prompt/

News, Tipps und Debatten zu rechtlichen Fragen rund um generative KI

Juli 2025

Liebe Leser*innen,
schnell den nächsten Urlaub planen und buchen, die Terminkoordination automatisieren oder die personalisierte Aktienstrategie umsetzen: **KI-Agenten** sollen all das (und mehr) möglich machen. Anders als Chatbots erledigen sie nicht nur einzelne Aufgaben, sondern können ganze Prozesse steuern – zumindest theoretisch. Manche Expert*innen haben das Jahr 2025 deshalb schon zum „Jahr der KI-Agenten" ernannt.

Es ist längst nicht klar, ob sie den Erwartungen in der Praxis standhalten. So oder so werfen KI-Agenten schon jetzt eine Menge Fragen auf – vor allem rechtliche. Wer haftet, wenn durch ihren Einsatz ein Schaden entsteht? Die Entwickler*innen? Der Anbieter? Die Nutzerin?

Die Wertschöpfungskette hinter KI-Agenten ist komplex und Verantwortung nur schwer zuzuweisen. Welche Rechtsrahmen können für Klarheit sorgen? Welche Regeln sieht die KI-Verordnung (KI-VO) für den Einsatz von KI-Agenten vor?
Diesen Fragen gehen wir in dieser Ausgabe von **prompt/** nach.

Wir wünschen, wie immer, eine erhellende Lektüre,
Ihr **prompt/**-Team

Und wenn Ihnen unser Newsletter **prompt/** gefällt, sagen Sie's gerne weiter – Dankeschön :-)

prompt/ empfehlen

Wenn Sie den **prompt/**-Newsletter regelmäßig bekommen möchten, können Sie sich über diesen Button anmelden:

prompt/ bestellen

Ebenso freuen wir uns über **Themenvorschläge, Fragen oder Erfahrungen aus Ihrem beruflichen oder persönlichen Alltag mit generativer KI.**
Schreiben Sie uns eine E-Mail an prompt@irights-lab.de

Auf bald und mit besten Grüßen,
Ihr **prompt/**-Team

/Inhalt

/Monatsthema

KI-Agenten im Fokus

KI-Agenten erweitern die Möglichkeiten von KI-Systemen erheblich, doch die rechtlichen Rahmenbedingungen dafür sind noch im Aufbau. Dadurch ergeben sich neue Fragen zu Verantwortung und Kontrolle.

„2025 ist das Jahr der KI-Agenten“, prognostizierte Kevin Weil, Chief Product Officer von OpenAI, Anfang Januar beim Weltwirtschaftsforum in Davos. Das US-amerikanische Cloud-Computing-Unternehmen Salesforce kündigt an, bis Ende des Jahres „eine Milliarde Agenten“ zu aktivieren. In vielen Wirtschaftsbereichen werden KI-Agenten bereits eingesetzt: in der Kundenbetreuung (Service-Bots), in der Softwareentwicklung (Automatisierung von Code-Reviews und Dokumentationen) sowie im Marketing (Automatisierung von Kampagnen) etwa. Auch im Finanzwesen werden sie zunehmend genutzt. Sie helfen dabei, wiederkehrende oder komplexe Aufgaben effizienter zu erledigen. Doch was sind KI-Agenten und welche rechtlichen Fragen werfen sie auf?

Was KI-Agenten von Chatbots unterscheidet

Im Gegensatz zu klassischen Chatbots können KI-Agenten „eigenständig“ mal mehr, mal weniger komplexe Ziele verfolgen und dabei sowohl in virtuellen als auch in realen Umgebungen handeln. Ein KI-Agent kann etwa Websites bedienen, E-Mails schreiben, in Software navigieren, Datenbanken durchsuchen oder Code generieren. Über programmierbare Software-Schnittstellen (auf Englisch: Application Programming Interfaces, kurz: API) lassen sich auch alltägliche Aktionen auslösen: Geld überweisen, Bestellungen aufgeben, Termine buchen oder Smart-Home-Geräte steuern zum Beispiel.

Technisch bestehen KI-Agenten aus einem leistungsstarken Sprachmodell und zusätzlicher Software, die Zugang zu externen Werkzeugen, Datenbeständen oder Dateien ermöglicht. Diese Systeme werden unter anderem von großen Technologieanbietern wie OpenAI, Google oder Microsoft als Zusatzdienste angeboten. Firmen können KI-Agentensysteme auch selbst entwickeln, indem sie beispielsweise vortrainierte große Sprachmodelle mit unternehmenseigenen Daten feinjustieren (Fine-Tuning). Oder indem sie spezielle Agentensysteme auf Basis von Retrieval Augmented Generation (RAG)-Methoden bauen (siehe auch Glossarbeitrag zu RAG weiter unten), die beispielsweise auf firmenspezifische Informationen zugreifen. Es gibt auch KI-Plattformen, die sogenannte Low-Code-Ansätze anbieten. Diese ermöglichen es Unternehmen, eigene KI-Agenten zu konfigurieren, ohne selbst KI-Modelle programmieren zu müssen.

Rechtliche Risiken und regulatorische Lücken

Rechtlich gesehen sind KI-Agenten grundsätzlich KI-Systeme im Sinne der EU-weiten KI-Verordnung (KI-VO). Die Einordnung in Risikoklassen der KI-VO hängt jedoch vom konkreten Einsatzzweck ab. Da KI-Agenten oft auf leistungsstarken Sprachmodellen basieren, die potenziell systemische Risiken bergen, unterliegen sie häufig den erweiterten Anforderungen zur Risikoabschätzung und -minderung. Beispielsweise bei Langzeitplanungen eines KI-Agenten, der Finanztransaktionen eigenständig ausführt. Kommt ein KI-Agent in kritischen Bereichen wie Medizin, Bildung oder Infrastruktur zum Einsatz, greifen strengere Anforderungen (Kapitel III der KI-VO). Anbieter von KI-Modellen mit vielseitigen Fähigkeiten (sogenannte General-Purpose-AI-Models, GPAI) müssen systematisch alle Risiken, die durch den Einsatz ihrer Modelle entstehen können, identifizieren, bewerten und geeignete Maßnahmen zu deren Beherrschung umsetzen.

Wer trägt welche Verantwortung?

Die KI-VO weist Pflichten entlang der gesamten Wertschöpfungskette zu. Setzt ein Unternehmen etwa einen KI-Agenten für die Personalauswahl ein – ein Hochrisikobereich laut KI-VO – muss es sicherstellen, dass er diskriminierungsfrei beziehungsweise diskrimierungsmindernd funktioniert. Die Verantwortung bleibt beim Unternehmen, auch wenn die Technik von einem externen Anbieter stammt. Sofern KI-Agenten nachweislich Schäden verursachen, ist allerdings oft unklar, wer

dafür haftet: War es ein Modellfehler seitens des KI-Herstellers, ein Konfigurationsproblem oder mangelnde Überwachung seitens des KI-Betreibers? Effektive Kontrolle erfordert eine detaillierte Protokollierung. Wenn KI-Agenten nämlich für persönliche Aufgaben eingesetzt werden, kann Systemüberwachung schnell die Privatsphäre verletzen. Die aktuellen Regelungen der KI-VO zu Hochrisiko-Anwendungen erfassen die spezifischen Gefahren von KI-Agenten – zumindest in ihrer derzeitigen Fassung – noch nicht vollständig.

Wie kann sich die Rechtslage weiterentwickeln?

Die rechtliche Einordnung wird sich schrittweise klären – durch Gerichtsurteile in bedeutenden Verfahren, Praxisleitfäden und regulatorische Konzepte. Eine mögliche Lösung sind sogenannte Sandboxes: kontrollierte Umgebungen, in denen Regulierungsbehörden, Unternehmen und mitunter Verbraucher*innen gemeinsam neue Lösungen, Produkte oder Dienstleistungen testen können, ohne die regulatorischen Anforderungen vollständig erfüllen zu müssen. Allerdings ist umstritten, wie übertragbar die Ergebnisse und Erkenntnisse aus den Tests in der Sandbox sind und ob sie sich hinreichend skalieren lassen.

Nicht zuletzt sollten auch Expert*innengremien eingesetzt und die Zivilgesellschaft eingebunden werden, um die Regulierung von KI-Agenten weiterzuentwickeln. Zudem geht es auch darum, die komplexen rechtlichen und ethischen Fragestellungen zu klären und praxisnah zu gestalten. Stand Juni 2025 arbeitet die EU-Kommission intensiv an den in der KI-VO festgelegten Schritten zur Umsetzung der einzelnen Regelungskapitel. Allerdings hinkt sie bei den Praxisleitfäden bereits hinter dem Zeitplan her. Durch Interventionen von Verbänden und Politiker*innen aus zahlreichen EU-Mitgliedsländern wird zudem schon diskutiert, ob und wie die Fristen und Ziele der KI-VO-Umsetzung gelockert beziehungsweise geändert werden sollten.

Verantwortung von Anfang an mitdenken

Die rechtlichen Rahmenbedingungen für KI-Agenten entstehen gerade erst. Viele zentrale Fragen sind offen – beispielsweise die genaue Definition von „Allzweck-Systemen“ oder die praktische Umsetzung menschlicher Aufsicht. Die KI-VO bietet zwar eine Grundlage, doch viele Details fehlen noch. Umso wichtiger ist es daher für Organisationen und Unternehmen, diese Entwicklungen aufmerksam zu verfolgen. Zudem sollten sie beim Einsatz und Entwickeln von KI-Agenten von Anfang an auf ethische Standards setzen und gesetzeskonform handeln (Compliance). KI-Agenten können Fehler machen, unerwünschte Aktionen ausführen oder Sicherheitslücken verursachen. Deshalb ist es unerlässlich, Mechanismen zur Überwachung zu etablieren, Eingriffe in die KI-gestützten Abläufe zu ermöglichen und diese transparent und nachvollziehbar zu gestalten.

Wie funktionieren KI-Agenten technisch?

KI-Agentensysteme bestehen aus einem großen Sprachmodell (Large Language Model, kurz: LLM, wie GPT-4 von Open AI oder Claude von Anthropic) plus einer Software, die das Modell mit der Außenwelt verbindet, was man als „Scaffolding“ bezeichnet (zu deutsch: Gerüstbau). Dieses Scaffolding umfasst: (1) Sogenannte „Reasoning-Frameworks“, die komplexe Aufgaben in Schritte zerlegen, (2) allgemeine Tools, wie Browser oder Code-Interpreter, (3) spezialisierte Schnittstellen, beispielsweise zu Banken, E-Mail-Systemen oder anderen Services. Diese Kombination ermöglicht mehrstufige, vom KI-Agenten autonom ausführbare Aktionen, von der Planung bis zur Umsetzung – die Herausforderung liegt in der Kontrolle dieser Autonomie.

/Hintergrund

KI-Agenten aus rechtlicher Perspektive

Problem of many hands

Im Zusammenhang mit KI-Agenten sprechen manche vom „Problem of many hands“, ein Konzept aus Philosophie und Ethik. Es fragt danach, wie in einem komplexen System mit vielen verschiedenen Akteuren – auch rechtliche – Verantwortung zugewiesen werden kann. Im Zusammenhang mit KI-Agenten bezieht sich das „Problem der vielen Hände“ darauf, dass vom Training des zugrunde liegenden Modells über die Entwickler*innen des KI-Agenten bis zu dessen Einsatz viele Akteure an der Wertschöpfungskette beteiligt sind und darin auch miteinander interagieren.

Hinzu kommen mögliche, von KI-Agenten ausgelöste Interaktionen, etwa die Delegation von Aufgaben an andere KI-Agenten oder Menschen. Ein wesentlicher Unterschied zwischen KI-Agenten und anderen Produkten, die ebenfalls das Ergebnis einer komplexen Wertschöpfungskette sind (Autos oder Maschinen zum Beispiel), ist also, dass die Agenten autonome Entscheidungen treffen und ihr Verhalten in Echtzeit ohne menschliche Intervention ändern können. Geht dabei etwas schief, ist es schwer zu bestimmen, wo genau der Fehler lag und wer dafür verantwortlich ist.

Haftungsfragen und mögliche Lösungsansätze

Will man klären, wer im Schadensfall haftet, gibt es gleich mehrere Herausforderungen. Zunächst ist es nicht einfach, angemessene Sorgfaltspflichten festzulegen. Wie entscheidet man beispielsweise ob ein*e Nutzer*in einem KI-Agenten sorgfältigere Anweisungen hätte geben müssen? Oder ob ein*e Entwickler*in hätte sicherstellen müssen, dass der KI-Agent auch zweideutige

Anweisungen interpretieren kann? Wo können unterschiedliche Akteure entlang der Wertschöpfungskette also wirklich Kontrolle ausüben und wo können sie entsprechend auch verantwortlich – und haftbar – gemacht werden?

Eine weitere Herausforderung ist, dass Schäden, die durch den Einsatz von KI-Agenten entstehen können, nicht immer materieller Natur sind. Auch immaterielle Schäden, wie Diskriminierung oder die Verletzung von Rechten, kommen infrage. Denkbar sind auch systemische Schäden, die erst nach längerer Zeit erkennbar werden: Die Folgen von Misinformation oder makroökonomische Schäden wie Marktverzerrungen zum Beispiel. Wie man damit umgeht, ist bislang nicht ganz klar.

Die KI-Verordnung kann ein Teil der Lösung sein. Sie weist verschiedenen Akteuren in der Wertschöpfungskette, abhängig von ihren individuellen Möglichkeiten, unterschiedliche Pflichten zu: Beispielsweise hinsichtlich erforderlicher Risikobewertungen und Transparenzgebote, oder auch zu technischen Einsatzkontrollen und menschlicher Aufsicht bestimmter Arbeitsgänge.

Die Verordnung wurde aber nicht explizit für KI-Agenten konzipiert und weist deshalb einige Schwachstellen auf: Es ist zum Beispiel nicht ganz klar, in welche der drei Risikokategorien, die die KI-VO definiert, KI-Agenten fallen. Bei Haftungsfragen ist die KI-VO zudem wenig konkret und verweist meist auf die Produkthaftungsrichtlinie (siehe dazu auch die Mai-Ausgabe von **prompt/**) und nationales Recht.
Erste Ergänzungsvorschläge, etwa ein Regulierungsmodell, das sich an den Regeln für autonome Fahrzeuge orientiert, gibt es bereits (siehe dazu das nachfolgende Interview in dieser Ausgabe).

/Nachfrage

„Klare Haftungsregeln würden die Akzeptanz von KI-Agenten fördern“

Je autonomer KI-Agenten funktionieren, desto mehr sollen ihre Entwickler*innen für etwaige Folgen verantwortlich sein. So der Vorschlag einiger Forscher*innen vom Ada Lovelace Institut und der Denkfabrik Interface. Wie man KI-Agenten regulieren könnte und was autonome Autos damit zu tun haben, erklärt uns Julia Smakman.*

Julia Smakman ist wissenschaftliche Mitarbeiterin im Forschungsbereich „Recht und Politik” des Ada Lovelace Institute. Der Fokus von Julia Smakmanliegt auf der wirksamen Regulierung von KI-Systemen einschließlich

Foundation Models im Vereinigten Königreich und in der Europäischen Union (EU).

Foto: Karla Gowlett

iRights.Lab: 2025 soll das „Jahr der KI-Agenten“ sein. Was bedeutet das eigentlich – und wie realistisch ist das?

Julia Smakman: Die Behauptung ist, dass KI-Agenten uns von Chatbots zu etwas führen werden, das in der realen Welt (einigermaßen) autonom handeln kann. Ein Chatbot könnte Ihre Urlaubsplanung erstellen, aber ein Agent könnte auch Ihre Flüge und Hotels buchen. Das klingt großartig, aber es ist nicht klar, wann oder ob KI-Agenten in der Lage sein werden, komplexe, mehrstufige Aufgaben zuverlässig zu erledigen. Wir erleben immer noch, dass Agenten bei relativ einfachen Aufgaben stecken bleiben oder mit unerwarteten Aktionen „aus der Reihe tanzen“. Wir sind noch nicht an dem Punkt angelangt, an dem Sie einem KI-Agenten Ihre Kreditkartendaten anvertrauen können. Es gibt jedoch Agenten, die in der Lage sind, enger definierte Aufgaben in begrenzteren Umgebungen auszuführen. Die Technologie könnte sich rasch weiterentwickeln und zuverlässiger werden, aber es gibt immer noch große Bedenken hinsichtlich Sicherheit, Transparenz und Fairness. Insbesondere besteht die Sorge, dass unsere Rechtssysteme nicht in der Lage sein werden damit umzugehen, wenn viele KI-Agenten untereinander interagieren.

iRights.Lab: Sind wir in Europa rechtlich gut auf den Einsatz von KI-Agenten vorbereitet? Zum Beispiel durch die KI-Verordnung und die kürzlich aktualisierte Produkthaftungsrichtlinie – oder gibt es hier Lücken?

Julia Smakman: Leitplanken wie die KI-Verordnung und die aktualisierte Produkthaftungsrichtlinie sind definitiv ein Schritt in die richtige Richtung. Die Richtlinie entlastet betroffene Verbraucher*innen, indem sie die Beweislast in Haftungsfällen etwas erleichtern (siehe dazu auch die Mai-Ausgabe von **prompt/**). Bei der KI-Verordnung wird viel von der tatsächlichen Umsetzung und Durchsetzung abhängen. Und obwohl KI-Agenten laut Verordnung definitiv „KI-Systeme“ sind, kann es vom Kontext, in dem sie eingesetzt werden, abhängen, ob sie den strengeren Anforderungen für „KI-Systeme mit hohem Risiko“ unterliegen.

iRights.Lab: Sie haben kürzlich zusammen mit Kolleg*innen der Berliner Denkfabrik *Interface* einen möglichen Regulierungsrahmen für KI-Agenten vorgeschlagen. Als Vorbild diente Ihnen die Regulierung selbstfahrender Autos in Großbritannien. Warum haben Sie gerade dieses Beispiel gewählt, und was konnten Sie davon übernehmen?

Julia Smakman: Die britischen Rechtsvorschriften für selbstfahrende Autos sind ein interessantes Beispiel für einen Rechtsrahmen, bei dem sich die Haftung mit zunehmender Autonomie und abnehmender Kontrolle durch den Nutzer, immer weiter von ihm weg verlagert. Obwohl selbstfahrende Autos in einem spezifischeren physischen Umfeld operieren – Straßen im Gegensatz zum Internet – bieten sie eine nützliche Analogie für KI-Agenten, insbesondere wenn man über Haftung und Verantwortung nachdenkt. Die Analogie hat jedoch auch Grenzen. Nach britischem Recht benötigen selbstfahrende Autos eine vorherige Genehmigung, KI-Agenten nicht. Künftige Nutzer müssen also bei der Auswahl eines Agenten möglicherweise mehr Sorgfalt walten lassen. Auch die Versicherung ist für selbstfahrende Autos sehr wichtig. Der Markt für KI-Versicherungen beginnt sich gerade erst zu entwickeln.

iRights.Lab: In Ihrem Ansatz unterscheiden Sie zwischen fünf Autonomiestufen, je nachdem, wie selbstständig ein KI-Agent handelt. Was genau verbirgt sich hinter diesen Stufen – und wie helfen sie bei der Frage, wer im Falle eines Schadens haftet?

Julia Smakman: Wir wollten differenzieren und zeigen, dass „nicht alle Agenten gleich sind“. Es gibt verschiedene Autonomiestufen, und wir sollten unterschiedliche Erwartungen an Entwickler und Nutzer stellen, je nachdem, welchen Einfluss sie auf das Verhalten der Agenten haben und ob sie in der Lage sind, Schäden vorherzusehen und zu verhindern. Auf niedrigeren Autonomiestufen sind die Benutzer eher in der Lage, den Agenten zu kontrollieren, so dass sie mehr rechtliche Verantwortung für die Handlungen des Agenten übernehmen. Bei einem höheren Grad an Autonomie, bei dem die Benutzer weniger in der Lage sind, den KI-Agenten zu kontrollieren, tragen die Entwickler möglicherweise eine größere rechtliche Verantwortung. Dies könnte ein Anreiz für die Entwickler sein, nur hochgradig autonome Agenten mit angemessenen Sicherheitsvorkehrungen zu veröffentlichen, um vorhersehbare Schäden zu verhindern.

iRights.Lab: Welche politischen, rechtlichen oder praktischen Hürden müssten überwunden werden, damit Ihr Modell tatsächlich umgesetzt werden kann?

Julia Smakman: Wir sehen nur begrenztes politisches Interesse daran, neue rechtliche Rahmenbedingungen für KI zu schaffen. Die Industrie wiederum könnte zögern, mehr rechtliche Verantwortung für hochgradig autonome Agenten zu übernehmen, mit dem Argument, dass Regulierung die Innovation hemmt. Wir glauben jedoch, dass klare Haftungsregeln die Akzeptanz von KI-Agenten fördern würden. Die Motivation, den Autonomous Vehicles Act in Großbritannien auf den Weg zu bringen, war sicherzustellen, dass Menschen sich in und mit selbstfahrenden Autos wohl fühlen, wenn diese auf die Straße kommen. Wer möchte schon ein autonomes System nutzen, das er nicht richtig kontrollieren kann, und gleichzeitig rechtlich für alles haften, was schiefgeht – persönlich oder als Teil seines Unternehmens?

iRights.Lab: Solange es noch keinen festen rechtlichen Rahmen gibt: Welchen Rat

würden Sie jemandem geben, der bereits heute mit KI-Agenten arbeitet?

Julia Smakman: Gehen Sie mit der gebotenen Sorgfalt vor und vergewissern Sie sich, dass jeder KI-Agent, den Sie einsetzen, ordnungsgemäß für den Zweck getestet wurde, für den Sie ihn verwenden. Verstehen Sie die Nuancen: Die Risiken von weniger autonomen Agenten, die für spezifischere Aufgaben eingesetzt werden, sind geringer als die Risiken, die mit hochgradig autonomen Agenten mit Internetzugang verbunden sind, die „open-end" Aufgaben ausführen. Wie bei jedem KI-System können Sie haftbar gemacht werden, wenn die KI einen Schaden verursacht. Vergewissern Sie sich also, ob Ihnen wohl dabei ist, dieses Risiko einzugehen. Wenn nicht, sollten Sie überdenken, ob ein KI-Agent die beste Lösung für Sie ist.

**Wir haben die Fragen schriftlich gestellt und Julia Smakman hat schriftlich geantwortet. Das Interview wurde mit DeepL aus dem Englischen ins Deutsche übersetzt und redaktionell auf Richtigkeit überprüft.*

/Glossar

RAG – Retrieval Augmented Generation

Bei der sogenannten Retrieval Augmented Generation (RAG) wird ein großes KI-Sprachmodell mit einer intelligenten Suche kombiniert, um aktuelle und kontextbezogene Antworten zu liefern. Sprachmodelle können – je nach Version – nur das Wissen nutzen, mit dem sie trainiert wurden. RAG erweitert dieses Wissen um zusätzliche, interne oder externe Datenquellen, Dokumente oder Datenbestände.

Beispiel: *Der Chatbot X soll für ein Bauprojekt bestimmte Arbeitsschritte und Klauseln zusammenfassen und dabei automatisch auf bisherige Projektbeschreibungen oder Verträge des Unternehmens zugreifen. Der Zugriff erfolgt über ein technisches Retrieval-Modul im Hintergrund, nicht direkt im Prompt. Der Prompt beschreibt den Auftrag.*

RAG unterstützt außerdem sogenannte Multi-Hop-Fragen. Sie erfordern, dass das System logisch zwischen mehreren Fakten oder Datenquellen „springt" und diese Informationen sinnvoll kombiniert.

Beispiel: *„Wer war der Trainer des Teams, das 2023 die europäische Champions League gewonnen hat?" Um diese Frage zu beantworten, muss das System herausfinden, welches Team 2023 die Champions League gewonnen hat (erster „Hop") und herausfinden, wer zu diesem Zeitpunkt der Trainer dieses Teams war (zweiter „Hop")*

Multi-Hop Reasoning hilft KI-Systemen, komplexe Fragen zu beantworten und menschliches Schlussfolgern besser zu imitieren. Werden hierbei allerdings zu viele, oder wenig relevante Informationen eingebunden, kann das die Antwortqualität verschlechtern.

Neuere oder umfangreichere Pro-Versionen von Chatbots, wie ChatGPT oder Perplexity, können auf aktuelle Quellen im Internet zugreifen. Diese Funktionalität ist kein Teil von RAG im engeren Sinne, sondern basiert auf integriertem Web-Browsing. Das ist funktional ähnlich, aber technisch nicht identisch mit einer RAG-Architektur, bei der gezielt eigene Datenbanken oder Dokumentenbestände eingebunden werden. Zudem bieten manche KI-Plattformen vorgefertigte RAG-Dienste an. Diese standardisierten Aufträge müssen allerdings immer noch individuell angepasst werden, um die gewünschten Ergebnisse zu liefern. Zum Beispiel indem konkrete Ziele definiert und relevante – oft interne – Datenbestände eingebunden werden.

Und rechtlich gesehen? RAG lässt sich als Schnittstelle zwischen generativer KI und Zugriff auf interne Daten beschreiben. Beim Einsatz von RAG ist aus rechtlicher Perspektive der Datenschutz zu beachten – insbesondere bei Behandlung personenbezogener oder sensibler Daten – aber auch der Umgang mit Geschäftsgeheimnissen sowie die von der KI-VO vorgegebenen Transparenzpflichten.

/Weiterlesen

Artikel und Quellen zu KI-Agenten und zu aktuellen Rechtsfragen bei generativer KI

Ein White Paper des World Economic Forum und Capgemini – **the AI Frontier: A Primer on the Evolution and Impact of AI Agents** – beschreibt zentrale Bestandteile von KI-Agenten und erklärt, warum die Kommunikation unter ihnen oft fehlerhaft ist. Es zeigt, wie sie in den Bereichen Bildung, Finanzen, Softwareentwicklung und Gesundheit eingesetzt werden und analysiert die technischen, sozio-ökonomischen und ethischen Risiken, die dabei entstehen. Die Autor*innen schlagen Maßnahmen vor, um diese Risiken zu verringern: ethische Leitlinien, Informationspflichten und aktive Überwachung des Verhaltens von KI-Agenten.

Mit ihrem Bericht „How AI Agents Are Governed Under the EU AI Act“ liefert die Organisation **The Future Society** einen tieferen Einblick plus eingehender Analysen zur Regulierung von KI-Agenten in der EU. Die 60 Seiten umfassende, in Englisch vorliegende Bestandsaufnahme erschien im Juni 2025.

In mehreren **gerichtlichen Verfahren um Rechtsverletzungen bei Trainings von großen Sprachmodellen beziehungsweise KI-Systemen** ergingen kürzlich Entscheidungen, die sich als wegweisend herausstellen könnten. In einem Fall urteilte das Oberlandesgericht Köln (Pressemitteilung des OLG Köln) zu einer Klage gegen den Facebook-Konzern Meta, demzufolge dieser Nutzer*inneninhalte auf Facebook und Instagram für KI-Trainings einsetzen darf. In den USA erhielten der KI-Hersteller Anthropic und ebenfalls Meta teilweise Recht, Kopien und Digitalisate geschützter Bücher massenhaft für KI-Trainings nutzen zu dürfen, wie die Zeit und heise.de berichten.

Ein **ungarisches Gericht** hat hingegen einen urheberrechtsrelevanten Fall gegen einen KI-Anbieter, in dem es ebenfalls um KI-Trainings mit geschützten Werken geht, dem Europäischen Gerichtshof (EuGH) vorgelegt.

Beatriz Botero Arcila ist Rechtsprofessorin an der Pariser Law School und Mitarbeiterin am Berkman Klein Center for Internet and Society an der Harvard University. In ihrem Policy Report: AI Liability Along the Value Chain erörtert sie Fragen der Haftungszuweisung entlang der KI-Wertschöpfungskette, die angesichts der Beteiligung zahlreicher Akteure an der Konzeption, Entwicklung und dem Einsatz von KI-Systemen komplex sind. Zudem untersucht sie verschiedene politische Optionen für die Gestaltung von Haftungsregelungen. Der Report entstand mit Unterstützung von Mozilla und ist unter der freien Lizenz CC-BY 4.0 veröffentlicht.

/Über uns | Impressum

Diesen Newsletter erstellen und produzieren wir, das iRights.Lab**,** im Rahmen des Forschungsprojekts *Generative KI – Innovation und Recht in Arbeitsprozessen (GenKI-IR)*, gefördert vom Bundesministerium für Bildung und Forschung (BMBF).

Das Projekt hat zum Ziel, rechtliche und gesellschaftliche Rahmen zu beschreiben, mit denen Anwendungen generativer KI unterschiedliche Arbeitsprozesse sinnvoll unterstützen können.

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

Texte: Elena Kalogeropoulos, Merlin Münch, Henry Steinhau, Ella Jordan
Redaktion: Henry Steinhau (verantwortlich), Merlin Münch

Mitarbeit und Lektorat: Solvejg Gunkel
Rechtehinweis: Alle Texte und Bilder stehen unter der offenen Lizenz CC-BY 4.0, außer das Porträt Julia Smakman, Foto: Karla Gowlett.

iRights.Lab GmbH

www.irights-lab.de

Oranienstr. 185

10999 Berlin

prompt@irights-lab.de

Tel: +49 30 40 36 77 230

Fax: +49 30 40 36 77 260

Steuernr. 37/359/5262 | USt-ID DE311181302

Registergericht:

Amtsgericht Charlottenburg Nr. HRB 185640 B

Geschäftsführer*innen:

Philipp Otto, Dr. Wiebke Glässer