheberrechtliche Fragen. Einer der ersten Kritikpunkte bei der Nutzung von GenAI seit 2022 und dem Hereinbrechen von GenAI-Tools in den Markt war insbesondere die Verwendung bereits vorhandener Inhalte für das massive Training von KI-Systemen. So zeigte beispielsweise schon 2023 der Streik der Drehbuchautoren in den USA, dass der Einsatz von KI-Tools in der audiovisuellen Wirtschaft schnell zu urheberrechtlichen Streitigkeiten zwischen Produzenten und Urhebern führen dürfte.³⁸

Im Zusammenhang mit dem Europäischen Schutzschild für die Demokratie und dem Mediensektor haben die EU-Institutionen gemeinsam vor den Auswirkungen des KI-Trainings gewarnt. Sie loben zwar die Entwicklung der KI, mahnen aber auch: „Online-Piraterie und die unerlaubte Nutzung urheberrechtlich geschützten Materials zum Trainieren von KI-Modellen stellen nach wie vor eine erhebliche Bedrohung für den

³⁷ Acampora G., Chiatto A. et al. [Quantum artificial intelligence: a survey](#) (auf Englisch) [Künstliche Quantenintelligenz: Ein Überblick], *Computer Science Review* 59, Februar 2026.

³⁸ Cullins A., Kilkenny K, [As Writers Strike, AI Could Covertly Cross the Picket Line](#) (auf Englisch) [Wenn Autoren streiken, könnte KI heimlich die Streikpostenkette durchbrechen], *The Hollywood Reporter*, 3. Mai 2023.

Mediensektor dar, denn sie führen zu sinkenden Einnahmen und beeinträchtigen die Qualität und Vielfalt der Medien.“³⁹

Wie es im bereits erwähnten Voss-Bericht des Europäischen Parlaments heißt, stützt sich das Training von KI meistens auf bereits bestehende Inhalte, „zu denen auch urheberrechtlich geschütztes Material gehören könnte“. Der Bericht mahnt, dass „die Entwicklung, der Einsatz und die Nutzung von KI uneingeschränkt mit dem derzeitigen Rechtsrahmen im Einklang stehen müssen“. Er zeichnet ein weltweit düsteres Bild, insbesondere für die EU, wo es „Beweise für die weitverbreitete Verletzung der Urheberrechtsvorschriften durch Anbieter von generativer KI gibt, etwa das unbefugte Sammeln von Werken aus dem Internet, die Missachtung des Rechtsvorbehalts der Rechteinhaber hinsichtlich Text- und Data-Mining, das Erlangen von Werken über Raubkopien sowie das Versäumnis der Beantragung von Lizenzen“.⁴⁰

Auch die Parlamentarische Versammlung des Europarates (PACE) hat kürzlich wegen ähnlicher Bedenken gemahnt. Laut PACE hat „das Aufkommen des Zeitalters der künstlichen Intelligenz eine Reihe besonders schwieriger Probleme für den Kreativsektor mit sich gebracht, weil KI-Unternehmen auf der Grundlage gesetzlicher Bestimmungen, die weder eindeutig noch zweckmäßig sind, ohne vorherige Genehmigung und ohne Vergütung der Urheber von Inhalten das Internet durchforsten, um ihre datenhungrigen Systeme zu füttern“.⁴¹

Rechtlich gesehen erfordert diese Situation in der Tat die Anwendung des Urheberrechts. Das Training von KI-Systemen umfasst neben anderen Vorgängen und Techniken in der Regel umfangreiche Vervielfältigungen und letztlich die Entnahme von Inhalten aus Datenbanken.⁴² Es steht in engem Zusammenhang mit dem Text- und Data-Mining (TDM), mit dem es möglich ist, große Informationsmengen zu verarbeiten, um neues Wissen zu gewinnen und neue Trends zu entdecken⁴³ und dadurch an spezifische Informationen zu gelangen, die für das Training nützlich sind.⁴⁴ Unklar ist insbesondere noch, ob die Bereitstellung eines Inputs in Form eines vom Nutzer eingegebenen Prompts eine Vervielfältigung urheberrechtlich geschützten Materials darstellt, wenn er auf urheberrechtlich geschützten Inhalten – z. B. Presseveröffentlichungen – basiert, die beim Training eingespeist wurden.⁴⁵

³⁹ Europäische Kommission, [JOIN\(2025\) 791 final](#), Gemeinsame Mitteilung an das Europäische Parlament, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen. Europäischer Schutzschild für die Demokratie: Förderung starker und widerstandsfähiger Demokratien, Brüssel, 12 November 2025.

⁴⁰ Erwägungsgrund L des Voss-Berichts.

⁴¹ PACE, [Resolution 2654 \(2026\)](#), *Copyright enforcement in the artificial intelligence environment* (auf Englisch) [Durchsetzung des Urheberrecht im Umfeld von künstlicher Intelligenz], verabschiedet am 23. April 2026.

⁴² Einige Online-Plattformen bieten Nutzern die Möglichkeit, die Verwendung von bereits vorhandenem Material beim Training von KI-Systemen teilweise nachzuvollziehen. So bietet beispielsweise der [AI Watchdog](#) von The Atlantic eine „Untersuchung der Bücher, Videos und sonstigen Medien, die von den mächtigsten Technologieunternehmen der Welt zum Training ihrer KI-Modelle verwendet werden“.

⁴³ Erwägungsgrund 8 CDSM-Richtlinie.

⁴⁴ Rosati, E., [The future of the movie industry in the wake of generative AI: A perspective under EU and UK copyright law](#) (auf Englisch) [Die Zukunft der Filmindustrie angesichts von generativer KI: Eine Betrachtung aus Sicht des EU- und des britischen Urheberrechts], *Computer Law & Security Review* (Zeitschrift für Computerrecht und -sicherheit), Vol. 59 in *Science Review*, November 2025.

⁴⁵ Europäische Union, [Rechtssache C-250/25](#), op. cit.

Die Abwägung aller beteiligten Rechte und Interessen wurde kürzlich in der Folgenabschätzung der britischen Regierung zum Thema Urheberrecht und KI zusammengefasst. Darin heißt es, das Urheberrechts sei wichtig für den Erfolg der Kreativwirtschaft, weil es „Schöpfern ermöglicht, Anerkennung und finanziellen Gewinn aus ihrer Arbeit zu ziehen und gleichzeitig Investitionen für die Zukunft anzulocken“. Allerdings wird auch ausdrücklich eingeräumt, dass „Rechteinhaber von Schwierigkeiten berichten, die Nutzung ihrer Werke und den Zugriff auf sie zu kontrollieren. KI-Entwickler berichten von Problemen, auf große Datenmengen zuzugreifen, um Foundation-Modelle zu trainieren [...]“.⁴⁶

Die Urheberrechtsregulierung spielt daher eine fundamentale Rolle an der Schnittstelle zwischen den Interessen der Rechteinhaber und denen der KI-Entwickler sowie der Plattformindustrie, die sich um die Bereitstellung von KI-Diensten herum formiert hat. Generative und agentische KI, aber auch andere verwandte Technologien, stehen im Mittelpunkt des Interesses der Gesetzgeber, weil ihre Auswirkungen auf KI-Training und, in geringerem Maße, Prompting erhebliche Folgen haben können, nicht nur mit Blick auf die Rechtssicherheit, sondern auch wirtschaftlich und gesellschaftlich.

⁴⁶ HM Government, [Copyright and AI Impact Assessment](#), (auf Englisch) [Folgenabschätzung Urheberrecht und KI] dem Parlament gemäß § 135 des Gesetzes über Daten (Nutzung und Zugang) von 2025 vorgelegt. *Department for Science, Innovation and Technology and Department for Culture, Media and Sport*, März 2026.

2. Die Rolle von Ausnahmen für Text- und Data-Mining beim KI-Training

2.1. KI-Regulierung: allgemeiner europäischer Rechtsrahmen für KI

Das Sonder-Eurobarometer zur Digitalen Dekade 2026 zeigt, dass 80 % der Europäer eine umsichtige KI-Regulierung, die für Sicherheit sorgt, gegenüber einer schrankenlosen Entwicklung den Vorzug geben.⁴⁷ Diese Einstellung der Europäer steht im Einklang mit der Begründung der KI-Verordnung, in der es heißt: „Angesichts der großen Auswirkungen, die KI auf die Gesellschaft haben kann, und der Notwendigkeit, Vertrauen aufzubauen, ist es von entscheidender Bedeutung, dass KI und ihr Regelungsrahmen im Einklang mit den [...] Werten der Union [...] entwickelt werden.“⁴⁸ Auf einer übergeordneten Ebene geht es auch im Rahmenübereinkommen des Europarats über künstliche Intelligenz und Menschenrechte, Demokratie und Rechtsstaatlichkeit darum, „sicherzustellen, dass die Tätigkeiten im Lebenszyklus von Systemen künstlicher Intelligenz uneingeschränkt mit den Menschenrechten, der Demokratie und der Rechtsstaatlichkeit vereinbar sind“.⁴⁹

In der EU ist die KI-Verordnung als Teil einer globalen Strategie zu verstehen, die über die Regulierung hinausgeht und Innovationspläne sowie eine breitere Strategie auf politischer Ebene umfasst.⁵⁰

Die KI-Verordnung, deren vollständige Umsetzung noch nicht abgeschlossen ist,⁵¹ ist heute das grundlegende Instrument der KI-Regulierung in Europa. Ihr Schwerpunkt liegt nicht auf dem Urheberrecht, sondern auf dem Funktionieren des Binnenmarkts und dem Schutz von Gesundheit, Sicherheit und Grundrechten.⁵² Auf nationaler Ebene haben mehrere

⁴⁷ Europäische Kommission: Generaldirektion Kommunikation und Generaldirektion Kommunikationsnetze, Inhalte und Technologien), [Digital decade – Eurobarometer report – February – March 2026](#) (auf Englisch) [Digitale Dekade – Eurobarometer-Bericht – Februar – März 2026], *Directorate-General for Communications Networks, Content and Technology*, 2026. Das Ergebnis zeigt regionale Unterschiede. Während die Bürger in Griechenland, Luxemburg und Irland eine KI-Regulierung befürworten, ist die Unterstützung in Rumänien, Polen und Estland geringer, auch wenn sich dort immer noch mehr Menschen für eine Regulierung aussprechen als für einen freizügigeren Ansatz. Der Bericht zeigt außerdem, dass Frauen eine KI-Regulierung eher befürworten als Männer und dass jüngere Europäer einer Regulierung im Allgemeinen weniger positiv gegenüberstehen. Siehe Seite 60 f.

⁴⁸ Erwägungsgrund 6 KI-Verordnung.

⁴⁹ Europarat, [Rahmenübereinkommen des Europarats über künstliche Intelligenz und Menschenrechte, Demokratie und Rechtsstaatlichkeit](#), Artikel 1.

⁵⁰ Europäische Kommission-Website, [Europäischer Ansatz für künstliche Intelligenz](#).

⁵¹ Die KI-Verordnung gilt ab dem 2. August 2026, wobei einige Kapitel bereits seit 2025 in Kraft sind, während Artikel 6 Absatz 1 (Einstufungsvorschriften für Hochrisiko-KI-Systeme) erst ab dem 2. August 2027 gelten soll.

⁵² Artikel 1 KI-Verordnung.

Länder entweder bereits KI-Gesetze in Kraft gesetzt (Dänemark, Finnland, Italien und Slowenien) oder Gesetzentwürfe vorgelegt.⁵³

Beim Thema Urheberrecht verweist die KI-Verordnung – abgesehen von einigen spezifischen Teilen, die sich auf Allzweck-KI-Systeme beziehen, wie beispielsweise Artikel 53 – direkt auf die CDSM-Richtlinie. In den Erwägungsgründen 105 und 106 der KI-Verordnung wird ausdrücklich auf die CDSM-Richtlinie Bezug genommen, mit der „Ausnahmen und Beschränkungen eingeführt [wurden], um unter bestimmten Bedingungen Vervielfältigungen und Entnahmen von Werken oder sonstigen Schutzgegenständen für die Zwecke des Text- und Data-Mining zu erlauben“.⁵⁴ Die KI-Verordnung überträgt somit weitreichende Befugnisse in Fragen der KI-Regulierung zum Urheberrecht auf die CDSM-Richtlinie.

2.2. Urheberrechtsrahmen

Die CDSM-Richtlinie ist eines der jüngsten Elemente des komplexen europäischen Urheberrechtsrahmens.⁵⁵ Sie ergänzt eine Reihe bestehender Richtlinien, die die wichtigsten Bestimmungen enthalten, auf denen die nationalen Urheberrechtssysteme basieren, insbesondere die Richtlinie 2001/29/EG zur Harmonisierung bestimmter Aspekte des Urheberrechts und der verwandten Schutzrechte in der Informationsgesellschaft (InfoSoc-Richtlinie).⁵⁶ Diese hat schon 2001 einige der wichtigsten strukturellen Elemente eingeführt, auf die sich der europäische Urheberrechtsrahmen bis heute stützt, und die CDSM-Richtlinie baut darauf auf.

Die CDSM-Richtlinie ist das zentrale Element der urheberrechtlichen Regulierung im Zusammenhang mit KI, weil es keine spezifischen Vorschriften für Urheberrecht und KI in der KI-Verordnung gibt, doch ihre Begründung folgt der Linie aller vorangegangenen Richtlinien zum Urheberrecht.⁵⁷

⁵³ Bashir, E., [Status of national AI laws across Europe](#) (auf Englisch) [Stand der nationalen KI-Gesetze in Europa] [am 12. Juni 2026 aktualisiert], *Cullen International*, 12 Juni 2026.

⁵⁴ Erwägungsgrund 105 KI-Verordnung.

⁵⁵ Dieser allgemeine Rahmen gilt für alle Branchen. Im audiovisuellen Sektor kann er beispielsweise für assistiertes Drehbuchschreiben, KI-assistierte Videobearbeitung, Untertitelung und Synchronisation, gepromptete Videoerzeugung, Inhaltsempfehlungen und vieles mehr gelten.

⁵⁶ [Richtlinie 2001/29/EG](#) des Europäischen Parlaments und des Rates vom 22. Mai 2001 zur Harmonisierung bestimmter Aspekte des Urheberrechts und der verwandten Schutzrechte in der Informationsgesellschaft.

⁵⁷ Neben der InfoSoc-Richtlinie gehören dazu auch folgende Richtlinien (die Auswirkungen auf KI haben können): Richtlinie [96/9/EG](#) des Europäischen Parlaments und des Rates vom 11. März 1996 über den rechtlichen Schutz von Datenbanken, Richtlinie [2006/115/EC](#) zum Vermietrecht und Verleihrecht sowie zu bestimmten dem Urheberrecht verwandten Schutzrechten im Bereich des geistigen Eigentums sowie Richtlinie [2009/24/EC](#) über den Rechtsschutz von Computerprogrammen.

2.2.1. Europäischer Urheberrechtsrahmen und geltende Vorschriften für KI-Training

Die CDSM-Richtlinie stammt aus der Zeit vor dem Aufkommen von GenAI um 2022–2023⁵⁸ und erwähnt KI auch nicht aber stellt lediglich fest: „Mit neuen, im Allgemeinen als Text- und Data-Mining bekannten Verfahren können in digitaler Form vorliegende Informationen wie Texte, Töne, Bilder oder Daten mit Computern automatisch ausgewertet werden.“ Außerdem erwähnt sie die „Entwicklung neuer Anwendungen oder Technologien“.⁵⁹ Außerdem heißt es in Erwägungsgrund 3 der CDSM-Richtlinie: „Die einschlägigen Rechtsvorschriften müssen zukunftstauglich sein, damit die technologische Entwicklung nicht behindert wird“. Unmittelbar nach der Verabschiedung der CDSMD erklärte das Europäische Parlament, dass „sich darüber hinaus die Mitgesetzgeber darauf einigten, zusätzlich zu der von der Kommission in ihrem Vorschlag vorgesehenen zwingenden Ausnahme für Text- und Datenauswertung zu Forschungszwecken eine weitere zwingende Ausnahme für die allgemeine Text- und Datenauswertung (Artikel 4) im EU-Recht zu verankern, um zur Entwicklung der Datenanalyse und der künstlichen Intelligenz beizutragen, sowie auf eine Reihe neuer EU-Vorschriften (z. B. einen neuen Lizenzierungsmechanismus für nicht mehr im Handel erhältliche Werke).“⁶⁰

Die Europäische Kommission schrieb kürzlich in ihrer Aufforderung zur Stellungnahme zur Überprüfung der CDSM-Richtlinie: „Der bestehende EU-Rechtsrahmen, einschließlich der CDSM-Richtlinie, der urheberrechtlichen Bestimmungen der KI-Verordnung und der anschließenden Durchführungsmaßnahmen, trägt dazu bei, Rechteinhaber bei der Ausübung ihrer Rechte zu unterstützen. Sie reichen [sic] jedoch möglicherweise nicht aus, um die Schwierigkeiten anzugehen, mit denen Rechteinhaber bei der Kontrolle und Lizenzierung ihrer Werke für KI-Nutzungen konfrontiert sind.“⁶¹ Daher kündigten die EU-Institutionen bereits Ende 2025 an, die Kommission werde eine Überprüfung der CDSM-Richtlinie vornehmen und erwägen, wie ihre Wirksamkeit vor dem Hintergrund von GenAI verbessert werden kann. Insbesondere heißt es in der Gemeinsamen Mitteilung an das Europäische Parlament, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen:⁶² „Die Entwicklung generativer KI bringt zwar Chancen mit sich, stellt das Medien-Ökosystem jedoch auch vor zahlreiche Herausforderungen. Online-Piraterie und die unerlaubte Nutzung urheberrechtlich geschützten Materials zum Trainieren von KI-Modellen stellen nach wie vor eine erhebliche Bedrohung für den Mediensektor dar, denn sie führen zu sinkenden Einnahmen und beeinträchtigen die Qualität und Vielfalt der Medien. Die Kommission wird

⁵⁸ Der erste Vorschlag der Europäischen Kommission stammt vom 14. September 2016: [COM\(2016\) 593 final](#). Auch in vorbereitenden Studien, die für die CDSM-Richtlinie herangezogen wurden, ist von KI keine Rede. Siehe Triaille J-P., de Meeûs, J., de Francquen A., [Study on the legal framework of text and data mining \(TDM\)](#) (auf Englisch) [Studie zum rechtlichen Rahmen von Text- und Data-Mining (TDM)], *Amt für Veröffentlichungen der Europäischen Union*, März 2014.

⁵⁹ Erwägungsgründe 8 und 18 CDSM-Richtlinie.

⁶⁰ Europäisches Parlament, [Legislative Train Schedule](#), *Modernisation of European copyright rules: directive on copyright in the digital single market* (Informationen aktualisiert am 20 November 2011).

⁶¹ Europäische Kommission, [Aufforderung zur Stellungnahme](#), Gezielte Initiative für ein besseres urheberrechtliches Umfeld für Kreativität und Innovation in Europa.

⁶² Europäische Kommission, [JOIN\(2025\) 791 final](#), op. cit.

eine Überprüfung der Richtlinie über Urheberrechte im digitalen Binnenmarkt vornehmen und erwägen, wie ihre Wirksamkeit vor dem Hintergrund dieser Entwicklungen verbessert werden kann.“

Damit das Urheberrecht anwendbar ist, müssen bereits bestehende Werke in den Geltungsbereich der nationalen Urheberrechtsgesetze fallen. Einige könnten also mit Blick auf das KI-Training irrelevant sein. So sind etwa Werke, die gemeinfrei sind (auch wenn Urheberpersönlichkeitsrechte bestehen bleiben können), oder Inhalte, die im Sinne des Urheberrechts nicht als Werk gelten, nicht vollständig geschützt.⁶³ In diesem Sinne gelten die allgemeinen urheberrechtlichen Bestimmungen zur Schutzfähigkeit von Werken.

2.2.1.1. Betroffene Rechte

Entsprechend den von Anbietern von KI-Modellen im Trainingsprozess angewandten Verfahren, die das Urheberrecht an bereits bestehenden Werken berühren und somit das KI-Training oder Teile davon verhindern können, lassen sich folgende Rechte identifizieren:⁶⁴

Tabelle 1. Beim Training von KI-Systemen involvierte Urheberrechte

Bereich		Zusammenhang mit dem KI-Training
	Vervielfältigungsrecht	Das Vervielfältigungsrecht ist das ausschließliche Recht der Rechteinhaber, die unmittelbare oder mittelbare, vorübergehende oder dauerhafte Vervielfältigung von Werken, Tonträgern und Darbietungen auf jede Art und Weise und in jeder Form ganz oder teilweise zu erlauben oder zu verbieten. ⁶⁵
	Datenbankrechte	Datenbanken genießen zwei Arten von Schutz: - <u>Urheberrechtlicher Schutz</u>: Urheber einer Datenbank haben das ausschließliche Recht, die Vervielfältigung dieser Datenbank zu genehmigen, unabhängig davon, ob diese Vervielfältigung vorübergehend oder dauerhaft ist.⁶⁶ - <u>Sui-generis-Recht</u>: Urheber einer Datenbank haben das Recht, die Entnahme und/oder die Weiterverwendung (der Gesamtheit oder eines in qualitativer oder

⁶³ Wie in Erwägungsgrund 53 der CDSM-Richtlinie erläutert, hat der Ablauf der Schutzdauer eines Werks zur Folge, dass das Werk gemeinfrei wird und die Rechte erlöschen, die das EU-Urheberrecht in Bezug auf dieses Werk gewährt.

⁶⁴ European Union Intellectual Property Office (EUIPO), [The development of Generative Artificial Intelligence from a Copyright Perspective](#), op. cit., S. 35.

⁶⁵ Artikel 2 [InfoSoc-Richtlinie](#).

⁶⁶ Artikel 3 Richtlinie [96/9/EC](#).

Bereich		Zusammenhang mit dem KI-Training
		quantitativer Hinsicht wesentlichen Teils) des Inhalts dieser Datenbank zu untersagen. ⁶⁷
	Rechte an Computerprogrammen	Aufbauend auf dem allgemeinen Urheberrechtsschutz gemäß der InfoSoc-Richtlinie besteht ein spezifisches Recht auf dauerhafte oder vorübergehende Vervielfältigung sowie auf Umarbeitung, das Übersetzungen, Bearbeitungen und Arrangements umfasst. ⁶⁸
	Rechte von Presseverlagen	Die CDSM-Richtlinie sieht einen spezifischen Schutz in Form eines verwandten Schutzrechts gegen bestimmte Online-Nutzungen durch Anbieter von Diensten der Informationsgesellschaft vor. Nach der jüngsten Auslegung des GHEU ist es den nationalen Gesetzgebern unter anderem gestattet, für Presseverlage einen Anspruch auf eine angemessene Vergütung als Gegenleistung für die den Anbietern von Diensten der Informationsgesellschaft erteilte Erlaubnis zur Nutzung ihrer Veröffentlichungen vorzusehen, sofern die entsprechende Regelung - den Presseverlagen nicht die Möglichkeit nimmt, eine solche Erlaubnis zu verweigern oder unentgeltlich zu erteilen, - die Anbieter von Diensten der Informationsgesellschaft nicht zu einer Zahlung verpflichtet, die in keinem Zusammenhang mit der Nutzung solcher Presseveröffentlichungen steht, - und die Verpflichtungen und etwaigen Sanktionen, die diesen Anbietern auferlegt werden, den Grundsatz der Verhältnismäßigkeit wahren.⁶⁹
	Technische Schutzmaßnahmen	Die InfoSoc-Richtlinie verlangt, dass die nationalen Rechtsvorschriften der EU-Mitgliedstaaten einen angemessenen rechtlichen Schutz gegen die Umgehung wirksamer technischer Maßnahmen bieten, die von Rechteinhabern eingesetzt werden, sofern sie die Begünstigten bestimmter Beschränkungen und Ausnahmen des Urheberrechts und der verwandten Schutzrechte nicht daran hindern, diese in Anspruch zu nehmen.

Quelle: Europäische Audiovisuelle Informationsstelle und EUIPO⁷⁰

⁶⁷ Ebd., Artikel 7.

⁶⁸ Siehe Richtlinie [2009/24/EC](#).

⁶⁹ Siehe [Rechtssache C-797/23](#), op. cit.

⁷⁰ EUIPO, [The development of Generative Artificial Intelligence from a Copyright Perspective](#), op. cit.

2.2.1.2. Relevante Ausnahmen

Dem Urheberrecht in Europa stehen jedoch eine Reihe von Ausnahmen gegenüber. Dieses Konzept, das insbesondere in der InfoSoc-Richtlinie verankert ist, soll ein ausgewogenes Verhältnis zwischen den Rechten und Interessen der verschiedenen Kategorien von Rechteinhabern im urheberrechtlichen Umfeld gewährleisten. Im Einklang mit dieser Argumentation, hat die CDSM-Richtlinie neue Ausnahmen und Einschränkungen für Text- und Data-Mining, Lehrzwecke und die Erhaltung des kulturellen Erbes eingeführt.⁷¹

Ausgehend von der Annahme, dass bei bestimmten Organisationen und Institutionen eine Rechtsunsicherheit hinsichtlich der Frage besteht, inwieweit sie bei Inhalten TDM einsetzen dürfen, hat die CDSM-Richtlinie die TDM-Ausnahme eingeführt, auch um sicherzustellen, dass diese Technik rechtmäßig ist, wenn sie außerhalb des Rahmens der wissenschaftlichen Forschung angewendet wird.

Daher führt die CDSM-Richtlinie für TDM eine Ausnahme für den Fall ein, dass der Einsatz von TDM-Techniken von privaten wie auch von öffentlichen Einrichtungen zur Analyse großer Datenmengen in verschiedenen Lebensbereichen und für verschiedene allgemeine Zwecke erfolgt. Diese Ausnahme oder Beschränkung gilt aber nur, wenn der Begünstigte rechtmäßigen Zugang zu dem Werk oder sonstigen Schutzgegenstand hat, wozu auch gehört, dass es bzw. er der Öffentlichkeit im Internet zugänglich gemacht wurde, und soweit die Rechteinhaber sich nicht in angemessener Weise das Recht, Vervielfältigungen und Entnahmen zum Zwecke des Text- und Data-Mining anzufertigen, vorbehalten haben.⁷²

Mit anderen Worten: Die CDSM-Richtlinie führt die „Opt-out-Regelung“ ein, die sich rasch zu einem der zentralen Themen im Bereich des KI-Trainings entwickelt hat. Einige Stimmen stellen die uneingeschränkte Anwendbarkeit auf das Training von KI-Systemen in Frage.⁷³

Auch sieben Jahre nach ihrer Einführung und vier Jahre nach der Popularisierung von GenAI in der breiten Öffentlichkeit sind diese Fragen noch offen.⁷⁴

In diesem Zusammenhang wird TDM in den Artikeln 3 und 4 der CDSM-Richtlinie behandelt. Dabei steht Artikel 4 im Mittelpunkt der politischen Diskussion, da er eine

⁷¹ Erwägungsgrund 5 CDSM-Richtlinie.

⁷² Erwägungsgrund 18 CDSM-Richtlinie.

⁷³ Dornis T., [The Training of Generative AI is not Text and Data Mining](#) (auf Englisch) [Das Training von generativer KI ist kein Text- und Data-Mining], European Intellectual Property Review (E.I.P.R. – Europäische Zeitschrift für geistiges Eigentum), 19. Oktober 2024. Der Argumentation des Verfassers zufolge liegt der Schwerpunkt eher auf Urheberrechtsverletzungen, weil das Training von Modellen der generativen KI ohne Lizenzen für die als Trainingsdaten verwendeten Werke eigentlich nicht in den Anwendungsbereich der europäischen TDM-Ausnahme fällt. Siehe auch: [Urheberrecht und generative künstliche Intelligenz – Chancen und Herausforderungen](#), Entschließung des Europäischen Parlaments vom 10. März 2026 zu Urheberrecht und generativer künstlicher Intelligenz – Chancen und Herausforderungen (2025/2058(INI)), Absatz K.

⁷⁴ Keller P. (IViR), [A Blind Spot at the Heart of EU Copyright and AI Policymaking?](#) (auf Englisch) [Ein blinder Fleck im Herzen der EU-Politik zu Urheberrecht und KI?], *Kluwer Copyright Blog*, 27. Mai 2026. Der Verfasser ist der Ansicht, dass die Opt-out-Regelung nicht nur mit Unsicherheiten hinsichtlich ihrer Umsetzung behaftet ist, sondern auch nicht die erwarteten Ergebnisse liefert, weil „die Anreize für die massenhafte Nutzung öffentlich zugänglicher Informationen und Inhalte von Dynamiken geprägt sind, die weit über den Geltungsbereich des EU-Urheberrechtsrahmens hinausreichen“.

weitreichende Ausnahme vorsieht, gleichzeitig eine Opt-out-Regelung einführt und die Rechteinhaber verpflichtet, das Opt-out im Falle von online öffentlich zugänglichen Inhalten auf maschinenlesbare Weise auszuüben.

Tabelle 2. Für den KI-Trainingsprozess relevante TDM-Ausnahmen

Artikel der CDSM-Richtlinie		Ausnahme
	3	Ausnahme für Vervielfältigungen und Entnahmen, die durch Forschungsorganisationen und Einrichtungen des Kulturerbes von Werken oder sonstigen Schutzgegenständen, zu denen sie rechtmäßig Zugang haben, zum Zwecke der wissenschaftlichen Forschung für Text- und Data-Mining vorgenommen werden.
	4	Ausnahme oder Beschränkung für zum Zwecke des Text- und Data-Mining vorgenommene Vervielfältigungen und Entnahmen von rechtmäßig zugänglichen Werken und sonstigen Schutzgegenständen. - Die Ausnahme bzw. Beschränkung findet Anwendung, sofern die jeweiligen Rechteinhaber die in diesem Absatz genannten Werke und sonstigen Schutzgegenstände nicht ausdrücklich in angemessener Weise, etwa mit maschinenlesbaren Mitteln im Fall von online veröffentlichten Inhalten, mit einem Nutzungsvorbehalt versehen haben [Opt-out-Regelung].

Quelle: Europäische Audiovisuelle Informationsstelle

2.2.2. Text- und Data-Mining und Transparenzpflichten

Text- und Data-Mining ist in der CDSM-Richtlinie definiert als Technik für die automatisierte Analyse von Texten und Daten in digitaler Form, mit deren Hilfe Informationen unter anderem – aber nicht ausschließlich – über Muster, Trends und Korrelationen gewonnen werden können.⁷⁵

Im Hinblick auf das Urheberrecht kann das Text- und Data-Mining Handlungen umfassen, die durch das Urheberrecht, das Sui-generis-Recht an Datenbanken oder mitunter beides geschützt sind. Dies kann insbesondere bei der Vervielfältigung von Werken (oder auch anderen Schutzgegenständen) der Fall sein, beispielsweise wenn die Entnahme von Inhalten erfolgt, wenn die Daten während des Text- und Data-Minings normalisiert⁷⁶ werden.⁷⁷ Dies führt zur automatischen computergestützten Auswertung von in digitaler Form vorliegenden Informationen wie Texten, Tönen, Bildern oder Daten mit Computern, die im Allgemeinen in den Anwendungsbereich des TDM fällt.

⁷⁵ Artikel 2 Absatz 2 CDSM-Richtlinie.

⁷⁶ Organisiert, bereinigt und standardisiert.

⁷⁷ Erwägungsgrund 8 CDSM-Richtlinie.

Der Voss-Bericht stellt fest, dass die derzeitigen Systeme zur Erklärung eines Rechtsvorbehalts bei urheberrechtlich geschützten Inhalten auch sechs Jahre nach der Einführung der CDSM-Richtlinie (laut dieser „sollte es nur als angemessen erachtet werden, einen Rechtsvorbehalt mit maschinenlesbaren Mitteln auszusprechen“⁷⁸) „oft in der Praxis schwierig sind, möglicherweise nicht alle relevanten Handlungen des Text- und Data-Minings abdecken und ihnen die für eine wirksame Umsetzung und Durchsetzung erforderliche Transparenz fehlt“. Zudem räumt der Bericht ein, dass es Beweise für die weitverbreitete Verletzung der Urheberrechtsvorschriften durch Anbieter von generativer KI gibt, etwa die Missachtung des Rechtsvorbehalts der Rechteinhaber hinsichtlich Text- und Data-Mining.⁷⁹ Die Parlamentarische Versammlung des Europarats betont außerdem, dass „die sogenannten ‚Ausnahmen für Text- und Data-Mining (TDM)‘, die vor dem Aufkommen der generativen KI verabschiedet wurden, den Urheberrechtsinhabern die Last auferlegen, sich gegen diese Ausnahme zu entscheiden, und keine Vergütungsbestimmungen enthalten“.⁸⁰ Dennoch ermöglicht die Opt-out-Regelung in Verbindung mit den ausschließlichen Rechten die Existenz von nachgelagerten Lizenzierungs- und Vergütungsmodellen.

Unterdessen veröffentlichte die Europäische Kommission im Jahr 2025 den Allgemeinen KI-Verhaltenskodex,⁸¹ der von einigen der bedeutendsten KI-Entwickler oder -Anbieter (darunter Amazon, Anthropic, Google, IBM, Microsoft und Mistral AI) unterzeichnet wurde. Die Unterzeichner erkannten darin die Ausnahme oder Beschränkung für Text- und Data-Mining an, „die unter der Voraussetzung gelten soll, dass der Zugang rechtmäßig erfolgt und die jeweiligen Rechteinhaber die in diesem Absatz genannten Werke und sonstigen Schutzgegenstände nicht ausdrücklich in angemessener Weise [...] mit einem Nutzungsvorbehalt versehen haben“.⁸²

In diesem Kodex verpflichten sich die Unterzeichner, sicherzustellen, dass nur rechtmäßig zugängliche Werke und sonstige Schutzgegenstände vervielfältigt und entnommen werden. Zu diesem Zweck sind sie zwei Verpflichtungen eingegangen, die den Einsatz von Web-Crawlern zum Scraping oder Zusammenstellen von Daten zum Zwecke des Text- und Data-Minings regeln:

- **Keine Umgehung wirksamer technischer Maßnahmen** im Sinne der InfoSoc-Richtlinie, die dazu dienen, unbefugte Handlungen in Bezug auf Werke und sonstige Schutzgegenstände zu verhindern oder einzuschränken, insbesondere durch die Beachtung jeglicher technischer Zugangsverweigerungen oder -beschränkungen aufgrund von Abonnementmodellen oder Bezahlschranken.
- **Ausschluss des Web-Crawlings auf Websites**, die Inhalte der Öffentlichkeit zugänglich machen und zum Zeitpunkt des Web-Crawlings von Gerichten oder Behörden in der EU und im Europäischen Wirtschaftsraum als Websites ausgewiesen sind, die dauerhaft und wiederholt in gewerbsmäßigem Umfang gegen Urheberrechte und verwandte Schutzrechte verstoßen.

⁷⁸ Erwägungsgrund 18 CDSM-Richtlinie.

⁷⁹ Abschnitt K des Voss-Berichts.

⁸⁰ Erwägungsgrund 6, [Resolution 2654 \(2026\)](#), op. cit.

⁸¹ Europäische Kommission, [Allgemeiner KI-Verhaltenskodex jetzt verfügbar](#), Pressemitteilung, 10. Juli 2025.

⁸² Ebd., Kapitel Urheberrecht.

Um beim Web-Crawling Rechtsvorbehalte zu identifizieren und zu beachten, wurden von den Unterzeichnern im Sinne von Artikel 4 der CDSM-Richtlinie gleichzeitig zwei weitere technische Verpflichtungen angenommen:

- **Einsatz von Web-Crawlern**, die Anweisungen lesen und befolgen, die gemäß dem Robot Exclusion Protocol (robots.txt) formuliert sind, wie im Request for Comments (RFC – Bitte um Kommentare) Nr. 9309 der Internet Engineering Task Force (IETF – Arbeitsgruppe Internettechnik) festgelegt.⁸³
- **Identifizierung und Befolgung anderer geeigneter maschinenlesbarer Protokolle** zur Formulierung von Rechtsvorbehalten gemäß Artikel 4 Absatz 3 der CDSM-Richtlinie. Dies kann beispielsweise durch objekt- oder standortbasierte Metadaten erfolgen, die entweder von internationalen oder europäischen Normungsorganisationen übernommen wurden oder dem neuesten Stand der Technik entsprechen. Dazu können solche gehören, die technisch umsetzbar sind und von Rechteinhabern unter Berücksichtigung verschiedener Kulturbereiche auf breiter Front eingesetzt werden und die im Rahmen eines inklusiven Prozesses auf Basis redlich geführter Gespräche, die auf EU-Ebene zu fördern sind, allgemein vereinbart wurden.

Zur Transparenz beim Training von GenAI führt die KI-Verordnung für Anbieter von KI-Systemen, darunter auch KI-Systeme mit allgemeinem Verwendungszweck, die synthetische Audio-, Bild-, Video- oder Textinhalte erzeugen, unter anderem die Verpflichtung ein, die technische Dokumentation des Modells zu erstellen und zu aktualisieren, einschließlich seines Trainings- und Testverfahrens und der Ergebnisse seiner Bewertung.⁸⁴ Diese Dokumentation muss die Modalität (z. B. Text, Bild) und das Format der Eingaben und Ausgaben, die Entwurfsspezifikationen des Modells und des Trainingsverfahrens einschließlich Trainingsmethoden und -techniken sowie die wichtigsten Entwurfsentscheidungen mit den Gründen und getroffenen Annahmen umfassen.⁸⁵ Zudem verpflichtet Art. 53 Abs. 1 Buchstabe d die Anbieter von Allzweck-KI-Modellen dazu, eine hinreichend detaillierte Zusammenfassung der für das Training des Allzweck-KI-Modells verwendeten Inhalte zu erstellen und öffentlich zugänglich zu machen.

Klarstellen müssen die Anbieter darüber hinaus gegebenenfalls auch Informationen über die für das Trainieren, Testen und Validieren verwendeten Daten, einschließlich der Art und Herkunft der Daten und der Aufbereitungsmethoden (z. B. Bereinigung, Filterung usw.), der Zahl der Datenpunkte, ihres Umfangs und ihrer Hauptmerkmale, gegebenenfalls aber auch die Art und Weise, wie die Daten erlangt und ausgewählt wurden, sowie alle anderen Maßnahmen zur Feststellung, ob Datenquellen ungeeignet sind, und Methoden zur Erkennung ermittelbarer Verzerrungen.⁸⁶

Wie in Artikel 53 Absatz 2 und Artikel 54 Absatz 6 der KI-Verordnung geregelt und in den Leitlinien der Kommission zum Umfang der Pflichten für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck gemäß der Verordnung (EU) 2024/1689 präzisiert, gelten besondere Ausnahmen für Anbieter von KI-Modellen, die im Rahmen einer freien

⁸³ Siehe Datatracker-Website, [Robots Exclusion Protocol RFC 9309](#), op. cit.

⁸⁴ Artikel 53 Absatz 1 Buchstabe a KI-Verordnung.

⁸⁵ Anhang XI KI-Verordnung.

⁸⁶ Ebd.

und quelloffenen Lizenz bereitgestellt werden, die den Zugang, die Nutzung, die Änderung und die Verbreitung des Modells ermöglicht.⁸⁷

2.2.3. Ausgewählte Beispiele für Ansätze zum Urheberrechtsschutz bei KI

Urheberrechtliche Regelwerke sind älter als die KI-Ära (und andere technologische Entwicklungen der Vergangenheit). Für die Gesetzgeber stellt sich damit die Frage, ob die europäischen und nationalen urheberrechtlichen Rahmenbedingungen für das Zeitalter der generativen KI geeignet sind oder nicht.

Die jüngste Folgenabschätzung zum Thema Urheberrecht und KI des Vereinigten Königreichs⁸⁸ beispielsweise gibt einen Überblick über die verschiedenen Ansätze zur Urheberrechtsgesetzgebung, die im heutigen technologischen Kontext zu erkennen sind, und berücksichtigt dabei wichtige Entwicklungen wie den EU-Urheberrechtsrahmen der CDSM-Richtlinie.

Im Vereinigten Königreich gibt es derzeit keine weitreichende Ausnahme, die die Verwendung urheberrechtlich geschützter Werke für kommerzielles TDM oder das Training von GenAI stützen könnte.⁸⁹ Bei der Diskussion über die künftige Politik werden in der britischen Folgenabschätzung drei Ziele benannt: 1) die Erhöhung der Transparenz hinsichtlich der Verwendung urheberrechtlich geschützter Werke zum Training von KI-Modellen und KI-generierter Inhalte, 2) die Ausstattung der Inhaber von Rechten an kreativen Werken mit größerer Kontrolle über die Verwendung ihres Materials zum Trainieren von KI-Modellen sowie die Unterstützung ihrer Möglichkeiten, bei einer solchen Nutzung vergütet zu werden, und 3) die Verbesserung der Möglichkeiten für KI-Entwickler, auf das Material zuzugreifen, das sie benötigen, um weltweit führende KI-Modelle im Vereinigten Königreich zu trainieren. Mit diesen Zielsetzungen ergeben sich in der Regel vier Optionen:

⁸⁷ Europäische Kommission, [Guidelines on the scope of obligations for providers of general-purpose AI models under the AI Act](#) (auf Englisch) [Leitlinien zum Umfang der Pflichten für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck im Rahmen der KI-Verordnung], C(2025) 7719 final, Brüssel, 18. Juli 2025.

⁸⁸ HM Government, [Copyright and AI Impact Assessment](#), (auf Englisch) [Folgenabschätzung Urheberrecht und KI] Ministerium für Wissenschaft, Innovation und Technologie. Ministerium für Kultur, Medien und Sport. Dem Parlament gemäß § 135 des Gesetzes über Daten (Nutzung und Zugang) von 2025 vorgelegt, 18 März 2026.

⁸⁹ UK parliament publications, [AI, copyright and the creative industries](#), (auf Englisch) [KI, Urheberrecht und die Kreativwirtschaft] Ausschuss für Kommunikation und Digitales, 4. Bericht der Legislaturperiode 2024–2026, Oberhaus, 6. März 2026.

Tabelle 3. Alternativen im Umgang mit dem KI-Training

Option	Richtlinie
0	Nichtstun. Bei dieser Option wäre in der Regel eine Genehmigung erforderlich, um geschützte Werke in verschiedenen Phasen des KI-Trainings und der KI-Entwicklung zu kopieren.
1	Stärkung des Urheberrechts, sodass für die Nutzung urheberrechtlich geschützter Inhalte beim KI-Training Lizenzen benötigt werden, verbunden mit Transparenzanforderungen und Maßnahmen zum Marktzugang. Da das Urheberrecht im Vereinigten Königreich bereits bestimmte Rechte vorsieht, würde das Lizenzierungssystem bei dieser Option weitgehend dem Status quo entsprechen.
2	Einführung einer weitreichenden urheberrechtlichen Ausnahme für TDM. Die Ausnahme wäre nicht völlig schrankenlos, da Data-Mining-Aktivitäten nur dann zulässig wären, wenn der Nutzer rechtmäßigen Zugang hat. Ein Rechtsvorbehalt und ein Opting-out von der Ausnahme wären für die Rechteinhaber jedoch nicht möglich.
3	Einführung einer TDM-Urheberrechtsausnahme mit Mechanismus für den Rechtsvorbehalt und Transparenzanforderungen. Die Ausnahme würde nur gelten, wenn ein Rechteinhaber seine Rechte nicht ausdrücklich vorbehalten hat. Mit diesem Vorbehalt könnten Rechteinhaber die Ausnahme für ausgewählte Werke „abschalten“.

Quelle: *AI, copyright and the creative industries. Communications and Digital Committee. 4th Report of Session 2024–2026*⁹⁰ (auf Englisch) [KI, Urheberrecht und die Kreativwirtschaft. Ausschuss für Kommunikation und Digitales. 4. Bericht der Legislaturperiode 2024–2026]

Da sich die urheberrechtlichen Regelwerke und die Ansätze mit Blick auf das KI-Training von Land zu Land unterscheiden, ist auch der Umgang mit KI-Training in der Urheberrechtspolitik weltweit verschieden. So wird beispielsweise in Frankreich derzeit über die mögliche Einführung einer Vermutung der Nutzung von Werken diskutiert, um die Transparenz hinsichtlich der Nutzung urheberrechtlich geschützter Inhalte im Zusammenhang mit KI zu erhöhen.⁹¹

Zur Harmonisierung der KI-Gesetzgebung entstehen gerade allgemeine Instrumente wie das Rahmenübereinkommen des Europarats über künstliche Intelligenz, Menschenrechte, Demokratie und Rechtsstaatlichkeit, doch auf globaler Ebene gibt es kein bedeutendes Übereinkommen, das sich mit KI und Urheberrecht befasst. Das Urheberrecht

⁹⁰ [Copyright and AI Impact Assessment](#), op. cit., S. 15.

⁹¹ Siehe di Französischer Sénat -Website, [Proposition de loi relative à l'instauration d'une présomption d'exploitation des contenus culturels par les fournisseurs d'intelligence artificielle](#) (auf Französisch) [Entwurf des Gesetzes zur Einführung einer Vermutung der Verwertung kultureller Inhalte durch Anbieter künstlicher Intelligenz], verabschiedet am 8. April 2026.

unterliegt im Allgemeinen weiterhin der Berner Übereinkunft⁹², anderen internationalen Verträgen wie dem WCT⁹³ und dem WPPT,⁹⁴ sowie nationalen (und europäischen) Instrumenten.

Im Jahr 2026 werden unter anderem folgende Ansätze zu KI und Urheberrecht verfolgt:

Tabelle 4. Europäische und nationale Ansätze im Umgang mit KI-Training

	EU	TDM-Ausnahmen gemäß Artikel 3 und 4 der CDSM-Richtlinie. Einziges Regelwerk, das einen „Opt-out“-Mechanismus vorsieht.
	UK	Gesetzliche Ausnahme, die TDM ausschließlich für nichtkommerzielle Forschung zulässt, sofern der Nutzer rechtmäßigen Zugang zum Werk hat. Politische Diskussionen laufen.
	USA	Kein spezifisches Instrument für TDM. Es gilt die Fair-Use-Doktrin. ⁹⁵
	KANADA	Kein spezifisches Instrument für TDM. Politische Diskussionen laufen. ⁹⁶
	JAPAN	Umfassende gesetzliche Ausnahme, die TDM für kommerzielle und nichtkommerzielle Zwecke erlaubt, insbesondere für „Nichtvergnügungszwecke“. ⁹⁷

⁹² Weltorganisation für geistiges Eigentum (WIPO), [Berne Convention for the Protection of Literary and Artistic Works](#) (auf Englisch) [Berner Übereinkunft zum Schutz von Werken der Literatur und Kunst], World Intellectual Property Office. Eine deutsche Fassung ist verfügbar [hier](#).

⁹³ WIPO, [WIPO Copyright Treaty](#) (WCT), Genf, verabschiedet am 20. Dezember 1996.

⁹⁴ WIPO, [WIPO Performances and Phonograms Treaty](#), Genf, verabschiedet am 20. Dezember 1996.

⁹⁵ Der amerikanische Umgang mit KI-Training ist besonders wichtig, da einige der größten KI-Anbieter aus Amerika kommen. Gemäß der Fair-Use-Doktrin, die in Section 107 des 1976 Copyright Act (Abschnitt 107 des Urheberrechtsgesetzes von 1976) festgeschrieben ist, stellt die faire Nutzung eines urheberrechtlich geschützten Werks keine Rechtsverletzung dar, sofern bestimmte Voraussetzungen erfüllt sind. Dabei sind vier Kriterien zu prüfen: 1. Zweck und Art der Nutzung des Werks (mit Prüfung der Gewerbsmäßigkeit). 2. Art des urheberrechtlich geschützten Werks. 3. Umfang und Wesentlichkeit des genutzten Teils im Verhältnis zur Gesamtheit des urheberrechtlich geschützten Werks. 4. Auswirkungen der Nutzung auf den potenziellen Markt für einen Wert des urheberrechtlich geschützten Werks. Zur offiziellen Position des US-Urheberrechtsamtes siehe [Copyright and Artificial Intelligence](#). Part 3. Generative AI training (auf Englisch) [Urheberrecht und künstliche Intelligenz. Teil 3. Training generativer KI]. US Copyright Office (US-Urheberrechtsamt), Mai 2025.

⁹⁶ Siehe *Government of Canada*-Website, [Consultation on Copyright in the Age of Generative Artificial Intelligence](#) (auf Englisch) [Konsultation zum Urheberrecht im Zeitalter der generativen künstlichen Intelligenz].

⁹⁷ [AI, copyright and the creative industries](#), op. cit., S. 12.

	CHINA	Kein spezifischer Rahmen. Das Regelwerk zum Urheberrecht ist eine Mischung aus dem Urheberrechtssystem (mit Fair Use) und dem Urheberschaftssystem. ⁹⁸
---	--------------	---

Quelle: Europäische Audiovisuelle Informationsstelle.

Bei den europäischen Vorschriften ist zu berücksichtigen, dass einige der Regeln zum räumlichen Geltungsbereich⁹⁹ in der KI-Verordnung vorsehen, dass diese Verordnung sowohl für Betreiber von KI-Systemen gilt, die ihren Sitz in der Union haben oder sich in der Union befinden, als auch für Anbieter und Betreiber von KI-Systemen, die ihren Sitz in einem Drittland haben oder sich in einem Drittland befinden, wenn die vom KI-System hervorgebrachte Ausgabe in der Union verwendet wird.¹⁰⁰

Die CDSM-Richtlinie enthält hierzu keine spezifischen Vorschriften.¹⁰¹

⁹⁸ Zhang J., Li X., [Fair Use in Training AI Models: A Review and Prospect of the Relevant Legal Development in China](#) (auf Englisch) [Fair Use beim Training von KI-Modellen: Überblick und Ausblick zu den einschlägigen rechtlichen Entwicklungen in China], *Cambridge University Press*, 24. April 2026.

⁹⁹ Erwägungsgrund 106 CDSM-Richtlinie: „Anbieter, die KI-Modelle mit allgemeinem Verwendungszweck in der Union in Verkehr bringen, sollten die Erfüllung der einschlägigen Pflichten aus dieser Verordnung gewährleisten. Zu diesem Zweck sollten Anbieter von KI-Modellen mit allgemeinem Verwendungszweck eine Strategie zur Einhaltung des Urheberrechts der Union und der verwandten Schutzrechte einführen, insbesondere zur Ermittlung und Einhaltung des gemäß Artikel 4 Absatz 3 der Richtlinie (EU) 2019/790 durch die Rechteinhaber geltend gemachten Rechtsvorbehalts. Jeder Anbieter, der ein KI-Modell mit allgemeinem Verwendungszweck in der Union in Verkehr bringt, sollte diese Pflicht erfüllen, unabhängig davon, in welchem Hoheitsgebiet die urheberrechtlich relevanten Handlungen, die dem Training dieser KI-Modelle mit allgemeinem Verwendungszweck zugrunde liegen, stattfinden. Dies ist erforderlich, um gleiche Wettbewerbsbedingungen für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck sicherzustellen, unter denen kein Anbieter in der Lage sein sollte, durch die Anwendung niedrigerer Urheberrechtsstandards als in der Union einen Wettbewerbsvorteil auf dem Unionsmarkt zu erlangen.“

¹⁰⁰ Artikel 2 KI-Verordnung.

¹⁰¹ Zapala I., [Territorial scope of the authorization requirement and liability regime under Directive \(EU\) 2019/790 on copyright and related rights in the digital single market](#), *Journal of Intellectual Property Law & Practice* (auf Englisch) [Räumlicher Geltungsbereich der Genehmigungspflicht und der Haftungsregelung gemäß der Richtlinie (EU) 2019/790 über das Urheberrecht und die verwandten Schutzrechte im digitalen Binnenmarkt, 16/8, S. 827–849, August 2021.

3. Urheberrecht bei KI-Training und KI-Prompting

3.1. KI-Trainingsaktivitäten

Häufig wird davon ausgegangen, dass das Training von KI-Systemen unter die TDM-Ausnahme im Rahmen des europäischen Regelwerks fällt, aber es gibt auch abweichende Stimmen, die bezweifeln, dass der gesamte Trainingsprozess tatsächlich durch die Artikel 3 und 4 der CDSM-Richtlinie gedeckt ist.

Zwar erscheint der Rechtsrahmen für die audiovisuelle Industrie insgesamt solide,¹⁰² doch die akademische Lehre deutet darauf hin, dass KI-Training über TDM hinausgeht, weil TDM Teil des KI-Trainings ist, aber nicht umgekehrt, und weist warnend darauf hin, dass bislang kein Rechtsinstrument KI-Training und TDM als gleichwertig betrachtet hat.¹⁰³ Zudem warnt die Rechtslehre, dass es bei TDM um die Sammlung, Verarbeitung und Analyse von Informationen geht, aber wichtige – etwa inhaltsbezogene – Aspekte des KI-Trainings nicht darunter fallen.¹⁰⁴ Das Amt der Europäischen Union für geistiges Eigentum (EUIPO) vertritt aus institutioneller Sicht eine differenziertere Position, die die verschiedenen Standpunkte gegeneinander abwägt.¹⁰⁵ Auch der Voss-Bericht hat, neben anderen Zweifeln an den tatsächlichen Auswirkungen der TDM-Ausnahme, kürzlich gewarnt, dass die Opt-out-Regelung „möglicherweise nicht alle relevanten Handlungen des Text- und Data-Minings“ abdeckt. Aktuell erscheinen in Europa einige wichtige Urteile innerstaatlicher Gerichte. Zu den relevantesten Fällen auf nationaler Ebene gehört der Rechtsstreit der deutschen Verwertungsgesellschaft GEMA mit OpenAI, in dem das Landgericht München urteilte, dass ChatGPT geschützte Liedtexte vervielfältigt habe.¹⁰⁶ In anderen Fällen, wie beispielsweise Getty Images / Stability AI im Vereinigten Königreich, gab der britische High

¹⁰² Siehe z. B., die ACT-Website, [AI Principles for Audiovisual Media Services](#) (auf Englisch) [KI-Prinzipien für audiovisuelle Mediendienste], Brüssel, 16. Juni 2026: „Der Rechtsrahmen für Urheberrecht und KI der EU – insbesondere die CDSM-Richtlinie und die KI-Verordnung – legt Regeln für die Nutzung geschützter Werke zum Trainieren von KI fest. Bei der Umsetzung bestehen zwar noch erhebliche Unsicherheiten, doch der Rechtsrahmen ist robust genug, um sich an technologische Veränderungen anzupassen, und bietet eine solide Grundlage für die Beurteilung von Rechtsverletzungen sowohl in der Eingabe- als auch in der Ausgabephase von KI-Systemen.“

¹⁰³ Siehe Rosati, E., [The future of the movie industry in the wake of generative AI: A perspective under EU and UK copyright law](#), op. cit. Rosati betont, dass im Rahmen des TDM „die wichtigsten Handlungen nach britischem und EU-Recht die Entnahme (sofern sich die Daten in einer Datenbank befinden) und die Vervielfältigung sind“, nicht andere.

¹⁰⁴ Siehe Dornis, T., [The Training of Generative AI is not Text and Data Mining](#) (auf Englisch) [Das Training von generativer KI ist kein Text- und Data-Mining], *European Intellectual Property Review* (E.I.P.R.), 19. Oktober 2024.

¹⁰⁵ Siehe EUIPO-Website, [The development of Generative Artificial Intelligence from a Copyright Perspective](#), op. cit., S. 67.

¹⁰⁶ Z. B. mit Blick auf KI-Training [GEMA / OpenAI](#), 42 O 14139/24, Landgericht München I.

Court (Obergericht) zwar einige Hinweise, doch relevante Fragen wurden noch nicht geklärt.¹⁰⁷

Da es sich beim Training von GenAI um einen komplexen technischen Prozess handelt, gibt es verschiedene Formen und Teilprozesse. Die neuesten GenAI-Modelle sind multimodal, d. h. sie funktionieren mit verschiedensten Inhaltstypen.¹⁰⁸ In der Regel stützen sich diese Systeme auf Datensätze aller Art, darunter Videos, Texte oder Bilder, die von privaten und öffentlichen Institutionen oder Einzelpersonen stammen, sowie auf synthetische Daten. Eine der gängigen Methoden für den Zugriff auf diese Daten erfolgt über Web-Scraping und Crawling.¹⁰⁹ Die gewonnenen Daten werden anschließend erneut verarbeitet, um ihre Qualität sicherzustellen, also Richtigkeit, Konsistenz, Relevanz und andere wichtige Aspekte. Zu guter Letzt umfasst das Training noch eine Feinabstimmung, bevor ein Output generiert wird. Unter Urheberrechtsaspekten können sich daher in den verschiedenen Phasen des KI-Trainings unterschiedliche Auswirkungen ergeben, weil in den einzelnen Phasen auch unterschiedliche Aktivitäten stattfinden. Der Wissenschaft zufolge gibt es im KI-Training fünf relevante Phasen mit unterschiedlichen technischen Prozessen:

Abbildung 3. Der technische Prozess des Trainierens von KI-Systemen.

- 
- 1. Erfassung und Vorverarbeitung:** Inhalte werden gesammelt.
 - 2. Trainingsschleife:** KI-Systeme generieren bereinigte Daten.
 - 3. Modellartefakt:** Darstellungen der gewonnenen Daten werden generiert, können aber nicht angesehen werden.
 - 4. Generierung und Abruf:** Benutzeranfragen werden tokenisiert.
 - 5. Protokolle und Caches:** Informationen werden ausgewählt und teilweise für weiteres Training zwischengespeichert.

Quelle: Copyright in Formaldehyde: How GEMA v OpenAI Freezes Doctrine and Chills AI – Part 1 (auf Englisch) [Urheberrecht in Formaldehyd: Wie GEMA / OpenAI die Rechtslehre einfriert und die KI abschreckt – Teil 1].¹¹⁰

¹⁰⁷ Stokes G. et al., [Getty Images v Stability AI: the UK High Court decision offers guidance, but critical questions on AI and copyright infringement remain](#) (auf Englisch) [Getty Images / Stability AI: Das Urteil des britischen High Court gibt Hinweise, doch wichtige Fragen zu KI und Urheberrechtsverletzungen bleiben offen], *DLA Piper*, 5. November 2025.

¹⁰⁸ Marcelin T. and Cassetti F., [AI and copyright: The training of general-purpose AI](#), (auf Englisch) [KI und Urheberrecht: Das Trainieren von KI mit allgemeinem Verwendungszweck], *European Parliamentary Research Service*, April 2025.

¹⁰⁹ Das EUIPO erklärt: „Beim Crawling werden in der Regel öffentlich zugängliche Webseiten systematisch durchsucht, um Metadaten und Inhalte für Suchmaschinen zu indexieren. Dieser Vorgang gilt allgemein als weniger starker Eingriff und dient in erster Linie der Auffindbarkeit. Beim Scraping werden dagegen spezifische Daten oder Inhalte entnommen, wobei es oft um detailliertere Informationen geht und Teile von Werken so vervielfältigt werden, dass urheberrechtliche Bedenken wahrscheinlicher sind.“ Siehe EUIPO-Website, [The development of Generative Artificial Intelligence from a Copyright Perspective](#), op. cit., S. 139.

¹¹⁰ Siehe Frosio G., [Copyright in Formaldehyde: How GEMA v OpenAI Freezes Doctrine and Chills AI – Part 1](#) (auf Englisch) [Urheberrecht in Formaldehyd: Wie GEMA / OpenAI die Rechtslehre einfriert und die KI abschreckt –

3.2. KI-Prompting

Der Begriff „Prompting“ wird in den Rechtsvorschriften nicht definiert, lässt sich aber als Vorgang beschreiben, bei dem einem KI-System spezifische Anweisungen oder Informationen gegeben werden, um ein bestimmtes Ergebnis zu erhalten, wobei zu berücksichtigen ist, dass die Klarheit und Präzision der Anweisungen oder Prompts positiven Einfluss auf die Qualität des Outputs hat.¹¹¹

Prompts bilden den Anstoß, um in dialogorientierten Chatbots einen Output zu erhalten. Abgesehen von der Debatte um die Urheberrechtsfähigkeit des Outputs (für die sich Prompts bei GenAI und chatbotähnlicher KI als unverzichtbares Tool erweisen und die im zweiten Teil dieses Berichts behandelt werden), besteht ein zweiter Aspekt in der Schutzfähigkeit von Prompts als solchen, die sowohl den allgemeinen urheberrechtlichen Vorschriften (wonach bloße Ideen nicht unter Schutz stehen) als auch möglicherweise anderen Rechtsgebieten wie dem Geschäftsgeheimnisrecht unterliegt.¹¹²

In den USA ist das Copyright Office (Urheberrechtsamt) zu dem Schluss gekommen, dass „Prompts allein nicht genug menschliche Kontrolle bieten, um Nutzer eines KI-Systems zu Urhebern des Outputs zu machen. Prompts fungieren im Wesentlichen als Anweisungen, die nicht schutzfähige Ideen vermitteln. Sehr detaillierte Prompts könnten zwar die vom Nutzer gewünschten Äußerungselemente enthalten, bestimmen aber derzeit nicht, wie das KI-System sie bei der Erzeugung des Outputs verarbeitet.“ In diesem Sinne hat das US-Urheberrechtsamt gewarnt, dass „die Diskrepanz zwischen Prompts und resultierenden Outputs zeigt, dass der Nutzer keine Kontrolle über die Umsetzung seiner Ideen in eine feste Äußerung hat und für die Bestimmung der Äußerungselemente im Output maßgeblich das System verantwortlich ist“.¹¹³

Teil 1], *Kluwer Copyright Blog*, 10. Dezember 2025. Einen umfassenden Überblick über den Trainingsprozess bietet zudem die Abbildung „Generative AI tech stack from data collection to end-user interaction“ (Tech-Stack für generative KI – von der Datensammlung bis zur Interaktion mit dem Endnutzer) in der [Guidance note on the implications of generative artificial intelligence for freedom of expression](#) (auf Englisch) [Leitfaden zu den Auswirkungen von generativer KI auf die Meinungsfreiheit] des Europarats], S. 44 f.

¹¹¹ Siehe die KU Leuven-Website, [Guiding AI with prompting](#) (auf Englisch) [KI mit Prompts führen].

¹¹² Für eine Auseinandersetzung mit dem Thema Geschäftsgeheimnisse und KI siehe Hrdy, C. [Trade Secrecy Meets Generative AI](#) (auf Englisch) [Geschäftsgeheimnisrecht trifft generative KI], *Chicago Kent Law Review* Nr. 100., 3. Februar 2025.

¹¹³ US-Urheberrechtsamt, [Copyright and Artificial Intelligence. Part 2. Copyrightability](#) (auf Englisch) [Urheberrecht und künstliche Intelligenz. Teil 2. Urheberrechtsfähigkeit], Januar 2025.

4. Plattformen und Nutzer: die Nutzungsbedingungen

Das Training von KI ist (zumindest teilweise) gesetzlich geregelt, und Plattformen nehmen in der Regel bestimmte rechtliche Bestimmungen in ihre Nutzungsbedingungen auf, denen die Nutzer bei der Nutzung zustimmen. In diesem Zusammenhang begründen die Nutzungsbedingungen ein privatrechtliches Verhältnis zwischen Nutzer und Plattform, das verschiedene Aspekte wie Urheberrecht und Training regelt. Nach europäischem Recht kann eine Plattform jedoch in jedem Fall nach den jeweils geltenden rechtlichen Regelungen haften.

So kann im europäischen Urheberrecht die Anwendung der TDM-Ausnahme dazu führen, dass ein KI-Anbieter haftet, wenn er gegen die Opt-out-Regelung verstößt. In anderen Fällen können KI-Plattformen für Verstöße gegen andere Rechtsvorschriften wie die KI-Verordnung und den DSA haften, weil KI-Dienste über digitale Plattformen bereitgestellt werden, die in deren Anwendungsbereich fallen können.¹¹⁴

Wo Nutzungsbedingungen vorliegen, müssen diese umfassend berücksichtigt werden, aber die Haftung des Anbieters kann nicht allein deshalb ausgeschlossen werden, weil die Nutzungsbedingungen die Verantwortung auf den Nutzer übertragen.¹¹⁵

Einige der beliebtesten GenAI-Plattformen enthalten tatsächlich Hinweise auf den Input und auf die Möglichkeit, dass sie Prompts und anderes von Nutzern eingegebenes Material weiterverwenden können. Soweit dies auf die Zustimmung des Nutzers zu den Nutzungsbedingungen zurückgeht und somit ein durchsetzbares privatrechtliches Verhältnis darstellt, können urheberrechtliche Auswirkungen auf den Input bestehen, und der Nutzer akzeptiert in der Regel, dass die Inhalte zu Trainingszwecken verwendet werden dürfen.

Als Beispiel seien hier einige der häufigsten Fälle genannt, die bei beliebten GenAI-Tools beobachtet wurden:

Tabelle 5. Bestimmungen zum KI-Training in den Nutzungsbedingungen einiger der wichtigsten KI-Plattformen.

Plattform	Nutzungsbedingungen (relevante Erwähnungen)
	Ihre Prompts und Eingaben an generative KI-Funktionen in Adobe-Produkten sowie die von diesen generierten Ergebnisse können sowohl durch

¹¹⁴ [Verordnung \(EU\) 2022/2065](#) des Europäischen Parlaments und des Rates vom 19. Oktober 2022 über einen Binnenmarkt für digitale Dienste und zur Änderung der Richtlinie 2000/31/EG (DSA). Der DSA gilt laut Artikel 2 für Vermittlungsdienste, die für Nutzer mit Niederlassungsort oder Sitz in der Union angeboten werden, ungeachtet des Niederlassungsortes des Anbieters dieser Vermittlungsdienste.

¹¹⁵ Siehe Rosati, E., [The future of the movie industry in the wake of generative AI: A perspective under EU and UK copyright law](#), op. cit.

Plattform	Nutzungsbedingungen (relevante Erwähnungen)
Adobe ¹¹⁶	automatisierte (wie maschinelles Lernen) als auch durch manuelle Methoden zur Missbrauchsprävention sowie zur Inhaltsfilterung überprüft werden. Die Verwendung der generativen KI-Funktionen von Adobe zur Erstellung, zum Hochladen oder zur Weitergabe von Inhalten, die Urheber-, Marken-, Datenschutz-, Persönlichkeits- oder andere Rechte Dritter verletzen, ist untersagt. Dies kann Folgendes umfassen: das Eingeben von Prompts zur Generierung urheberrechtlich geschützter, markenrechtlich geschützter oder anderweitig rechtsverletzender Inhalte, das Hochladen von Eingaben oder Referenzbildern, die urheberrechtlich geschützte Inhalte eines Dritten enthalten, das Generieren von Text, der Inhalte Dritter plagiiert, oder die Verwendung persönlicher Informationen eines Dritten unter Verletzung seiner Datenschutz- oder Datenschutzrechte [sic].
 ChatGPT ¹¹⁷	Sie können Eingaben in die Dienste („Eingaben“) vornehmen und auf der Grundlage dieser Eingaben Ausgaben von den Diensten erhalten („Ausgaben“). Widerspruch (Opt-out). Wenn Sie nicht möchten, dass wir Ihre Inhalte zum Trainieren unserer Modelle verwenden, haben Sie die Möglichkeit, dies zu verhindern, indem Sie die Anweisungen in diesem Artikel befolgen. Bitte beachten Sie, dass dies in einigen Fällen die Fähigkeit unserer Dienste einschränken kann, besser auf Ihren spezifischen Anwendungsfall einzugehen.
 Claude ¹¹⁸	Es ist Ihre Entscheidung, ob Sie die Verwendung Ihrer Daten zur Verbesserung neuer Claude-Modelle zulassen möchten, und Sie können Ihre Auswahl jederzeit in Ihren Datenschutzeinstellungen ändern.
 Copilot ¹¹⁹	Wir sind nicht Eigentümer Ihrer Inhalte, aber wir können Ihre Inhalte verwenden, um Copilot zu betreiben und zu verbessern. Durch die Nutzung von Copilot erteilen Sie uns die Erlaubnis, Ihre Inhalte zu verwenden, was bedeutet, dass wir sie kopieren, verteilen, übertragen, öffentlich ausstellen, öffentlich aufführen, bearbeiten, übersetzen und neu formatieren können, und wir können dieselben Rechte auch anderen gewähren, die in unserem Namen arbeiten. Wir können entscheiden, ob wir Ihre Inhalte verwenden, und wir müssen Sie nicht bezahlen, Sie um Erlaubnis bitten oder Ihnen sagen, wann wir dies tun. Aber das bedeutet nicht, dass wir es verwenden können, wie wir wollen. Die [sic] In den Microsoft-Datenschutzbestimmungen wird erläutert, wie wir Ihre Inhalte verwenden, und die Datenschutzooptionen in Copilot geben Ihnen die Kontrolle über einige dieser Verwendungszwecke.

¹¹⁶ Siehe die Adobe-Website, [Benutzungsrichtlinien zur generativen KI von Adobe](#).

¹¹⁷ Siehe die OpenAI-Website, [Nutzungsbedingungen](#).

¹¹⁸ Siehe die Anthropic-Website, [Updates to Consumer Terms and Privacy Policy](#) (auf Englisch) [Aktualisierungen der Nutzungsbedingungen für Verbraucher und der Datenschutzerklärung].

¹¹⁹ Siehe die Microsoft-Website, [Microsoft Copilot-Nutzungsbedingungen](#).

Plattform	Nutzungsbedingungen (relevante Erwähnungen)
 Midjourney ¹²⁰	Durch die Nutzung der Dienste gewähren Sie Midjourney sowie seinen verbundenen Unternehmen, Nachfolgern und Abtretungsempfängern eine unbefristete, weltweite, nicht-exklusive, unterlizenzierbare, unentgeltliche, lizenzgebührenfreie und unwiderrufliche urheberrechtliche Lizenz, die von Ihnen in die Dienste eingegebenen Inhalte sowie alle von Ihnen durch den Dienst erzeugten Objekte zu vervielfältigen, abgeleitete Werke davon zu erstellen, öffentlich auszustellen, öffentlich aufzuführen, unterzulizenzieren und zu verbreiten. Diese Lizenz bleibt auch nach Beendigung dieser Vereinbarung durch eine der Parteien aus gleich welchem Grund bestehen.

¹²⁰ Siehe die Midjourney-Website, [Terms of Service](#) (auf Englisch) [Nutzungsbedingungen].

5. Wichtigste Erkenntnisse

Die KI-Entwicklung schreitet weiterhin rasant voran, wobei GenAI zunehmend in berufliche und private Nutzungen integriert wird und agentische KI rapides Wachstum verzeichnet. Besonders stark sind die Auswirkungen in der Kreativwirtschaft, etwa in der audiovisuellen Produktion, wo KI Texte, Bilder, Audio und Video erzeugen kann. Diese rasante Entwicklung eröffnet neue Möglichkeiten, wirft aber auch, insbesondere unter dem europäischen Urheberrechtsrahmen, komplexe urheberrechtliche Herausforderungen auf, denen sich die Legislative mit neuen Rechtsinstrumenten zu stellen versucht.

Das KI-Training steht im Mittelpunkt des Zusammenspiels zwischen Urheberrecht und KI, denn dazu werden umfangreiche Datensätze verwendet, die urheberrechtlich geschützte Werke und sonstige Schutzgegenstände enthalten können. Die unbefugte Nutzung urheberrechtlich geschützter Inhalte stößt bei Urhebern, Verlagen und der Kulturwirtschaft auf Kritik, nicht zuletzt im audiovisuellen Sektor, wo kognitive Aufgaben eine wichtige Rolle spielen. Für die Politik steht zurzeit im Fokus, den Zugang zu Daten mit dem Schutz der Rechteinhaber in Einklang zu bringen.

In Europa gibt es spezifischen urheberrechtlichen Vorschriften für GenAI in einigen Bestimmungen in der KI-Verordnung und der späteren Anwendung der CDSM-Richtlinie. Die KI-Verordnung führt Transparenzpflichten ein, verweist aber bei urheberrechtlichen Fragen auf die CDSM-Richtlinie. Verbleibende Unsicherheiten werden wohl durch zukünftige Gerichtsentscheidungen und Gesetzesüberprüfungen geklärt.

Beim KI-Training stehen die Vorschriften zum Text- und Data-Mining (TDM) im Mittelpunkt der Debatte, weil sie eine mögliche Rechtsgrundlage für bestimmte KI-Trainingsaktivitäten bieten. Offen bleibt jedoch die Frage, ob diese Ausnahmen die Komplexität des KI-Trainings mit seinen vielen technischen Prozessen voll und ganz erfassen. Auch die Wirksamkeit von Mechanismen für den Rechtsvorbehalt (wie dem europäischen „Opt-out“-Mechanismus) ist Thema von Diskussionen, zuletzt beispielsweise im Voss-Bericht.

Das Prompting wirft ebenfalls urheberrechtliche Fragen auf, insbesondere nach dem Schutz von Prompts und der Verwendung urheberrechtlich geschützter Materialien. Da Prompts in erster Linie Anweisungen geben und nicht die endgültige Äußerung bestimmen, wird bei aktuellen Ansätzen in Frage gestellt, ob die Nutzer eine ausreichende kreative Kontrolle über ihre Materialien beim KI-Training gemäß den in den Nutzungsbedingungen beliebter KI-Plattformen festgelegten Vorgaben ausüben. Das Verhältnis zwischen menschlichem Input und KI-generiertem Inhalt bleibt ein rechtliches Thema, das sich fortwährend weiterentwickelt.

Eine Publikation
der Europäischen Audiovisuellen Informationsstelle

