prompt/

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

April 2025

Liebe Leser*innen,
dass **Chatbots** systembedingt dazu neigen, sich bei Antworten, Texten, oder Bildern im Wortsinn etwas zusammenzureimen, ist bekannt. Das kann in harmlosen Fällen amüsieren und nicht selten nerven – aber mitunter auch **ernsthafte Schäden** anrichten. Etwa, weil die KI Falschaussagen über Menschen generiert oder irreführende Empfehlungen halluziniert, was zu Mobbing oder vermeidbaren Fehlinvestitionen führen kann. Wer haftet dann dafür? Wer muss etwaige Entschädigungen leisten? Kann man sich gegen solche Schäden versichern – mit einer KI-Haftpflicht- oder Schadens-Versicherung zum Beispiel?
Diese und ähnliche Fragen sind der Ausgangspunkt für unser **Monatsthema, KI und Haftung**. Stand Frühjahr 2025 können wir – Sie ahnen es vielleicht – noch wenig konkrete Antworten geben. Gleichwohl zeigten unsere Recherchen und Gespräche, dass hier in Wirtschaft und Politik viel in Bewegung ist. Die EU kippte zwar kürzlich die geplante **KI-Haftungsrichtlinie**, bezieht aber dafür KI-Systeme in die überarbeitete Produkthaftungsrichtlinie ein; siehe dazu unter anderem unser Interview mit IT-Rechtsanwalt Niko Härting. Mehrere Versicherungsunternehmen dienen sich mit spezifischen Produkten den KI-Herstellern an, um sie bei Gewährleistungen und Risikomanagement zu unterstützen.
Außerdem fassen wir zusammen, wen die **Verbote der KI-Verordnung (KI-VO)** betreffen, die seit Anfang Februar in Kraft sind. Und wir erklären ausführlich, was die EU mit den sogenannten **Praxisleitfäden** beabsichtigt, die Ende Mai an die KI-VO angedockt werden.
Wie gewohnt empfehlen wir abschließend aufschlussreiche Beiträge zu Rechtsfragen rund um generative KI. Darunter zwei **Policy-Publikationen zu Haftungsfragen bei KI-Systemen und KI-Agenten** sowie ein Artikel über die kuriose Entdeckung eines **„digitalen Fossils“**.

Wir wünschen Ihnen viel Freude und Erkenntnisgewinne beim Lesen.
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-Systeme und Haftung

Worum geht's?

Wer haftet, wenn durch den Einsatz von KI ein Schaden entsteht? Diese Frage ist nicht

immer eindeutig zu beantworten. Sie wird aber zunehmend relevanter, je komplexer KI-Systeme werden, je häufiger sie in unterschiedlichen, auch sensiblen Bereichen zum Einsatz kommen und je autonomer sie eine ganze Kette von Aufgaben erledigen, etwa durch den Einsatz von sogenannten KI-Agenten (siehe dazu auch das Policy Briefing von interface).

Niemand will Schuld haben

Stellen Sie sich folgendes Szenario vor: Eine bekannte Investmentfirma verschickt jeden Morgen ein KI-generiertes Börsen-Briefing. Die Firma bewirbt ihr Briefing aggressiv als Must-read, mit dem die Leser*innen stets auf dem neuesten Stand seien. Sie setzt einen KI-Agenten ein, um das Briefing zu erstellen. Laut Anbieter liefert der Agent besonders verlässliche und akkurate Ergebnisse. Er erstellt täglich eine Übersicht über die wichtigsten Entwicklungen am Markt und versendet das Briefing an eine ausgewählte Gruppe zahlender Abonnent*innen.

Nun unterläuft der KI-Software ein folgenschwerer Fehler: Sie berichtet fälschlicherweise vom bevorstehenden Wertverlust der Aktie eines großen Automobilherstellers. Nervöse Leser*innen sind alarmiert, verkaufen daraufhin ihre Anteile, verlieren viel Geld und verklagen schließlich die Investmentfirma. Die aber sieht die Schuld beim KI-Anbieter: Schließlich hatte der versprochen, dass die KI-Software verlässliche Ergebnisse liefert! Der Anbieter weist jegliche Verantwortung von sich: In den Nutzungsbedingungen sei ausdrücklich erwähnt, dass man keine Gewähr für die Korrektheit der Ergebnisse gebe. Außerdem sollten diese selbstverständlich immer von einem Menschen überprüft werden. Und überhaupt müsse man erstmal nachweisen, dass der Fehler wirklich im System liege.

Bei Schäden, die durch den Einsatz von KI entstehen, wird die Verantwortung gerne mal weitergereicht.
Bild: imgflip.com und eigene Bearbeitung

Wie ist die Lage?

Der fiktive Fall soll zeigen: Haftungsfragen rund um den Einsatz von KI sind komplex. Theoretisch ist es möglich, nach den so genannten allgemeinen Haftungsgrundsätzen des Rechts zu ermitteln, wer in einem solchen Fall zur Verantwortung gezogen werden könnte – und wer nicht. Speziell auf KI bezogene Regelungen gibt es bisher allerdings nicht. Klar ist nur: Ein KI-Programm selbst kann *nicht* haftbar gemacht werden. Welche Gesetze greifen also?

- Das Bürgerliche Gesetzbuch (BGB): Es regelt zum Beispiel die *vertragliche Haftung* (die KI macht nicht, was vertraglich vereinbart ist) und die *deliktische Haftung* (ein Schaden entsteht zum Beispiel durch eine vernachlässigte Sorgfaltspflicht oder unzureichende Sicherheitsvorkehrungen).
- Das deutsche Produkthaftungsgesetz (ProdHaftG): Hier ist umstritten, ob und unter welchen Bedingungen KI-Anwendungen als „Produkt" eingestuft werden können. Die Überarbeitung der EU-weiten Produkthaftungsrichtlinie soll für Klarheit sorgen (siehe unten).

Was macht die EU?

Die EU Kommission hat eine lange geplante KI-Haftungsrichtlinie jüngst zurückgezogen. Sie sollte Haftungsfragen mit Blick auf KI-Systeme regeln, die laut KI-Verordnung als Hochrisiko-Systeme eingestuft werden. Außerdem sollte es für Geschädigte leichter werden, Schadensersatzansprüche durchzusetzen. Stattdessen wurde nun die sogenannte Produkthaftungsrichtlinie überarbeitet. Software – und damit auch KI-Anwendungen – wird darin künftig eindeutig als Produkt eingestuft. Haftungsfragen sollen so leichter geklärt werden können. (Siehe hierzu das Interview mit IT-Rechtsanwalt Niko Härting.)

Was heißt das für Verbraucher*innen und Unternehmen?

KI-Anbieter*innen und Hersteller*innen sollen nach dieser Novellierung leichter haftbar gemacht werden können. Außerdem kehrt die Richtlinie die Beweislast um: Es wird per Default davon ausgegangen, dass KI-Systeme Schäden verursachen können. Bisher mussten Kläger*innen das eindeutig nachweisen. Ab sofort sind Hersteller*innen und Anbieter*innen in der Pflicht, dies zu widerlegen. Für Unternehmen, die KI-Anwendungen anbieten – laut KI-Verordnung sowohl KI-System-Hersteller als auch KI-System-Betreiber – bedeutet das unter anderem, dass sie Veränderungen an ihren Modellen für den Schadensfall dokumentieren müssen; darunter fallen jegliche Modifikationen, Updates und das Fine Tuning.

Die Rechtsanwaltskanzlei Härting kommt daher zu der Einschätzung, dass

Anbieter*innen „ihre Entwicklungs- und Dokumentationsprozesse überprüfen und gegebenenfalls anpassen [müssen], um den neuen Anforderungen gerecht zu werden und Haftungsrisiken zu minimieren und die [...] im Einzelfall ggf. treffende Beweislast erfüllen zu können.“ Um die KI-Unternehmen darin zu unterstützen, dass deren Modelle verlässlich funktionieren und performen, bieten mehrere Versicherungen bereits KI-bezogene Produkte an. Darunter die **Munich Re** und die **Swiss Re** (siehe dazu auch der Beitrag zu Versicherungen für KI-Unternehmen weiter unten).

Die EU-Kommission zieht die geplante KI-Haftungsrichtlinie zurück – und nun?

IT-Rechtsanwalt Niko Härting über die Kehrtwende der EU in Sachen Produkthaftung bei KI-Systemen

***Niko Härting** ist Fachanwalt für IT-Recht und Gründer der Berliner Kanzlei Härting Rechtsanwälte. 2012 wurde er zum Honorarprofessor an der Hochschule für Wirtschaft und Recht ernannt, an der er seit dem Jahr 1991 Lehrbeauftragter ist.*
Foto: Mit freundlicher Genehmigung von Niko Härting

prompt/: Die EU arbeitete mehrere Jahre daran, die sogenannte KI-Haftungsrichtlinie (KI-HaftRL) auf den Weg zu bringen. Damit wollte sie Haftungsfragen zu KI-Systemen regeln, die laut KI-Verordnung als Hochrisiko-Systeme eingestuft werden. Nun zog die EU-Kommission diese Haftungsrichtlinie vor kurzem zurück. Warum?
Niko Härting: Aus Brüssel hört man, dass sich die Regierungen der Mitgliedsstaaten von dem Richtlinienentwurf nicht überzeugen ließen. Zum einen fragte man sich – wohl nicht ganz zu Unrecht – ob das bestehende, strenge europäische Produkthaftungsrecht nicht ausreicht. Zum anderen sind die Mitgliedsstaaten immer skeptisch bei europäischer Regulierung, die in das Zivilrecht der Mitgliedsstaaten eingreift.
prompt/: Ist die Richtlinie damit endgültig vom Tisch?
Niko Härting: Einige Abgeordnete, beispielsweise Axel Voss (CDU/EVP, Anmerkung der Redaktion) und einige Verbände, sind sehr unzufrieden über den Rückzug des Richtlinienentwurfs. Aber in Brüssel ist immer die Kommission Herrin des Verfahrens. Wenn die Kommission einen Regulierungsentwurf wie die KI-Richtline zurückzieht, ist der Entwurf vom Tisch. Ob und wann die Kommission einen neuen Anlauf nimmt, ist nicht absehbar.
prompt/: Es gibt ja noch die sogenannte Produkthaftungsrichtlinie (ProdHaftRL),

die bald auch für KI-Anwendungen gelten soll (siehe nachfolgender Text). Warum wollte man überhaupt noch eine weitere KI-spezifische Richtlinie auf den Weg bringen?
Niko Härting: Die KI-Haftungsrichtlinie sollte die Produkthaftungsrichtlinie im Hinblick auf die spezifischen Risiken und Herausforderungen von Hochrisiko-KI-Systemen ergänzen. So sollte beispielsweise der Nachweis von Kausalität und Verschulden im Rahmen von Schadensersatzansprüchen bei Hochrisiko-KI-Systemen vereinfacht werden. Ein ,,Flickenteppich" sollte verhindert werden.
prompt/: Was bedeutet der Rückzieher der EU-Kommission für Verbraucher*innen und Unternehmen in der EU?
Niko Härting: Es bleibt abzuwarten, ob die nationalen Gesetzgeber jetzt Hochrisiko-KI regulieren werden. Praxisfälle, an denen sich Regulierungsbedarf ableiten lässt, gibt es nach meiner Kenntnis nicht.

Die Produkthaftungsrichtlinie der EU und wie sie bezüglich KI verändert werden soll

Die in der EU geltende Produkthaftungsrichtlinie wurde zuletzt 1999 überarbeitet. Sie enthält Regelungen zur sogenannten verschuldensunabhängigen Haftung des Herstellers für fehlerhafte Produkte und regelt insbesondere die Voraussetzungen, unter denen ein Hersteller für Schäden haftet, die durch ein fehlerhaftes Produkt verursacht wurden. Im Zuge der vorbereiteten Novellierung soll Software – und damit auch KI-Anwendungen – künftig eindeutig als Produkt eingestuft werden. Ziel ist, Rechtsklarheit zu schaffen und den Schutz der Verbraucher*innen zu stärken, womit sich Haftungsfragen leichter klären lassen sollen.

Versicherungen für KI-Unternehmen

Wenn KI-gestützte Werkzeuge nicht das leisten, was die Hersteller*innen ihren Kund*innen vertraglich zusichern, kann das für beide Seiten negative Folgen haben. Das ruft Versicherungsunternehmen auf den Plan, die den Unternehmen hierfür neue Produkte anbieten, etwa um Leistungsgarantien geben zu können.

Das Münchener Unternehmen Munich Re bietet für Hersteller*innen, Anbieter*innen und Betreiber*innen von KI-Systemen schon länger ein solches Produkt an. Die Versicherung soll garantieren, dass es bei Nutzung von KI-Modellen weder zu plötzlichen Leistungseinbußen kommt, noch zu Verzögerungen bei der unternehmensweiten Einführung von KI. Die Munich Re setzt dabei auf mehrere Softwareprodukte. Beispielsweise ein sogenanntes Cybersicherheits-Framework, das

Schadsoftware identifiziert und blockiert oder eine Software zur Früherkennung von Betrugsversuchen und Cyberangriffen.

Einen ähnlichen Ansatz verfolgt der Schweizer Rückversicherer Swiss RE, der mit dem kanadischen Start-up Armilla kooperiert, das bestimmte Leistungsgarantien für KI-Anwendungen bietet. Das von Armilla entwickelte System bewertet anhand einer Kombination aus acht Faktoren die Fähigkeiten und Merkmale des KI-Modells: Beispielsweise den Umgang mit Trainingsdaten, wer das System entwickelt hat, wie es in Tests abschneidet und wie der*die Kund*in das Modell nutzt. So will Armilla garantieren, dass die KI-Modelle den versprochenen Nutzen liefern. Sollte das Modell nicht den Anforderungen entsprechen und eine Nachbesserung nicht möglich sein, soll den Kund*innen eine Kostenerstattung zustehen.

Das kalifornische Unternehmen Vouch geht noch einen Schritt weiter. Es bietet eine spezielle KI-Versicherung für KI-Unternehmen an, die auf zentrale Risikofelder zielt und finanzielle Schäden durch fehlerhafte KI-Leistungen absichern will. Ein enthaltener Bias- und Diskriminierungsschutz soll etwa bei Klagen wegen unfairer Entscheidungen durch KI greifen. Zudem wird eine Absicherung gegen Urheber- oder Patentrechtsverletzungen versprochen. Ergänzend soll eine Cyber-Versicherung vor KI-gestützten Angriffen, Datenlecks oder Systemmanipulationen schützen.

Fazit

Die genannten Versicherungsunternehmen positionieren sich auf dem Feld KI mehrheitlich mit Risikoanalysen und Leistungsgarantien für KI-Hersteller und KI-Betreiber. Sie bieten damit Dienstleistungen im Sinne von Qualitätssicherung, um damit die KI-Unternehmen davor zu bewahren, dass ihre KI-Produkte Schäden verursachen, für die sie womöglich haften müssen. Ob es künftig von Versicherungsunternehmen auch Policen gibt, die derartige Produkthaftungsfälle kompensieren sollen, bleibt abzuwarten – scheint aber wahrscheinlich. Denn es ist damit zu rechnen, dass die Fälle zunehmen, in denen Menschen, Firmen oder Organisationen durch KI-Fehler, Bias oder KI-vermittelte Diskriminierung geschädigt werden und die KI-Firmen dafür eine Lösung brauchen.

Drei Haftungsverfahren

Man könnte sie als Vorboten betrachten für vermutlich kommende juristische Auseinandersetzungen: Verfahren, in denen Haftung und Entschädigung verhandelt werden, die auf falsche Ergebnisse von oder fehlerhaften Umgang mit KI-Systemen zurückgehen. Zumindest lassen drei Fälle aus Deutschland, den USA und Kanada Rückschlüsse auf diesbezügliche Kernfragen zu.

Haftung für Falschinformationen

Wer klagt?
Ein mittelständisches Unternehmen klagt am Landgericht Kiel gegen einen Wirtschaftsauskunftsdienst.
Was ist passiert?
Der Wirtschaftsauskunftsdienst veröffentlichte eine KI-gestützte Analyse, in der fälschlicherweise behauptet wurde, dass ein Unternehmen wegen Vermögenslosigkeit aus dem Handelsregister zu löschen sei. Das betroffene Unternehmen klagte daraufhin auf Unterlassung.
Wie lautet das Urteil?
Der Beklagte wurde zur Unterlassung und zur Zahlung vorgerichtlicher Rechtsanwaltskosten verurteilt.
Wie lautet die Begründung?
Der Auskunftsdienst haftet für die durch die KI erzeugten Falschinformationen. Laut Landgericht Kiel liegt eine Verletzung des Unternehmenspersönlichkeitsrechts vor. Der Dienst kann sich nicht darauf berufen, nicht an dem automatisierten Verfahren beteiligt gewesen zu sein, da er das KI-System bewusst zur Beantwortung von Suchanfragen eingesetzt hat. Der Beklagte habe sich die KI-generierten Inhalte zu eigen gemacht und nach außen erkennbar die inhaltliche Verantwortung übernommen.
Quelle: BBS Rechtsanwälte

Haftung für seelische und körperliche Schäden

Wer klagt?
Das Social Media Victims Law Center und das Tech Justice Law Project klagen in Texas gegen Character Technologies.
Was ist passiert?
Die Eltern zweier texanischer Teenager klagen gegen Character.AI, den Anbieter eines Chatbots, der es Usern ermöglicht, mit prominenten oder fiktiven Personas – so genannten Charakteren – zu interagieren. Die Kläger*innen behaupten, ihre Kinder hätten nach längerer Interaktion mit dem Chatbot problematische Verhaltensweisen entwickelt. Eines der Kinder sei etwa gewalttätig gegenüber seinen Eltern geworden, nachdem diese versucht hatten, seine Bildschirmzeit zu verringern. Ein Screenshot von einer Konversation zwischen dem 17-Jährigen und dem Chatbot zeigt, wie dieser dem Jungen vorschlägt, dass er doch seine Eltern töten könne, um die Bildschirmzeit nicht verringern zu müssen. Die Kläger*innen argumentieren, dass das Unternehmen in mehrfacher Hinsicht fahrlässig gehandelt hätte. Zum Beispiel sei es zu fahrlässigen Fehlern im Design des Systems gekommen zudem hätten Warnhinweise zu den möglichen Risiken gefehlt, die mit der Nutzung des Chatbots verbunden sind.
Wie lautet das Urteil?
Das Verfahren läuft noch.

Quelle: The National Law Review

Schadensersatz für fehlerhaften Chatbot

Wer klagt?
Ein Passagier gegen Air Canada
Was ist passiert?
Ein KI-Chatbot von Air Canada lieferte einem Passagier falsche Informationen zu einem vermeintlichen Rabatt auf Flugreisen. Der Chatbot hatte behauptet, dass Reisende, die wegen eines Todesfalls in der Familie reisen, bis zu 90 Tage nach ihrer Reise einen sogenannten Trauerrabatt beantragen können. Zwar gibt es diesen Rabatt tatsächlich, allerdings kann er nicht rückwirkend geltend gemacht werden. Air Canada wies den Antrag des Klägers auf Rückerstattung also ab.
Wie lautet das Urteil?
Das Gericht entschied zugunsten des Klägers auf Schadensersatz. Air Canada sei für alle Informationen auf seiner Website verantwortlich, unabhängig davon, ob sie von einem Chatbot oder einer statischen Website stammen. Vom Kunden könne nicht erwartet werden, dass dieser die Informationen, die ihm das Unternehmen auf seiner Webseite anzeigt, nochmals eigenständig überprüft.
Quelle: American Bar Association

`/Hintergrund`

Was sind die Praxisleitfäden der EU zur KI-Verordnung?

Ursprünglich wollte die EU mit der KI-Verordnung solche KI-Systeme adressieren, die zu einem eindeutigen Zweck entwickelt und auf den Markt gebracht werden. Dass auch KI-Modelle reguliert werden müssen, deren Einsatzzweck vorab nicht eindeutig bestimmt werden kann, wurde erst später klar – im eigentlichen Gesetzgebungsverfahren zwischen dem Europäischen Parlament und den Europäischen Mitgliedsstaaten. Einige der in der KI-VO vorgesehenen Regulierungsmechanismen, beispielsweise zu Risikomanagement, Transparenz, Menschenrechtsschutz oder Urheberrecht, erschienen dem EU-Gesetzgeber daher ohne konkrete Beispiele für mögliche Einsatzzwecke unbrauchbar. Also fanden EU-Parlament und -Mitgliedsstaaten eine Kompromisslösung: Die in der KI-VO vorgesehen gesetzlichen Vorgaben sollten nachträglich in einem sogenannten **Code**

of Practice genauer definiert werden. Dieser **Praxisleitfaden** enthält Pflichten und Auflagen für Unternehmen, ist jedoch rechtlich nicht bindend (mehr dazu im Glossar-Text unter diesem Beitrag). Gleichwohl sollen die darin festgehaltenen Vorgaben als umsetzbare Anreize dienen, rechtliche Mindeststandards einzuhalten. Weil diese Regelungen mit Beteiligung der Unternehmen entwickelt wurden, könnten sie der EU außerdem als belastbare Blaupause für künftige gesetzliche Regulierungen dienen.

Vier Arbeitsgruppen

Im September 2024 begannen vier Arbeitsgruppen – aus internationalen Fachleuten bestehend – mit der Entwicklung der thematisch unterschiedlichen Praxisleitfäden. Unternehmen, NGOs und Gemeinwohlverbände konnten mit Stellungnahmen zu den Inhalten beitragen. Am 11. März 2025 veröffentlichte die EU-Kommission den dritten und voraussichtlich letzten Entwurf dieser Leitfäden. Diesem Third Draft gingen Eingaben und Diskussionsbeiträge von deutlich mehr als 1.000 Unternehmen, NGOs und Gemeinwohlverbänden voran. Nach seiner Veröffentlichung kam es zu zahlreichen, deutlich vernehmbaren Reaktionen.

Kritik aus der Zivilgesellschaft

Zivilgesellschaftliche Organisationen, wie das Centre for Democracy & Technology Europe kritisieren den aktuellen Entwurf als unzureichend: Die Anbieter von KI-Modellen würden nicht hinreichend präzise und verbindlich auf die Achtung und den Schutz von Menschenrechten verpflichtet. (Siehe auch hier.)

- Das gemeinwohlorientierte AppliedAI Institute for Europe bemängelt, dass bestimmte Risiken als weniger wichtig eingestuft würden – zum Beispiel hinsichtlich IT-Sicherheit und in Bezug auf Menschenrechte.
- Die Europäische Verbraucherschutzorganisation BEUC (Bureau Européen des Unions de Consommateurs) betont unzureichende Transparenzverpflichtungen der KI-Anbieter. Das führe letztlich zu Unsicherheiten beim Schutz personenbezogener Daten.
- Verlage, Urheber*innen und Verwertungsgesellschaften kritisieren, dass der finale Entwurf keine hinreichenden Möglichkeiten zur Durchsetzung von Urheberrechten böte. Dafür brauche es einen effektiven Zugang zu Trainingsdaten.
- Die Organisation Reporter ohne Grenzen (Reporters sans Frontières, RSF) hat sich im März 2025 aus dem Beteiligungsprozess zurückgezogen, da sie den zu großen Einfluss der Tech-Industrie kritisiert.
- Eine Gruppe von namhaften EU-Parlamentarier*innen schloss sich dieser Kritik in einem Brief an die zuständige EU-Kommissarin an: Der AI Act drohe zugunsten von vor allem US-amerikanischen Tech-Unternehmen zu verwässern.

Industrie und Wissenschaft bemängeln strenge Regeln

Manche KI-Anbieter halten die im Praxisleitfaden vorgesehenen Selbstverpflichtungen und Umsetzungsmaßnahmen hingegen für zu streng:

- Die Computer and Communications Industry Association (CCIA) – internationaler Handelsverband und Interessenvertretung für Kommunikations- und Technologieunternehmen – formuliert hinsichtlich der diskutierten Transparenzverpflichtungen Bedenken mit Blick auf eigene Geschäftsgeheimnisse. Problematisch sei auch die vorgesehene Risikobewertung und Auditierung durch Drittstellen.
- Der französische Anbieter Mistral AI betonte, dass die Einhaltung des AI Acts in der Praxis schon jetzt schwierig sei.
- Auch seitens der Wissenschaft melden sich Stimmen, die die Umsetzbarkeit des finalen Entwurfs kritisieren – sei es mit Blick auf das Urheberrecht oder hinsichtlich des Risikomanagements.

Die finale Version soll im Mai vorliegen

Die bis Ende März bei der Kommission eingereichten Reaktionen zum Third Draft werden von der Europäischen Kommission in die Erstellung des finalen Code of Practice einbezogen. Ziel ist es, diesen im Mai 2025 zu veröffentlichen – einen Monat später als vorgesehen.

/Glossar

Code of Practice (deutsch: Praxisleitfaden)

Innerhalb der Regelungsmechanismen der EU meint ein *Code of Practice* ein rechtlich nicht bindendes Dokument mit Pflichten und Auflagen für Unternehmen, auch *Soft Law* genannt. Er wird in Abstimmung mit den Unternehmen erstellt, die den Code umsetzen sollen. Im Kern dient er ihnen als Selbstverpflichtung. Das heißt, Unternehmen, die die Vorgaben des Codes nicht umsetzen, können nicht sanktioniert, sprich mit Strafen belegt werden. *Codes of Practice* sind vergleichsweise neu und erst in wenigen Politikbereichen der EU etabliert. Der erste *Code of Practice* erschien 2018 im Rahmen der Bekämpfung von Desinformation auf Onlineplattformen.
Der Rückgriff auf *Soft Law* soll der EU die Chance zu flexibler Regulierung bieten, die sie sich an Praxiserfahrungen orientiert und schrittweise entsteht. Tatsächlich wurde der Code zu Desinformation evaluiert und bereits zweimal überarbeitet.

/KI-Verordnung

Verbote bestimmter KI-Systeme: Was seit Februar gilt – und für wen

Seit dem 2. Februar 2025 sind KI-Unternehmen in den EU-Mitgliedstaaten verpflichtet, verbotene Praktiken nach Artikel 5 der KI-Verordnung einzustellen. Gemeint sind Praktiken, die als unannehmbares Risiko für die Sicherheit, die Gesundheit und die Grundrechte von Personen gelten.

Welche KI-Praktiken sind jetzt (teilweise) verboten?

Das Verbot umfasst das Inverkehrbringen, die Inbetriebnahme oder Verwendung folgender KI-Systeme:
(Die folgenden Absätze sind direkt dem Artikel 5 der KI-VO entnommen, zum besseren Verständnis jedoch sprachlich angepasst.)

- Manipulative oder täuschende KI-Systeme, die das Ziel oder die Wirkung haben, das Verhalten einer Person oder einer Gruppe von Personen wesentlich zu beeinflussen – mit dem Ziel oder der Wirkung, dass den Personen ein erheblicher Schaden zugefügt wird oder zugefügt werden kann.
- KI-Systeme, die Schwachstellen einer natürlichen Person oder einer Gruppe von Personen aufgrund ihres Alters, ihrer Behinderung oder ihrer spezifischen sozialen oder wirtschaftlichen Situation ausnutzen – mit dem Ziel oder der Wirkung, dass diesen Personen ein erheblicher Schaden zugefügt wird oder zugefügt werden kann.
- KI-Systeme zur Bewertung oder Klassifizierung natürlicher Personen oder Personengruppen mit dem Ergebnis der Schlechterstellung oder Benachteiligung bestimmter Personen (Social Scoring).
- KI-Systeme zur Bewertung oder Vorhersage des Risikos, dass eine natürliche Person eine Straftat begeht (Predictive Policing) (Ausnahmen gelten für KI-Systeme, die menschliche Bewertung unterstützen, die bereits auf objektiven und nachprüfbaren Fakten beruht, die direkt mit krimineller Aktivität verbunden sind).
- KI-Systeme, die durch das ungezielte Auslesen von Gesichtsbildern aus dem Internet oder von Überwachungskameraaufnahmen Gesichtserkennungsdatenbanken erstellen oder erweitern.
- KI-Systeme zur Ableitung von Emotionen einer natürlichen Person am Arbeitsplatz und in Bildungseinrichtungen (Ausnahme: medizinische oder Sicherheitsgründe).
- Biometrische Kategorisierungssysteme, um politische Meinungen, die sexuelle Orientierung, religiöse Überzeugungen und ähnliches abzuleiten (Ausnahme gelten für die Kennzeichnung beziehungsweise Filterung rechtmäßig erworbener Daten, wie beispielsweise zur Strafverfolgung).
- KI-Systeme, die in Echtzeit eine Fernbiometrie-Identifizierung in öffentlich zugänglichen Räumen zu Strafverfolgungszwecken durchführen (Ausnahmen gelten bei bestimmten, schweren Straftaten).

Worauf müssen Unternehmen jetzt achten?

Die eben erläuterten Verbote der KI-Verordnung sind nicht auf eine bestimmte Branche begrenzt. Die Geldbuße bei einem Verstoß kann unterschiedlich hoch sein. Sie wird im Einzelfall und unter anderem an der Größe und am Jahresumsatz des Unternehmens bemessen. Unternehmen, die von der Regelung betroffen sein könnten, sollten also sicherstellen, dass die von ihnen entwickelten oder eingesetzten KI-Systeme nicht gegen die Verbote der KI-Verordnung verstoßen. Unternehmen können sich an den Leitlinien der EU-Kommission orientieren, um festzustellen, ob eine der Verbotsregeln auf sie zutrifft.

Was passiert, wenn Unternehmen ein Verbot nach der KI-Verordnung nicht einhalten?

Sofern ein Unternehmen ein Verbot nach Artikel 5 der KI-Verordnung missachtet, werden Geldbußen von bis zu 35 Millionen Euro oder von bis zu sieben Prozent des gesamten weltweiten Jahresumsatzes (des vorangegangenen Geschäftsjahres) verhängt, je nachdem, welcher Betrag höher ist. Handelt es sich um EU-Organe, Einrichtungen oder sonstige Stellen der Union, beträgt eine solche Geldbuße bis zu 1,5 Millionen Euro.

Wer überwacht die Umsetzung der Verbote und verhängt die Geldbußen?

Die EU-Mitgliedsstaaten müssen bis spätestens August 2025 benennen, welche Behörde die Umsetzung der Verbote im Land überwachen soll. In Deutschland wird diese Aufgabe voraussichtlich die Bundesnetzagentur übernehmen.

`/Weiterlesen`

Artikel und Quellen zu generativer KI

Wer noch mehr über die **KI-Haftungsrichtlinie** wissen will, warum es sie gebraucht hätte und was ohne sie gilt, kann das im Gutachten von Philipp Hacker für das Europäische Parlament (PDF) sehr genau nachlesen.

In ihrem Bericht zur Haftung von KI (AI Liability) (PDF) beleuchtet die **Mozilla Foundation** die Herausforderungen, die sich den Gesetzgebenden hinsichtlich der zivilrechtlichen Haftung von KI-Herstellern stellen. Außerdem erörtern die Autor*innen die Möglichkeiten der Geschädigten, Wiedergutmachung zu verlangen

und sie geben zahlreiche Empfehlungen, wie ein wirksamer KI-Haftungsrahmen aussehen könnte.

Ebenfalls mit Haftungsfragen, insbesondere auf KI-Agenten bezogen, beschäftigt sich das Policy Briefing des **Thinktanks Interface**. Darin beleuchten die Autor*innen unter anderem näher, wie sich Schäden identifizieren und Verantwortungen zuweisen ließen – was gerade bei KI-Agenten, angesichts komplizierter Wertschöpfungsketten und dem Problem der „vielen Hände", als besondere Herausforderung gilt.

Ein digitales Fossili entdeckte das **Wissenschaftsmagazin „The Conversation"**. Durch einen schlechten Scan und eine fehlerhafte Übersetzung gelangte vor vielen Jahren der Nonsensbegriff „vegetative Elektronenmikroskopie" in wissenschaftliche Arbeiten. Nach und nach verbreitete er sich in wissenschaftlichen Veröffentlichungen und wird nun „als Verunreinigung" von großen Sprachmodellen in KI-Outputs gespült. Fazit der Recherche, „es ist nicht einfach, Fehler dieser Art zu finden. Sie zu beheben, kann fast unmöglich sein." *(Danke für den Hinweis auf dieses Fundstück, Georg.)*

Von einem tragischen Fall vermeintlichem KI-Versagens berichtet die Wochenzeitung **Die Zeit**. Es geht darum, dass die spanische Polizei mit Hilfe eines KI-gestützten Systems die Risiken bewertet, die potenziellen Opfern häuslicher Gewalt ausgesetzt sind. Doch nachdem mehrere Bewertungen nachweislich falsch waren und es zu einem Todesfall kam, ist das eingesetzte britische KI-System in der Kritik.

/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).

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

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

Texte: Merlin Münch, Anissa Tammoui, Ella Jordan, Solvejg Gunkel, Matthieu Binder,

Henry Steinhau
Redaktion: Henry Steinhau (verantwortlich), Merlin Münch
Mitarbeit und Lektorat: Solvejg Gunkel, Katja Berg

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