

Rechtliche Rahmenbedingungen in der Schweiz und EU

Mario Marti und Leandra Gafner

KI existiert schon lange. Sie filtert unsere Spam-Mails, übersetzt auf DeepL unsere Texte und entscheidet auf Social Media, was uns angezeigt wird. Mit der Veröffentlichung von ChatGPT Ende November 2022 wurde der Öffentlichkeit gezeigt, wozu generative künstliche Intelligenz fähig ist: Da gab es plötzlich eine generative KI, die einfach in der Anwendung war und auf sämtliche Fragen eine Antwort wusste – oftmals sogar eine korrekte. Die Entwicklung verlief seither rasant. Generative KI wurde in Windeseile Teil des Alltags, wurde in Arbeitsabläufe integriert und prägt bereits heute verschiedenste Branchen. Es scheint, als seien generativer KI keine Grenzen gesetzt.

Genau hiermit setzt sich dieser Beitrag auseinander: Welche gesetzlichen Grenzen und sonstige Rahmenbedingungen bestehen für KI und die Verwendung von KI-Tools?

Regulierung in der Schweiz

Gleich zu Beginn: Die Schweiz kennt bisher kein «KI-Gesetz». Trotzdem muss künstliche Intelligenz den bereits geltenden Gesetzen entsprechen. Die Schweizer Gesetzestexte sind typischerweise technologieneutral verfasst, um auch zukünftige Entwicklungen abzubilden. Beispielhaft sei hier das Datenschutzgesetz erwähnt, welches direkt auf KI anwendbar ist. Vor diesem Hintergrund werden im letzten und umfangreichsten Teil dieses Artikels einige ausgewählte rechtliche Aspekte zum Umgang mit KI in der Schweiz erläutert.

Dass die Schweiz bisher kein «KI-Gesetz» kennt, bedeutet nicht, dass der Gesetzgeber bisher untätig war. So hat der Bundesrat dem eidgenössischen Departement für Umwelt, Verkehr, Energie und Kommunikation UVEK am 22. November 2023 den Auftrag gegeben, bis Ende 2024 eine Übersicht möglicher Regulierungsansätze zu erstellen. Die Analyse des UVEK soll auf bestehendem Schweizer Recht aufbauen und mögliche Regulierungsansätze aufzeigen, die mit den EU-Regulierungen («AI Act», hierzu unten mehr) kompatibel sind. Mit dieser Analyse will der Bundesrat die Basis schaffen, damit er 2025 einen konkreten Auftrag für eine Regulierungsvorlage erteilen kann.¹

Bis die Auslegeordnung des UVEK vorlag, wurden in der juristischen Lehre insbesondere die folgenden Ansätze für eine Regulierung von künstlicher Intelligenz diskutiert, sowie eine Mischung davon:

- **Horizontale Regelung:** Dies würde bedeuten, dass allgemeingültige Vorschriften für den Umgang mit KI eingeführt werden, wie dies die EU mit dem AI-Act bereits gemacht hat (hierzu unten mehr).
- **Vertikale Regelung:** Damit ist gemeint, dass in gewissen Branchen aufgrund eines erhöhten Risikos spezielle KI-Regelungen eingeführt werden, diese jedoch nicht allgemeingültig werden.
- **Selbstregulierung:** Schliesslich gäbe es auch die Option, KI nicht gesetzlich zu regulieren und zu hoffen, dass die Wirkung des AI Acts über die EU-Grenzen zu einer Selbstregulierung der Privatwirtschaft führt. Eine solche Tendenz konnte bei der Einführung der Datenschutzgrundverordnung (DSGVO) in der EU bereits festgestellt werden, als viele Schweizer Unternehmen ihr Bewusstsein für den Datenschutz schärfen und gewisse Massnahmen umsetzen, lange bevor das revidierte Schweizer Datenschutzgesetz in Kraft trat.

Am 12. Februar 2025 legte das UVEK dem Bundesrat seine Auslegeordnung zur Regulierung von künstlicher Intelligenz vor.¹¹ Darin skizziert das UVEK die erwähnten Möglichkeiten und weist insbesondere auf die KI-Konvention des Europarats¹¹¹ (dazu gleich mehr) hin. Gleichentags kommunizierte der Bundesrat, welchen Regulierungsansatz er verfolgen wird.^{11V} Der Bundesrat hat sich wie folgt entschieden:

- Die Schweiz soll die KI-Konvention ratifizieren und in das Schweizer Recht übernehmen.
- Die Schweiz soll, wo Gesetzesanpassungen nötig sind, einen möglichst sektorbezogenen Ansatz verfolgen. In zentralen, grundrechtsrelevanten Bereichen wie dem Datenschutz werden jedoch auch allgemeine Regulierungen angestrebt.
- Neben der Gesetzgebung sollen auch rechtlich nicht verbindliche Massnahmen zur Umsetzung der KI-Konvention erarbeitet werden, beispielsweise Selbstdeklarationsvereinbarungen oder Branchenlösungen.

Insgesamt soll KI so reguliert werden, dass ihr Potential für den Wirtschafts- und Innovationsstandort Schweiz nutzbar gemacht wird und die Risiken für die Gesellschaft möglichst klein bleiben. Der Bundesrat hat einen pragmatischen Mittelweg der oben erwähnten Regulierungsmöglichkeiten gewählt und will nur in wenigen Bereichen auf eine horizontale Regelung setzen. Da die horizontale Regelung insbesondere die KI-Konvention abbilden soll, lohnt sich ein Blick in ebendiese.

Das Ziel der KI-Konvention ist, dass KI-Systeme während ihres gesamten Lebenszyklus mit den Menschenrechten, der Demokratie und der Rechtsstaatlichkeit vereinbar sind. Dafür setzt die KI-Konvention einige Grundsätze fest, welche die beitretenden Länder in Bezug auf KI-Systeme anwenden sollen. Zu diesen Grundsätzen gehören Menschenwürde und Selbstbestimmung, Transparenz und Überwachung, Gleichstellung und Nichtdiskriminierung sowie Privatsphäre und Datenschutz. Um diese Grundsätze umzusetzen, sollen die Mitgliedsstaaten der KI-Konvention einige Rechtsbehelfe, Verfahrensrechte und Schutzmechanismen einführen. Insbesondere sollen die relevanten Informationen über KI-Systeme, welche das Potential haben, die Menschenrechte zu beeinträchtigen, und ihre Nutzung dokumentiert und den Betroffenen zur Verfügung gestellt werden. Damit soll es den Betroffenen ermöglicht werden, die Entscheidungen anzufechten, die durch den Einsatz solcher KI-Systeme getroffen wurden. Weiter soll eine Beschwerdemöglichkeit eingeführt werden, bei welcher wirksame Verfahrensgarantien, Schutzmassnahmen und Rechte für betroffene Personen bestehen, wenn ein KI-System erhebliche Auswirkungen auf die Wahrnehmung der Menschenrechte hat. Schliesslich sollen betroffene Personen darüber aufgeklärt werden, dass sie mit einem solchen KI-System interagieren.

Die KI-Konvention ist vorab auf Staaten und deren Institutionen ausgerichtet, ebenso wie auf private Akteure, welche im Namen von Staaten handeln. Dennoch sollen die Mitglieder der KI-Konvention auch die Risiken und Auswirkungen der Nutzung von KI durch private Akteure regeln. Genau dies sind die Aspekte, welche der Bundesrat horizontal regeln will.

Zum Ende dieses Zukunftsausblickes sei noch darauf hingewiesen, dass der Bundesrat bereits einen Fahrplan für die KI-Regulierung definiert hat. So sollen das Eidgenössische Justiz- und Polizeidepartement (EJPD), das UVEK und das Eidgenössische Departement für auswärtige Angelegenheiten (EDA) bis Ende 2026 eine Vernehmlassungsvorlage erstellen, welche die KI-Konvention umsetzt. Ebenfalls bis Ende 2026 sollen das UVEK, das EJPD, das EDA und das

Eidgenössische Departement für Wirtschaft, Bildung und Forschung (WBF) einen Plan für die weiteren Massnahmen im Bereich der Selbstregulierung erarbeiten.

Regulierung in der EU

Anders als die Schweiz hat sich die EU für eine gesamtheitliche, horizontale Regelung von KI entschieden. Sie hat sehr schnell auf die Entwicklung von KI reagiert und bereits am 1. August 2024 den «AI Act»^V in Kraft gesetzt. Im Folgenden erläutern wir die wichtigsten Aspekte des AI Acts. Damit wollen wir nicht bloss eine mögliche Regulierung von KI aufzeigen: Der AI Act ist unter gewissen Umständen auch direkt auf Schweizer Unternehmen anwendbar.

Risikobasierte Einteilung der KI-Anwendungen

Der AI Act verfolgt einen risikobasierten Ansatz der Regulierung von KI und teilt die KI-Systeme und -Praktiken in vier Risikokategorien ein:

- **Verbotene KI-Praktiken:** Die EU stuft einige KI-Praktiken als so stark risikobehaftet ein, dass sie deren Anwendung gänzlich verbietet. Dazu gehören gemäss Art. 5 des AI Acts beispielsweise Verhaltensmanipulation und gezielte Ausnutzung von Schwächen, Social Scoring, Risikobewertung und Profiling betreffend Straffälligkeit, Erstellung und Erweiterung von Datenbanken zur Gesichtserkennung, Ableitung von Emotionen am Arbeitsplatz und in Bildungseinrichtungen sowie Echtzeit-Fernidentifizierungssysteme zu Strafverfolgungszwecken in öffentlich zugänglichen Räumen (wobei einige Ausnahmen möglich sind).
- **Hochrisiko-KI-Systeme:** KI-Systeme mit hohem Risiko müssen bestimmte Anforderungen erfüllen, wozu insbesondere die Einrichtung und die Umsetzung eines Risikomanagementsystems gehören. Als mit hohem Risiko behaftet gelten gemäss Art. 6 Abs. 1 des AI Acts vorab KI-Systeme, die selbst Produkte oder Sicherheitskomponenten von Produkten sind, die gemäss den in Anhang I aufgeführten Harmonisierungsrechtsvorschriften einer Konformitätsbewertung unterliegen (bspw. gewisse Maschinen, Spielzeuge, Aufzüge, Medizinprodukte, Fahrzeuge). Weiter wird in Art. 6 Abs. 2 des AI Acts auf Anhang III verwiesen, welcher weitere KI-Systeme als mit hohem Risiko behaftet definiert. Dazu gehören beispielsweise KI-Systeme zur biometrischen Fernidentifizierung, zur biometrischen Kategorisierung oder für die Erkennung von Emotionen, KI-Systeme als Sicherheitskomponenten bei der Verwaltung oder in sicherheitskritischen Bereichen, KI-Systeme im Bildungsbereich oder im Arbeitnehmermanagement sowie KI-Systeme für die Kreditwürdigkeitsprüfung oder für die Risikobewertung im Zusammenhang mit Lebens- und Krankenversicherungen.

- **KI-Systeme mit begrenztem Risiko:** Die EU auferlegt einigen KI-Systemen gewisse Transparenzpflichten. Ziel davon ist, dass die Benutzerinnen und Benutzer darüber informiert sind, dass sie ein KI-System verwenden. Diese KI-Systeme werden in Art. 50 des AI Acts definiert und umfassen insbesondere KI-Systeme für die direkte Interaktion mit natürlichen Personen (bspw. Chatbots), KI-Systeme, die synthetische Audio-, Bild-, Video- oder Textinhalte generieren, sowie KI-Systeme zur Erstellung von Deep Fakes.
- **KI-Systeme mit minimalem Risiko:** Der AI Act reguliert nicht sämtliche KI-Systeme. Fällt ein KI-System unter keine der erläuterten Risikokategorien, wird es nicht vom AI Act erfasst.

Neben den KI-Systemen reguliert die EU mit dem AI Act auch Allzweck-KI-Modelle (General Purpose AI Models [GPAIM]; vgl. Art. 51 ff. AI-Act). GPAIM können für viele verschiedene Zwecke eingesetzt werden und ein breites Spektrum unterschiedlicher Aufgaben kompetent erfüllen; als Paradebeispiel für ein GPAIM sei an dieser Stelle ChatGPT genannt. Birgt ein GPAIM systemische Risiken, gelten zusätzliche Pflichten für die Anbieter.

Geltungsbereich des AI Acts

Im Folgenden wird der Geltungsbereich des AI Acts in drei Aspekten beleuchtet. Damit sollen die dringendsten Fragen beantwortet werden: Ab wann gilt der AI Act (zeitliche Geltung)? Für wen gilt der AI Act (räumliche Geltung)? Und gilt der AI Act auch in der Schweiz?

- **Zeitliche Geltung:** Der AI Act entfaltet in vier Schritten seine gesamte zeitliche Geltung. Bereits seit dem 2. Februar 2025 gelten die Regelungen für verbotene KI-Praktiken. Ein Jahr nach dem Inkrafttreten, am 2. August 2025, beginnen weitere Regeln zu gelten, insbesondere jene zu den GPAIM. Am 2. August 2026 finden schliesslich, mit Ausnahme von Art. 6 Abs. 1 (Regeln für KI-Systeme, die als Sicherheitsbauteil eines Produkts bestimmt oder selbst ein Produkt sind), sämtliche Bestimmungen des AI Acts Anwendung. Ein weiteres Jahr später, am 2. August 2027, beginnt auch Art. 6 Abs. 1 des AI Acts zu gelten, womit ab diesem Zeitpunkt der gesamte AI Act gilt.
- **Persönliche Geltung:** Doch wen betrifft der AI Act eigentlich konkret? Der AI Act erfasst in erster Linie Anbieter (Provider) und Betreiber (Deployer) von KI-Systemen. Anbieter ist, wer ein KI-System oder ein GPAIM entwickelt oder entwickeln lässt und es unter dem eigenen Namen oder einer eigenen Marke in der EU in Verkehr bringt oder in Betrieb nimmt (Art. 3 Abs. 3 AI Act). Davon erfasst ist beispielsweise auch ein Unternehmen, das eine bestehen-

de KI-Anwendung auf die eigenen Bedürfnisse anpassen lässt, sofern sie diese dann unter eigenem Namen oder eigener Marke in Verkehr bringt oder in Betrieb nimmt. Der Anbieter trägt den grössten Teil der Pflichten aus dem AI Act.

Betreiber ist, wer ein KI-System in eigener Verantwortung in der EU einsetzt. Explizit nicht als Betreiber gilt, wer ein KI-System im Rahmen einer persönlichen, nicht beruflichen Tätigkeit verwendet (Art. 3 Abs. 4 AI-Act). Der Gebrauch von KI in der Freizeit, ohne gewerbliche Zwecke, wird somit vom AI Act nicht reguliert. Wenn ein Unternehmen aber ein KI-System auch nur für interne Zwecke einsetzt, ohne es weiterzuentwickeln oder als eigenes Produkt anzubieten, ist es ein Betreiber und hat den AI Act einzuhalten. Der Betreiber trägt die Verantwortung, das KI-System regelkonform zu nutzen. Er muss aber nicht die umfassenden Pflichten eines Anbieters übernehmen. Der Übergang vom Betreiber zum Anbieter kann schleichend erfolgen. Es empfiehlt sich deshalb eine regelmässige Überprüfung, ob die Kategorisierung noch korrekt ist.

- **Geltung in der Schweiz:** Der AI Act ist eine Verordnung der EU. Ist er denn für die Schweiz überhaupt relevant? Ja, denn der AI Act kann auch für Schweizer Unternehmen Geltung entwickeln. Sein Anwendungsbereich ist – ähnlich wie jener der DSGVO – extraterritorial. Für die Anwendbarkeit des AI Acts genügt es, dass das KI-System oder das GPAIM in der EU in Verkehr gebracht oder in Betrieb genommen wird, ohne dass der Anbieter in der EU niedergelassen sein muss (vgl. Art. 2 Ziff. 1 lit. a AI-Act). Der AI Act ist zudem für Anbieter und Betreiber von KI-Systemen unabhängig vom Niederlassungsort auch dann anwendbar, wenn die vom KI-System hervorgebrachte Ausgabe beziehungsweise das Resultat in der EU verwendet wird (vgl. Art. 2 Ziff. 1 lit. c AI Act).

Vor diesem Hintergrund müssen auch schweizerische Unternehmen prüfen, ob die von ihnen eingesetzten KI-Anwendungen unter den AI Act fallen. Um diese Prüfung zu vereinfachen, stellt *economiesuisse* in Zusammenarbeit mit Kellerhals Carrard ein KI-Selbsteinschätzungstool zur Verfügung.^{VI} Trotz EU-Bezug den AI Act zu ignorieren empfiehlt sich nicht: Je nach Verstoß drohen Sanktionen in der Höhe von bis zu CHF 35 000 000.00 oder 7 % des weltweiten Jahresumsatzes (vgl. Art. 99 Abs. 3 AI Act).

Einzelfragen aus dem Schweizer Recht

Wie bereits zu Beginn dieses Beitrages festgehalten, kennt die Schweiz bisher kein KI-Gesetz. In diesem Kapitel werden deshalb einige bestehende Gesetze beleuchtet, welche bereits heute für KI relevant sind und es wird ausgeführt, was das für die Anwendung von KI bedeutet.

Datenschutzrecht

Mit einer Kurzmeldung hat der eidgenössische Datenschutz- und Öffentlichkeitsbeauftragte (EDÖB) am 9. November 2023 kommuniziert, dass das (damals noch sehr neue) Datenschutzgesetz direkt auf KI-gestützte Datenbearbeitungen anwendbar ist.^{VII} In dieser Kurzmeldung hebt der EDÖB weiter einige datenschutzrechtliche Grundsätze hervor, welche bei der Datenbearbeitung mittels KI besonders zu beachten sind.

Der EDÖB nimmt vorab die Hersteller, Anbieter und Verwender von KI-Anwendungen in die Pflicht, bereits bei der Entwicklung neuer Technologien und bei der Planung ihres Einsatzes sicherzustellen, dass den betroffenen Personen (also den Personen, deren Personendaten mittels KI bearbeitet werden) ein möglichst hohes Mass an digitaler Selbstbestimmung zukommt. Ohne dies explizit zu erwähnen, stützt er sich dabei auf das Prinzip «Privacy by Design» beziehungsweise «Datenschutz durch Technik», welches in Art. 7 Abs. 1 und 2 DSGVO verankert ist. Zentral für die Einhaltung sind die Datenvermeidung und die Datensparsamkeit – die technische Umsetzung des Grundsatzes «so viel wie nötig, so wenig wie möglich».

Weiter betont der EDÖB die Wichtigkeit der Transparenz. Er schreibt, dass die Hersteller, Anbieter und Verwender von KI-Systemen den Zweck, die Funktionsweise und die Datenquellen der KI-gestützten Datenbearbeitungen transparent machen müssen. Damit soll es ermöglicht werden, dass die betroffenen Personen einer automatischen Datenbearbeitung widersprechen oder eine automatisierte Einzelentscheidung von einem Menschen überprüfen lassen können. Der EDÖB führt weiter aus, dass die Benutzerinnen und Benutzer von intelligenten Sprachmodellen (wie bspw. Chatbots) ein Recht darauf haben, zu wissen, ob sie mit einem Menschen oder einer Maschine sprechen und ob die eingegebenen Daten verwendet werden, bspw. zur Verbesserung der selbstlernenden Programme. Schliesslich weist der EDÖB im Rahmen seiner Ausführungen noch darauf hin, dass die Verwendung von Deep Fakes deutlich erkennbar sein muss.

Zum Schluss seiner Kurzmeldung geht der EDÖB noch auf KI-gestützte Datenbearbeitungen mit hohem Risiko ein. Diese seien zulässig, wenn angemessene Massnahmen ergriffen und eine Datenschutz-Folgenabschätzung durchgeführt werde. Höchstwahrscheinlich hat sich der EDÖB vom AI Act inspirieren lassen, denn er erwähnt an dieser Stelle einige Anwendungen, welche datenschutzrechtlich verboten seien. Dazu gehören die flächendeckende Gesichtserkennung in Echtzeit und das Social Scoring, welche auch unter dem AI Act verboten sind. Der EDÖB begründet das Verbot da-

mit, dass mit diesen Anwendungen die Privatsphäre und die informationelle Selbstbestimmung gefährdet werden.

Worauf der EDÖB kaum eingeht, ist die Frage, inwiefern Personendaten für das Maschinenlernen verwendet werden dürfen. KI benötigt eine riesige Menge an Daten, um zu funktionieren, beziehungsweise um zu «lernen». Das Training der KI kann grundsätzlich entweder durch den Verantwortlichen selbst oder alternativ durch einen Auftragsbearbeiter erfolgen. Dabei ist Folgendes zu beachten^{VIII}:

- Verwendet der oder die Verantwortliche selbst Personendaten (bspw. Kundendaten) für das Training einer KI, liegt eine sogenannte Sekundärnutzung von Personendaten vor. Diese ist – in aller Kürze und vereinfacht gesagt – zulässig, wenn die betroffenen Personen in der Datenschutzerklärung darüber informiert wurden und die Daten nicht in personenbezogener Form weitergegeben werden. Widerspricht die betroffene Person einer solchen Verwendung ihrer Personendaten, muss die Verwendung durch ein überwiegendes Interesse des Verantwortlichen gerechtfertigt werden. Ein solches überwiegendes Interesse im Sinne von Art. 31 Abs. 2 lit. e DSGVO dürfte regelmässig vorliegen, weil die Daten nicht in personenbezogener Form weitergegeben werden. Es ist jedoch zu beachten, dass auch eigentlich anonyme KI-Anwendungen Datenschutzrisiken bergen und dass das überwiegende Interesse nur dann bejaht werden kann, wenn diese Risiken minimiert werden.
- Trainiert ein Auftragsbearbeiter seine eigene KI mit den Personendaten seiner Kundinnen und Kunden (bzw. der Verantwortlichen), stellt dies eine Zweckentfremdung der Personendaten dar und der Auftragsbearbeiter wird selbst Verantwortlicher für diese Datenbearbeitung. Damit das Training mit «fremden» Personendaten zulässig ist, muss die Verwendung der Personendaten für das KI-Training in den Datenschutzerklärungen sowohl des Auftragsbearbeiters als auch dessen Kunden angegeben sein. Sollen auch ältere, bereits bestehende Daten verwendet werden, müssen die entsprechenden betroffenen Personen über diese Datenbearbeitung informiert werden. Eine Rechtsgrundlage für die Weiterverwendung der Daten ist unter dem DSGVO (im Gegensatz zur DSGVO) nicht notwendig. Die Bearbeitung muss aber mit dem Zweck vereinbar sein, zu welchem die Personendaten beschafft wurden. Ist die Vereinbarkeit nicht gegeben – was für KI-Training in der Regel der Fall sein dürfte –, kann diese Verletzung des Grundsatzes der Zweckgebundenheit mit einem überwiegenden privaten oder öffentlichen Interesse oder einer Einwilligung der betroffenen Person gerechtfertigt werden. Vorliegend am ehesten einschlägig ist das überwiegende

private Interesse, was dann gegeben sein dürfte, wenn das KI-Training keine negative Auswirkung auf die betroffenen Personen hat, keine besonders schützenswerten Personendaten betrifft und die Daten nicht personenbezogen bearbeitet werden.

Bei der Zusammenarbeit mit Auftragsbearbeitern, welche KI-Tools anbieten oder entwickeln, sollten die Vertragsbestimmungen des Auftragsbearbeiters genau geprüft werden. Wenn darin vorgesehen ist, dass die Daten weiterverwendet werden, sind die notwendigen Massnahmen (insb. die Anpassung der Datenschutzerklärung) zu ergreifen, sofern der Vertrag trotzdem abgeschlossen wird.

In datenschutzrechtlicher Hinsicht steht KI vor verschiedenen praktischen Herausforderungen: Wie kann geprüft und dargelegt werden, dass automatisch getroffene Entscheidungen einer KI korrekt sind? Wie wird verhindert, dass die KI Verzerrungen oder blinden Flecken ausgesetzt ist und deshalb Personen vor- oder nachteilig behandelt? Und wie kann sichergestellt werden, dass sich die für das Training von KI-Modellen verwendeten Personendaten nicht aus diesen Modellen extrahieren lassen? Werden diese Fragen sorgfältig geprüft und entsprechende Massnahmen umgesetzt, ist die Verwendung von KI datenschutzrechtlich nicht problematischer oder an höhere Voraussetzungen gebunden als eine «normale» Bearbeitung von Personendaten.

Urheberrecht

Die zentralen urheberrechtlichen Fragen im Zusammenhang mit KI sind: Darf ich urheberrechtlich geschützte Werke für das Training der KI benutzen?^{1x} Und wem gehört, was die KI erstellt hat?^x

Das Urheberrechtsgesetz gewährt dem Urheber ein absolutes Ausschliesslichkeitsrecht an seinem Werk. Der Urheber hat unter anderem das ausschliessliche Recht, zu bestimmen, ob wann und wie das Werk verwendet wird (vgl. Art. 10 Abs. 1 URG). Es stellt sich deshalb die Frage, ob urheberrechtlich geschützte Werke für das KI-Training verwendet werden dürfen. Eine gerichtliche Entscheidung dazu, ob das KI-Training einen Eingriff in das Ausschliesslichkeitsrecht des Urhebers darstellt, gibt es bisher soweit ersichtlich nicht. Die juristische Lehre hat sich jedoch bereits mit dieser Frage befasst und tendiert dazu, einen Eingriff in das Vervielfältigungsrecht anzunehmen. Dies wird insbesondere damit begründet, dass beim KI-Training, insbesondere bei der Erstellung des Datensets, unkörperliche Kopien entstehen, welche den Konsum des Werkes ermöglichen. Für die Verwendung von urheberrechtlich geschützten Werken ist somit

entweder die Zustimmung des Urhebers bzw. der Urheberin oder die Anwendbarkeit einer Ausnahme notwendig.

Zu denken wäre insbesondere an folgende Ausnahmen:

- Verwendung zum Eigengebrauch im betriebsinternen Bereich für die interne Information oder Dokumentation (Art. 19 Abs. 1 lit. c URG): Die Ausnahme ist eng auszulegen und bezieht sich nur auf die Verwendung einzelner Auszüge ohne kommerziellen Zweck. Da für das KI-Training in der Regel nicht bloss Auszüge, sondern gesamte veröffentlichte Werke verwendet und die Resultate der KI üblicherweise kommerziell verwertet werden, dürfte diese Schranke des Urheberrechts für das KI-Training nur äusserst selten einschlägig sein.
- Vorübergehende Vervielfältigungen (Art. 24a URG): Die vorübergehende Vervielfältigung eines Werkes ist nur dann zulässig, wenn sie flüchtig oder begleitend ist, einen integralen und wesentlichen Teil eines technischen Verfahrens darstellt, ausschliesslich der Übertragung in einem Netz zwischen Dritten durch einen Vermittler oder einer rechtmässigen Nutzung dient und keine eigenständige wirtschaftliche Bedeutung hat. Diese Ausnahme ist für das KI-Training aus mehreren Gründen nicht einschlägig. So dürfte die Datenbasis für das KI-Training in der Regel weder flüchtig noch begleitend sein, und auch damit, dass das KI-Training keine eigenständige wirtschaftliche Bedeutung hat, lässt sich angesichts der hohen Investitionssummen im KI-Bereich kaum argumentieren.
- Verwendung von Werken zum Zweck der wissenschaftlichen Forschung (Art. 24d URG): Zentral ist bei dieser Schranke des Urheberrechts die Frage, ob ein konkreter und vorherrschender Forschungszweck besteht. Dies sollte im Einzelfall differenziert betrachtet werden – genauso wie die Frage, ob ein rechtmässiger Zugang zu den für das KI-Training verwendeten Werken besteht. Diese Ausnahme vom Urheberrecht dürfte in einigen Fällen (namentlich bei einer Ausrichtung des KI-Systems auf die wissenschaftliche Forschung) erfüllt sein, es empfiehlt sich jedoch eine einzelfallweise Prüfung.

Neben der Zustimmung der Urheberinnen und Urheber gibt es somit keine Ausnahme des Urheberrechts, welche für alle oder zumindest einen Grossteil der KI-Systeme greifen würde. Das Training von KI mit urheberrechtlich geschützten Werken ist somit schnell problematisch. Es gibt deshalb Bemühungen, Lizenzverträge abzuschliessen, um das KI-Training urheberrechtlich abzusichern und die Urheberinnen und Urheber zu vergüten, beispielsweise von der SUIISA (Genossenschaft der Urheber und Verleger von Musik) und der SSA (société suisse des auteurs).

Betreffend urheberrechtlichen Schutz von KI-generiertem Inhalt ist vorab Folgendes festzuhalten: Wenn urheberrechtlich geschützte Werke für das KI-Training oder den Prompt verwendet werden und die Werke im Output (also im KI-generierten Inhalt) erkennbar sind, gilt der urheberrechtliche Schutz auch für den Output.

In der Schweiz ist ein Werk urheberrechtlich nur dann geschützt, wenn es eine «geistige Schöpfung» ist (vgl. Art. 2 Abs. 1 URG). Das Urheberrecht geht grundsätzlich und ausschliesslich von einer menschlichen Schöpfung aus – für nicht menschengeschaffene Inhalte besteht somit kein Urheberrechtsschutz. Ein urheberrechtlicher Schutz kommt vor diesem Hintergrund nur dann in Frage, wenn die KI als «Werkzeug» verwendet wird, ähnlich einer Kamera oder einem Stift. Dies kann insbesondere dann der Fall sein, wenn der Mensch erheblich zum Output beigetragen hat (bspw. sich die Kreativität eines Prompts im Output niederschlägt) oder er den Output abgeändert hat (Output nur als Grundlage für das danach entstehende Werk).

Ein Urheberrechtsschutz für Output von KI besteht somit nicht per se, sondern kann nur bei einer Bearbeitung oder einer erheblichen Beeinflussung des Outputs durch den Menschen im Einzelfall bejaht werden.

Haftung

Da KI sehr unterschiedlich ausgestaltet und angewendet werden kann, gibt es auch keine «one size fits all»-Antwort auf die Frage, wer für die KI und deren Fehler haftet. Im Folgenden werden deshalb sowohl die Haftung des Herstellers einer KI als auch die Haftung des Anwenders bzw. der Anwenderin einer KI analysiert.

Wie in der Einleitung dieses Artikels festgehalten, werden die Schweizer Gesetze technologieneutral ausgestaltet. Trotzdem werden teilweise neue Entwicklungen erst verzögert abgebildet. So ist das Schweizer Produkthaftungsgesetz bis heute gemäss Wortlaut nur auf bewegliche Sachen und Elektrizität anwendbar (Art. 3 Abs. 1 PrHG). Ob und inwiefern eine Software als Produkt gilt, ist umstritten und wird in der Lehre stark diskutiert.^{XI}

Da zwischen Hersteller und Anwenderin bzw. Anwender in der Regel ein Vertragsverhältnis besteht, ist neben der produkthaftungsgesetzlichen auch an die vertragsrechtliche Haftung zu denken. So muss der Hersteller sicherstellen, dass bei der Entwicklung des KI-Systems mit der angemessenen Sorgfalt gearbeitet wurde. Wie hoch diese Hürde ist, kann nicht allgemein gesagt werden und hängt insbesondere

vom Risiko des jeweiligen KI-Systems ab. So ist diese Hürde bei einer KI, welche zur Diagnostik von Krankheiten eingesetzt wird, zweifellos höher als bei einer KI, welche süsse Katzenbilder generiert. Ist eine KI tatsächlich fehlerhaft beziehungsweise kann einem Hersteller eine Verletzung der Sorgfaltspflicht bewiesen werden, ist die Haftung des Herstellers eines KI-Systems denkbar.

In der Praxis wohl häufiger relevant wird mit der fortschreitenden Digitalisierung die Frage sein, ob der Anwender bzw. die Anwenderin einer KI für die Resultate der KI haftet. In diesem Zusammenhang hat das Bundesgericht ein algorithmisches System einer Bank zum «Market Making» als Hilfsmittel eingeordnet, da dem Algorithmus keine Rechtspersönlichkeit zukommt. Stand heute gilt die sogenannte Werkzeugtheorie auch für künstliche Intelligenz: Die Entscheidungen und Handlungen der Maschine gelten als diejenigen des Anwenders. Sofern also ein vertragliches Verhältnis zwischen dem Anwender und der geschädigten Person vorliegt, muss geprüft werden, ob der Anwender bei der Erfüllung seines Vertrages die angemessene Sorgfalt gewahrt hat. Gerade bei generativer KI im Dienstleistungsbereich (Rechtsberater, Architektinnen, Ingenieure) dürfte mit einer direkten Weiterverwendung des Outputs einer KI zum heutigen Stand die Sorgfaltspflicht in jedem Fall verletzt werden. Der Output ist stets kritisch zu prüfen und wo notwendig anzupassen, bevor er der Kundschaft vorgelegt wird.^{XII}

Rechtlich komplizierter wird es bei der ausservertraglichen Haftung, weil KI äusserst flexibel angewendet werden kann. Eine gewisse Fehleranfälligkeit liegt deshalb in der Natur einer KI. Mit diesen Überlegungen lässt sich argumentieren, dass die Verwendung einer KI naturgemäss ein Gefährdungspotential mit sich bringt und den Anwendenden deshalb eine Schadenminderungspflicht trifft. Es sind deshalb beim Einsatz von KI die für das betroffene System angemessenen Sicherheitsmassnahmen zu treffen.^{XIII}

Weisungswesen

KI wird heute bereits von vielen Arbeitnehmenden im Arbeitsalltag verwendet, ob mit oder ohne Erlaubnis der Arbeitgeberin. Die Verwendung von KI ist eine Tatsache – ein Verbot dürfte kaum durchsetzbar sein. Um einen Wildwuchs zu verhindern, ist es sinnvoller und zielführender, wenn Arbeitgeberinnen selbst bestimmte KI-Systeme zur Verfügung stellen.

Für die Verwendung dieser KI-Systeme sollten Arbeitgeberinnen zudem eine Weisung im Sinne von Art. 321d OR er-

lassen. Eine KI-Weisung muss nicht lang sein, sie sollte aber folgende zentrale Elemente enthalten:

- **Information über erlaubte Nutzung:** Welche KI-Tools stellt die Arbeitgeberin zur Verfügung und wie dürfen diese verwendet werden? Bei welchen KI-Tools dürfen welche Daten (bspw. Personendaten, Daten, die einem Berufsgeheimnis unterliegen, Geschäftsgeheimnisse) eingegeben werden? Ist etwas Zusätzliches zu beachten (bspw. das zwingende Login bei der Verwendung von DeepL Pro, um den Datenschutz zu gewährleisten)?
- **Wahrung von Betroffenenrechten:** Wann müssen die Kundinnen und Kunden über den Einsatz von KI informiert werden? Wie detailliert muss diese Information sein?
- **Haftung:** KI-Tools sind hilfreich, aber nicht fehlerfrei. Um Haftungsfälle zu vermeiden, sind sämtliche Resultate der KI auf ihre Richtigkeit hin zu überprüfen, bevor sie verwendet werden.
- **Ansprechperson:** An wen können sich die Mitarbeitenden melden, wenn sie eine Frage zur Verwendung von KI-Tools haben?

Mit diesen Anweisungen und Informationen erhalten die Mitarbeitenden die notwendigen Werkzeuge, um rechtskonform mit KI-Tools arbeiten zu können. Im Zusammenhang mit dem Erlass einer KI-Weisung empfiehlt es sich weiter, gewisse Standards und Werte zur Verwendung von KI innerhalb des Unternehmens festzusetzen. Dazu können beispielsweise Aspekte wie die Zuverlässigkeit, die Transparenz, die Betroffenenrechte und die Risikominimierung gehören.

Das wichtigste in Kürze

KI ist aus dem Alltag nicht mehr wegzudenken. Beim beruflichen Umgang damit gilt es jedoch, einige Kernfragen zu beachten:

- Ist der AI Act anwendbar und wenn ja, halten wir die entsprechenden Vorschriften ein?
- Wissen wir, welche Daten unsere KI-Anbieter für das Training der KI verwenden und ob sie das Datenschutzrecht einhalten?
- Verwendet der KI-Anbieter unsere Daten für das KI-Training? Wenn ja, wird dies in unserer Datenschutzerklärung abgebildet?
- Welche Daten und vor allem welche Personendaten dürfen wir für welche KI-Systeme verwenden?
- Haben wir den Output der KI auf dessen Richtigkeit hin geprüft und bei Bedarf angepasst?
- Entspricht der Output einem urheberrechtlichen Werk und verletzt er dadurch das Urheberrecht?
- Sind unsere Mitarbeitenden über die korrekte Anwendung von KI-Systemen informiert und wissen sie, an wen sie sich bei Fragen wenden können?

Schliesslich ist es zu empfehlen, sich regelmässig über die Entwicklung von KI und der entsprechenden Gesetzgebung zu informieren, um auf dem aktuellen Stand zu bleiben.

- I Medienmitteilung des UVEK vom 22.11.2023, «Bundesrat prüft Regulierungsansätze für Künstliche Intelligenz», abrufbar unter: www.uvek.admin.ch/uvek/de/home/uvek/medien/medienmitteilungen.msg-id-98791.html, zuletzt besucht am 10.02.2025.
- II UVEK, Bundesamt für Kommunikation BAKOM, Auslegeordnung zur Regulierung von künstlicher Intelligenz, Bericht an den Bundesrat vom 12.02.2025, abrufbar unter: www.bakom.admin.ch/bakom/de/home/digital-und-internet/strategie-digitale-schweiz/ki_leitlinien.html, zuletzt besucht am 13.02.2025.
- III Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law vom 05.09.2024, Council of Europe Treaty Series – No. 225, abrufbar unter: www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence, zuletzt besucht am 13.02.2025.
- IV Medienmitteilung des Bundesrates vom 12.02.2025, «KI-Regulierung: Bundesrat will Konvention des Europarats ratifizieren», abrufbar unter: www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-104110.html, zuletzt besucht am 13.02.2025.
- V Verordnung (EU) 2024/1689 des europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für Künstliche Intelligenz, abrufbar unter: eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=OJ:L_202401689, zuletzt besucht am 10.02.2025.
- VI www.economiesuisse.ch/de/ai-act-self-assessment-tool oder ai.kellerhals-carrard.ch.
- VII Kurzmeldung des Eidgenössischen Datenschutz- und Öffentlichkeitsbeauftragten (EDÖB) vom 09.11.2023, «Geltendes Datenschutzgesetz ist auf KI direkt anwendbar», abrufbar unter: www.edoeb.admin.ch/de/09112023-geltendes-dsg-ist-auf-ki-anwendbar, zuletzt besucht am 10.02.2025.
- VIII David Rosenthal, Datenschutz und KI: Worauf in der Praxis zu achten ist, in: Jusletter IT 22. April 2022
- IX Für Ausführungen zum Input im urheberrechtlichen Kontext: Marmy-Brändli Sandra/Oehri Isabelle, Das Training künstlicher Intelligenz, sic! 2023, S. 655–666.
- X Für Ausführungen zum Output im urheberrechtlichen Kontext: Hilty Reto M., Urheberrecht, 2. Aufl., Bern 2020 und Semmelmann Constanze, Künstliche Intelligenz und Urheberrecht: Stand zu Training und Output 2024, sic! 2024 S. 637–647.
- XI Für Ausführungen zur Produkthaftpflicht: Wildhaber Isabelle, Eine Einführung in die ausservertragliche Haftung für Künstliche Intelligenz (KI), Fellmann Walter (Hrsg.), Haftpflichtprozess 2021, S. 28–70, Zürich – Basel – Genf 2021.
- XII Für Ausführungen zur vertraglichen Haftung: BGer, Urteil 4A-305/2021 vom 02.11.2021, E. 7.3.1; Lohmann Melinda F./Pressler Theresa, Algorithmische Vertragserfüllung (Teil 1), SJZ 119/2023 S. 879–888.
- XIII Für Ausführungen zur ausservertraglichen Haftung: Quadroni Mauro, Künstliche Intelligenz – praktische Haftungsfragen, HAVE 2021 S. 345–354.