Themenschwerpunkt KI

David Rosenthal und Livio Veraldi

Training von KI-Sprachmodellen: Was das geltende Urheberrecht & Co. erlauben

Training AI language models: How to Handle Copyright and Other Restrictions

https://doi.org/10.1515/abitech-2025-0064

Zusammenfassung: KI-Modelle und allen voran große Sprachmodelle (sog. Large Language Models – LLM) müssen im Rahmen ihres Trainings viele Daten sehen, damit sie die für ihre Funktion nötigen Konzepte in ihren Parametern abbilden können. Hierfür werden vor allem öffentlich zugängliche und zwangsläufig auch urheberrechtlich geschützte Inhalte (Werke) verwendet – in der Regel ohne Zustimmung der Rechteinhaber*innen. Doch ist das erlaubt? Im Schweizer Urheberrecht müssen dafür zwar spezielle Pfade beschritten werden, grundsätzlich lautet die Antwort aber: ja. Die Schweiz ist – Stand heute – ein gutes Pflaster für das Training solcher KI-Modelle. Gewisse Kreise wollen das allerdings ändern.

Schlüsselwörter: Training von KI-Sprachmodellen, Urheberrecht, IT-Recht

Abstract: AI models – especially Large Language Models (LLM) – need to see a lot of data as part of their training so that they can map the concepts required for their function within their parameters. For this purpose, the companies that train the LLM primarily use publicly accessible content, and they do so without asking for permission. But is this allowed? Under Swiss copyright law the answer is a positive one, based on various legal concepts, including limitations in the scope of copyright and an exemption for research. There are, however, initiatives on the way to change this – initiatives that some believe could have serious negative effects for AI use in Switzerland.

Keywords: Training of AI language models, Copyright law, IT law

1 Einleitung

Die Vorstellung, dass KI-Sprachmodelle sich alle Inhalte merken, die ihnen beim Training gezeigt werden, ist ein großes Missverständnis. Das ist weder das Ziel noch wäre es besonders effizient. Vielmehr merkt sich ein Modell beim Training jene sprachlichen und inhaltlichen Konzepte, die es besonders häufig sieht. Dabei kann es durchaus zur inhaltlichen oder gar wortwörtlichen „Memorisierung“ kommen, wenn das Modell bspw. immer wieder demselben Slogan begegnet oder eine bestimmte Information in zahlreichen Unterlagen immer wieder sieht. Im Falle der wortwörtlichen Memorisierung ist das KI-Sprachmodell in der Lage, die Trainingsinhalte in mehr oder weniger gleicher Form wiederzugeben. Die Memorisierung von einzelnen Inhalten stellt allerdings nicht den Normalfall dar; schließlich soll ein Modell neue Inhalte generieren, nicht einzelnes Gesehenes rezitieren. Doch insbesondere dort, wo es zur Memorisierung kommt, stellt sich die Frage, was dies aus rechtlicher Sicht für das Training und die Weitergabe solcher Modelle bedeutet.

2 Geltendes Urheberrecht

Die Frage, ob und gegebenenfalls unter welchen Voraussetzungen urheberrechtlich geschützte Werke beim Training von KI-Sprachmodellen unter geltendem Recht verwendet werden dürfen, wird seit einiger Zeit intensiv diskutiert und ist umstritten. Unsere Analyse des geltenden Urheberrechts, deren Details wir in einem kürzlich erschienenen wissenschaftlichen Beitrag[1] ausführlich dargelegt haben, ergab jedoch, dass sich die Zulässigkeit des Trainings etwa

1 Veraldi, Livio und David Rosenthal. „Das Training von KI-Sprachmodellen mit fremden Inhalten und Daten aus rechtlicher Sicht.“ In: *Jusletter* 3 (3. Februar 2025). https://www.rosenthal.ch/downloads/Rosenthal-Veraldi_Training_LLM_Schweizer_Recht_Jusletter.pdf. Zuletzt geprüft am 11.09.2025.

von großen Sprachmodellen mit öffentlichen Inhalten mit und ohne Memorisierung auf drei Ebenen ergeben kann, die hier im Folgenden erläutert werden.

2.1 Das Training stellt keine urheberrechtlich relevante Handlung dar.

Diese – auf den ersten Blick kontraintuitive Schlussfolgerung – ergibt sich in einem ersten Ansatz, wenn wir uns bewusst werden, dass das Urheberrecht nur Werkverwendungen erfasst, die letztlich dem Werkgenuss, also der sinnlichen Wahrnehmung von Werken durch Menschen, dienen. Dies geschieht jedoch beim Training mit urheberrechtlich geschützten Werken gerade nicht. Auch im Falle einer Memorisierung liegen diese Werke im Modell nicht in einer Form vor, die den Werkgenuss ermöglicht. Erst in Kombination mit einem passenden Prompt ist es denkbar, dass ein bestimmter Output erzeugt wird, bei dem ein Werkgenuss überhaupt ein Thema sein kann. In dem Falle werden die Werke aber – wenn überhaupt – nicht einfach aus dem Modell abgerufen, sondern vielmehr neu generiert. Aus urheberrechtlicher Perspektive ist mithin von entscheidender Bedeutung, dass die Werke beim Training von KI-Modellen nicht verwendet werden, um sie später als solche aus dem Modell abzurufen und zu genießen, sondern um die in den Werken enthaltenen sprachlichen und inhaltlichen Konzepte zu analysieren und im Modell zu repräsentieren, um dieses sodann zu befähigen, aufgrund eines Prompts selbst neue Werke zu generieren.

Ein zweiter, damit zusammenhängender Ansatz basiert auf der Erkenntnis, dass das, was nach dem Training im Modell enthalten ist, nicht mehr als Werk gelten kann, selbst wenn es zu einer Memorisierung kommt. Das liegt daran, dass jedes für das Training verwendete Werk in seine (urheberrechtlich nicht mehr relevanten) Bestandteile zerlegt wird, und diese Bestandteile mit den Bestandteilen aller anderen zerlegten Inhalte quasi zu einer Suppe vermischt werden. Auch hier vermag die Tatsache, dass mit dem passenden Prompt die einzelnen Teile unter Umständen wieder zu einem Werk zusammengefügt werden können, an der rechtlichen Beurteilung des Trainings nichts zu ändern. Ein solches Zusammenfügen bedeutete vielmehr ein neues und eigenständiges Generieren eines Werks. Selbst wenn davon auszugehen wäre, dass im Modell ein Werk gewissermaßen auf einer Meta-Ebene noch vorhanden ist, müsste zugestanden werden, dass ein solches Werk bei der Betrachtung eines großen Sprachmodells als Ganzes so sehr in den Hintergrund rückt, dass es in Bezug auf seine individuellen Züge verblasst und daher urheberrechtlich nicht mehr relevant ist. Mindestens aber wird das Vorkommen des Werks im Modell als Repräsentation seiner sprachlichen und inhaltlichen Konzepte den nötigen „inneren Abstand“ zum Originalwerk aufweisen, um aus diesem Grund nicht mehr urheberrechtlich relevant zu sein.

Ein dritter Ansatz geht dahingehend, dass wir es mit dem eigentlichen „Werkgenuss“ der KI zu tun haben, der – analog zum menschlichen Werkgenuss – urheberrechtlich freigestellt ist. Wie auch ein Mensch zuerst diverse Werke konsumiert haben muss, um überhaupt im Stande zu sein, eigene Werke zu schöpfen, ist auch die KI darauf angewiesen, im Rahmen des Trainings (mitunter urheberrechtlich geschützte) Inhalte zu konsumieren, um überhaupt einen Output generieren zu können. Die Frage, wie ein so trainiertes Modell später benutzt wird, ist – wie beim Menschen – eine davon zu trennende Frage.

2.2 Es greift (meist) die Wissenschaftsschranke

Selbst wenn das Training eines großen Sprachmodells als urheberrechtlich relevante Vervielfältigung zu qualifizieren ist, wird es regelmäßig durch die Schranke für die Verwendung von Werken zum Zweck der wissenschaftlichen Forschung (Art. 24d URG) freigestellt sein. Diese Schranke setzt zwar einen rechtmäßigen Zugang zu den Inhalten voraus, aber etwaige vertragliche Beschränkungen, die eine Verwendung zu Trainingszwecken untersagen, dürften regelmäßig unwirksam sein. Zudem gibt es in der Schweiz kein Opt-out-Recht, wie es das EU-Recht mit der „Text and Data Mining“-Regelung (TDM) vorsieht. Die Kernfrage im Rahmen der Wissenschaftsschranke, nämlich die Frage, ob das Training eines Sprachmodells einen Forschungszweck darstellt, ist nach unserem Dafürhalten klar zu bejahen, wird Forschung doch als systematische, methodische Suche nach neuen Erkenntnissen verstanden. Genau darum geht es beim Training eines KI-Modells: Es sollen Erkenntnisse aus den dem Modell vorgelegten Inhalten gewonnen werden. Kommerzielle Forschung ist in der Schweiz im Übrigen mitgemeint. Allerdings gilt die Wissenschaftsschranke nicht für Computerprogramme.

Die andere Schrankenbestimmung, die greifen kann, ist der betriebliche Eigengebrauch (Art. 19 Abs. 1 lit. c URG), der Vervielfältigungen zur betriebsinternen Information und Dokumentation erlaubt. Eine Herausforderung ist dabei allerdings, dass diese Schranke für bestimmte Werkkategorien nicht zur Verfügung steht und im Handel erhältliche Werkexemplare für den betrieblichen Eigengebrauch nicht vollständig oder weitgehend vollständig vervielfältigt werden dürfen. Letzterem kann in der Praxis mit Anti-Memorisierungs-Techniken begegnet werden, die für einen

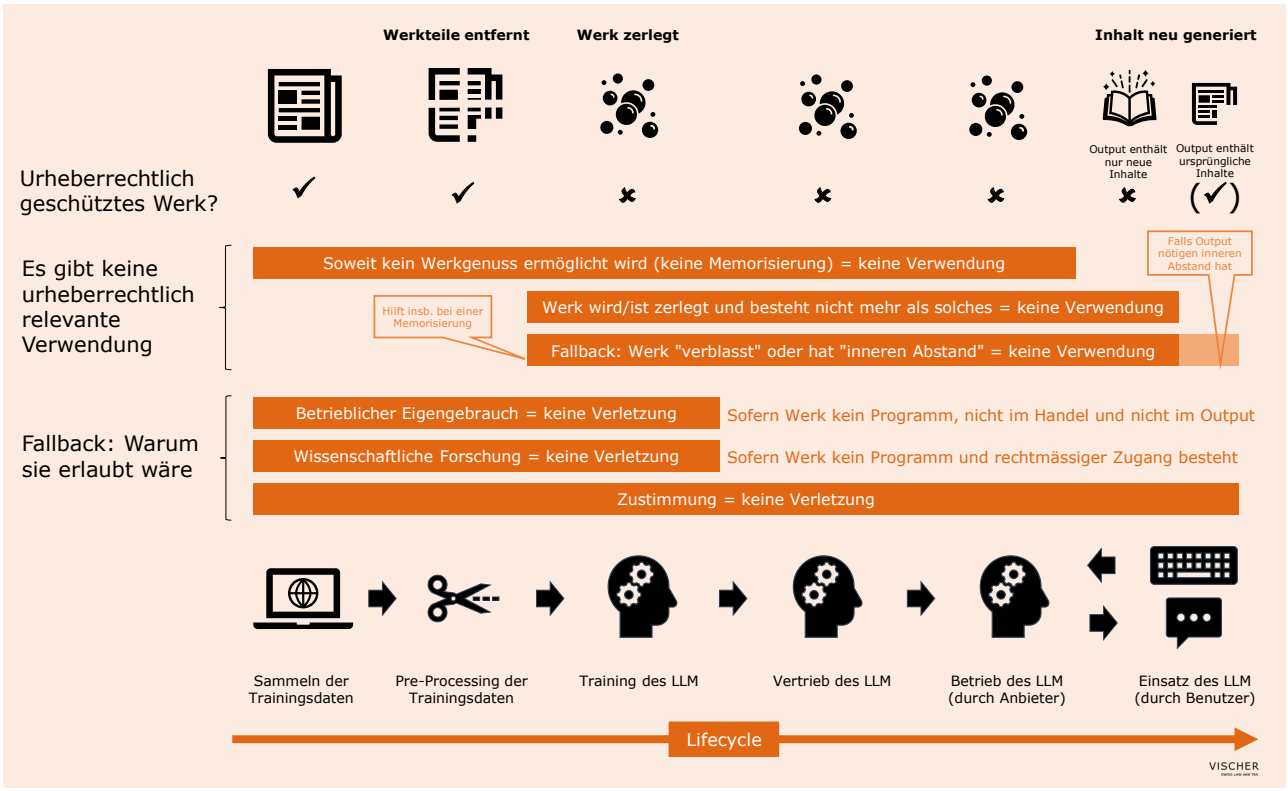

Abb. 1: Training eines LLM aus Sicht des Schweizer Urheberrechts (Quelle: David Rosenthal)

menschlichen Nutzer*innen relevante Teile eines Texts entfernen, ohne seinen Wert für die Maschine wesentlich zu beeinträchtigen.

2.3 Es liegt (oft) eine implizite Zustimmung der Rechteinhaber*innen vor

Eine implizite Zustimmung der Rechteinhaber*innen kann dort gegeben sein, wo Inhalte im Internet ohne entsprechende Gegenmaßnahmen publiziert werden. Es kann vertreten werden, dass inzwischen damit gerechnet werden muss, dass sie für KI-Trainings verwendet werden. Diese Rechtsgrundlage ist in der Praxis jedoch schwach, da ihr Rechteinhaber*innen aktiv entgegenwirken können. Das geht mangels gesetzlicher Regelung nicht nur über die oft zitierten maschinenlesbaren Opt-outs, sondern schon mit bloßen Hinweisen auf der Website. Zudem kann von einer impliziten Zustimmung nur dort ausgegangen werden, wo Rechteinhaber*innen den Inhalt selbst publiziert haben. Alle, die ihr KI-Modell mit Internet-Inhalten trainieren wollen, wissen also bei der schieren Menge an Websites nie, woran sie sind.

2.4 EU-Recht und Schweizer Recht

Das Schweizer Urheberrecht weist große Parallelen auf zu den urheberrechtlichen Harmonisierungsbestrebungen in der EU, sodass die vorstehenden Überlegungen auch für die Beurteilung der Rechtslage in der EU Anregungen liefern können. Eine Ausnahme ist immerhin der Entscheid der Schweiz, die TDM-Regelung und insbesondere das Opt-out-Recht nicht nachzubilden. Das macht den Standort Schweiz für das Training von KI-Sprachmodellen attraktiver als die EU. Daran ändern auch die Vorgaben des EU AI Act für Allzweck-KI-Modelle nichts. Der EU-Gesetzgeber hatte zwar im Sinn, eine Regelung zu erlassen, die Nicht-EU-Anbieter von solchen Modellen zwingen sollte, sich beim Trainieren ihrer Modelle an die TDM-Regelung zu halten. Beim Verfassen der einschlägigen Bestimmung (Art. 53 Abs. 1 Bst. c AI Act) wurde jedoch nachlässig gearbeitet und übersehen, dass das EU-Urheberrecht bei einem Training in der Schweiz (oder in den USA) nicht zur Anwendung kommen wird. Daher kann das Training in der Schweiz unter Nicht-Beachtung der TDM-Regelung das EU-Urheberrecht nicht verletzen.

3 Ist das geltende Urheberrecht anzupassen?

Das Urheberrecht sorgt für die meisten Diskussionen, was vor allem daran liegen dürfte, dass es hier auf den ersten Blick um besonders viel geht: Inhaber*innen von Inhalten und Kreativschaffende wollen nicht, dass die Früchte ihrer Arbeiten ohne ihre explizite Zustimmung und vor allem ohne Bezahlung einer Vergütung durch die großen KI-Player verwendet werden. Den Kreativschaffenden dürfte es aber auch um die KI als Konkurrenz gehen, die Inhalte günstiger herzustellen vermag und dadurch auf Kreativschaffende wirtschaftlichen Druck ausübt. Das tun andere Künstler*innen zwar auch, und auch sie müssen anhand bestehender Werke lernen, wie dies geht bzw. werden in ihrem Wirken von bestehenden Werken inspiriert, aber das geschieht alles mit sehr viel weniger Hebelwirkung und ist weniger effektiv.

Letztere Bedenken betreffen im Kern jedoch nicht die Frage der Verletzung des Urheberrechts, sondern sind wirtschaftspolitischer Natur – jedenfalls insoweit, als der Output der KI keine Urheberrechte verletzt. Außerdem wird in der Schweizer Medienszene die Bedrohung durch US-Tech-Konzerne weniger im Training von Sprachmodellen mit „alten" Inhalten gesehen, sondern vielmehr in Geschäftsmodellen, bei denen KI-gestützte Anwendungen aktuelle Informationen – wie etwa Medienberichte – als Rohstoff ohne eine aus ihrer Sicht angemessene Vergütung zu neuen, urheberrechtlich nicht mehr geschützten Produkten verarbeiten, die in Konkurrenz zu den traditionellen Medienprodukten stehen. Bedienen sich KI-Anbieter ohne Erlaubnis an Inhalten hinter Bezahlschranken, gibt es jedoch bereits rechtliche Handhabe.

Dennoch sieht der Schweizer Gesetzgeber Handlungsbedarf: Im März 2025 nahm der Ständerat auf Empfehlung des Bundesrats die Motion „Besserer Schutz des geistigen Eigentums vor KI-Missbrauch" von FDP-Ständerätin Petra Gössi an. Nach unserem Dafürhalten beschränkt sich die Motion allerdings nicht auf die Missbrauchsbekämpfung (wobei sich sowieso die Frage stellt, was in diesem Bereich wirklich Missbrauch und was einfach auf veränderte Verhältnisse zurückzuführen ist). Gemäß den sehr konkreten, von ihr verlangten Regelungen wären faktisch manche KI-Dienste in der Schweiz rechtswidrig geworden, weil sie von der Zustimmung aller tangierten Rechteinhaber*innen abhängig gemacht werden und Schrankenbestimmungen nicht mehr gelten sollen. Das schreckte auch den Nationalrat auf: Er strich die konkrete Forderung und verlangte vom Bundesrat stattdessen die Ausarbeitung einer Regelung, die sowohl die Medien und Kulturschaffenden als auch den Wirtschafts- und Innovationsstandort Schweiz schützt. Wir rechnen mit einer Zustimmung des Ständerats.

Gesetzgebungsbestrebungen wie diese nähren daher die Befürchtungen der KI-Forschung und -Entwicklung, wonach ohne Zugang zu den vorgenannten Inhalten die Weiterentwicklung der Technologie gefährdet ist. Diese Stimmen führen weiter ins Feld, dass die Schweiz durch eine strenge Regulierung nur verlieren kann: Für Kreativschaffende ist dadurch wenig gewonnen, denn die Mehrheit der Modelle stammt aus Ländern, in denen kaum Einschränkungen existieren (z. B. USA, China). Zudem droht die Schweiz als Standort für die KI-Forschung und damit für den Nukleus der Innovation an Attraktivität zu verlieren.

Sollen Rechteinhaber*innen besser für die KI-Nutzung ihrer Inhalte abgegolten werden, was derzeit wohl dem politischen Willen entspricht, sehen Fachleute in der kollektiven Rechteverwertung einen gangbaren Weg, der im geltenden Recht bereits angelegt ist. Denkbar wäre, dass die Schweiz eine Schranke in Form einer gesetzlichen Lizenz erlässt, welche die Nutzung von Werken für das Training von KI-Modellen erlauben und einer Vergütungspflicht (etwa analog dem „Kopierrappen") unterstellen würde. Eine solche Regelung könnte mit einem Opt-out-Recht verbunden werden. Eine weitere Möglichkeit wäre die sogenannte erweiterte Kollektivlizenz, die es schon heute gibt und die speziell für jene Fälle gedacht ist, in denen die Einholung der Zustimmung aller Rechteinhaber*innen nicht möglich ist. Diese Lizenz wird aber nur einzelfallweise vergeben und gilt nur hinsichtlich bereits veröffentlichter Werke. Rechteinhaber*innen, die das nicht wollen, haben ein Opt-out-Recht. Ein Opt-out-Recht bei der Wissenschaftsschranke scheint sich derzeit denn auch als Minimalkonsens herauszukristallisieren, da das EU-Urheberrecht – wie bereits erwähnt – ein solches vorsieht. Die Nutzung von Inhalten wäre somit grundsätzlich erlaubt, aber es stünde jedem frei, seinen Opt-out zu erklären. Die praktische Umsetzung dieser Regulierungsansätze würde allerdings diverse neue, ungelöste Fragen aufwerfen, und inwieweit sich die Anbieter von KI-Modellen an ein Opt-out halten würden, wäre eine andere Frage. Wie sich heute zeigt, halten sich viele nicht daran. Eine gute Nachricht für die Rechteinhaber*innen ist hingegen, dass sich die meisten der großen KI-Anbieter kürzlich im Rahmen des EU AI Acts zur Einhaltung bestimmter Transparenz-Regeln im Zusammenhang mit Modellen verpflichtet haben.

Zwingend erforderlich wäre das alles allerdings nicht. Das geltende Urheberrecht weist im Grunde keine Schlupflöcher in Bezug auf KI-Trainings auf. Sollten zukünftig Gesetzesanpassungen vorgenommen werden, dann werden diese vor allem auch wirtschaftspolitisch moti-

viert sein und dem Schutz bestimmter Wirtschaftszweige vor Marktentwicklungen dienen, wie sie heute bereits weltweit stattfinden. Dieser Schutz dürfte auf nationaler Ebene sodann auch vorderhand greifen, der Macht des Faktischen wird damit aber kaum entgegenzuwirken sein. Die Schweiz ist bisher jedenfalls gut gefahren, wenn sie gerade nicht übereilt reguliert hat. Wichtig erscheint, dass nicht einfach reflexartig versucht wird, vermeintlich bestehende Verhältnisse zu schützen. In einem ersten Schritt sollte vielmehr sachlich diskutiert werden, welche Balance das Urheberrecht zwischen den Interessen der Rechteinhaber*innen einerseits und jenen der KI-Anbieter andererseits angesichts des Wandels der Zeit fairerweise vorsehen sollte. Dass Rechteinhaber*innen besser geschützt bzw. abgegolten werden sollen im Hinblick auf die Verwendung ihrer Werke für KI, scheint jedoch ebenfalls weitgehend unbestritten.

4 Datenschutz: Trainingsdaten finden meist nur anonymisiert Eingang ins Modell

Auch die datenschutzrechtliche Analyse ergibt bei genauer Betrachtung des Sachverhalts, dass ein Training von KI-Sprachmodellen nach dem geltenden Recht im Großen und Ganzen unproblematisch ist. Kommt es zu einer Memorisierung, ergibt sich die datenschutzrechtliche Zulässigkeit – jedenfalls in der Schweiz – spätestens auf der Ebene der Rechtfertigung, und zwar aus demselben Grund, der für die Memorisierung ursächlich ist: Eine Memorisierung erfolgt normalerweise nur bei öffentlichen Personen, an denen ein entsprechendes öffentliches oder privates Interesse besteht, und zwar in Bezug auf jene Personendaten, über die entsprechend zahlreich berichtet wird. Nur in diesem Fall wird eine Information normalerweise statistisch so relevant bzw. häufig, dass sie Eingang ins Modell finden kann. Schwindet dieses Interesse, wird auch die Wahrscheinlichkeit schwinden, dass überhaupt je ein Prompt formuliert wird, der zu einer Generierung von entsprechendem Output führt. Ohne einen solchen Prompt gibt es keinen Output und damit keine Personendaten. Kommt es hingegen nicht zu einer Memorisierung, ist die Datenschutzkonformität noch offenkundiger: Beim Training mit öffentlichen Inhalten werden zwar Personendaten verwendet, aber die einzelnen Personendaten spielen keine eigenständige Rolle, sondern werden im Sprachmodell quasi zu übergeordneten sprachlichen und inhaltlichen Konzepten aggregiert, faktisch also anonymisiert.

5 Nicht zu vergessen: Lauterkeitsrecht und Vertragsrecht

Das Lauterkeitsrecht sieht Regelungen gegen die unbefugte Übernahme von fremden Arbeitsergebnissen vor. Die Erkenntnis, dass das Training von KI-Sprachmodellen zur Abstrahierung von Informationen aus dem Trainingsmaterial führt, spielt auch hier eine entscheidende Rolle. Dieser Umstand grenzt nämlich das Training von der Übernahme eines konkreten geistigen Produkts als solches ab. Beim Training von KI-Sprachmodellen werden also nicht fremde Arbeitsergebnisse verwertet, sondern „nur“ die abstrakten Muster bzw. Konzepte daraus – und auch diese nur, wenn sie statistisch relevant sind, also zum allgemeinen Sprach- und Sachwissen gehören, das nicht monopolisierbar ist.

Bleibt noch die Vertragsverletzung als Einfallstor derjenigen, die sich gegen eine Verwendung ihrer Inhalte für das Training eines KI-Sprachmodells schützen wollen. Hier fällt auf, dass sehr viele Standardgeheimhaltungsklauseln nicht nur Vertraulichkeitspflichten vorsehen, sondern auch den zulässigen Zweck, zu welchem Inhalte genutzt werden dürfen, einschränken. Dies kann einem Training tatsächlich direkt entgegenstehen, spielt aber freilich bei öffentlich verfügbaren Inhalten grundsätzlich keine Rolle.

Autoren

David Rosenthal
Schützengasse 1
8021 Zürich
Schweiz
david.rosenthal@vischer.com

Livio Veraldi
Schützengasse 1
8021 Zürich
Schweiz
livio.veraldi@gmail.com