bitkom

Whitepaper zu Urheberrecht und generativer KI

Technische und rechtliche Hintergründe und Möglichkeiten für Urheberinnen und Urheber

Einleitung

Künstliche Intelligenz, vor allem generative Künstliche Intelligenz, ist zunehmend präsent und beeinflusst unser tägliches Leben immer stärker. Zuweilen nehmen die Diskussionen rund um generative KI – insbesondere im Verhältnis zu Urhebern und dem Urheberrecht – emotionale Züge an. Es ist deshalb wichtig zu wissen, wie generative KI technisch funktioniert, welche rechtlichen Aspekte dabei eine Rolle spielen und was das Urheberrecht schützt – oder eben nicht schützt. Neue Technologien können aufgrund der Neuartigkeit und dem anfangs fehlendem Verständnis naturgemäß bei einigen Menschen Skepsis hervorrufen. Dieses Whitepaper soll einen leicht verständlichen Einblick in die Entwicklung von Large Language Models (LLM) und Large Multimodal Models (LMM) geben. Es handelt sich dabei um zwei der wesentlichen Technologien, die Anwendungen generativer KI zugrunde liegen.

LLM und LMM fußen in ihrer Entwicklung auf einem Trainingsschritt, in welchem sich die LLM und LMM die Regeln der Informationsverarbeitung anlernen.[1] Grundlage dieses Trainings sind Trainingsdaten. Deshalb beginnt dieses Whitepaper mit den trainingsvorbereitenden Prozessen, in welchen die Trainingsdaten gesammelt und aufbereitet werden. Dabei steht die rechtliche Zulässigkeit der Trainingsdatenzusammenstellung im Mittelpunkt, insbesondere im Hinblick auf automatisierte Ansätze. Hierbei stellen sich Fragen nach dem Urheberrecht und den geltenden Schrankenregelungen, wie etwa der Text and Data Mining (TDM)-Schranke gemäß § 44b des Urheberrechtsgesetzes. Im Hinblick auf diesen Entwicklungsschritt wird dargelegt, welche Maßnahmen Urheberinnen und Urheber ergreifen können, um einer Verwendung ihrer Werke für das Training von generativer KI zu widersprechen.

In einem zweiten Schritt wird der eigentliche Trainingsprozess von KI-Modellen betrachtet. Dabei liegt ein Fokus auf der Parametrisierung des Modells und den verschiedenen Phasen des Trainings selbst. Besonderes Augenmerk liegt auf der rechtlichen Perspektive während des Trainings – vor allem darauf, dass im Training in der Regel keine urheberrechtlich relevanten Vervielfältigungshandlungen stattfinden.

Abschließend werden die rechtlichen Aspekte nach dem Training und im Betrieb von KI-Modellen untersucht. Hierbei wird untersucht, inwieweit KI-Output möglicherweise Urheberrechte verletzen könnte und wie und ob KI-Output urheberrechtlich geschützt ist.

[1] Im Gegensatz zu herkömmlichen Programmierweisen, bei welchen die Regeln zur Informationsverarbeitung mehrheitlich direkt vorgegeben werden.

1 Vor dem Training

Daten als Grundbedingung für maschinelles Lernen

Das Sammeln von Daten und die Erstellung eines Datenkorpus für das Trainieren eines Large Language Models (LLM) oder eines Large Multimodal Models (LMM) ist entscheidend, da Daten die Grundbedingung für das maschinelle Lernen darstellen. Ein Trainingsdatenkorpus, eine Sammlung von Datenpunkten, die für das Training verwendet werden, ist hierbei zentral. Datenpunkte sind Tatsachen oder Informationen, die eine Beobachtungseinheit zu einem bestimmten Zeitpunkt auf der Ebene der Datenerfassung beschreiben. Die Datenpunkte werden zum Zweck des Trainings des KI-Modells in den Datenkorpus übertragen. Die Größe und die Anzahl der Datenpunkte in diesen Datensätzen können stark variieren. Sie hängen vom Einsatzgebiet des LLM/LMM ab.

Die Größe der Datensätze für führende LLM und LMM wie GPT 4 oder LLaMA-2 können mehrere hundert Gigabyte bis hin zu mehreren Tera- oder Petabytes an Text- und Bilddaten umfassen. Die Anzahl der Datenpunkte, also individueller Textstücke oder Beispiele, die in diesen Datensätzen enthalten sind, erreicht oft mehrere Milliarden bis Billionen. Diese immense Datenmenge ist erforderlich, um etwa die Komplexität der menschlichen Sprache abzubilden und umfassende, genaue und vielseitige Sprachmodelle zu trainieren. Die genauen Zahlen variieren je nach spezifischem Modell und dessen Anwendungsbereich.

Ansätze zur Erhebung von Daten

Zur Erhebung dieser Datenpunkte gibt es mehrere Ansätze. Unternehmen haben zum einen die Möglichkeit, ihre Daten selbst zusammenzustellen, was allerdings gerade bei allgemeinen generativen KI-Modellen aufgrund der benötigten Datenmenge und -vielfalt nicht praktikabel ist. In diesen Fällen werden häufig Bots (sog. Crawler) eingesetzt, um große Datenmengen aus dem frei zugänglichen Internet zu sammeln.

Im Zusammenhang mit dem Training von LLM und LMM werden jedoch regelmäßig vorgefertigte Datensets von Drittanbietern („Prepackaging“), wie beispielsweise das Common Crawl-Dataset, erworben. Dieses Dreiparteienverhältnis – zwischen dem Entwickler des KI-Systems, dem Sammler der Daten und der Nutzerin bzw. dem Nutzer des KI-Modells hat selbstredend einen Einfluss vor allem auf die Dokumentations- bzw. Transparenz- und Nachverfolgungsmöglichkeiten der jeweiligen Parteien hinsichtlich der Trainingsdaten. Auf Grundlage der erhobenen Datenpunkte werden die LLM und LMM trainiert. In einem zweiten Schritt – dem sogenannten Fine-Tuning – trainieren die Entwicklerinnen und Entwickler das Modell auf spezifischeren Datensets, um das Modell für den jeweiligen Zweck funktionsfähig zu machen. Je nach Modell erfolgt auch sogenannten nachgelagerten Anbietern oder von Endnutzern ein weiteres bzw. eigenes Fine Tuning.

Zu beachten ist dabei stets, dass die Quantität, Qualität und Vielfalt der Daten direkte Auswirkungen auf die Leistungsfähigkeit und etwaige Verzerrungen des trainierten KI-Modells haben. Je höher die Quantität und Qualität der Daten, desto besser ist auch die Leistung des LLM/LMM. Ethische und datenschutzrechtliche Überlegungen und Aspekte spielen ebenfalls eine wesentliche Rolle für die Phase vor dem Training, da die gesammelten Daten personenbezogene Informationen enthalten können. Darüber hinaus kann ein Bias in den gesammelten Daten zu einem Bias im späteren Output des KI-Systems führen. Daher ist ein verantwortungsvoller Umgang mit den Daten unerlässlich, um einerseits ungewollten Bias zu minimieren und andererseits den Schutz personenbezogener Daten zu gewährleisten.

Insgesamt sind die Kuration und Erstellung eines Datenkorpus ein komplexer, zeit- wie kostenintensiver Prozess, der sorgfältige Planung und Kenntnis der jeweiligen Anforderungen erfordert.

Rechtliche Zulässigkeit von Scraping und Crawling

Vor dem Hintergrund der benötigten Datenmenge werden die meisten LLM und LMM heute weitgehend auf öffentlich zugänglichen Daten aus dem Internet trainiert. Zur Vorbereitung dieses Trainings auf öffentlich zugänglichen Daten aus dem Internet sowie zur Sammlung der Trainingsdaten wird oft auf Bots zurückgegriffen, die Web-Inhalte automatisiert herunterladen. Man spricht dabei vom sog. Web Crawling und Web Scraping.[2]

Sowohl das Herunterladen als auch das Zusammenführen in einen Datenkorpus stellen Vervielfältigungshandlungen dar. Zwar genießt bei Weitem nicht jeder Web-Inhalt urheber- oder leistungsschutzrechtlichen[3] Schutz[4], es können sich in den Web-Inhalten aber auch urheberrechtlich geschützte Werke bzw. Werkteile und -sammlungen befinden. Bei diesen stellt sich die Frage nach der urheberrechtlichen Zulässigkeit des automatischen Web Crawlings bzw. -Scrapings. Warum? Das Urheberrecht sieht in § 16 Abs. 1 Ziff. 1 UrhG vor, dass Vervielfältigungen von urheberrechtlich geschützten Werken grundsätzlich der Zustimmung des Urhebers bedürfen. Grundsätzlich, weil auch Ausnahmen gelten, bei welchen zustimmungsfrei urheberrechtlich relevante Handlungen vorgenommen werden dürfen. Man spricht dabei von Schranken.[5]

[2] Beim Scraping handelt es sich um einen gezielten Ansatz, um bestimmte Teile von Webseiten zu extrahieren und in ein neues Dateiformat zu überführen. Beim Crawling hingegen wird der gesamte Inhalt einer Webseite gelesen und gespeichert. (https://www.it-daily.net/it-management/big-data-analytics/web-scraper-web-crawler, zuletzt besucht am 27. Mai 2024)
[3] Bspw. das Leistungsschutzrecht des Datenbankherstellers; vgl. betreffend die verschiedenen Leistungsschutzrechte auch §§ 70 ff. UrhG
[4] Siehe zu den Voraussetzungen für urheberrechtlichen Schutz § 2 UrhG.
[5] Der Begriff der Schranke kommt daher, dass die Verwertungsrechte des Urhebers eingeschränkt werden. Siehe auch https://de.wikipedia.org/wiki/Schranken_des_Urheberrechts, zuletzt besucht am 27. Mai 2024)

Text and Data Mining (TDM)-Schranke (§ 44b UrhG)

Eine dieser Schrankenregelungen ist die Text and Data Mining-Schranke in § 44b UrhG („TDM-Schranke“). Diese Regelung geht auf Art. 4 der DSM-Richtlinie (Richtlinie (EU) 2019/790) zurück und trat zum 7. Juni 2021 in Deutschland in Kraft. Diese Vorschrift erlaubt die Vervielfältigung urheberrechtlich geschützter Werke für das Text- und Data Mining für kommerzielle Zwecke.[6] § 44b Abs. 1 UrhG definiert das Text- und Data Mining als „automatisierte Analyse von einzelnen oder mehreren digitalen oder digitalisierten Werken, um daraus Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen.“ Der Gesetzgeber entschied sich dafür, diese Schrankenregelung vergütungsfrei auszugestalten, sodass also weder eine Zustimmung des Urhebers notwendig, noch eine Vergütung an diesen geschuldet ist.

Beim Training von generativer KI handelt es sich um Text- und Data Mining[7]. Genau wie beim Text- und Data Mining geht es beim Training von generativer KI nicht darum, über die Vervielfältigung Zugang zu dem zu erhalten, was das Urheberrecht schützt: die Ausdrucksform. Es geht vielmehr darum, über das Mining an die zugrundeliegende Information – wie Ideen, Fakten, Strukturen, Muster, Logiken und Zusammenhänge – zu gelangen. Dabei ist wesentlich zu wissen und zu verstehen, dass das Urheberrecht diese Information – wie bspw. auch den künstlerischen Stil – nicht schützt, sondern nur die konkrete, wahrnehmbare Ausdrucksform. Das Urheberrecht schützt also lediglich die Form, mit der die Information kommuniziert wird, nicht die Information selbst.[8]

Das Gewinnen dieser Information ist Ziel des Trainings von generativer KI und erfolgt als Form des Text- und Data Minings. Es existieren jedoch Meinungen, dass das Training generativer KI nicht von der Text- und Data Mining-Schranke erfasst sei. Diese Meinungen dürften heute die Mindermeinung darstellen. Gute Gründe und damit die herrschende Lehre gehen davon aus, dass die TDM-Schranke auf das Training generativer KI anwendbar ist: Die Erwägungsgründe zur DSM-Richtlinie führen aus, dass Text- und Data Mining vor allem zur Entwicklung neuer Anwendungen oder Technologien eingesetzt werde. Diese Entwicklung sollte mit der Richtlinie gerade gefördert werden, indem Innovationsanreize gesetzt und Rechtssicherheit geschaffen werden. Das wird zusätzlich dadurch unterstrichen, dass verschiedene Positionierungen im Hinblick auf die DSM-Richtlinie die Berücksichtigung von Datengewinnung für die KI-Entwicklung als Anwendungsbereich der Schrankenbestimmung forderten.[9] Darüber hinaus verweist der AI Act im Rahmen des Trainings von KI-Modellen auf die TDM-Schranke nach Art. 4 der DSM-Richtlinie bzw. die damit zusammenhängenden Möglichkeiten zum Opt-out (siehe dazu sogleich vertiefter).

[6] Auch die Schranke in § 60d UrhG erlaubt das Text and Data Mining. §60d UrhG für wissenschaftliche Zwecke und ist weitreichender als § 44b UrhG.
[7] Statt vieler: De la Durantaye, ZUM 2023, 651
[8] Statt vieler: De la Durantaye, ZUM 2023, 651
[9] Erwgr. 18, S. 1 & 4 DSM-Richtlinie; vgl. auch de la Durantaye, ZUM 2023, 651.

Voraussetzungen TDM-Schranke nach § 44b UrhG

Rechtlich zulässiges Text- und Data Mining setzt gemäß § 44b Abs. 2 UrhG erstens voraus, dass die Nutzerin bzw. der Nutzer (die Person, die das Mining durchführt) rechtmäßigen Zugang zu den Daten hatte. Das dürfte bei frei im Internet zugänglichen Informationen regelmäßig der Fall sein. Zweitens sind die Daten zu löschen, wenn sie für das Text and Data Mining nicht mehr erforderlich sind. Die dritte Voraussetzung ist eine Negative: Die am Werk berechtigte Person darf nicht der Nutzung ihres Werkes zum Text und Data Mining widersprochen haben (sog. Opt-out bzw. Nutzungsvorbehalt). Bei online zugänglichen Werken muss dieser Nutzungsvorbehalt in maschinenlesbarer Form erfolgen.

Möglichkeiten von Urheberinnen und Urhebern bzw. Rechteinhaberinnen und Rechteinhabern

Genau diese Möglichkeit zum Nutzungsvorbehalt ist der Steigbügel für Urheberinnen und Urheber bzw. Rechteinhaberinnen und Rechteinhaber, die nicht wollen, dass ihr online zugängliches Werk für Text- und Data Mining genutzt wird: Durch Erklärung des Nutzungsvorbehalts signalisieren sie, dass ihr einzelnes Werk oder ihre Werke nicht für das Training von generativer KI verwendet werden soll. Deshalb sieht das Gesetz in § 44b Abs. 3 UrhG auch vor, dass der Nutzungsvorbehalt bei online zugänglichen Werken in maschinenlesbarer Form erfolgt. Der Nutzungsvorbehalt muss deshalb maschinenlesbar sein, damit der Vorbehalt von den Bots, die beim Crawling und Scraping eingesetzt werden[10], erkannt und beachtet werden kann. Trifft der Bot beim Crawling oder Scraping auf einen maschinenlesbaren Nutzungsvorbehalt, verwendet er das betroffene Datum oder die betroffenen Daten nicht.

Sind vom Opt-out Werke erfasst, die den Entwicklern der KI-Technologie als für das Training relevant genug erscheinen, können die Parteien – Entwickler und Urheber bzw. Rechteinhaber – in Verhandlungen treten und die Entwickler können sich die entsprechenden Rechte u. a. vertraglich einräumen lassen und das Werk bzw. die Werke entsprechend zum Training verwenden. Dabei ist in Mehrparteienverhältnissen[11] selbstverständlich wesentlich, dass jede Partei den Nutzungsvorbehalt (erneut) äußert, damit keine Lücke entsteht. Für die Äußerung des maschinenlesbaren Nutzungsvorbehaltes bestehen zahlreiche technisch mehr oder minder einfach umsetzbare Möglichkeiten. Einige davon setzen am Werk selbst an, andere wiederum an dem „Ort“, an dem die Datei hinterlegt ist, zB die Domain, spezifische URL, das Verzeichnis, Dateiformat etc.[12]

Dabei stellt die robots.txt-Datei eine einfach umzusetzende und wirksame Variante dieser Möglichkeiten dar, die an den Ort des Datums anküpft. Da sie an der Domain ansetzt, kann sie Bots gänzlich davon abzuhalten, den Inhalt der Webseite (des Hosts) als Ganzes oder bestimmte Unterseiten oder Verzeichnisse zu scrapen und zu crawlen. Anweisungen in der robots.txt-Datei können sich auf bestimmte Crawler oder auf alle

[10] Vgl. oben.
[11] Bspw. Urheberschaft-Verlag-Contentprovider
[12] Entsprechend der gewünschten Granularität.

nicht näher bezeichneten Crawler beziehen. Da sie dem Robots Exclusion Standard entspricht, ist sie nicht nur vergleichsweise einfach zu implementieren, sondern auch effektiv und transparent.[13]

Jedenfalls können Urheberinnen und Urheber bzw. Rechteinhaberinnen und Rechteinhaber[14] speziell darauf achten, eine der Lösungen zu wählen, die keine Suchmaschinen-Crawler ausschließt bzw. dass diese Crawler dann explizit wieder zugelassen werden. So können sie verhindern, dass der Nutzungsvorbehalt einen negativen Einfluss auf das Suchmaschinenranking des Werkes hat.

Trotz der Transparenz und Klarheit, die Opt-out-Lösungen, z. B. auf der Grundlage von robots.txt, mit sich bringen, sind einige Rechteinhaber nach wie vor besorgt darüber, ob ihre Opt-outs respektiert wurden. Wie können Urheberinnen und Urheber also nachvollziehen, ob ihr Nutzungsvorbehalt respektiert wurde? Bei den KI-Entwicklern nachzufragen bzw. eine Liste der Trainingsdaten zu verlangen, ist aufgrund der Datenmengen nicht nur technisch nicht umsetzbar, sondern ist vor allem deshalb für die Urheberinnen und Urheber nicht zielführend, weil sie innerhalb eines immens großen Datensatzes mit Milliarden bis Billionen Datenpunkten[15] ihr einzelnes Werk und damit ihren einzelnen Datenpunkt heraussuchen müssten.[16] Hinzu kommt der Umstand, dass die Vorstellung in den meisten Fällen unzutreffend ist, dass es sich bei dem Anbieter der KI-Technologie auch um die Person handelt, die die Trainingsdaten gesammelt und aufbereitet hat – in manchen Fällen wird, wie schon zuvor erwähnt, auf sogenanntes Prepackaging zurückgegriffen. In solchen Fällen werden vorgefertigte (bereits gescrapte/gecrawlte) Datensets aus bspw. dem sogenannten Common Crawl verwendet und die Technologie nach dem Training mit einem solchen Prepackage mit spezielleren bzw. für den Verwendungszweck geeigneteren Daten trainiert und verfeinert (Fine-Tuning). Darüber hinaus können auch Entwickler und Anbieter auseinanderfallen.

Über die Äußerung des Nutzungsvorbehaltes lässt sich im Onlinezusammenhang einfacher Transparenz für die Urheberinnen und Urheber sowie Rechteinhaber und Rechteinhaberinnen herstellen: Webseitenbesuche – egal ob Mensch oder Bot – werden geloggt.[17] Seriöse Crawler verwenden klar erkennbare Namen ("HTTP User-Agent"-Texte), die es dem Webseitenbetreiber ermöglichen, sie zu ihrer Quelle zurückzuverfolgen. Die meisten seriösen Crawler kommen von IP-Adressen, die unmissverständlich auf ihre Quelle verweisen, so dass ein Webseitenbetreiber ihre Berechtigung überprüfen kann. Sowohl der HTTP- User-Agent als auch die IP-Adresse werden in der Regel für Webanfragen protokolliert und ermöglichen es dem Webseitenbetreiber zu überprüfen, welche Seiten von einzelnen Crawlern angefordert wurden. So können Webseitenbetreiber überprüfen, ob ihre robots.txt-Richtlinien

[13] Dabei wird die robots.txt-Datei ins Stammverzeichnis der Webseite hochgeladen. Der Befehl kann entweder enumerativ die einzelnen Bots ausschließen, die ausgeschlossen werden sollen oder aber mit dem Setzen eines Sternchens (*) beim user-agent werden sämtliche Bots ausgeschlossen. Hier können über Rückausnahmen unter „allow:“ in der Datei wiederum einzelne Bots zum crawling und scraping zugelassen werden – bspw. Suchmaschinencrawler, damit das Werk weiterhin in den Suchmaschinen indexiert werden.
[14] Je nach Einzelfall die zur Äußerung des opt-outs berechtigte Person.
[15] S. oben.
[16] Darüber hinaus würde eine solche Werk-für-Werk Listenlösung auch signifikante Risiken hinsichtlich der Modellsicherheit und den Schutz der Geschäftsgeheimnisse von Modellentwicklern bergen.
[17] https://www.crowdstrike.com/cybersecurity-101/observability/web-server-logs/

befolgt werden. Als Webseitenbetreiber ist es also auch nachvollziehbar, ob ein Bot die Webseite besucht hat, die er nicht hätte besuchen sollen/dürfen. Die Webseitenbetreiber verfügen mit dem Webserverlog damit über die Information, ob sich an einen Nutzungsvorbehalt gehalten wurde, indem sie nachvollziehen können, ob ein Bot oder Crawler die Webseite bzw. Web-Unterseite besucht hat, obwohl Bots und Crawler ausgeschlossen waren.

Es wird demnach deutlich, dass der Gesetzgeber durch die gesetzlichen Regelungen bereits einen sachgerechten Ausgleich zwischen den Interessen der Urheberinnen und Urheber und den Interessen der Gesellschaft und Industrie an Innovation geschaffen hat. Die bestehenden technischen Möglichkeiten erlauben es dem Urheber und Rechteinhaber zudem, einen Nutzungsvorbehalt zu erklären, der auch technisch im Rahmen des Web Scrapings oder -Crawling berücksichtigt werden kann. Darüber hinaus liegt Transparenz mit den bestehenden technischen Möglichkeiten bereits ausreichend vor, die durch zusätzliche Transparenzvorschriften im AI Act[18] noch ausgebaut wurden. Weitere Transparenzvorschriften, für die etwa gar nicht die Datenstrukturen vorliegen, sind also nicht nur nicht zielführend[19], sondern auch nicht notwendig.

2 Während des Trainings

Parametrisierung

Das Training eines LLM/LMM ist ein komplexer Prozess, der in verschiedenen Phasen abläuft und dabei eine Reihe technischer Herausforderungen mit sich bringt. Im Kern des Trainings steht die Parametrisierung der gesammelten und bereinigten Daten. Bei der Parametrisierung passt das Modell seine Parameter an, um die gegebenen Daten bestmöglich auf Bedeutungsebene zu repräsentieren und Muster sowie Zusammenhänge zu lernen und zu identifizieren. Dies geschieht implizit, das heißt, die Daten werden nicht explizit in Form von wahrnehmbaren expliziten Kopien gespeichert (etwa in Form von Bits, Codes oder gar Text), sondern ihr Inhalt, bzw. ihre Struktur und Zusammenhänge werden semantisch auf einer Bedeutungsebene und implizit in den Gesamtparametern des Systems repräsentiert. Dieser Prozess wird durch Optimierungsalgorithmen ermöglicht, die die Fehler zwischen den vom Modell vorhergesagten und den tatsächlichen Werten minimieren. Parameter sind die einstellbaren Elemente in einem Modell, die auf Basis der Trainingsdaten justiert werden. Sie bilden somit die Wissensbasis des Modells. Dazu gehören z. B. Gewichte in neuronalen Netzen und Einstellungen in maschinellen Lernalgorithmen. Parameter

[18] Wie die „sufficiently detailed summary" bzw. die „hinreichend detaillierte Zusammenfassung der für das Training des KI-Modells mit allgemeinem Verwendungszweck verwendeten Inhalte." Nach Artikel 53 Abs. 1 Bst. d) AI Act.
[19] Siehe bspw. oben zur geringen Sinnhaftigkeit umfänglicher Trainingsdatenlisten.

beeinflussen das spätere Verhalten von KI-Modellen und bestimmen, wie sie Vorhersagen oder Entscheidungen treffen.

Phasen des Trainings

Das Training von LLM/LMM kann dabei in verschiedene Phasen unterteilt werden: Im sogenannten „Pre-Training“ wird ein Sprachmodell allgemein auf Texte trainiert, ohne dass eine spezifische Aufgabe zu lösen ist. Beim „Supervised Fine-Tuning“ wird das Modell auf spezifische Aufgaben vorbereitet und das Modell-Wissen auf diese angepasst. Das Reinforcement Learning basiert auf menschlichem Feedback, in dem KI-Trainer die Antworten des KI-Modells bewerten und das KI-Modell durch dieses Feedback lernt und sich verbessert.[20]

Vor diesem Hintergrund ist ein zentraler Aspekt beim Training von LLMs und LMMs die Loslösung des Trainingsdatenkorpus vom Trainingsergebnis. Während des Trainings findet keine Vervielfältigung der Trainingsdaten oder andere urheberrechtliche relevante Nutzungshandlung statt. Das bedeutet, dass im Rahmen des Trainings keine Trainingsdaten in das LLM kopiert und dort eben nicht wie bei einer Datenbank strukturiert oder unstrukturiert abgelegt werden. Stattdessen geht es beim Training darum, Zugang zu den hinter den Daten liegenden, schutzfreien Inhalten, wie Strukturen, Mustern, Ideen und Fakten, zu erhalten und diese zur Entwicklung des KI-Modells zu nutzen. Eine urheberrechtlich relevante Vervielfältigung von Daten liegt gerade nicht vor.

Zusammenfassend ist das Training eines LLMs bzw. LMMs ein vielschichtiger Prozess, der technisches Know-how, sorgfältige Planung und ein Verständnis für rechtliche Rahmenbedingungen erfordert. Durch die Berücksichtigung dieser Aspekte können Entwickler leistungsfähige, faire und rechtlich konforme KI-Systeme schaffen.

3 Nach dem Training

Nach dem Training bzw. im Betriebsmodus ergeben sich nur in Ausnahmefällen urheberrechtlich relevante Vorgänge bzw. Urheberrechtsverletzungen. Dies liegt an der Funktionsweise generativer KI-Technologien: Es handelt sich bei KI-Modellen im Wesentlichen um statistische Modelle. Sie geben Information nicht wieder, sondern ermitteln ausgehend von der Eingabe der Nutzerin oder des Nutzers Wahrscheinlichkeiten für bestimmte Ergebnisse. Hierfür erschließen die KI-Systeme Zusammenhänge aus den verschiedenen im Training gewonnen Informationen auf statistischer Basis. Im Rahmen des Outputs geben die KI-Systeme die Trainingsinhalte jedoch nicht wieder, sondern ermitteln vielmehr, welches Ergebnis auf Grundlage der verschiedenen Trainingsinhalte am wahrscheinlichsten ist. Wie bereits erläutert,

[20] Daneben existieren noch weitere Trainingsformen und Mischformen.

werden diese Trainingsinhalte nicht im Modell gespeichert. D. h. es wird ein Output generiert, der auf den Stilen, Zusammenhängen, Strukturen bzw. die Ideen der ursprünglichen Werke basiert. Künstlerischer Stil und abstrakte Ideen sind aber urheberrechtlich gerade nicht geschützt.[21] Deshalb liegt bei KI-Output im Regelfall keine Urheberrechtsverletzung vor.

Schafft die KI einen Output, der einem bestehenden Werk ähnlich ist, ohne mit diesem Werk trainiert worden zu sein, so liegt ebenfalls keine urheberrechtliche Verletzung vor – es handelt sich um eine (zulässige) sogenannte Doppelschöpfung. Dabei hat aber die anwendende Person die Doppelschöpfung – und damit das Nichtvorhandensein des Werkes in den Trainingsdaten – zu beweisen.[22]

Einzelfall „Erinnerung“

Sollte es sich beim Werk ausnahmsweise aber um eines handeln, mit dem die KI trainiert wurde – „erinnert“ sich also die KI gewissermaßen an das Werk und gibt es 1:1 und nicht bloß stilistisch wieder, ist eine genauere Betrachtung notwendig:

Bei dieser Betrachtung sind die wiedererkennbaren Merkmale des ursprünglichen Werkes im KI-Output entscheidend. Rechtlich ist dabei dreistufig vorzugehen:[23]

1. Liegen Ähnlichkeiten zwischen KI-Output und Originalwerk vor?
2. Handelt es sich bei den Ähnlichkeiten um Werkteile des Originals, die urheberrechtlichen Schutz genießen?
3. Handelt es sich bei diesen urheberrechtlich geschützten Merkmalen im Originalwerk um Merkmale, die auch im KI-Output erkennbar sind?

Erst wenn diese drei Fragen mit Ja beantwortet werden können, kommt eine Urheberrechtsverletzung unter Umständen infrage. Der BGH nimmt dabei eine Gesamtbetrachtung vor. Der KI-Output muss also in seiner Gesamtheit betrachtet werden. Abweichungen in anderen Werkteilen sprechen gegen eine Urheberverletzung insgesamt.

Selbstverständlich handelt es sich um eine Urheberrechtsverletzung, wenn die nutzende Person Eingabeinformationen in die KI so gestaltet, dass absichtlich eine Kopie hergestellt wird.

Liegt eine Urheberrechtsverletzung vor, stehen der berechtigten Person Unterlassungs- und u. U. auch Schadensersatz- und Auskunftsansprüche zu.

[21] Bspw. Baumann, NJW 2023, 3673
[22] Aufgrund der Menge der Trainingsdaten sowie den Mehrparteienverhältnisse dürfte ein solcher Beweis nicht gerade leicht zu erbringen sein.
[23] Baumann, NJW 2023, 3673

Zusammenfassung

Das bestehende Urheberrecht bietet mit § 44b UrhG[24] eine Grundlage, die das Training generativer KI-Modelle mit urheberrechtlich geschützten Werken für kommerzielle Zwecke grundsätzlich ermöglicht, indem es die zur Trainingsvorbereitung notwendige Vervielfältigung von urheberrechtlich geschützten Werken zustimmungsfrei erlaubt. Der AI Act verweist darüber hinaus auf Art. 4 DSM-Richtlinie, der § 44b UrhG zugrunde liegt.

Möchten Urheber oder Rechteinhaber nicht, dass ihr Werk für das Training von generativer KI verwendet wird, es aber dennoch online zur Verfügung stellen, müssen sie nach § 44b Abs. 3 UrhG einen maschinenlesbaren Nutzungsvorbehalt wirksam äußern. Die Entwickler von KI-Modellen müssen diesen Nutzungsvorbehalt respektieren und dürfen ihn nicht umgehen und das Werk dennoch zur Trainingsvorbereitung vervielfältigen. Soll das Werk trotz Vorbehalt Teil des Trainingsdatenkorpus werden, kann der Entwickler in Lizenzverhandlungen mit dem Rechteinhaber treten.

Über das Webseitenlogging im Zusammenspiel mit den Transparenzpflichten im AI Act verfügen die Webseitenbetreiber zudem über die Information, ob ein Bot die Webseite besucht hat und auf welche IP der Bot zurückzuführen ist. Mit anderen Worten können die Webseitenbetreiber die notwendige Transparenz für die Urheber schaffen und auch aufgrund der Transparenzvorschriften im AI Act können diese nachvollziehen, ob ihr Nutzungsvorbehalt respektiert wurde und gegen unrechtmäßige Verwendungen ihres geschützten Werkes vorgehen.

Nach dem Training bzw. im Betriebsmodus ergeben sich nur ganz ausnahmsweise urheberrechtlich relevante Vorgänge. Dies liegt an der Funktionsweise generativer KI-Technologien – die in ihrer Natur keine Kopiermaschinen sind.

[24] Und mit § 60b Urhg für wissenschafltiche Zwecke im Besonderen.

Bitkom vertritt mehr als 2.200 Mitgliedsunternehmen aus der digitalen Wirtschaft. Sie generieren in Deutschland gut 200 Milliarden Euro Umsatz mit digitalen Technologien und Lösungen und beschäftigen mehr als 2 Millionen Menschen. Zu den Mitgliedern zählen mehr als 1.000 Mittelständler, über 500 Startups und nahezu alle Global Player. Sie bieten Software, IT-Services, Telekommunikations- oder Internetdienste an, stellen Geräte und Bauteile her, sind im Bereich der digitalen Medien tätig, kreieren Content, bieten Plattformen an oder sind in anderer Weise Teil der digitalen Wirtschaft. 82 Prozent der im Bitkom engagierten Unternehmen haben ihren Hauptsitz in Deutschland, weitere 8 Prozent kommen aus dem restlichen Europa und 7 Prozent aus den USA. 3 Prozent stammen aus anderen Regionen der Welt. Bitkom fördert und treibt die digitale Transformation der deutschen Wirtschaft und setzt sich für eine breite gesellschaftliche Teilhabe an den digitalen Entwicklungen ein. Ziel ist es, Deutschland zu einem leistungsfähigen und souveränen Digitalstandort zu machen.

Herausgeber
Bitkom e.V.
Albrechtstr. 10 | 10117 Berlin

Ansprechpartner
Dr. Pablo Schumacher | Referent für Digital Content & Recht
T 030 27576-289 | p.schumacher@bitkom.org

Verantwortliches Bitkom-Gremium
AK Intellectual Property

Titelbild
stock.adobe.com – eteri

Copyright
Bitkom 2024

Diese Publikation stellt eine allgemeine unverbindliche Information dar. Die Inhalte spiegeln die Auffassung im Bitkom zum Zeitpunkt der Veröffentlichung wider. Obwohl die Informationen mit größtmöglicher Sorgfalt erstellt wurden, besteht kein Anspruch auf sachliche Richtigkeit, Vollständigkeit und/oder Aktualität, insbesondere kann diese Publikation nicht den besonderen Umständen des Einzelfalles Rechnung tragen. Eine Verwendung liegt daher in der eigenen Verantwortung des Lesers. Jegliche Haftung wird ausgeschlossen. Alle Rechte, auch der auszugsweisen Vervielfältigung, liegen beim Bitkom oder den jeweiligen Rechteinhabern.