Inkl.
EU Artificial
Intelligence Act

Niklas Mühleis • Nick Akinci

RECHTSLEITFADEN KI IM UNTERNEHMEN

- Datenschutz, Haftungsrisiken, Urheberrecht, Compliance
- Einsatz von ChatGPT, Midjourney und Co.
- Einführung von KI im Unternehmen, Content Creation, Datenanalyse, Softwareentwicklung, Sprachassistenten, Human Resources

Rheinwerk Computing

Vorwort

Wenn Sie dieses Buch gekauft haben, planen Sie wahrscheinlich den Einsatz von KI in Ihrem Unternehmen oder haben bereits damit begonnen. Zumindest aber haben Sie davon gehört, was Künstliche Intelligenz, und insbesondere generative Künstliche Intelligenz, mittlerweile zu leisten vermag. Schon jetzt sind aktuelle Modelle in der Lage, Tätigkeiten zu übernehmen, die bisher von Menschen ausgeführt werden mussten – mit beeindruckenden Ergebnissen! KI kann die Effizienz von Arbeitsprozessen steigern sowie die Kosten für Personal und externe Dienstleister reduzieren. Insbesondere die großen Sprachmodelle, die sogenannten Large Language Models (LLM), sind mächtige Werkzeuge, die auf menschlichem Niveau Zusammenhänge erkennen und auch komplexe Aufgaben lösen können.

So verlockend die Versprechungen dieser neuen Welt auch sein mögen, es gibt auch einiges zu beachten. Wie so oft birgt der Einsatz neuer Technologien auch rechtliche Risiken. Zudem müssen gängige Compliance-Vorgaben auch bei der Nutzung künstlicher Intelligenz beachtet werden. Gerade hier ergeben sich technologiespezifische Anforderungen, die Unternehmen vor neue Herausforderungen stellen.

Juristen wiederum stehen vor zwei wesentlichen Herausforderungen: Zum einen erwartet die Gesellschaft, dass die (Grund-)Rechte durch gut durchdachte KI-Regelungen geschützt werden. Gleichzeitig müssen diese Regelungen aber auch Raum für Innovationen lassen. Eine Überregulierung gilt es deshalb zu vermeiden. Auf der anderen Seite müssen sich Juristen mit KI und ihren Auswirkungen auseinandersetzen, um die rechtlichen Implikationen abschätzen zu können. Hier müssen derzeit noch oft bestehende, nicht auf KI zugeschnittene Gesetze auf die neuen Technologien angewandt werden.

Was erwartet Sie in diesem Buch?

Kernthema des Buches sind – Sie ahnen es bereits – die rechtlichen Implikationen beim Einsatz von KI im Unternehmen. Dabei wollen wir Sie nicht mit juristischem Fachwissen langweilen, sondern Ihnen verständlich erklären, wo aktuell Probleme liegen und Ihnen Lösungsmöglichkeiten an die Hand geben. Dadurch sollen Sie in die Lage versetzt werden, KI in Ihrem Unternehmen einzusetzen, ohne dabei größere Risiken einzugehen. Sie sollten sich aber bewusst sein, dass derzeit an vielen Stellen noch große Rechtsunsicherheit herrscht. Der Gesetzgeber und die Gerichte können mit der rasanten technologischen Entwicklung schlicht nicht Schritt halten.

Natürlich dürfen in einem Buch, in dem es um Technologie geht, auch die technischen Aspekte nicht zu kurz kommen. Deshalb haben wir neben juristischen Autoren auch solche mit technischer Expertise für unser Buch gewonnen. Insbesondere in Kapitel 1 werden wir Ihnen daher die Grundlagen der Künstlichen Intelligenz und insbesondere der Generativen KI erläutern. Dies soll Ihnen ein besseres Verständnis der nachfolgenden rechtlichen Aspekte geben, aber auch ein tieferes technisches Verständnis ermöglichen.

Wie sollten Sie dieses Buch lesen?

Auch wenn Sie keinen Roman vor sich haben, ist dieses Buch so aufgebaut, dass sie es gut von vorne nach hinten durcharbeiten können. Sie können aber auch zu den Kapiteln springen, die die für Sie relevanten Informationen enthalten. Durch das Stichwortverzeichnis und die zahlreichen Querverweise haben wir dafür gesorgt, dass Sie mühelos durch die verschiedenen, zusammenhängenden Themen navigieren können. Um eine gute Ausgangsbasis zu schaffen, empfehlen wir ihnen jedoch, zumindest das einführende Kapitel 1 im Ganzen zu lesen.

1.2 KI versus Urheberrecht

In diesem Abschnitt erhalten Sie einen ersten Überblick über die urheberrechtlichen Problemstellungen im Zusammenhang mit generativer KI. Zudem erläutern wir einige Grundprinzipien des Urheberrechts.

Generative KI bietet Ihnen die Möglichkeit, komplexe kreative Inhalte mit geringem Aufwand und vor allem in kurzer Zeit zu erstellen. Es eröffnen sich Anwendungsmöglichkeiten für Laien in Unternehmen, die bisher von externen Dienstleistern wie Grafikern, Mediendesignern und ähnlichen kreativen Berufsbildern abgedeckt wurden. Doch so praktisch diese neuen Möglichkeiten für den Einsatz in Ihrem Unternehmen auch sein mögen: Aus rechtlicher Sicht birgt der unreflektierte Einsatz von KI mitunter auf dem Gebiet des Urheberrechts einige Fallstricke, die es zu vermeiden gilt.

Für Anbieter von KI und Rechteinhaber, aber auch für Sie als Nutzer, ergeben sich auf rechtlicher Ebene bedeutende Problemfelder. Wie Sie in Abschnitt 1.1.3 gelernt haben, benötigen Anbieter von KI-Anwendungen Trainingsdaten, also den notwendigen »Input«, um die Fähigkeiten der KI-Anwendung zielgerichtet weiterentwickeln zu können. Doch welche Trainingsdaten können verwendet werden und wo gibt es urheberrechtliche Hürden? Was bedeutet Gemeinfreiheit in diesem Zusammenhang und worum handelt es sich bei Text- und Data-Mining? (Lesen Sie hierzu Abschnitt 3.2.3.) Rechteinhaber – häufig Kreative, die ihren Lebensunterhalt durch den Verkauf ihrer Werke bestreiten – sehen aktuell ihre Rechte bedroht und fürchten zum Teil um ihre berufliche Zukunft. Die Folge ist, dass sich einige Rechteinhaber gegen jede Nutzung ihrer Werke im Zusammenhang mit KI wehren. So streikten bereits im Sommer 2023 eine Reihe Beschäftigter der Filmindustrie in Hollywood und versuchten auf diese Weise, dem Einzug von KI, was mit der sukzessiven Ersetzung von Drehbuchautoren und Schauspielern einhergehen könnte, einen Riegel vorzuschieben. Ist die Sorge begründet?

Eine weitere bedeutende Frage ist, wie das Urheberrecht mit KI-generierten Inhalten, dem »Output«, umgeht. Besteht an ihnen ein Urheberrecht? Welche Gefahren bestehen bei einer arglosen Verwendung KI-generierter Inhalte? Ein Aspekt, dem Sie als Nutzer besondere Beachtung schenken sollten.

Hinzu kommt die rasante technische Entwicklung im Bereich der KI. Eine KI, die heute Stand der Technik ist, kann in wenigen Wochen schon wieder veraltet sein.

Hintergrundwissen

Das von OpenAI entwickelte KI-Modell GPT ist vor allem für seine Fähigkeiten im Bereich der Textgenerierung bekannt. Die im November 2022 veröffentlichte Version GPT-3.5 wurde mit 175 Millionen Parametern trainiert. Die bereits vier Monate später veröffentlichte Version GPT-4 basiert auf 100 Billionen Parametern und kann so deutlich verbesserte Ergebnisse liefern, die kaum von menschlichen Inhalten zu unterscheiden sind. Darüber hinaus enthält GPT-4 noch eine Reihe weiterer signifikanter Verbesserungen, wie beispielsweise die neuerdings mögliche Interaktion mit Bildern, eine verbesserte Steuerbarkeit und damit individuellere Interaktion oder die Fähigkeit, Kontext zu verstehen.

Dieser Umstand und weitere Aspekte machen die rechtliche Betrachtung nicht einfach. Aus diesem Grund ist die Bildung eines Rechtsrahmens für den Umgang mit KI eine anspruchsvolle Aufgabe. Die im Entwicklungsprozess befindliche KI-Verordnung (AI Act) der Europäischen Union stellt eine erste Anstrengung in diese Richtung dar (hierzu mehr in Abschnitt 5.1).

Übersicht: Urheberrechtliche Problemstellungen bei generativer KI

1. **Training und Urheberrecht**: Für das Training von generativer KI werden große Mengen Inhalte benötigt, die fast immer urheberrechtlich geschützt sind. Mehr hierzu lesen Sie in Abschnitt 3.2.
2. **Schutz von KI-generierten Inhalten**: KI-generierte Inhalte werden nicht durch Menschen, sondern durch Software geschaffen. Das Urheberrecht basiert jedoch auf dem menschlichen Kreativprozess. Mehr in Abschnitt 1.2.1.
3. **Urheberrechtsverletzungen durch generative KI**: KI-generierte Inhalte können potenziell Plagiate urheberrechtlich geschützter Werke darstellen und damit rechtsverletzend sein. Lesen Sie hierzu Abschnitt 2.3.1.
4. **Übertragung von KI-generierten Inhalten**: Da KI-generierte Inhalte nicht vom Urheberrecht geschützt werden, können diese auch nicht wie üblich lizenziert werden. Es müssen kreative juristische Lösungen erarbeitet werden, bis der Gesetzgeber nachbessert. Lesen Sie mehr in Abschnitt 2.3.4.

Um die vorgenannten Problemstellungen zu durchdringen, ist ein Grundverständnis der Funktionsweise des Urheberrechts essenziell. Nachfolgend erhalten Sie daher zunächst einige einführende Erläuterungen, bevor wir im späteren Verlauf des Buches näher auf die Problemfelder eingehen.

1.2.1 Das Schöpferprinzip: warum KI-generierte Inhalte nicht vom Urheberrecht geschützt sind

Im Zentrum des deutschen Urheberrechts steht das Schöpferprinzip. Der deutsche Gesetzgeber folgt damit einem Ansatz, der den Menschen als Schaffenden in den Mittelpunkt stellt. Urheber ist stets der Schöpfer eines Werkes, also diejenige natürliche Person, die das Werk durch eine persönliche geistige Leistung selbst geschaffen hat. Werke sind also immer persönliche geistige Schöpfungen.

Sie sind Urheber!

Sie selbst sind in Ihrem Leben schon oft Urheber geworden. Schon der Druck auf den Auslöser einer Kamera ist ein schöpferischer Akt, dessen Ergebnis die Fotografie ist – das geschützte Werk, an dem Sie als Schöpfer im Moment der Schöpfung ein Urheberrecht erwerben.

Die Bindung an das Schöpferprinzip bedeutet auch, dass Ihr Unternehmen niemals Urheber, sondern stets nur Inhaber vom Urheberrecht abgeleiteter Nutzungsrechte sein kann, da ein Unternehmen, also eine juristische Person, selbst keine geistige Leistung entfalten kann. Das Urheberrecht ist immer an einen Schöpfer, den Menschen, gebunden. Bezogen auf Unternehmensstrukturen erwirbt also grundsätzlich der Mitarbeiter, der das Werk geschaffen hat, auch das Urheberrecht an diesem Werk. Dies gilt auch, wenn das Werk im Rahmen der geschuldeten Arbeitsleistung im Unternehmen erschaffen wird.

Praxistipp: Nutzungsrechte regeln

Sofern Ihr Unternehmen im kreativen Bereich agiert und Ihre Mitarbeiter im Rahmen ihrer Tätigkeit – unabhängig von KI – urheberrechtlich geschützte Werke erstellen, erhalten Sie als Arbeitgeber meist »automatisch« ein Nutzungsrecht an diesen Werken – jedenfalls in gewissem Umfang. Es ist aber sinnvoll, im Arbeitsvertrag zu regeln, welche Nutzungsrechte an diesen Werken dem Arbeitgeber eingeräumt werden. Auf diese Weise kann der Umfang der Nutzungsrechteübertragung genau definiert und somit Streitpotenzial im Keim erstickt werden.

Aus dem Schöpferprinzip folgt damit auch, dass eine KI selbst nicht als Urheber in Betracht kommt, da nur ein Mensch Urheber sein kann. Was für das Zusammenwirken zwischen Mensch und KI gilt, klären wir im weiteren Verlauf.

Achtung, nicht jede Schöpfung eines Menschen ist auch ein urheberrechtlich geschütztes Werk. Sofern Sie ein Urheberrecht begründen möchten, muss Ihr geschaffenes Werk die sogenannte Schöpfungshöhe erreichen, also das normale, alltägliche Maß an Kreativität überschreiten und eine besondere Individualität aufweisen. Nicht jede Aneinanderreihung von Worten, nicht jede Tonfolge oder auch jeder Pinselstrich vermag diese Voraussetzung zu erfüllen. Die Abgrenzung ist dabei nicht immer einfach und muss in Zweifelsfällen von den Gerichten vorgenommen werden. Gleichwohl sind die Anforderungen für die Erreichung der Schöpfungshöhe im deutschen Recht nicht allzu hoch.

Hintergrundwissen

Die Abgrenzung, ob ein Werk vorliegt, erfolgt anhand des Begriffs der »kleinen Münze«. Die »kleine Münze« hielt bereits vor gut einhundert Jahren Einzug in das deutsche Urheberrecht und wurde von der deutschen Gerichtsbarkeit bis heute in einer Vielzahl von Entscheidungen bestätigt. Der Begriff beschreibt solche Schöpfungen, die an der untersten Grenze der gerade noch urheberrechtlich geschützten Werke liegen, und bringt zum Ausdruck, dass für die Annahme eines urheberrechtlichen Schutzes keine allzu großen Anforderungen zu stellen sind, denn auch die »kleine Münze«, die sich nach nicht viel anhört, ist urheberrechtlich geschützt. Als Beispiel für die »kleine Münze« dient der charakteristische Jingle der ARD-Tagesschau, der Ihnen sicherlich direkt in den Ohren klingt, wenn Sie diesen Text lesen. Diese kurze Folge von sechs Tönen genießt bereits urheberrechtlichen Schutz.

Praxistipp: Schöpfungshöhe einschätzen

Eine einfache Frage, die Sie sich zur groben Einordnung einer Schöpfung selbst stellen können, lautet »Würde das jeder so machen?«. Sofern Sie diese Frage mit ja beantworten können, besteht die Möglichkeit, dass es sich nicht um ein geschütztes Werk im Sinne des Urheberrechtsgesetzes handelt. In relevanten Fällen sollten Sie jedoch stets einen qualifizierten Rechtsrat einholen.

Der Zweck des Erfordernisses einer Schöpfungshöhe ist klar: Wäre jede noch so geringfügige menschliche Schöpfung urheberrechtlich geschützt, könnten wir im Alltag kaum agieren, ohne die Urheberrechte von anderen zu verletzen und uns möglichen rechtlichen Konsequenzen auszusetzen. Die Folgen für das gesellschaftliche Zusammenleben wären katastrophal.

Welche Arten von urheberrechtlich geschützten Werken gibt es?

Das Urheberrechtsgesetz (UrhG) listet hierzu unter § 2 einige Beispiele auf. Demnach zählen zu den geschützten Werken

- Sprachwerke, wie Schriftwerke, Reden und Computerprogramme
- Werke der Musik
- pantomimische Werke einschließlich der Werke der Tanzkunst
- Werke der bildenden Künste einschließlich der Werke der Baukunst und der angewandten Kunst und Entwürfe solcher Werke
- Lichtbildwerke einschließlich der Werke, die ähnlich wie Lichtbildwerke geschaffen werden
- Filmwerke einschließlich der Werke, die ähnlich wie Filmwerke geschaffen werden
- Darstellungen wissenschaftlicher oder technischer Art, wie Zeichnungen, Pläne, Karten, Skizzen, Tabellen und plastische Darstellungen

Das Gesetz listet die Arten von geschützten Werken bewusst nicht abschließend auf, sodass auch nicht genannte Arten von Werken urheberrechtlichen Schutz genießen können. Die Entscheidung darüber, ob ein nicht zuordenbares Werk urheberrechtlichen Schutz genießt, hängt dabei vom Einzelfall ab und muss im Zweifelsfall durch die Gerichte beurteilt werden.

Nun gibt es viele Möglichkeiten, wie sich Kreative technischer Geräte bedienen, um ihre urheberrechtlich geschützten Werke zu schaffen, es also zu einem Zusammenwirken zwischen Mensch und Maschine bei der Erstellung von Werken kommt. Auch der Fotograf benutzt ein technisches Gerät – die Kamera –, um sein Werk zu erschaffen. Trotz des Umstands, dass die Kamera das eigentliche Werk – die Fotografie – durch die Bildverarbeitung in den verbauten Mikrochips erstellt, erhält der Fotograf ein Urheberrecht an dem Ergebnis. Ist das auch bei der Nutzung einer KI der Fall?

Die Antwort nach derzeitiger Rechtslage lautet: Nein! Im Unterschied zur Nutzung von KI ist der Mensch bei der Fotografie erheblich am Schöpfungsprozess beteiligt. So wählt er den Bildausschnitt, die Tageszeit und das damit verbundene Licht, das Motiv und weitere Parameter, die das Endprodukt zu dem machen, was es am Ende ist. Die eher geringfügige Nachbearbeitung durch die Kamera genügt nicht, um die Voraussetzung der persönlich geistigen Schöpfung entfallen zu lassen.

Anders sieht es bei der Verwendung generativer KI aus. Zwar gibt der Mensch mit seinen Anweisungen, den »Prompts«, einen vorherigen Rahmen für die KI vor. Jedoch erstellt die KI das Werk im Anschluss vollkommen selbstständig. So erzeugt Midjour-

ney aus dem einfachen Prompt `Photo of a cute cat, reading a book with the title "GESETZ" on the cover` die folgenden Bilder:

Abbildung 1.44 Vier auf Basis eines Prompts von Midjourney generierte Bilder einer Katze

Das Ergebnis ist kaum vorhersehbar. Der entscheidende Teil der Schöpfung geschieht dabei gänzlich ohne Einfluss des Menschen. Das Schöpferprinzip greift somit nicht. KI-generierte Inhalte sind also keine persönlich geistigen Schöpfungen, an denen ein Urheberrecht entsteht.

Etwas anderes würde gelten, wenn im deutschen Urheberrecht das Verursacherprinzip zur Anwendung käme. Urheber wäre derjenige, der ein Werk verursacht hat, indem er die Schöpfungskette, an deren Ende das Werk als Ergebnis steht, in Gang gesetzt hat. Das Eingeben von Prompts in einen KI-Chat würde demnach als Start dieser Kettenreaktion aufgefasst werden. Jedoch hat sich der deutsche Gesetzgeber bewusst gegen das Verursacherprinzip und für das Schöpfungsprinzip entschieden, um den Schöpfer als Einzelperson in den Mittelpunkt zu stellen.

Der Umstand, dass keine Urheberrechte an KI-generierten Inhalten entstehen, führt dazu, dass auch keine Einräumung von Nutzungsrechten erfolgen kann. Eine Folge, die vielen Anbietern von Stockfotos nicht bewusst zu sein scheint.

Hintergrundwissen: Stockfotografie und KI

Das Geschäftsmodell der Stockfotografie besteht darin, dass große Bildagenturen Internet-Plattformen bieten, auf denen sowohl Freizeit- als auch professionelle Fotografen ihre Bilder zum Download anbieten. Gegen Zahlung eines Geldbetrages werden die entsprechenden Nutzungsrechte eingeräumt und der Kunde kann die Bilder im Rahmen der eingeräumten Nutzungsrechte frei verwenden.

Mehr und mehr werden auch KI-generierte Bilder auf den Plattformen angeboten. Auch hier versprechen die Anbieter, den Kunden gegen Gebühr Nutzungsrechte im Sinne des UrhG an den KI-Bildern einzuräumen. Dass das mangels Urheberrechts jedoch gar nicht möglich ist, wird dabei oftmals verkannt. Nach derzeitigem Stand können Sie KI-generierte Bilder von jedermann frei verwenden, ohne eine Lizenzgebühr

zu schulden. Gleichwohl können KI-generierte Bilder verkauft werden. Hier wird aber vertraglich nur die Einräumung der faktischen Nutzungsmöglichkeit versprochen. Eine Einräumung von Nutzungsrechten im urheberrechtlichen Sinne kann dagegen nicht stattfinden.

Nach alledem wird der Gesetzgeber zukünftig gefordert sein, die sich in Bezug auf generative KI ergebenden Lücken im Urheberrecht zu schließen und so für Rechtssicherheit zu sorgen. Konkrete Pläne oder gar Entwürfe liegen dafür nicht vor. Die weitere Entwicklung bleibt also abzuwarten.

1.2.2 Schutzfunktionen des Urheberrechts

Sicherlich fragen Sie sich, welche Schutzfunktionen das Urheberrecht bietet. Jeder, der ein Werk mit einer entsprechenden Schöpfungshöhe erschafft, kann sich selbst als Urheber dieses Werkes bezeichnen und seine Rechte an diesem Werk gegenüber Dritten behaupten. Dabei beginnt der Schutz mit dem Akt der Schöpfung an sich – also der Vollendung des Werkes oder eines signifikanten Teils. Bekannt ist die Praxis des Eintragens eines Urheberrechts vorwiegend aus den Vereinigten Staaten.

Das Gesetz beschreibt den Schutzumfang des Urheberrechts in § 11 Urhebergesetz (UrhG) wie folgt: Das Urheberrecht schützt den Urheber in seinen geistigen und persönlichen Beziehungen zum Werk und in der Nutzung des Werkes. Es dient zugleich der Sicherung einer angemessenen Vergütung für die Nutzung des Werkes. Einfach gesagt: Der Urheber kann mit seinem Werk machen, was er möchte und alle anderen Personen von der Nutzung des Werkes ausschließen beziehungsweise für die Nutzung durch Dritte eine angemessene Vergütung verlangen.

Der eigentliche Schutz des Urhebers und seines Werkes wird insbesondere durch das sogenannte Urheberpersönlichkeitsrecht und die Verwertungsrechte erreicht. Das Urheberpersönlichkeitsrecht umfasst das alleinige Recht des Urhebers, zu bestimmen, ob sein Werk veröffentlicht werden soll und ob es mit einer Bezeichnung des Urhebers zu versehen ist. Hinzu tritt das Recht, Entstellungen des Werkes, die nicht mit dem ursprünglichen Gedanken der Schöpfung vereinbar sind, zu unterbinden.

Hintergrundwissen: Entstellung eines Werkes

Ein berühmtes Beispiel für eine Entstellung ist die Abänderung des Entwurfes des Berliner Hauptbahnhofes des Architekten Meinhard von Gerkan. Die Deutsche Bahn wich bei der Bauausführung vom urheberrechtlich geschützten Entwurf ab und ließ statt der vorgesehenen Innendeckenkonstruktion eine flache Innendecke einbauen.

Hiergegen wehrte sich von Gerkan erfolgreich mit einer Klage. Aufgrund der Urheberrechtsverletzung wurde die Deutsche Bahn verurteilt, die bestehende Deckenkonstruktion zurückzubauen und entsprechend dem Entwurf neu zu realisieren.

Zu den Verwertungsrechten, die ebenfalls allein dem Urheber zustehen, gehören insbesondere:

- das Vervielfältigungsrecht oder einfach ausgedrückt das Recht, exakte Kopien des Werkes anzufertigen
- das Verbreitungsrecht, also das Recht das Original oder Vervielfältigungsstücke des Werkes in den öffentlichen Verkehr zu bringen
- das Ausstellungsrecht, also das Recht das Original oder Vervielfältigungsstücke des Werkes öffentlich zur Schau zu stellen
- das Vortrags-, Aufführungs- und Vorführungsrecht, also das Recht auf persönliche Darbietung des Werkes
- das Recht der öffentlichen Zugänglichmachung, was insbesondere die Veröffentlichung über das Internet betrifft
- das Senderecht, also das Recht, das Werk durch Funk, Satellit oder ähnliche Weise der Öffentlichkeit zugänglich zu machen
- das Recht der Wiedergabe des Werkes durch Bild- oder Tonträger
- das Recht der Wiedergabe von Funksendungen und von öffentlicher Zugänglichmachung, also einer Wiedergabe des Werkes durch Bildschirm, Lautsprecher oder ähnliche technische Einrichtungen
- das Bearbeitungsrecht, also das Recht, das ursprüngliche Werk abzuändern beziehungsweise umzugestalten

Zwar stehen die aufgezählten Verwertungsrechte grundsätzlich ausschließlich dem Urheber zu, jedoch kann der Urheber frei entscheiden, Dritten durch die Einräumung von Nutzungsrechten einzelne Verwertungsrechte an seinem Werk zuzugestehen. Dabei kann der Urheber beispielsweise die Nutzung auf einzelne Verwertungsrechte beschränken, zeitliche und örtliche Eingrenzungen vornehmen oder auch umfassende Nutzungsrechte einräumen und sogar sich selbst von der Nutzung des eigenen Werkes ausschließen. Selbstredend sollte die Einräumung solcher Lizenzen nur gegen angemessene Vergütung geschehen.

Wie Sie sehen, gibt das Urheberrecht dem Schöpfer eines urheberrechtlich geschützten Werkes zahlreiche Werkzeuge an die Hand, um das eigene Werk zu schützen und die Früchte aus der eigenen geistigen Leistung zu ziehen. Vielfach besteht bei Urhe-

bern – beispielsweise Fotografen – die Sorge, dass ihre Rechte durch die immer weiter verbreitete Nutzung von KI unterwandert werden könnten. Ob diese Sorge berechtigt ist, werden wir im Folgenden klären.

1.2.3 Wie Rechte Dritter verletzt werden können

Bei der Verwendung von KI können an vielen Stellen Rechte Dritter relevant werden. Zu diesen Rechten zählen insbesondere das Urheberrecht und davon abgeleitete Nutzungsrechte, aber auch Marken-, Design-, Patentrechte und Gebrauchsmuster. Für den Schutz des Urheberrechts ist keine Eintragung oder Registrierung notwendig bzw. vorgesehen. Die Schutzwirkung beginnt bereits mit dem reinen Akt der Schöpfung. Ein Urheberrecht kann also entstehen, ohne dass irgendjemand hiervon etwas mitbekommt. Dies ist bei Marken-, Design- und Patentrechten sowie Gebrauchsmustern anders, sodass bezüglich dieser Schutzrechte auf einfachere Weise recherchiert werden kann, ob eine Verletzung dieser Rechte in Betracht kommt.

Hintergrundwissen: DPMA

Der Schutz dieser Rechte beginnt anders als beim Urheberrecht mit der Eintragung der Marke oder des Designs oder auch der Anmeldung eines Patents oder eines Gebrauchsmusters beim Deutschen Patent- und Markenamt (DPMA). Das DPMA führt für alle genannten gewerblichen Schutzrechte ein Register, in welchem detailliert recherchiert werden kann. Durch diese Recherche bietet sich Ihnen die Möglichkeit, in einem frühen Stadium möglichst weitgehend festzustellen, ob KI-generierte Inhalte ein bereits geschütztes Recht verletzen könnten. Gleichwohl bietet die Recherche beim DPMA nur einen Anhaltspunkt. Denn es wird nicht in allen Fällen möglich sein, jede in Betracht kommende Ähnlichkeit oder sonstiges Potenzial für eine Rechtsverletzung aufzufinden.

Praxistipp: Recherche-Tool des DPMA

Eine solche Recherche kann von Ihnen auf einfache Weise online durchgeführt werden. Auf der Website des DPMA unter dem Link *https://www.dpma.de/index.html* findet sich ein Recherche-Tool, welches relevante Parameter abfragt und anhand dessen Suchergebnisse ausgibt. Probieren Sie es direkt aus und suchen Sie beispielsweise nach bekannten Marken, die Ihnen in den Sinn kommen.

Im Folgenden wollen wir uns auf das eigentliche Thema dieses Abschnitts, das Urheberrecht, beschränken. Es stellt sich also die Frage, wann Urheberrechte berührt sein können.

Urheberrechte können an vielen Arten von Werken bestehen. Wie bereits festgestellt, zählen zu den schutzfähigen Werken insbesondere Sprachwerke, Werke der Musik, pantomimische Werke, Werke der bildenden Künste, Lichtbildwerke, Filmwerke oder auch Darstellungen wissenschaftlicher oder technischer Art, wobei diese Liste nicht abschließend ist. Für fast alle dieser Werkarten existieren bereits KI-Anwendungen, die diese generieren können.

Sofern Sie eine solche KI verwenden, können auch Rechte Dritter berührt werden. Die Feststellung, ob Rechte Dritter berührt oder gar verletzt werden, ist ohne konkretes Werk, auf das bei der Verwendung der KI Bezug genommen wird, jedoch nicht immer ganz einfach. Recht leicht ist die Feststellung, wenn allgemein bekannte Werke Teil des generierten Inhalts sind oder zumindest starke Ähnlichkeiten aufweisen. Wenn Sie eine KI dazu bringen, die Comic-Figur »Micky Mouse« darzustellen, so ist recht wahrscheinlich, dass das generierte Bild eine solch starke Ähnlichkeit aufweist, dass eine Urheberrechtsverletzung bejaht werden muss, wenn Sie den generierten Inhalt in einer Weise verwerten, die nicht von den gesetzlich erlaubten Nutzungen gedeckt ist. Näheres zu den unzulässigen Verwertungen urheberrechtlich geschützter Werke finden Sie in Abschnitt 2.3.1.

Hintergrundwissen

Vielleicht haben Sie in den Medien davon gehört, dass das US-amerikanische Urheberrecht an der Figur »Micky Mouse« erloschen sei. Dies betrifft allerdings nur die 1928 erschienene schwarz-weiße Urfassung aus dem Zeichentrickfilm »Steamboat Willie«. Diese Version ist seit dem 1. Januar 2024 gemeinfrei geworden und darf daher von jedermann verwendet werden, ohne dass eine Urheberrechtsverletzung droht oder Lizenzen eingeholt werden müssten. Gemeinfreiheit bedeutet also, dass ein Werk keinem Urheberrecht mehr unterliegt. Nach deutschem Urheberrecht erlischt das Urheberrecht an einem bestimmten Werk grundsätzlich siebzig Jahre nach dem Tod des Urhebers (§ 64 UrhG). Zu diesem Zeitpunkt endet also die Schutzdauer und es tritt Gemeinfreiheit ein. Anders als in den USA ist eine Registrierung des Urheberrechts hierzulande nicht möglich, weshalb auch eine Verlängerung ausgeschlossen ist.

Schwerer wird die Feststellung, wenn ein unbekannteres Werk betroffen ist. Doch je unbekannter das Werk, desto größer ist auch die Wahrscheinlichkeit, dass der Urheber gar nicht bekannt ist bzw. auch durch eine sorgfältige Suche nicht festgestellt oder ausfindig gemacht werden kann. In diesem Fall handelt es sich um ein sogenanntes verwaistes Werk. Gemäß § 61 UrhG können bestimmte Verwertungen, wie die Vervielfältigung oder öffentliche Zugänglichmachung, in Bezug auf solche Werke

gestattet sein. Im Zweifel sollte hier jedoch immer ein qualifizierter Rechtsrat eingeholt werden.

Doch an welcher Stelle werden die Rechte Dritter genau berührt? Eine KI muss in jedem Fall trainiert werden, um die gewünschten Ergebnisse liefern zu können. Die Trainingsdaten, die hierfür zur Anwendung kommen, basieren – zumindest zum jetzigen Zeitpunkt – in den meisten Fällen auf Werken echter Künstler. Es wird sich daher kaum vermeiden lassen, dass auch urheberrechtlich geschützte Werke in den Datenpool der Trainingsdaten einfließen. Schon an dieser Stelle, dem sogenannten »Input«, kann es also durch die Verwendung der Werke zu Urheberrechtsverletzungen kommen, sofern es sich nicht um eine zulässige Nutzung handelt. Näheres hierzu finden Sie in Abschnitt 3.2.

Doch auch von KI generierte neue Inhalte, der sogenannte »Output«, können unter Umständen Urheberrechte verletzen. Einzelheiten zu den möglichen Urheberrechtsverletzungen durch KI finden Sie in Abschnitt 2.3.1 und in Abschnitt 2.3.2. Grundsätzlich lässt sich sagen, dass KI-generierte Inhalte dann gegen bestehende Urheberrechte verstoßen, wenn diese bereits vorhandenen, geschützten Werken sehr stark ähneln oder ganze Elemente aus solchen Werken enthalten. Ob dies im Einzelfall vorliegt, haben in künftig zu führenden Verfahren die Gerichte zu entscheiden.

Weiterlesen

In Abschnitt 2.3 gehen wir im Detail auf urheberrechtliche Problemstellungen beim Einsatz von KI im Unternehmen ein. In Abschnitt 3.2 lesen Sie mehr zum Thema KI-Training und Urheberrecht.

1.3 KI und Datenschutz: was Sie beachten müssen

In diesem Abschnitt erhalten Sie einen ersten Überblick über die problematischen Aspekte im Bereich des Datenschutzes im Zusammenhang mit der Nutzung generativer KI.

KI-Systeme verarbeiten in der Regel große Mengen sensibler Informationen, um Muster zu erkennen, Vorhersagen zu treffen und Entscheidungen zu automatisieren. Diese Daten umfassen dabei nicht selten auch Angaben über natürliche Personen (meint einen Menschen, in Abgrenzung zu Unternehmen als »juristische Personen«), sodass die Regeln des deutschen und europäischen Datenschutzregimes – allen voran die Datenschutzgrundverordnung (DSGVO) – Anwendung finden. Mit diesen Regelungen kommen auf KI-Entwickler und -Anwender strenge Auflagen und Pflichten

Kapitel 3
Individuallösung: Finetuning und Training eigener Modelle

In diesem Kapitel gehen wir näher auf die Individualisierung von KI-Modellen und -diensten sowie die damit zusammenhängenden spezifischen Rechtsprobleme ein.

Im zweiten Kapitel haben wir den Fokus auf bestehende KI-Modelle und -Systeme gelegt, die in Ihrem Unternehmen eingesetzt werden können. Das vorliegende Kapitel widmet sich demgegenüber KI, die speziell an die Bedürfnisse Ihres Unternehmens und an die vorhandenen Prozesse angepasst wird. Für solche individuellen Anpassungen gibt es verschiedene Möglichkeiten und Vorgehensweisen, die wir Ihnen zu Anfang dieses Kapitels aufzeigen werden. Zudem erfahren Sie, was die konkreten Vorteile individueller Lösungen sind und für wen sich der Aufwand lohnt.

Bei KI, die z. B. durch Training oder Finetuning angepasst wird, ergeben sich zudem gegenüber Standardlösungen einige spezifische rechtliche Problemstellungen, die es zu beachten gilt. Denn schon das Training selbst birgt einige rechtliche Stolperdrähte.

3.1 Warum Sie eigene Modelle betreiben sollten!

Die Welt der künstlichen Intelligenz (KI) entwickelt sich fortwährend mit hoher Geschwindigkeit weiter, und Unternehmen aller Größenordnungen experimentieren mit der Einführung von KI-Modellen, um ihre Prozesse zu verbessern, Kundenerfahrungen zu personalisieren und neue Erkenntnisse aus ihren Daten zu gewinnen. Doch während die Nutzung von cloud-basierten KI-Services ein sinnvoller Anfang sein kann, gibt es gute Gründe, warum ein Unternehmen es in Betracht ziehen sollte, eigene oder bestehende Modelle direkt selbst zu betreiben.

Ob Sie dabei technische Argumente wie Anpassbarkeit und Spezialisierung abwägen oder den Datenschutz und die Datensouveränität als entscheidendes Kriterium betrachten – dieses Kapitel wird Ihnen helfen, fundierte Entscheidungen darüber zu treffen, ob und wie Sie KI-Modelle in Ihrem Unternehmen einsetzen können.

3.1.1 Argumente für eigene Lösungen

Ein wesentliches Argument für den Betrieb eigener Modelle ist die Kontrolle über den Datenschutz und die Datensicherheit. Wenn sensible Unternehmensdaten ins Spiel kommen, ist es in der Regel nicht wünschenswert, diese an externe Cloud-Dienstleister zu übermitteln. Durch lokale Installation und Verwendung von KI-Modellen behalten Sie die vollständige Kontrolle über Ihre Daten und schützen sie vor unbefugtem Zugriff.

Darüber hinaus bietet das selbstständige Betreiben von Modellen ein großes Maß an Flexibilität und Anpassbarkeit. Sie können dabei Modelle an ihre spezifischen Anforderungen und Besonderheiten anpassen, statt sich auf Standardangebote der großen Anbieter verwiesen zu sehen. Dies ist besonders dann von Vorteil, wenn branchenspezifische Eigenheiten oder seltene Sprachen eine Rolle spielen, für die kommerzielle Cloud-Dienste oft keine optimierte Lösung bieten.

Weiterhin dreht es sich um die Thematik der Zensur bzw. Einschränkungen, die bei kommerziellen Anbietern greifen könnten. Bei der Nutzung eigener Modelle sind Unternehmen nicht den möglicherweise willkürlichen oder unpassenden Zensurvorgaben von Dienstleistern unterworfen. Dies ist insbesondere relevant für kreative Industrien, etwa beim Verfassen von Texten oder Erstellen von Kunstwerken, wo Freiheit im Ausdruck entscheidend ist.

Ein weiterer gewichtiger Aspekt ist die Kostenkontrolle. Es gibt Szenarien, in denen die lokal betriebenen KI-Modelle langfristig kostengünstiger sein können, vor allem wenn sie in großem Umfang genutzt werden. Während Cloud-Dienste oft auf einer Pay-per-use-Basis abgerechnet werden, können eigene Modelle nach der Anfangsinvestition ohne zusätzliche laufende Kosten pro User genutzt werden. Zu beachten ist auch die Unabhängigkeit von etwaigen strategischen Veränderungen der Anbieter, die Dienste ohne viel Vorlaufzeit einschränken, kostenpflichtig erweitern oder gar einstellen können. Wer eigene Modelle betreibt, ist vor solchen Unwägbarkeiten geschützt.

Obwohl es auf den ersten Blick so scheinen mag, als wäre das Selbstbetreiben von KI-Modellen nur etwas für Unternehmen mit umfassenden technischen Ressourcen, wird die Einstiegshürde auch für kleinere Organisationen fortlaufend geringer. Die Verfügbarkeit von Tools und Ressourcen zum Betrieb lokal verwendeter Modelle ist in den letzten Jahren deutlich gewachsen, und auch mit bescheideneren Mitteln lassen sich mittlerweile KI-Modelle betreiben. Im Vergleich zu teuren KI-Cloud-Diensten können lokale Modelle auf Hardware betrieben werden, die bereits vorhanden ist oder die bei geringen Kosten einen guten Einstieg ermöglicht. Beispielsweise bieten mittelpreisige Grafikkarten bereits ausreichend Leistung, um kleinere Modelle in einem bestimmten Rahmen zu betreiben.

Hinzu kommt eine Welle von Innovationen im Bereich der Modell-Optimierung, die Modelle effizienter machen und somit auch auf weniger leistungsstarker Hardware betriebsfähig werden. In einer Zeit, in der die Anforderungen an Grafikspeicher durch Spiele und die Nutzung größerer Modelle steigen, wird auch die Verfügbarkeit von Hardware mit ausreichend VRAM zunehmen.

Lokal betriebene KI-Modelle bieten auch eine höhere Zuverlässigkeit im Vergleich zum Cloud-Service, der unter Umständen durch Überlastungen temporär nicht verfügbar sein kann. Zu guter Letzt ist es oftmals auch ein Aspekt der Unabhängigkeit und des Empowerments. Unternehmen und Entwickler möchten nicht vollständig von einem Dienstleister abhängig sein, dessen Geschäftspolitik sich ändern und damit die eigenen Prozesse beeinträchtigen könnte. Ein selbst betriebenes Modell kann vor solchen Unabwägbarkeiten schützen und damit Souveränität bieten.

3.1.2 Individualisierungsmöglichkeiten

Die Einführung und das Betreiben eigener KI-Modelle im Unternehmen können auf unterschiedlichen Ebenen erfolgen. Die Bandbreite reicht von Lösungen für Nichttechniker, die eine einfache, nutzerfreundliche grafische Benutzeroberfläche (GUI) bevorzugen, bis hin zu umfassenderen Plattformen für Techniker/Entwickler, die eine tiefer gehende Steuerung und Anpassungsfähigkeit der Modelle ermöglichen.

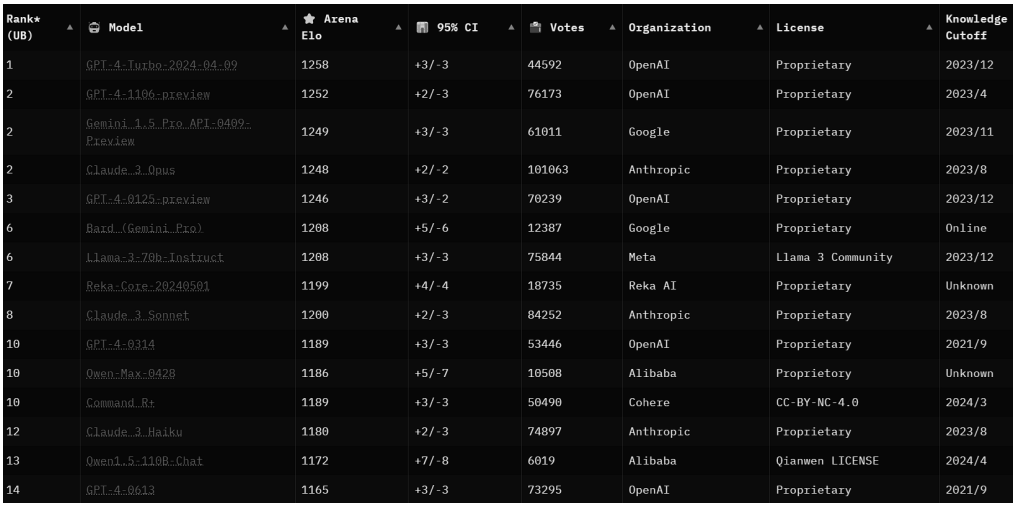
Für Nichttechniker stehen Tools wie LM-Studio oder jan (Open Source) zur Verfügung, die es ermöglichen, große Sprachmodelle (LLMs) direkt auf dem eigenen performanten Rechner vollständig offline zu betreiben. Beide bieten eine intuitive Chat-Benutzeroberfläche und eine Schnittstelle, die das Verhalten der OpenAI-API emuliert. Modelle können nahtlos von Hugging Face heruntergeladen und betrieben werden.

Webseiten wie *chat.lmsys.org* erlauben sogar den kostenlosen Test und Vergleich direkt im Browser ganz ohne Set-up.

Für Techniker oder fortgeschrittene Anwender bietet das quelloffene Projekt FastChat eine offene Plattform zum Betreiben und Bewerten von Chatbots auf Basis großer Sprachmodelle. Es bietet eine leistungsfähige Infrastruktur mit einem verteilten Multi-Modell-System, einer Web-Benutzeroberfläche und APIs, die ebenfalls mit OpenAI kompatibel ist.

Angesichts der Geschwindigkeit, mit der sich KI-Entwicklungen vollziehen, ist es ratsam, sich regelmäßig über Fortschritte und Neuerscheinungen zu informieren. Bevor

Sie diesen Schritt gehen, sollten Sie daher sorgsam planen und Zeit investieren, um die Aktualität und Relevanz von Tools und Methoden zu prüfen, bevor Sie sich für den Betrieb eigener Modelle entscheiden.



Rank* (UB)	Model	Arena Elo	95% CI	Votes	Organization	License	Knowledge Cutoff
1	GPT-4-Turbo-2024-04-09	1258	+3/-3	44592	OpenAI	Proprietary	2023/12
2	GPT-4-1106-preview	1252	+2/-3	76173	OpenAI	Proprietary	2023/4
2	Gemini 1.5 Pro API-0409-Preview	1249	+3/-3	61011	Google	Proprietary	2023/11
2	Claude 3 Opus	1248	+2/-2	101063	Anthropic	Proprietary	2023/8
3	GPT-4-0125-preview	1246	+3/-2	70239	OpenAI	Proprietary	2023/12
6	Bard (Gemini Pro)	1208	+5/-6	12387	Google	Proprietary	Online
6	Llama 3 70b Instruct	1208	+3/-3	75844	Meta	Llama 3 Community	2023/12
7	Reka Core-20240501	1199	+4/-4	18735	Reka AI	Proprietary	Unknown
8	Claude 3 Sonnet	1200	+2/-3	84252	Anthropic	Proprietary	2023/8
10	GPT-4-0314	1189	+3/-3	53446	OpenAI	Proprietary	2021/9
10	Qwen-Max-0428	1186	+5/-7	10508	Alibaba	Proprietory	Unknown
10	Command R+	1189	+3/-3	50490	Cohere	CC-BY-NC-4.0	2024/3
12	Claude 3 Haiku	1180	+2/-3	74897	Anthropic	Proprietary	2023/8
13	Qwen1.5-110B-Chat	1172	+7/-8	6019	Alibaba	Qianwen LICENSE	2024/4
14	GPT-4-0613	1165	+3/-3	73295	OpenAI	Proprietary	2021/9

Abbildung 3.1 Die Oberfläche von chat.lmsys.org

Bestehende Modelle ergänzen

Das Anpassen von KI-Modellen an die Bedürfnisse Ihres Unternehmens ist ein Schlüsselaspekt beim Betreiben eigener KI-Systeme. Hierbei bietet das Prompt Engineering eine leichte und schnelle Möglichkeit, die Antworten der KI zu steuern. Für eine ausführliche Diskussion über die Vorstufen des Finetuning, einschließlich Prompt Engineering und Retrieval Augmented Generation (RAG), verweisen wir auf Abschnitt 2.1.2.

Die Entscheidung, wie KI-Modelle in Ihrem Unternehmen betrieben werden sollen, hängt von vielen Faktoren ab, darunter technisches Know-how, verfügbare Ressourcen und spezifische Unternehmensziele. Sowohl nichttechnische als auch technisch versierte Mitarbeiterinnen und Mitarbeiter finden jedoch zunehmend Zugang zu KI-Tools und Plattformen, die den breiten Einzug von KI in die Unternehmenswelt erleichtern und fördern.

Frameworks wie H2O LLM Studio erlauben ein ähnliches Low-Code-Finetuning-Erlebnis wie OpenAI, siehe Abschnitt 2.1.2.

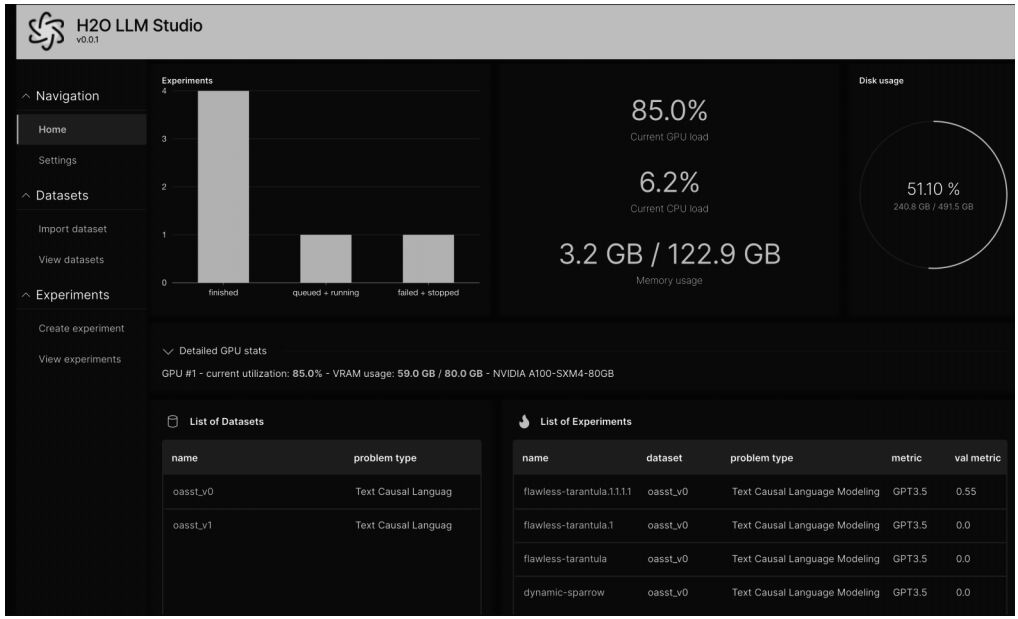


Abbildung 3.2 Die Oberfläche von H2O LLM Studio zum Finetuning lokaler Modelle

3.1.3 Neues eigenes Modell durch Training

In einer Zeit, in der künstliche Intelligenz (KI) viele unserer täglichen Prozesse zu transformieren vermag, mag es naheliegend erscheinen, eigene KI-Modelle von Grund auf zu trainieren. Doch ein solcher Ansatz ruft eine Reihe wichtiger Überlegungen hervor, von der Datenintegration bis hin zu den technischen und finanziellen Anforderungen. Die Entwicklung eigener KI-Modelle kann dann besonders sinnvoll sein, wenn Unternehmen spezifische, maßgeschneiderte Lösungen für eine einzigartige Problematik benötigen.

Beispielhaft angeführt werden kann hierfür der Finanzsektor: Hier können Banken oder Versicherungen von selbst entwickelten KI-Modellen profitieren, die Betrugserscheinungen (Fraud Detection) identifizieren. Angesichts des Umstandes, dass betrügerische Aktivitäten ständig wechselnde, ausgeklügelte Muster annehmen, kann ein maßgeschneidertes Modell, das auf den spezifischen Transaktionsdaten und Kundenprofilen des Unternehmens basiert, effektiver agieren als generische, vorgefertigte Lösungen.

Gleichermaßen kann im Bereich der Netzwerksicherheit das Training eines individuellen KI-Modells Unternehmen dazu befähigen, sich besser gegen zunehmend intel-

ligentere Cyberangriffe zu schützen. Ein auf den spezifischen Netzwerkverkehr und die Schutzbedürfnisse des Unternehmens abgestimmtes Modell kann dabei bösartige Aktivitäten präziser erkennen als standardisierte Lösungen.

Bevor Sie sich nun direkt dem Training eines eigenen KI-Modells widmen, sollten Sie daher zuerst sichergehen, dass Ihr Anwendungsfall dies wirklich erfordert und ob nicht Vorstufen wie Prompt-Engineering, Finetuning oder RAG hierzu ausreichen.

Die Bedeutung der Datenintegration

Die Grundlage jeden KI-Trainings ist der Datensatz. Die Beschaffung, Bereinigung und Strukturierung von Daten kann sich als eine Herkulesaufgabe herausstellen. Die Bewegung großer Datenmengen, insbesondere in die Cloud, kann nicht nur technisch komplex sein, sondern auch schnell zu hohen Kosten führen. Zum Training von GPT-2 beispielsweise wurden Daten aus über 7.000 selbst veröffentlichten Büchern und einer Sammlung von 8 Millionen Webseiten verwendet, ganz zu schweigen von GPT-3.5 und 4.

Eine herausfordernde Aufgabe, die nicht nur die Bereitstellung der Daten, sondern auch die Skizzierung einer klaren Datenstruktur erfordert, bevor mit dem eigentlichen Training begonnen werden kann.

Voraussetzungen für das Training eigener Modelle

Die technischen Voraussetzungen für das Training eines eigenen KI-Modells sind nicht zu unterschätzen. Umfangreiche Rechenkapazitäten sind erforderlich, um die massiven Berechnungen durchzuführen, die für das Training von Modellen erforderlich sind.

Ein Vergleich: Während für die Entwicklung von GPT-2 rund 50.000 Dollar anfielen, beliefen sich die Kosten für GPT-3 auf mehr als 4 Millionen Dollar, bei GPT-4 waren es anscheinend 100 Millionen Dollar. Unternehmen müssen letztendlich entscheiden, ob die Investition in die Entwicklung eigener Modelle den Aufwand wert ist. Die Erfahrungen von OpenAI zeigen, dass selbst ein Investment, das im Vergleich zu späteren Modellen moderat ausfiel, das Potenzial hatte, den Bereich der natürlichen Sprachverarbeitung zu revolutionieren. Das Training eigener KI-Modelle ist ein vielschichtiges und herausforderndes Projekt und kann in bestimmten Fällen durchaus sinnvoll sein.

Jedoch sollte die Entscheidung sorgfältig abgewogen werden, da es sowohl erhebliche finanzielle als auch technologische Investitionen erfordert.

3.1.4 Cloudlösungen

Die Implementierung von Künstlicher Intelligenz in Unternehmensprozesse ist ohne die angemessene technische Infrastruktur nicht denkbar. Cloud-Instanzen spielen dabei eine zentrale Rolle. Sie bieten die Rechenkraft und Flexibilität, die für das Trainieren und Betreiben von KI-Modellen erforderlich sind.

VM-Instanzen mit GPUs bieten gegenüber traditionellen CPU-Instanzen entscheidende Vorteile, indem sie speziell für rechenintensive Aufgaben des maschinellen Lernens und KI-Modelltrainings konzipiert wurden. Durch den Einsatz dieser GPU-Instanzen können Unternehmen die Effizienz ihrer KI-Operationen signifikant steigern und gleichzeitig ihre lokale Hardware von aufwendigen Rechenprozessen entlasten.

Mit Hybrid-, Private- oder Public-Cloud-Architekturen erhalten Sie unterschiedliche Stufen an Sicherheit, Skalierbarkeit und Kostenkontrolle – je nach Unternehmensbedürfnis und der gewünschten Integrationsstufe in Ihre IT-Infrastruktur.

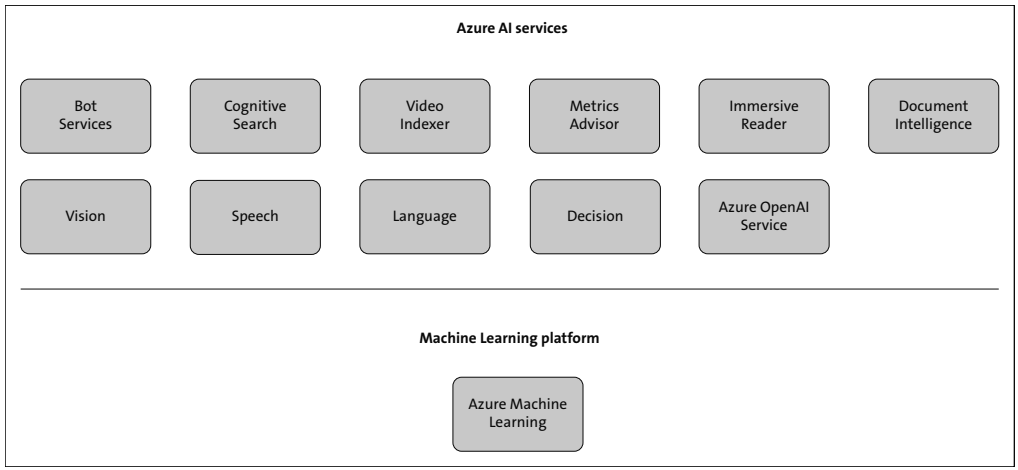


Abbildung 3.3 Überblick über die KI-Services von Microsoft Azure

Große Cloud-Anbieter wie Google Cloud stehen mit flexiblen und skalierbaren GPU-Services zur Verfügung, die auf unterschiedliche Anwendungsfälle und Unternehmensgrößen zugeschnitten sind.

Die Services bieten eine Auswahl von Instanz-Arten, die oft nach Zeit abgerechnet werden und je nach Bedarf hoch- oder runterskaliert werden können. Die Auswahl der Cloud-Infrastruktur kann über den Erfolg Ihrer KI-Projekte entscheiden. Eine sorgfältige Prüfung der Leistung, Flexibilität und Kosten der verschiedenen Cloud-

Services ist daher unerlässlich. Betrachten Sie die spezifischen Angebote der großen Anbieter und prüfen Sie, welcher Service die optimale Basis für Ihre KI-Initiativen bildet. Gleichzeitig bietet die Cloudlandschaft heute Alternativen für diejenigen, die eine auf bestimmte Zeiträume beschränkte oder finanziell sparsame Lösung suchen, ohne dabei Abstriche bei der Leistung machen zu müssen.

3.1.5 Eigene Hardware nutzen

Die Entscheidung, KI-Modelle lokal zu betreiben, ist ein bedeutsamer Schritt in Richtung vollständiger Autonomie und Kontrolle über Ihre maschinellen Lernumgebungen. Indem Sie sich von Cloud-Diensten unabhängig machen, erlangen Sie die Freiheit, die Modelle Ihren genauen Bedürfnissen anzupassen und dabei sensible Daten intern zu halten.

Um Ihnen einen umfassenden Einblick zu geben, wird in diesem Kapitel erörtert, wie Sie KI-Modelle auf eigene Faust betreiben und trainieren können.

Die wesentlichen Hardware-Komponenten umfassen dabei nicht nur eine ausreichende Prozessorleistung (CPU) und genügend Arbeitsspeicher (RAM), sondern auch eine leistungsstarke Grafikprozessoreinheit (GPU), die insbesondere für das maschinelle Lernen optimiert ist.

Hierbei ist die Menge des Videospeichers (VRAM) pro GPU von besonderer Bedeutung, da sie bestimmt, wie komplexe Modelle oder umfangreiche Datenmengen verarbeitet werden können.

Ressourcenverbrauch KI

Das vollständige Training eines KI-Modells ist am ressourcenintensivsten.

Finetuning erfordert im Vergleich zum bloßen Betreiben meist eine stärkere Hardware. Hier muss das Modell geringfügig angepasst werden, was wiederum zusätzliche Rechenkapazität erfordert.

Das Betreiben vortrainierter KI-Modelle erfordert eine Hardwareumgebung, die sich nicht ausschließlich durch hohe Rechenleistung, sondern insbesondere durch ausreichenden Videospeicher (VRAM) auszeichnet.

Der VRAM ist häufig der begrenzende Faktor bei der Auswahl der GPU für KI-Modelle, da diese tendenziell mehr von der Speicherkapazität als von der reinen Rechenleistung abhängen. Bei begrenzter GPU-Anzahl kann durch intelligentes Batching von Anfragen ein effizienter Parallelbetrieb ermöglicht werden, ohne pro Anfrage eine gesamte GPU zu blockieren.

Fertige Modelle finden sich inzwischen auf Plattformen wie Hugging Faces Modellhub, der eine umfassende Bibliothek zur Verfügung stellt. Die Sammlung reicht von Modellen kleinerer Größenordnungen bis hin zu sehr großen wie zum Beispiel einem 120B-Sprachmodell. Das »B« steht für Milliarden (Billions) und bezieht sich auf die Gesamtanzahl der Parameter, aus denen das KI-Modell besteht.

Parameter sind im Grunde die Elemente eines KI-Modells, die aus dem Trainingsprozess gelernt werden und entscheiden, wie das Modell Daten interpretiert und Antworten generiert. Je mehr Parameter ein Modell hat, desto umfassender ist seine Fähigkeit, Wissen zu speichern und komplexe Muster zu erkennen. Ein KI-Modell mit 13 Milliarden Parametern kann also eine enorme Menge an Informationen verarbeiten und ist somit in der Lage, subtile Nuancen in der Sprache zu erfassen. Das ermöglicht es dem Modell, Antworten zu generieren, die menschlicher Kommunikation näherkommen, und kann in komplexen Aufgaben wie der Sprachübersetzung, dem Textverständnis oder der automatischen Beantwortung von Fragen eingesetzt werden.

Die große Anzahl an Parametern eines solchen Modells erfordert jedoch entsprechend dimensionierte Hardware-Ressourcen, insbesondere hinsichtlich des VRAM, um die umfangreichen Berechnungen effektiv durchführen zu können.

Hintergrundwissen

Erforderliche technische Ressourcen für den lokalen KI-Betrieb (geschätzt):

- 7B-Modell: mindestens 13 GB VRAM, 1 × NVIDIA 4070 Ti SUPER (16 GB)
- 13B-Modell: mindestens 26 GB VRAM, 1 × NVDIA RTX A6000 (48 GB)
- 30B-Modell: mindestens 65 GB VRAM, 1 × NVDIA H100 (80 GB)
- 65B Modell: Mindestens 131 GB VRAM, 2x NVDIA A100(80 GB)
- 175B-Modell (~GPT-3): mindestens 350 GB VRAM, 5 × NVDIA A100 (80 GB)

Darüber hinaus spielen Hyperparameter eine wichtige Rolle, da diese bestimmen, ob ein Modell auf einer existierenden GPU betrieben werden kann. Durch Anpassungen, wie beispielsweise Quantisierung, Kontextlänge und effiziente Batch-Verarbeitungen, werden Modelle auch für Hardwareplattformen zugänglich gemacht, deren Spezifikationen geringer sind als die oben genannten Anforderungen. Somit lassen sich auch mit eingeschränkter VRAM-Ausstattung Modelle realisieren, die ursprünglich nicht dafür vorgesehen waren. Modelle lassen sich auch über mehrere GPUs verteilen, sodass man 2 GPUs mit jeweils 16 GB zusammenschalten kann, um ein 13B-Modell zu betreiben, ohne die Hyperparameter zu verändern.

Hintergrundwissen: Was ist Quantisierung?

Wenn wir über KI-Modelle sprechen, denken Sie vielleicht an eine immense Sammlung von Zahlen, die jede Verbindung und Gewichtung innerhalb des Modells darstellen. Normalerweise werden diese Zahlen mit hoher Präzision gespeichert, was bedeutet, dass sie viel Speicherplatz einnehmen.

Quantisierung reduziert diese Präzision gezielt, ähnlich dem Vorgang des Komprimierens einer Musikdatei, um Speicherplatz zu sparen. Dabei werden die Zahlen einer umfangreichen Datenbank in einer kompakteren Form ausgedrückt, indem zum Beispiel statt 32 Bits nur 16 oder sogar 8 Bits verwendet werden. Dies macht das Modell kleiner und leichter und somit weniger anspruchsvoll in Bezug auf die Speicher- und Verarbeitungskapazität einer GPU, jedoch sinkt damit auch die Qualität der Ausgaben.

Auf diese Weise ist es möglich, leistungsfähige KI-Modelle auf Hardware mit begrenztem VRAM zu betreiben, was besonders für Organisationen von Bedeutung ist, die ihre Modelle lokal und nicht in der Cloud betreiben möchten. Quantisierung ermöglicht es also, in einem gewissen Rahmen High-End-KI-Modelle einem breiteren Anwenderkreis zugänglich zu machen und bestehende Hardware besser zu nutzen.

3.2 Trainingsdaten und Urheberrecht

Wenn Sie ein bestehendes Modell durch Finetuning verbessern oder sogar ein eigenes Modell von Grund auf trainieren (lassen) wollen, benötigen Sie vor allem eines: Trainingsdaten. Denn insbesondere aktuelle Sprach- und Diffusionsmodelle sind gerade deshalb so leistungsfähig, weil sie mit riesigen Datenmengen trainiert wurden.

Sofern Sie nicht gerade selbst über die für das Training benötigten Daten(mengen) verfügen, liegt es nahe, sich an der wohl größten Datenquelle überhaupt zu bedienen: dem Internet. Und so dürften es auch bisher die meisten getan haben. Das Problem: Inhaber von Urheberrechten haben häufig etwas dagegen, wenn ihre im Internet verfügbaren Werke von Dritten genutzt werden.

3.2.1 Zustimmung als Ausgangspunkt

Im Grundsatz gilt im Urheberrecht die Regel, dass der Urheber allein darüber entscheiden darf, wer sein Werk in welcher Form nutzt. Das gilt insbesondere auch für Werke wie Bilder, Texte und Videos, die frei im Internet abrufbar sind. Diese Inhalte sind so gut wie immer vom Urheberrecht geschützt.

Inhalt

RECHTSLEITFADEN KI
IM UNTERNEHMEN

Für Unternehmen ist KI unverzichtbar geworden. Bei ihrem Einsatz müssen jedoch zahlreiche juristische Aspekte berücksichtigt werden. In präziser und zugänglicher Sprache vermittelt Ihnen dieser Leitfaden umfassendes Wissen über die rechtlichen Rahmenbedingungen, auf die es dabei ankommt. Mit vielen praktischen Beispielen, konkreten Anwendungsszenarien, Lösungsansätzen und Handlungsempfehlungen. Aktuell zum EU AI Act.

Hier finden Sie Antworten auf Ihre Fragen!

- Grundlagen KI, Datenschutz, Urheberrecht
- Vertragsgestaltung, Wettbewerbsrecht, Rechtsberatung
- Einführung im Unternehmen
- Alles Wichtige zum EU AI Act
- Einsatz von ChatGPT und Co.
- Softwareerstellung mit KI
- Human Resources und KI
- Unterstützung durch Sprachassistenten
- Rechtliche Risiken und Stolperfallen
- Potenzielle Haftungsrisiken
- Rechtsfragen beim Training eigener KI-Modelle

Niklas Mühleis und **Nick Akinci** sind Rechtsanwälte und Partner der Kanzlei Heidrich Rechtsanwälte. Gemeinsam mit einem interdisziplinären Autorenteam aus Technikern und Juristen machen sie Sie mit den komplexen rechtlichen Herausforderungen vertraut und bieten Orientierung für Ihre Pläne und Herausforderungen.

Rheinwerk Computing

Gedruckt in Deutschland
Papier aus nachhaltiger Waldwirtschaft
Mineralölfreie Druckfarben

€ 39,90 [D] € 41,10 [A]
KI/Jura
ISBN 978-3-367-10098-9