Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Justiz- und Polizeidepartement EJPD

Bundesamt für Justiz BJ
Direktionsbereich Öffentliches Recht
Fachbereich Rechtsetzungsprojekte I

31.08.2024

Rechtliche Basisanalyse im Rahmen der Auslegeordnung zu den Regulierungsansätzen im Bereich künstliche Intelligenz

BJ-D-F0D63401/159

Übersicht

Die vorliegende rechtliche Analyse wird vom EJPD (BJ) im Rahmen des Auftrags des Bundesrats vom 22. November 2023, eine Auslegeordnung zu den Regulierungsansätzen im Bereich künstliche Intelligenz (KI) zu erarbeiten, erstellt. Mit dieser Auslegeordnung sollen namentlich der Regulierungsbedarf sowie Regulierungsansätze aufgezeigt werden, die dem Rahmenübereinkommen des Europarats über künstliche Intelligenz und Menschenrechte, Demokratie und Rechtsstaatlichkeit (KI-Konvention) und der Verordnung der Europäischen Union zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (KI-Verordnung) Rechnung tragen. Die rechtliche Analyse zeigt, dass das Schweizer Recht im KI-Bereich vollumfänglich Anwendung findet, aber Anpassungen erforderlich wären, wenn die Schweiz die KI-Konvention ratifizieren sollte, und allgemein, um den Herausforderungen von KI zu begegnen. Eine Annäherung an die KI-Verordnung würde ein umfassendes Eingreifen des Gesetzgebers erfordern.

Ausgangslage

Die vorliegende rechtliche Analyse soll als Unterstützung für die Auslegeordnung zu den Regulierungsansätzen im KI-Bereich dienen, die vom Bundesrat am 22. November 2023 in Auftrag gegeben wurde. Der Auftrag des Bundesrats geht von der Prämisse aus, dass der gesetzgeberische Handlungsbedarf der Schweiz im KI-Bereich erneut geprüft werden sollte. Seit 2019 und der Herausgabe des Berichts der interdepartementalen Arbeitsgruppe «Künstliche Intelligenz», der zum Schluss kam, dass die geltende Gesetzgebung ausreichend sei, hat sich die Situation stark verändert. So werden KI-Systeme zunehmend in immer unterschiedlicheren Bereichen des privaten und des öffentlichen Sektors eingesetzt und auch auf internationaler Ebene wurden Texte erarbeitet, namentlich die KI-Konvention des Europarats und die KI-Verordnung der EU.

Die rechtliche Analyse hat zum Ziel, in Bezug auf die Herausforderungen, die sich aufgrund von KI ergeben, allfällige Lücken im Schweizer Rechtsrahmen aufzuzeigen und den gesetzgeberischen Handlungsbedarf vor dem Hintergrund des geltenden Rechts und der möglichen politischen Optionen des Bundesrats festzulegen. Sie wurde vom BJ erarbeitet und enthält Beiträge der DV (Ziff. 3 und 5.2.11), des BAKOM (Ziff. 5.2 und 5.3.2) und des IGE (Ziff. 6.2).

Inhalt

Die rechtliche Analyse stützt sich in erster Linie auf die KI-Konvention des Europarats, die für die Schweiz bei einer Ratifizierung verbindlich würde. Als erster völkerrechtlicher Vertrag über KI zeigt diese Konvention die wichtigsten rechtlichen Herausforderungen auf, die sich aufgrund von KI im Bereich des Schutzes der Menschenrechte, der Demokratie und der Rechtsstaatlichkeit ergeben. Eine Untersuchung dieser Konvention ermöglicht daher, allfällige Lücken im Schweizer Recht unter Berücksichtigung der auf internationaler Ebene erkannten Problembereiche zu ermitteln. Die Analyse betrifft das Völker- und das Landesrecht; auf das kantonale und kommunale Recht wird nicht eingegangen. In einem zweiten Teil wird die Situation dargestellt, die sich bei einer Annäherung des Schweizer Rechts an die KI-Verordnung der EU präsentieren würde. Weiter befasst sich die rechtliche Analyse mit weiteren ausgewählten Rechtsgebieten, die weder in der KI-Konvention noch in der KI-Verordnung umfassend behandelt werden, wie das Immaterialgüterrecht sowie der Bereich der zivilrechtlichen Haftung und strafrechtlichen Verantwortlichkeit.

Die Schlussfolgerungen der rechtlichen Analyse bestätigen, dass das Schweizer Recht bereits über auf KI anwendbare Rechtsnormen verfügt. Diese würden eine teilweise Übernahme der KI-Konvention ermöglichen. Es scheinen jedoch Ergänzungen nötig, insbesondere um die Transparenz der KI-Systeme zu verbessern, ihre Folgen auf die Grundrechte abzuschätzen und Kontrollmechanismen sicherzustellen. Würde eine Annäherung an die KI-Verordnung der EU in Betracht gezogen, so müssten ausführlichere Vorschriften verabschiedet werden. Bei der KI-Verordnung handelt es sich hauptsächlich um Produktsicherheitsvorschriften, wobei den verschiedenen Akteuren im KI-Bereich spezifische Pflichten auferlegt werden. Bei den weiteren ausgewählten Rechtsgebieten hat die Analyse gezeigt, dass sich gewisse Fragen stellen, die sich jedoch mit den geltenden Vorschriften grundsätzlich beantworten lassen. Ferner hätte die Verabschiedung allgemeiner Rechtsnormen für die Umsetzung der KI-Konvention, beispielsweise zur Stärkung der Transparenz der KI-Systeme, häufig zur Folge, dass der Schutz in diesen Bereichen, so etwa bei der zivilrechtlichen Haftung und strafrechtlichen Verantwortlichkeit, zusätzlich verbessert würde.

Die rechtliche Analyse konzentriert sich auf die grossen Herausforderungen und präsentiert meistens Schlussfolgerungen, die unter Berücksichtigung der getroffenen politischen Entscheidungen noch vertieft werden müssen, namentlich im Fall einer allfälligen Ratifizierung der KI-Konvention des Europarats oder einer eventuellen Annäherung an die Gesetzgebung der EU.

Die rechtliche Analyse wurde unter Berücksichtigung der Entwicklungen bis am 31. August 2024 durchgeführt.

Inhaltsverzeichnis

1 Einleitung

Mit Entscheid vom 22. November 2023 hat der Bundesrat das UVEK (BAKOM) und das EDA (Abteilung Europa) beauftragt, bis Ende 2024 eine Auslegeordnung zu den Regulierungsansätzen im Bereich künstliche Intelligenz (KI) in der Schweiz zu erarbeiten. Zur Steuerung der Arbeiten wurde ein zentraler Knotenpunkt geschaffen, der sich aus Vertreterinnen und Vertretern des UVEK (BAKOM), des EJPD (BJ) und des EDA (Abteilung Europa, DV) zusammensetzt. Das EJPD (BJ) wurde mit der Leitung der rechtlichen Basisanalyse betraut.

Zurzeit gibt es in der Schweiz keine Querschnittsgesetzgebung zu KI. Allerdings entwickelt sich die KI, wie der Bericht der interdepartementalen Arbeitsgruppe «Künstliche Intelligenz» von 2019 aufgezeigt hat, nicht im rechtsfreien Raum.[1] Der geltende Rechtsrahmen ist auf den Bereich der KI vollumfänglich anwendbar. Dabei handelt es sich namentlich um die Bundesverfassung, das DSG, das GIG, die Vorschriften über die zivilrechtliche Haftung und strafrechtliche Verantwortlichkeit usw.

Seit dem Bericht von 2019 haben im KI-Bereich bedeutende technologische Entwicklungen stattgefunden und auch der internationale Rechtsrahmen hat sich weiterentwickelt. Diese rechtliche Analyse hat zum Ziel, eine Einschätzung der gegenwärtigen Rechtslage vorzunehmen und zu untersuchen, wo im Schweizer Recht allenfalls ein Anpassungsbedarf besteht. Die Struktur und Methodik werden weiter unten vorgestellt (vgl. Ziff. 2).

Die rechtliche Analyse dient als Grundlage – neben der vom BAKOM erstellten Analyse der Regulierung von künstlicher Intelligenz in verschiedenen Ländern und Weltregionen und jener der Regulierungstätigkeiten im Zusammenhang mit KI in den verschiedenen Sektoren – für die Erarbeitung der vom Bundesrat verlangten Auslegeordnung. Diese Auslegeordnung soll dem Bundesrat als Entscheidungshilfe dienen, um 2025 einen konkreten Auftrag für eine allfällige KI-Regulierungsvorlage zu erteilen und die Zuständigkeiten in der Verwaltung zu regeln.

2 Ziel und Methodik der rechtlichen Basisanalyse

Im Rahmen dieser rechtlichen Basisanalyse soll untersucht werden, inwiefern KI-Systeme neue Herausforderungen mit sich bringen, und beurteilt werden, ob die geltende Schweizer Rechtsordnung Lücken aufweist, die geschlossen werden sollten, um diesen Herausforderungen zu begegnen. Angesichts der kurzen Frist für die Durchführung dieser Analyse konzentriert sich diese auf die wichtigsten rechtlichen Herausforderungen. Zudem konnte der gesetzgeberische Handlungsbedarf nicht immer eindeutig festgestellt werden und muss gegebenenfalls noch eingehender untersucht werden. Ferner betrifft die Analyse das Völker- und das Landesrecht; auf das kantonale und kommunale Recht wird nicht eingegangen (vgl. zu diesem Punkt Ziff. 4.1.3). Sie wurde vom BJ in Zusammenarbeit mit der DV (Ziff. 3 und 5.2.11), dem BAKOM (Ziff. 5.2 und 5.3.2) und dem IGE (Ziff. 6.2) erarbeitet.

---

[1] Herausforderungen der künstlichen Intelligenz, Bericht der interdepartementalen Arbeitsgruppe «Künstliche Intelligenz» an den Bundesrat, 2019, 91, abrufbar unter: www.sbfi.admin.ch > BFI-Politik > Bildungs-, Forschungs- und Innovationspolitik des Bundes 2025–2028 > Transversale Themen im BFI-Bereich > Digitalisierung im BFI-Bereich > Künstliche Intelligenz.

Als Ausgangspunkt für diese Arbeit wurde das Rahmenübereinkommen des Europarats über künstliche Intelligenz und Menschenrechte, Demokratie und Rechtsstaatlichkeit (im Folgenden: KI-Konvention)[2] ausgewählt. Der Auftrag des Bundesrats vom November 2023 sieht in der Tat vor, dass allfällige Regulierungsansätze dieses Übereinkommen berücksichtigen müssen. Die KI-Konvention wurde vom Ausschuss für künstliche Intelligenz («Committee on Artificial Intelligence», CAI) des Europarats ausgehandelt und am 17. Mai 2024 vom Ministerkomitee des Europarats offiziell verabschiedet (vgl. Ziff. 4). Die Schweiz ist Mitglied des Europarats und die Arbeiten des CAI fanden unter ihrem Vorsitz statt. Zudem beteiligte sich eine Schweizer Delegation ebenfalls aktiv an den Arbeiten.

Würde auf politischer Ebene entschieden, dieses Übereinkommen zu ratifizieren, so wäre dieser Text für die Schweiz verbindlich. Da er sehr allgemein formuliert ist, müsste er übernommen werden (vgl. Ziff. 4.1.2). Mit dieser Analyse soll ein allfälliger Anpassungsbedarf im Schweizer Recht im Fall einer Ratifizierung skizziert werden. Dazu wird zunächst der Inhalt der Verpflichtungen präsentiert, die sich aus dem Übereinkommen ergeben, und anschliessend wird die derzeitige Schweizer Rechtslage untersucht. Gegebenenfalls wird angegeben, ob das durch das Schweizer Recht gebotene Schutzniveau ausreichend ist. Aufgrund der verfügbaren Zeit für die Erarbeitung dieser Analyse handelt es sich nicht um eine abschliessende Untersuchung; in vielen Fällen wird eine Vertiefung nötig sein, um alle bestehenden rechtlichen Fragen zu klären.

Als erster verbindlicher völkerrechtlicher Vertrag über KI schafft die KI-Konvention einen allgemeinen Rechtsrahmen und benennt die wichtigsten rechtlichen Herausforderungen in diesem Bereich in Bezug auf den Schutz der Menschenrechte, die Demokratie und die Rechtsstaatlichkeit. Thematisiert werden insbesondere Fragen des Schutzes der Grundrechte, der Transparenz, des Datenschutzes, der Gleichstellung und Nichtdiskriminierung sowie Haftungs- und Verfahrensaspekte. Die Untersuchung dieses Textes ermöglicht somit auch, allfällige Lücken im Schweizer Recht unter Berücksichtigung dieser Problembereiche ganz allgemein zu erkennen.

In einem zweiten Teil befasst sich die Analyse mit der Verordnung der EU, die harmonisierte Vorschriften für künstliche Intelligenz (im Folgenden: KI-Verordnung)[3] festlegt (vgl. Ziff. 5). Dieser Text ist für die Schweiz nicht verbindlich. Es besteht somit keine rechtliche Pflicht, diesen in das innerstaatliche Recht zu überführen. Angesichts der geografischen Lage der

---

[2] Rahmenübereinkommen des Europarats über künstliche Intelligenz und Menschenrechte, Demokratie und Rechtsstaatlichkeit, SEV 225, verfügbar in französischer Sprache unter www.coe.int > Droits humains > Intelligence artificielle et droits humains > Convention-cadre, und in englischer Sprache unter www.coe.int > Human Rights > Artificial Intelligence and Human Rights > Framework Convention (abgerufen am 26. August 2024).

[3] Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828; ABl. L 2024/1689, 12. Juli 2024, abrufbar unter: https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=OJ%3AL_202401689 (abgerufen am 26. August 2024).

Schweiz und der wirtschaftlichen Beziehungen mit der EU ist eine Untersuchung dieser Verordnung jedoch unumgänglich. Zudem gilt diese für Schweizer Anbieter, die in der EU KI-Systeme in Verkehr bringen oder in Betrieb nehmen oder KI-Modelle mit allgemeinem Verwendungszweck in Verkehr bringen. Sie gilt auch für Schweizer Anbieter und Betreiber von KI-Systemen, wenn die erzeugten Ergebnisse in der EU verwendet werden. Ferner ergibt sich die Prüfung dieser Verordnung aus dem Auftrag des Bundesrats vom 22. November 2023.

Anschliessend wird eine Würdigung der KI-Verordnung vorgenommen, in der auf Fragen eingegangen wird, die aus Sicht des Schweizer Rechts relevant sind. Die Analyse beleuchtet auch das Verhältnis und die Synergien zwischen der KI-Verordnung und der KI-Konvention. Insbesondere wird untersucht, welchen Ansatz die KI-Verordnung verfolgt und inwiefern diese über die KI-Konvention hinausgeht oder sich von dieser unterscheidet.

Schliesslich geht die Analyse auf die Schweizer Rechtslage in Bezug auf Fragen der KI in Rechtsgebieten ein, die in der KI-Konvention oder der KI-Verordnung nicht umfassend behandelt werden, aber für welche die KI eine Herausforderung darstellt (vgl. Ziff. 6). Dabei handelt es sich insbesondere um relevante Aspekte des Immaterialgüterrechts, der zivilrechtlichen Haftung und der strafrechtlichen Verantwortlichkeit.

Sektorspezifische Fragen, wie die Regulierungsvorlagen in den Bereichen Energie und Gesundheit, werden in einer separaten Analyse unter der Leitung des BAKOM behandelt, wobei bestimmte transversale Aspekte, namentlich der Schutz vor Diskriminierungen, der Schutz der Privatsphäre oder die Haftungsfragen, vorbehalten sind.

Die rechtliche Analyse wurde unter Berücksichtigung der Entwicklungen bis am 31. August 2024 durchgeführt.

3 Instrumente auf internationaler Ebene[4]

Auf internationaler Ebene sind neben der KI-Konvention und der KI-Verordnung, die bereits erwähnt wurden und die gesondert untersucht werden, mehrere andere verbindliche (vgl. Ziff. 3.1) oder nicht verbindliche (vgl. Ziff. 3.2) Instrumente auf den KI-Bereich anwendbar.

3.1 Völkerrechtliche Verträge

Die KI entwickelt sich auf internationaler Ebene nicht im rechtsfreien Raum. Vielmehr ist der bestehende Rechtsrahmen auch auf den KI-Bereich anwendbar. So sind zahlreiche Verträge, Bräuche und allgemeine Grundsätze des Völkerrechts in diesem Bereich für die Schweiz bindend.

Als Beispiel seien hier die EMRK und die Konvention 108+[5] des Europarats erwähnt. Wie viele andere völkerrechtliche Verträge sind diese Konventionen auf die KI anwendbar, auch wenn sie sich nicht ausdrücklich darauf beziehen. Dank der Offenheit und der Flexibilität dieser Rechtsnormen können sich diese an die technologischen Entwicklungen anpassen.

In seinem Urteil Glukhin gegen Russland vom 4. Juli 2023[6] gelangte der EGMR beispielsweise zum Schluss, dass der Einsatz von Gesichtserkennungstechnologie eine Verletzung von Art. 8 (Recht auf Achtung des Privatlebens) und 10 EMRK (Freiheit der Meinungsäusserung) darstellen kann. Dieses Beispiel zeigt, dass gewisse Situationen, in denen ein KI-System rechtswidrig eingesetzt wird, mit dem bestehenden Recht bereits erfasst werden können (für weitere Erwägungen zu diesem Urteil vgl. Ziff. 4.3.1.2).

Der EGMR anerkennt zudem in bestimmten Fällen, dass die Staaten positive Verpflichtungen haben, Voraussetzungen zu schaffen, die den Schutz und die Ausübung der durch die EMRK garantierten Rechte ermöglichen. Somit könnte die Auslegung des bestehenden Völkerrechts unter bestimmten Voraussetzungen bereits eine positive Verpflichtung ergeben, im KI-Bereich Schutznormen zu erlassen.[7] Die internationalen Menschenrechtsnormen enthalten nicht nur die Verpflichtung, die Rechte des Einzelnen nicht einzuschränken; sie verlangen vom Staat auch, den Schutz des Einzelnen gegen Handlungen Dritter zu garantieren und positive Massnahmen zum Schutz der Menschenrechte zu ergreifen. Vor diesem Hintergrund könnte aus dem geltenden Völkerrecht eine positive Verpflichtung abgeleitet werden, im KI-Bereich Rechtsvorschriften zu erlassen.

---

[4] Dieses Kapitel wurde auf der Grundlage von Beiträgen der DV verfasst.

[5] Die Konvention 108 wurde mit dem Änderungsprotokoll zum Übereinkommen zum Schutz des Menschen bei der automatischen Verarbeitung personenbezogener Daten, das vom Ministerkomitee an seiner 128. Sitzung in Helsingör am 18. Mai 2018 verabschiedet wurde, modernisiert. Die Schweiz hat das Protokoll am 7. September 2023 ratifiziert, aber es ist noch nicht Kraft getreten.

[6] EGMR, Glukhin gegen Russland, 11519/20 (4. Juli 2023).

[7] WOLFGANG HOFFMANN-RIEM, Artificial Intelligence as a Challenge for Law and Regulation, in: Thomas Wischmeyer/Timo Rademacher (Hrsg.), Regulating Artificial Intelligence, Cham 2020, 1 ff., 8.

3.2 Nicht verbindliche Instrumente im KI-Bereich

Neben dem oben erwähnten Rechtsrahmen bestehen mehrere nicht verbindliche internationale Instrumente (*soft law*), die für die Schweiz im KI-Bereich relevant sind.

Dabei handelt es sich namentlich um folgende Texte: Empfehlung zu künstlicher Intelligenz der OECD vom 22. Mai 2019, überarbeitet am 8. November 2023, mit einer aktualisierten Definition des Begriffs «KI-System»; Erklärung des Ministerkomitees des Europarats zu den manipulativen Fähigkeiten algorithmischer Prozesse, Decl(13/02/2019)1; Empfehlung des Ministerkomitees des Europarats über die Auswirkungen algorithmischer Systeme auf die Menschenrechte, CM/Rec(2020)1; UNESCO-Empfehlung zur Ethik der Künstlichen Intelligenz, verabschiedet am 23. November 2021; Erklärung der Staats- und Regierungschefinnen und -chefs, verabschiedet am 4. Gipfeltreffen des Europarats in Reykjavik vom 16. und 17. Mai 2023; Erklärung von Bletchley der Länder, die am Gipfeltreffen zur Sicherheit künstlicher Intelligenz am 1. und 2. November 2023 teilnahmen; Resolution der UN-Generalversammlung vom 21. März 2024 zu künstlicher Intelligenz.

Zudem werden zurzeit weitere nicht verbindliche Instrumente erarbeitet. Insbesondere verfasst ein Ausschuss des Europarats bis Ende 2025 eine Empfehlung zu den Auswirkungen von KI-Systemen, ihrem Potenzial zur Förderung der Gleichstellung, einschliesslich der Geschlechtergleichstellung, und den Risiken, die sie in Bezug auf Diskriminierung mit sich bringen können. Ebenfalls bis 2025 werden von einem anderen Ausschuss des Europarats Leitlinien zu den Auswirkungen generativer KI auf die Freiheit der Meinungsäusserung fertiggestellt. Die Schweiz beteiligt sich aktiv in diesen beiden Ausschüssen.

Auch wenn nicht rechtlich verbindlich, sind diese Texte dennoch von Bedeutung, da sie einen gewissen Konsens auf internationaler Ebene im KI-Bereich widerspiegeln. Die Schweiz engagiert sich seit mehreren Jahren aktiv in diesen verschiedenen internationalen Prozessen. Zudem dienen diese Instrumente der Schweiz als Referenz. So werden sie im Bericht «Herausforderungen der künstlichen Intelligenz» der interdepartementalen Arbeitsgruppe von 2019 mehrmals erwähnt.[8]

Zum Schluss seien noch die internationalen Leitprinzipien und der internationale Verhaltenskodex des Prozesses von Hiroshima für Organisationen, die fortgeschrittene KI-Systeme entwickeln, vom 30. Oktober 2023 sowie die Erklärung der Staats- und Regierungschefinnen und -chefs der G7 zum KI-Prozess von Hiroshima vom 30. Oktober 2023 erwähnt. Die Schweiz hat sich nicht an diesen Initiativen beteiligt.

[8] Bericht «Herausforderungen der künstlichen Intelligenz» (Fn. 1).

4 Rahmenübereinkommen des Europarats über künstliche Intelligenz und Menschenrechte, Demokratie und Rechtsstaatlichkeit

4.1 Einleitende Bemerkungen

4.1.1 Ausgangslage

Der Ausschuss für künstliche Intelligenz («Committee on Artificial Intelligence», CAI) wurde im Juni 2022 vom Ministerkomitee des Europarats beauftragt, ein verbindliches, übergreifendes Rechtsinstrument zu KI zu entwickeln, das auf den Standards des Europarats für Menschenrechte, Demokratie und Rechtsstaatlichkeit aufbaut.

Der CAI setzt sich aus Vertretern aller Mitgliedstaaten des Europarats und der Europäischen Union zusammen. Darüber hinaus sind mehrere Staaten mit Beobachterstatus in den Prozess eingebunden: Argentinien, Australien, Kanada, Costa Rica, der Heilige Stuhl, Israel, Japan, Mexiko, Peru, die Vereinigten Staaten von Amerika und Uruguay. Daneben sind auch Vertreter der Zivilgesellschaft beteiligt.

Die Verhandlungen des CAI fanden unter der Leitung eines Schweizers als Vorsitzenden statt. Darüber hinaus war die Schweiz im CAI durch eine Delegation aus Repräsentanten des BAKOM, der DV und des BJ vertreten.

Die Verhandlungen zur KI-Konvention wurden im März 2024 abgeschlossen. Das Ministerkomitee hat den Text im Mai 2024 verabschiedet. Im September 2024 wurde die Konvention zur Unterzeichnung aufgelegt. Die Hauptaufgabe des CAI ist abgeschlossen, der Ausschuss arbeitet jedoch weiterhin an verwandten Themen, insbesondere an der Entwicklung einer fakultativen Methodik, die als Instrument für das Risiko- und Folgenmanagement von KI dienen soll.

4.1.2 Anwendbarkeit und Justiziabilität

Zunächst stellt sich die Frage, ob die KI-Konvention unmittelbar anwendbar (oder justiziabel bzw. *self-executing*) ist, d. h. unmittelbar Rechte und Pflichten für natürliche und juristische Personen begründet, oder ob sie sich nur an Staaten richtet, die sie noch umsetzen müssen, um ihr Wirkung zu verleihen.

Unmittelbar anwendbar sind Normen, die so konkret und bestimmt sind, dass natürliche oder juristische Personen aus ihnen Rechte und Pflichten ableiten können, und sich in Gerichts- oder Verwaltungsverfahren darauf berufen können. Entsprechend können rechtsanwendende Behörden und Gerichte solche internationalen Normen unmittelbar anwenden.[9]

[9] Das Verhältnis von Völkerrecht und Landesrecht, Bericht des Bundesrates vom 5. März 2010 in Erfüllung des Postulats 07.3764 der Kommission für Rechtsfragen des Ständerates vom 16. Oktober 2007 und des Postulats 08.3765 der Staatspolitischen Kommission des Nationalrates vom 20. November 2008, BBl **2010** 2263 2286.

Das BGer ist der Ansicht, dass die unmittelbare Anwendbarkeit einer völkerrechtlichen Bestimmung voraussetzt, dass diese nach übereinstimmender Auffassung im Gesamtzusammenhang sowie im Lichte von Ziel und Zweck des Vertrags betrachtet unbedingt und eindeutig genug formuliert sein muss, um eine direkte Wirkung erzeugen und in einem konkreten Fall angewendet werden bzw. die Grundlage für eine Entscheidung darstellen zu können.[10] Die unmittelbare Anwendbarkeit einer Norm ist letztlich eine Frage der Auslegung der betreffenden Bestimmung durch die staatlichen Behörden.

Zur Bestimmung der unmittelbaren Anwendbarkeit hat das BGer eine umfangreiche Rechtsprechung entwickelt. Für die direkte Anwendbarkeit spricht ein hohes Schutzbedürfnis des Einzelnen. Anders verhält es sich hingegen bei erheblichen Auswirkungen auf den Staat als Ganzes, bei einer komplexen Materie, wenn sich die Beurteilung des Einzelfalls bei der gerichtlichen Überprüfung als schwierig erweist, oder bei erheblichen finanziellen Auswirkungen. Zurückhaltung übt das BGer auch, wenn die unmittelbare Anwendung eine politische Beurteilung oder die Prüfung von Grundsatzfragen erfordert.[11]

Als nicht direkt anwendbar (d. h. nicht justiziabel oder *non-self-executing*) gelten Normen programmatischer Natur, bzw. Bestimmungen, die sich an den Staat als Ganzes richten. Sie bedürfen der Konkretisierung durch die zuständigen staatlichen Stellen, bevor sie für Private Rechte und Pflichten begründen.[12]

Die Prüfung der Normen der KI-Konvention nach den oben genannten Kriterien führt zu dem Ergebnis, dass sich die Konvention allgemein an Staaten richtet und nicht unmittelbar anwendbar ist. Dies ergibt sich auch aus dem Text der Konvention selbst, wo es in Artikel 1 Absatz 2 heisst, dass jede Vertragspartei geeignete Gesetzgebungs-, Verwaltungs- oder sonstigen Massnahmen trifft oder aufrechterhält, um diesem Übereinkommen Wirksamkeit zu verleihen. Es ist jedoch nicht auszuschliessen, dass ein Gericht im Einzelfall zu einem anderen Ergebnis kommt. Auch die Entwicklung der Rechtsprechung könnte eine Rolle spielen.

Ein Gericht könnte die Konvention im Rahmen einer Klage einer Einzelperson prüfen, die sich darüber beschwert, dass der Staat seinen Verpflichtungen nicht nachkommt. Ein Gericht könnte dann zum Schluss kommen, dass der Gesetzgeber aufgrund der Bestimmungen der Konvention tätig werden muss. Wenn die fragliche Norm jedoch nur eine Verpflichtung zum Tätigwerden auferlegt, aber nicht darlegt, wie dies zu geschehen hat, wird sich das Gericht darauf beschränken, den Gesetzgeber durch einen sogenannten Appellentscheid zum Handeln aufzufordern.[13]

---

[10] Das Verhältnis von Völkerrecht und Landesrecht (Fn. 9), 2303 f.

[11] Das Verhältnis von Völkerrecht und Landesrecht (Fn. 9), 2303 f.

[12] Das Verhältnis von Völkerrecht und Landesrecht (Fn. 9), 2286.

[13] Vgl. BGE **137** I 305, E. 6.6 f.

Eine Berufung auf die KI-Konvention vor dem EGMR wäre allenfalls im Zusammenhang mit der Verletzung eines durch die EMRK geschützten Rechts denkbar. Die Nichtumsetzung einer Bestimmung der KI-Konvention, die die Vertragsstaaten auffordert, in einem bestimmten Bereich gesetzgeberische oder andere Massnahmen zu ergreifen, könnte z. B. dazu führen, dass der EGMR eine Verletzung eines durch die EMRK garantierten Rechts feststellt, sofern die Untätigkeit als Verletzung positiver Verpflichtungen des Staates, z. B. aus Art. 8 EMRK, ausgelegt wird. Dies gilt selbstverständlich nur für diejenigen Vertragsstaaten, die Mitglied des Europarats sind. Dies könnte zu einer Zweiklassenstruktur führen, auf der einen Seite mit den Vertragsstaaten, die Mitglied des Europarats sind, und auf der anderen Seite den Vertragsstaaten, die nicht Mitglied des Europarats sind. Diese Situation ist jedoch allen Übereinkommen des Europarats gemein, die auch von Nichtmitgliedstaaten ratifiziert werden können, und ist daher nicht neu.

4.1.3 Umsetzung in innerstaatliches Recht und Föderalismus

Gemäss Art. 5 Abs. 4 BV haben Bund und Kantone das Völkerrecht zu beachten.

Wie bereits erwähnt (vgl. Ziff. 4.1.2) richtet sich die KI-Konvention an die Staaten und enthält allgemein formulierte Regeln und Grundsätze. Die darin enthaltenen Verpflichtungen müssen daher in innerstaatliches Recht umgesetzt werden.

In einem föderalistischen Staat wie der Schweiz kann diese Pflicht zur Umsetzung je nach Rechtsgebiet dem Bund, den Kantonen oder diesen beiden Staatsebenen gleichzeitig obliegen.[14] Bei einer allfälligen Ratifizierung der KI-Konvention wären für jede Verpflichtung die Regeln der Kompetenzverteilung zwischen Bund und Kantonen im betreffenden Bereich zu prüfen. Je nach Ergebnis läge die Pflicht zur Umsetzung beim Bund oder bei den Kantonen.

Im Bereich des Datenschutzes kann sich der Bund beispielsweise auf Artikel 95 Absatz 1, Artikel 97 Absatz 1 und Artikel 122 Absatz 1 BV sowie auf die ihm eigene Organisationskompetenz für seine Behörden stützen, um Vorschriften für den privaten Sektor und die Bundesorgane zu erlassen. Demgegenüber untersteht die Datenbearbeitung durch kantonale oder kommunale Organe unter Vorbehalt der Zuständigkeiten, die sich aus den materiellen Datenschutzbestimmungen in den Spezialgesetzen des Bundes ergeben,[15] dem kantonalen Recht.[16]

So liegt beispielsweise die Zuständigkeit für das Bildungswesen im Wesentlichen bei den Kantonen (vgl. Art. 62 BV).

---

[14] JUDITH WYTTENBACH, Umsetzung von Menschenrechtsübereinkommen in Bundesstaaten, Zürich/St. Gallen 2017, 299.

[15] So gelten beispielsweise die Artikel 89*a* ff. SVG, oder die Art. 49*a* ff. AHVG für alle rechtsanwendenden Organe oder Privatpersonen, einschliesslich kantonaler Organe.

[16] CR LPD-SYLVAIN MÉTILLE/LIVIO DI TRIA, Art. 2 N 14; PHILIPPE MEIER, Protection des données – Fondements, principes généraux et droit privé, Bern 2010, N 356 ff.

Die vorliegende Analyse konzentriert sich auf eine allfällige Umsetzung der KI-Konvention auf Bundesebene. Die Kantone wären jedoch, wie bereits erwähnt, für die Umsetzung in ihren jeweiligen Kompetenzbereichen zuständig.

4.2 Kapitel I: Allgemeine Bestimmungen

4.2.1 Artikel 1 – Ziel und Zweck

Artikel 1 regelt das Ziel und den Zweck der KI-Konvention.

Gemäss Absatz 1 soll die Konvention sicherstellen, dass die Aktivitäten im Lebenszyklus von KI-Systemen uneingeschränkt mit den Menschenrechten, der Demokratie und der Rechtsstaatlichkeit vereinbar sind. Dies entspricht der Aufgabe des Europarats. Der Begriff «KI-System» wird in Artikel 2 definiert (vgl. Ziff. 4.2.2).

Die Konvention ergänzt die in jedem Vertragsstaat bestehenden Mechanismen zum Schutz der Menschenrechte, einschliesslich der anwendbaren nationalen und völkerrechtlichen Verpflichtungen. Die Konvention schafft jedoch keine neuen Grundrechte. Sie regelt auch nicht die KI als solche, sondern nur Aspekte von KI, die einen Zusammenhang zum Schutz der Menschenrechte, der Demokratie und der Rechtsstaatlichkeit aufweisen.

Absatz 2 betrifft die Umsetzung der Konvention. Er sieht vor, dass jede Vertragspartei die erforderlichen gesetzgeberischen, verwaltungstechnischen oder sonstigen Massnahmen einführt oder beibehält, um den Bestimmungen der Konvention Wirksamkeit zu verleihen. In dieser Formulierung kommt der völkerrechtliche Grundsatz zum Ausdruck, dass die Staaten bei der Art und Weise der Umsetzung ihrer völkerrechtlichen Verpflichtungen frei sind. Bezüglich der Massnahmen zur Umsetzung der Konvention steht den Staaten ein grosser Spielraum zur Verfügung. Es handelt sich um eine Ergebnis- und nicht um eine Mittelverpflichtung. Theoretisch darf ein Staat somit frei entscheiden, ob eine Bestimmung in seiner internen Rechtsordnung durch die Legislative, die Exekutive oder die Judikative umgesetzt wird. Artikel 1 Absatz 2 sieht zudem vor, dass diese Massnahmen abgestuft und differenziert sein müssen, soweit erforderlich, je nach Schwere und Wahrscheinlichkeit des Auftretens negativer Auswirkungen auf die Menschenrechte, die Demokratie und der Rechtsstaatlichkeit während des gesamten Lebenszyklus von KI-Systemen. Dies bedeutet, dass die Massnahmen dem Risikograd eines KI-Systems anzupassen sind

Im schweizerischen Recht ist die jeweilige Normstufe für die Umsetzung der Konvention in innerstaatliches Recht anhand von Artikel 5 Absatz 1 und Artikel 164 BV zu prüfen. Letzterer Artikel bestimmt, dass alle wichtigen rechtsetzenden Bestimmungen in der Form eines Bundesgesetzes zu erlassen sind. Dazu gehören insbesondere die grundlegenden Bestimmungen über: a) die Ausübung der politischen Rechte; b) die Einschränkungen verfassungsmässiger Rechte; c) die Rechte und Pflichten von Personen; d) den Kreis der Abgabepflichtigen sowie den Gegenstand und die Bemessung von Abgaben; e) die Aufgaben und die Leistungen des Bundes; f) die Verpflichtungen der Kantone bei der Umsetzung und beim Vollzug des Bundesrechts; g) die Organisation und das Verfahren der Bundesbehörden.

Art. 164 BV schreibt vor, dass in bestimmten Bereichen gesetzgeberische Massnahmen zu treffen sind, wenn die Bedeutung der zu regelnden Materie keine andere Form der Regelung

(Verordnung, Beschluss usw.) zulässt. Es handelt sich dabei um den Grundsatz des allgemeinen Vorbehalts des Gesetzes im formellen Sinn[17], wonach wichtiges staatliches Handeln einer formell gesetzlichen Grundlage bedarf. Dieser Grundsatz stellt sicher, dass die grossen Werte- und Interessenabwägungen und die grossen politischen und rechtstechnischen Entscheidungen[18] von der Legislative getroffen werden, wodurch politische Repräsentativität und Partizipation gewährleistet werden.

Somit sind für die zur Umsetzung der Konvention im innerstaatlichen Recht zu erlassenden Rechtsakte die verfassungsrechtlichen Vorgaben zu berücksichtigen. Die wichtigsten Grundsätze und Grundregeln müssen in einem formellen Gesetz konkretisiert werden.

4.2.2 Begriffsbestimmungen

4.2.2.1 Artikel 2 – Begriffsbestimmung «System künstlicher Intelligenz»

Im Rahmen der Konvention wird ein KI-System als ein maschinengestütztes System definiert, das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können. Die verschiedenen Systeme künstlicher Intelligenz unterscheiden sich in ihrem Grad der Autonomie und der Anpassungsfähigkeit nach ihrer Betriebsaufnahme.

Diese Definition lehnt sich an die von der OECD am 8. November 2023 verabschiedete Definition[19] an und trägt der Notwendigkeit Rechnung, die internationale Zusammenarbeit im Bereich der KI zu stärken und die Bemühungen um eine weltweite Harmonisierung von KI-Governance, einschliesslich der Vereinheitlichung der Terminologie, zu unterstützen.

Diese Definition umfasst ein breites Verständnis von KI-Systemen, insbesondere in Abgrenzung zu anderen, einfacheren Arten von Softwaresystemen, die auf ausschliesslich von Menschen erstellten Regeln zur Automatisierung von Arbeitsschritten beruhen. Sie ist hinreichend abstrakt und flexibel, um zukünftige technologische Entwicklungen zu berücksichtigen. Die Definition ist auch insofern funktional, als sie speziell für die Zwecke der KI-Konvention entwickelt wurde und nicht darauf abzielt, eine universelle Bedeutung des Begriffs zu vermitteln.

---

[17] BSK BV-JUDITH WYTTENBACH/KARL-MARC WYSS, Art. 164 N 4; SGK BV-PIERRE TSCHANNEN, Art. 164 N 4; CR Cst.-JACQUES DUBEY, Art. 164 N 10.

[18] CR Cst.-JACQUES DUBEY, Art. 164 N 11.

[19] OECD, Empfehlung des Rates zu künstlicher Intelligenz, 2024, OECD/LEGAL/0449, Kapitel I, abrufbar unter: https://legalinstruments.oecd.org/fr/instruments/OECD-LEGAL-0449 (abgerufen am 26. August 2024).

Laut dem erläuternden Bericht zur KI-Konvention können die Vertragsparteien diese Definition im Interesse der Rechtssicherheit in ihrer nationalen Rechtsordnung präzisieren, ohne dabei deren Anwendungsbereich einzuschränken.[20]

Die Definition entspricht weitgehend derjenigen in der KI-Verordnung der EU (vgl. Ziff. 5.2.4.1).

Der erläuternde Bericht zur KI-Konvention enthält nur wenige Hinweise zur Auslegung der Definition. Er verweist auf die Begründung zur aktualisierten Definition von KI durch die OECD.[21] Für Details sei daher auf dieses Dokument verwiesen. Die wichtigsten Elemente sind folgende:

- Explizite oder implizite Ziele eines KI-Systems

  Die Ziele eines KI-Systems können explizit oder implizit sein. Explizite Ziele werden von einem Menschen direkt ins System eingebettet. Implizite Ziele ergeben sich aus einem von Menschen definierten Regelwerk (z. B. ist ein System für automatisiertes Fahren so programmiert, dass es das Strassenverkehrsgesetz einhält, aber implizit schützt es Leben); weiter gibt es Zwecke, die sich aus dem System, das selbst neue Zwecke lernt, abgeleitet werden (z. B. generative KI-Systeme wie ChatGPT).

- Ein KI-System leitet von erhaltenen Inputs ab, wie Ergebnisse zu erzeugen sind

  Das KI-System wird mit Inputs versorgt, d. h. Daten und Programmierregeln, die von Menschen oder Maschinen stammen können. Es verarbeitet diese Eingaben anhand eines algorithmischen Modells, das in der Lage ist, ein Ergebnis zu berechnen.

  Beispielsweise leitet ein System zur visuellen Objekterkennung ab, wie ein Ergebnis (d. h. die Identifizierung des Objekts im Bild) zu erzeugen ist, aus einer Eingabe, die aus den Pixeln des Bildes besteht.

- Ein KI-System produziert Ergebnisse, die physische oder virtuelle Umgebungen beeinflussen können

  Die verschiedenen von KI-Systemen ausgeführten Funktionen führen zu Ergebnissen. Diese lassen sich Kategorien wie Empfehlungen, Vorhersagen und Entscheidungen

---

[20] Erläuternder Bericht zum Rahmenübereinkommen des Europarats über künstliche Intelligenz und Menschenrechte, Demokratie und Rechtsstaatlichkeit, STCE 225, N 27 (verfügbar in französischer Sprache unter www.coe.int > Droits humains > Intelligence artificielle et droits humains > Convention-cadre, und in englischer Sprache unter www.coe.int > Human Rights > Artificial Intelligence and Human Rights > Framework Convention).

[21] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 24; OCDE, Explanatory memorandum on the update OECD definition of an AI system. OECD Artificial intelligence papers, Nr. 8, OECD Publishing, März 2024, abrufbar unter: https://doi.org/10.1787/623da898-en (abgerufen am 26. August 2024).

zuordnen. Wenn beispielsweise ein System für automatisiertes Fahren vorhersagt, dass es sich bei den über seine Kamera eingegebenen Pixeln um einen Fussgänger handelt, kann es entweder empfehlen zu bremsen, oder entscheiden zu bremsen.

Die Definition nennt explizit «Inhalte» als Ergebniskategorie. Der Begriff bezieht sich insbesondere auf sogenannte generative KI-Systeme wie ChatGPT, die Inhalte wie Texte, Bilder, Ton oder Videos produzieren.

Die von KI-Systemen beeinflussten Umgebungen können physisch oder virtuell sein. Die Ergebnisse werden somit entweder von Menschen oder von Maschinen- oder Gerätekomponenten wahrgenommen.

- Unterschiedliche Grade an Autonomie und Anpassungsfähigkeit nach Inbetriebnahme

  Die *Autonomie eines KI-Systems* bezieht sich auf die Fähigkeit dieses Systems, ohne menschliches Eingreifen zu lernen oder zu handeln. Bestimmte KI-Systeme sind in der Lage, Ergebnisse zu generieren, ohne dass dies explizit in den programmierten Zwecken des Systems vorgesehen ist und ohne dass ein Mensch spezifische Anweisungen gibt.

  Die Erwähnung der *Anpassungsfähigkeit* berücksichtigt die Tatsache, dass sich einige KI-Systeme auch nach Konzeption und Inbetriebnahme weiterentwickeln können. Beispiele hierfür sind Spracherkennungssysteme, die sich an die Stimme einer Person anpassen, oder personalisierte Musikempfehlungssysteme. KI-Systeme können einmalig, periodisch oder kontinuierlich trainiert werden und arbeiten, indem sie Muster und Verbindungen in den Daten ableiten. Aufgrund dieses Trainings können einige KI-Systeme die Fähigkeit erwerben, neue Arten von Schlussfolgerungen zu ziehen, die von den Programmierern ursprünglich nicht vorgesehen waren.

Aktuell findet sich im Schweizer Recht keine Definition von KI-Systemen. Aufgrund des technologieneutralen Ansatzes des schweizerischen Gesetzgebers deckt der geltende Rechtsrahmen jedoch auch ohne eine solche Definition Sachverhalte mit KI-Systemen ab.

Es sei darauf hingewiesen, dass das beim BFS angesiedelte CNAI eine harmonisierte Terminologie für den KI-Bereich bereitstellt, um eine gemeinsame Sprache und ein entsprechendes gemeinsames Begriffsverständnis auf Ebene der Bundesverwaltung zu erleichtern.[22] Diese Terminologie enthält auch eine Definition von KI-Systemen, die derjenigen in Artikel 2 der KI-Konvention und der OECD sehr nahe kommt.

[22] Vgl. www.cnai.swiss > Dienstleistungen > Terminologie (Stand 21. Dezember 2023), 7, abgerufen am 26. August 2024: «Ein KI-System ist ein maschinenbasiertes System, das für explizite oder implizite Ziele aus den empfangenen Inputs schlussfolgert, wie es Outputs wie Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erzeugen kann, welche die physische oder virtuelle Umgebung beeinflussen können. KI-Systeme können mit unterschiedlichem Ausmass an Autonomie ausgestattet werden.»

Einige der aus der KI-Konvention abgeleiteten Verpflichtungen beziehen sich spezifisch auf KI-Systeme, wie z. B. die Verpflichtung zur Schaffung eines Rahmens für das Risiko- und Folgenmanagement (vgl. Ziff. 4.3.4). Im Falle einer Ratifizierung der KI-Konvention wäre es daher wünschenswert, auch im schweizerischen Recht eine Definition von KI-Systemen einzuführen, um den Anwendungsbereich einzelner Bestimmungen der Konvention genauer abzugrenzen.

4.2.2.2 Lebenszyklus

Die KI-Konvention nimmt jeweils Bezug auf den Lebenszyklus der KI-Systeme. Damit bringt sie zum Ausdruck, dass die Bestimmungen der KI-Konvention in allen Tätigkeitsphasen von KI-Systemen umzusetzen sind, von der Konzeption bis zur Ausserbetriebnahme, und zwar unabhängig davon, wer an diesen Tätigkeiten beteiligt ist.[23]

Obwohl die Nutzungsphase von KI-Systemen häufig als die risikoreichste Phase angesehen wird, kann es auch in anderen Phasen zu Verletzungen der Menschenrechte, der Demokratie und der Rechtsstaatlichkeit kommen. So soll beispielsweise Artikel 10 der KI-Konvention sicherstellen, dass Gleichstellung und Nichtdiskriminierung in allen Phasen eines KI-Systems gewahrt werden, also nicht nur beim Einsatz eines KI-Systems, sondern auch in der Entwicklungsphase, z. B. bei den Anforderungen an den Datensatz für das Training des Systems.

Im Text der Konvention fehlt eine Definition des Begriffs «Lebenszyklus», der erläuternde Bericht enthält jedoch einige Klarstellungen. Ohne die Phasen im Lebenszyklus eines KI-Systems abschliessend aufzuzählen, wird unter Bezugnahme auf die Arbeiten der OECD ein Überblick über mögliche Phasen gegeben: 1) Planung und Entwurf, 2) Datensammlung und -bearbeitung, 3) Entwicklung von KI-Systemen, einschliesslich der Entwicklung von Modellen und/oder der Anpassung vorhandener Modelle an spezifische Aufgaben, 4) Testen, Verifizierung und Validierung, 5) Lieferung/Zugänglichmachung der Systeme, 6) Einsatz, 7) Betrieb und Überwachung, und 8) Ausserbetriebnahme. Diese Aktivitäten sind häufig iterativ, aber nicht notwendigerweise sequenziell.[24]

[23] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 15.

[24] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 15.

4.2.3 Artikel 3 – Anwendungsbereich

4.2.3.1 Öffentlicher und privater Sektor

4.2.3.1.1 Allgemeines

Die KI-Konvention hat einen weiten Anwendungsbereich, um Tätigkeiten im Lebenszyklus von KI-Systemen zu erfassen, die mit den Menschenrechten, der Demokratie und der Rechtsstaatlichkeit in Konflikt geraten können.[25]

Die Konvention gilt sowohl für den öffentlichen als auch für den privaten Sektor, wenn auch mit gewissen Unterschieden. Adressat der Verpflichtungen aus der Konvention ist in beiden Fällen der Staat, der auch in beiden Fällen über einen gewissen Ermessensspielraum bei der Umsetzung seiner Verpflichtungen verfügt. Die Verhandlungsführer wollten damit der Vielfalt der verschiedenen Rechtssysteme, Traditionen und Gepflogenheiten sowie den zahlreichen Einsatzmöglichkeiten von KI-Systemen sowohl im öffentlichen als auch im privaten Sektor gerecht werden.[26]

Im öffentlichen Sektor sind die Staaten verpflichtet, die Konvention auf die Tätigkeiten während des ganzen Lebenszyklus von KI-Systemen anzuwenden. Darunter fallen auch Tätigkeiten privater Akteure, die im Auftrag der öffentlichen Hand handeln (Art. 3 Abs. 1 Bst. a).[27] KI-Systeme werden im öffentlichen Sektor beispielsweise bei der Gewährung von Sozialleistungen, im Bereich der Steuern oder im Rahmen von Asylverfahren eingesetzt.[28] In der Praxis müssen die Vertragsstaaten nach Prüfung der Risiken und Auswirkungen der betreffenden Tätigkeiten entscheiden, wie sie die Bestimmungen der Konvention umsetzen wollen. Sie können ihre Gesetze beibehalten, anpassen oder neue Gesetze erlassen. Je nach Abstufung in Artikel 1 Absatz 2 können diese Massnahmen aus administrativen Massnahmen oder Kreisschreiben, der Rechtsprechung der Gerichte oder nicht zwingenden Massnahmen bestehen.[29] In der Schweiz sind, wie bereits erwähnt (vgl. Ziff. 4.2.1), in jedem Fall die verfassungsrechtlichen Vorgaben für die Form der zu erlassenden Rechtsakte zu beachten.

Während der Verhandlungen musste ein Kompromiss zwischen den Staaten gefunden werden, die sich für bzw. gegen den Einbezug des Privatsektors in den Geltungsbereich der KI-

---

[25] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 26.

[26] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 16.

[27] Dazu gehören die Tätigkeiten privater Akteure, die im Rahmen eines Vertrags mit einer Behörde oder anderen privaten Stelle zur Erbringung öffentlicher Dienstleistungen tätig sind, sowie die öffentliche Auftragsvergabe und das öffentliche Beschaffungswesen, vgl. Erläuternder Bericht zur KI-Konvention (Fn. 20), N 28.

[28] Vgl. NADJA BRAUN BINDER/NINA LAUKENMANN/LILIANE OBRECHT, KI in der Verwaltung: Entwicklungen und Herausforderungen, Jusletter IT 4. Juli 2024, 5 N 9.

[29] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 18 ff.

Konvention aussprachen. Nun sieht der Text vor, dass die Vertragsparteien in Übereinstimmung mit dem Ziel und dem Zweck der Konvention die Risiken und Auswirkungen angehen müssen, die sich aus den Tätigkeiten innerhalb des ganzen Lebenszyklus der im privaten Sektor verwendeten KI-Systeme ergeben. Die Bezugnahme auf das Ziel und den Zweck der Konvention bedeutet, dass alle in Artikel 1 enthaltenen Konzepte übernommen werden. Somit hat der Staat auch im Privatsektor die Risiken und Auswirkungen von KI-Tätigkeiten auf Menschenrechte, Demokratie und Rechtsstaatlichkeit zu prüfen und dann zu entscheiden, wie er diese angehen will.[30] Er kann entscheiden, die in den Kapiteln II bis VI der KI-Konvention genannten Grundsätze und Verpflichtungen anzuwenden, kann aber auch andere Massnahmen ergreifen.

Der Unterschied zwischen den beiden Sektoren besteht nicht so sehr im angestrebten Ziel – in beiden Fällen geht es darum, die mit KI-Systemen verbundenen Risiken und deren Auswirkungen auf Menschenrechte, Demokratie und Rechtsstaatlichkeit anzugehen – sondern in der Art und Weise, wie dieses Ziel erreicht werden soll. Im Privatsektor besteht diesbezüglich ein grösserer Spielraum.

Gemäss dem erläuternden Bericht können die in Artikel 3 Absatz 1 Buchstabe b genannten «anderen Massnahmen» für den Privatsektor Verwaltungsmassnahmen sein.[31] Gemeint sind Weisungen oder Kreisschreiben. Ob für die Schweiz der Rückgriff auf solche Rechtsakte ausreicht, ist jedoch fraglich. Die Auferlegung von Rechten und Pflichten im Bereich der KI im Privatsektor bedarf grundsätzlich einer formell gesetzlichen Grundlage, insoweit diese Grundrechte der Betroffenen (insbesondere die Wirtschaftsfreiheit) tangieren. Gemäss dem erläuternden Bericht können die Massnahmen auch freiwillig sein.[32] Dabei wird es sich aller Voraussicht nach um Selbstregulierungsmassnahmen handeln. Es ist jedoch zweifelhaft, dass die Ziele der Konvention in der Schweiz ohne staatliche Massnahmen erreicht werden können, die auch darin bestehen könnten, private Akteure zur Schaffung von Selbstregulierungsmechanismen zu verpflichten oder diese durch den Staat zu unterstützen.

In jedem Fall muss der Staat sicherstellen, dass der gewählte Ansatz die Risiken und Auswirkungen von KI im Privatsektor in einer Weise erfasst, die dem Ziel und dem Zweck der der KI-Konvention entspricht. Darüber hinaus dürfen die Vertragsparteien ihre völkerrechtlichen Verpflichtungen nicht unterschreiten (Art. 3 Abs. 1 Bst. b *in fine*).

Die Staaten müssen ihre Wahl, die jederzeit geändert werden kann, der Generalsekretärin oder dem Generalsekretär bei der Ratifikation, der Annahme oder der Genehmigung der Konvention, bzw. beim Beitritt zur Konvention mitteilen (Art. 3 Abs. 1 Bst. b). Jede Vertragspartei legt der Konferenz der Vertragsparteien innerhalb von zwei Jahren nach ihrem Beitritt und da-

---

30 Erläuternder Bericht zur KI-Konvention (Fn. 20), N 29.

31 Erläuternder Bericht zur KI-Konvention (Fn. 20), N 29.

32 Erläuternder Bericht zur KI-Konvention (Fn. 20), N 29.

nach in regelmässigen Abständen einen Bericht über die Tätigkeiten vor, die sie unternommen hat, um Artikel 3 Absatz 1 Buchstaben a und b Wirksamkeit zu verleihen (vgl. Art. 24 der Konvention, Ziff. 4.4.2).

4.2.3.1.2 Bedeutung für die Schweiz

Privatsektor

Wie oben erwähnt (vgl. Ziff. 4.2.3.1.1) ist es fraglich, ob rein administrative oder freiwillige Massnahmen im schweizerischen Recht ausreichen, um den Risiken und Auswirkungen, die sich aus den Tätigkeiten innerhalb des Lebenszyklus von KI-Systemen im Privatsektor ergeben, in einer Weise anzugehen, die mit dem Ziel und dem Zweck der KI-Konvention vereinbar ist. Mangels anderer in Frage kommender Massnahmen geht die Analyse daher davon aus, dass der schweizerische Gesetzgeber Artikel 3 Absatz 1 Buchstabe b der Konvention umsetzen wird, indem er die Bestimmungen der Konvention auch auf den privaten Sektor anwendet.

Für die Regulierung des privaten Sektors besteht jedoch ein grösserer Spielraum. In Betracht zu ziehen sind hier Unterscheidungen zwischen öffentlichem und privatem Sektor, die die Besonderheit des schweizerischen Rechts berücksichtigen, insbesondere angesichts der Tatsache, dass im öffentlichen Sektor die rechtlichen Anforderungen grundsätzlich höher sind (siehe unten) und dass die Grundrechte grundsätzlich nur vertikal eine Wirkung haben.

Konkret bedeutet das für das schweizerische Recht, dass die Bestimmungen und Grundsätze der Konvention im Privatsektor dort zur Anwendung gelangen sollten, wo eine direkte oder indirekte horizontale Wirkung der Grundrechte besteht oder in Zukunft erkannt wird. Die KI-Konvention hat keine Erweiterung des Wirkungsbereichs der Grundrechte zwischen Privaten zur Folge, die über die Bestimmungen von Artikel 35 Absätze 1 und 3 BV hinaus geht (vgl. Ziff. 4.3.1.1). Die Anerkennung einer direkten oder indirekten horizontalen Wirkung der Grundrechte zwischen Privaten ist somit Voraussetzung, dass die Konvention für privatrechtliche Beziehungen in der Schweiz anwendbar ist.

Die Auswirkung der Konvention auf die Beziehungen zwischen Privaten könnte sich jedoch weiterentwickeln, insbesondere aufgrund von völkerrechtlichen Verpflichtungen, die die Schweiz eingegangen ist (z. B. das Internationale Übereinkommen zur Beseitigung jeder Form von Rassendiskriminierung[33], das Internationale Übereinkommen zur Beseitigung jeder Form von Diskriminierung der Frau[34] oder das Übereinkommen über die Rechte von Menschen mit Behinderungen[35]). Auch die Entwicklung der Rechtsprechung könnte eine Rolle spielen, einschliesslich der Rechtsprechung des EGMR zur Auslegung der EMRK und der sich daraus ergebenden positiven Verpflichtungen.

---

[33] SR **0.104**.

[34] SR **0.108**.

[35] SR **0.109**.

An dieser Stelle ist darauf hinzuweisen, dass bestimmte Beziehungen zwischen Privatpersonen von der Umsetzung der KI-Konvention überhaupt nicht betroffen sein könnten. Dies ist dann der Fall, wenn keine Anknüpfung an ein Grundrecht besteht. Zu denken ist hier an die klassische Gewährleistung für die Funktionsfähigkeit eines Geräts (z. B. ein Mähroboter, der nur den halben Garten mäht oder ein intelligenter Kühlschrank, der die Temperatur falsch einstellt) oder die mietrechtlichen Ansprüche des Mieters bei Mängeln. Mögliche wirtschaftliche Schäden aufgrund von Mängeln solcher KI-Systeme werden a priori keine Grundrechtsverletzung darstellen. Diese Frage lässt sich jedoch nicht eindeutig beantworten: Die Wirksamkeit eines Grundrechts im Verhältnis zwischen Privaten hängt letztlich von der Auslegung des jeweiligen Rechts, dem geschützten Rechtsgut, den Funktionen, die es erfüllt und den Umständen seiner Anwendung ab.[36]

Es sei darauf hingewiesen, dass die vorliegende Analyse sich bei der Frage nach einem allfälligen gesetzgeberischen Handlungsbedarf im Privatsektor nicht zum Regelungsort äussert. Es müssen nicht zwingend privatrechtliche Vorschriften (ZGB, OR usw.) sein, sondern es können auch öffentlich-rechtliche Vorschriften sein, die auf Privatpersonen anwendbar sind, oder gemischte Regelungen wie das DSG oder das GlG, die beide Arten von Normen enthalten.

Öffentlicher Sektor

Beim öffentlichen Sektor ist zu beachten, dass der Staat aufgrund seiner Funktion bestimmten Verpflichtungen unterliegt, die für den privaten Sektor nicht gelten und unterschiedliche Regelungen rechtfertigen können. Der wichtigste Grundsatz im schweizerischen Recht ist sicherlich das Legalitätsprinzip, wonach staatliches Handeln einer gesetzlichen Grundlage bedarf (Art. 5 und 164 BV; vgl. auch Art. 36 zur Einschränkung von Grundrechten). Dieser Grundsatz ist auch beim Einsatz von KI-Systemen zu beachten.[37]

Der Staat kann KI-Systeme in verschiedenen Bereichen und in unterschiedlichem Umfang einsetzen. Die Einsatzmöglichkeiten reichen von rein interner Unterstützung mit keiner oder nur geringer Aussenwirkung (z. B. Anwendungen zur automatisierten Aufgabenverteilung unter Mitarbeitenden, Anwendungen zum Übersetzen von Dokumenten oder Erstellen von Sitzungsprotokollen) über Systeme, welche die Verwaltung bei der Entscheidungsfindung unterstützen (Teilautomatisierung), offenbar die häufigste Form, bis hin zu Systemen, die selbst Entscheidungen treffen (Vollautomatisierung).[38]

---

[36] Vgl. CR Cst.-VINCENT MARTENET, Art. 35 N 77.

[37] NADJA BRAUN BINDER/NINA LAUKENMANN/LILIANE OBRECHT, KI in der Verwaltung (Fn. 28), 7 N 17.

[38] BJ, Totalrevision des Datenschutzgesetzes (DSG) – Übersicht zu den wichtigsten Änderungen für die Erarbeitung der Rechtsgrundlagen betreffend Datenbearbeitungen durch Bundesorgane, Oktober 2022, 23 f., abrufbar unter: www.bj.admin.ch > Staat & Bürger > Datenschutz > Informationen für Bundesorgane > Totalrevision des Datenschutzgesetzes (DSG) – Übersicht zu den wichtigsten Änderungen für die Erarbeitung der Rechtsgrundlagen betreffend Datenbearbeitungen durch Bundesorgane (abgerufen am 27. August 2024).

Die Beantwortung der Frage, ob, auf welcher Normstufe und in welcher Normdichte eine ausdrückliche gesetzliche Grundlage erforderlich ist, hängt im Wesentlichen von der Bedeutung der behandelten Materie (Art. 164 BV) bzw. von allfälligen Grundrechtseinschränkungen (Art. 36 BV) ab.[39] Interne Assistenzsysteme ohne oder mit geringer Aussenwirkung sind grundsätzlich der Verwaltungsunterstützung zuzuordnen und haben ihre Legitimation in der gesetzlichen Grundlage, welche die Aufgaben der Behörde regelt (z. B. Anwendungen, welche die Aufgaben automatisch unter den Mitarbeitenden verteilen, Dokumente übersetzen oder Sitzungsprotokolle erstellen).[40] Dies dürfte auch für *Chatbots* gelten, sofern sie sich auf die Erteilung von Informationen und Auskünften beschränken und für die Bearbeitung von Personendaten eine gesetzliche Grundlage besteht. In diesem Fall kann davon ausgegangen werden, dass es sich um die Wahl eines Kommunikationskanals handelt, bei der die Behörden grundsätzlich frei sind.[41] Hingegen bedürfen ganz oder teilweise automatisierte Entscheide einer formell gesetzlichen Grundlage, wenn sie schwerwiegend in die Grundrechte der betroffenen Personen eingreifen können (Art. 36 Abs. 1 BV und Art. 164 Abs. 1 Bst. b BV)[42] und/oder wenn es sich um grundlegende Bestimmungen handelt, die die Organisation und das Verfahren der Bundesbehörden betreffen (Art. 164 Abs. 1 Bst. g BV). Eine Rechtsgrundlage in einem Gesetz im materiellen Sinn dürfte z. B. zumindest dann erforderlich sein, wenn die Unterstützung so weit geht, dass der Algorithmus, insbesondere im Rahmen eines Verwaltungsverfahrens, einen Text vorbereitet, den der Mensch nur noch zu überprüfen hat.[43]

Im Datenschutzbereich hat der Gesetzgeber das Erfordernis der gesetzlichen Grundlage in Artikel 34 DSG konkretisiert. Gemäss dieser Bestimmung, die eine Konkretisierung von Artikel 36 BV darstellt, dürfen Bundesorgane, von Ausnahmen abgesehen (vgl. Art. 34 Abs. 4 und Art. 36 Abs. 2 DSG), Personendaten nur bei Vorliegen einer gesetzlichen Grundlage bearbeiten. Absatz 2 sieht vor, dass eine Grundlage in einem Gesetz im formellen Sinn in folgenden Fällen erforderlich ist: bei der Bearbeitung von besonders schützenswerten Personendaten (Bst. a), beim Profiling (Bst. b) und wenn der Bearbeitungszweck oder die Art der Datenbearbeitung zu einem schwerwiegenden Eingriff in die Grundrechte der betroffenen Person führen können (Bst. c). KI-Systeme, die Personendaten bearbeiten, unterliegen selbstverständlich diesen Anforderungen. KI-Systeme, die eine vollständig oder teilweise automatisierte Einzelentscheidung generieren, können unter Buchstabe c fallen, sofern sie nicht bereits von Artikel 34 Buchstaben a und b erfasst sind. Dies ist abhängig von der Schwere des allfälligen Grundrechtseingriffs, insbesondere im Bereich des Schutzes der Privatsphäre und des Datenschutzes (Art. 13 BV). Generell gilt: Je intransparenter die Art der Bearbeitung, je grösser der Ermessensspielraum und je mehr Daten aus verschiedenen Quellen zusammengetragen werden, desto höher ist das Risiko einzuschätzen und desto eher bedarf es einer Grundlage in einem Gesetz im formellen Sinn.[44]

---

[39] NADJA BRAUN BINDER/MATTHIAS SPIELKAMP/CATHERINE EGLI et al., Einsatz Künstlicher Intelligenz in der Verwaltung: rechtliche und ethische Fragen, Schlussbericht vom 28. Februar 2021 zum Vorprojekt IP6.4, 34.

[40] CATHERINE REITER, Künstliche Intelligenz im Verwaltungsverfahren – Ermessen als Stolperstein?, AJP 2022, 984 ff., 988; PHILIP GLASS, Datenschutzrecht für künstliche Intelligenz in der öffentlichen Verwaltung: eine Auslegeordnung am Beispiel des Kantons Zürich, in: Michael Widmer (Hrsg.), Datenschutz: Rechtliche Schnittstellen, Zürich 2023, 177 ff., 208.

[41] NADJA BRAUN BINDER/MATTHIAS SPIELKAMP/CATHERINE EGLI et al. (Fn. 39), 58 f. Dies bedeutet jedoch nicht, dass sich nicht auch andere rechtliche Fragen stellen, insbesondere im Zusammenhang mit der Einhaltung von Treu und Glauben (Art. 9 BV), vgl. für weitere Entwicklungen zu diesem Thema NADJA BRAUN BINDER/LILIANE OBRECHT/GRACE WITTMER, Vertrauensschutz bei fehlerhaften Behördenauskünften durch Chatbots, IusNet DigR 2024; NADJA BRAUN BINDER/NINA LAUKENMANN/LILIANE OBRECHT, KI in der Verwaltung (Fn. 28), 5 N 10 ff., 18 N 45 ff.

[42] NADJA BRAUN BINDER/MATTHIAS SPIELKAMP/CATHERINE EGLI et al., Einsatz Künstlicher Intelligenz (Fn. 39), 59.

[43] CATHERINE REITER, Künstliche Intelligenz im Verwaltungsverfahren (Fn. 40), 984 ff., 988.

[44] MONIQUE COSSALI SAUVAIN, in: Yaniv Benhamou/Bertil Cottier (Hrsg.), Petit commentaire LPD, Art. 34 N 35–37; BJ, Gesetzgebungsleitfaden Datenschutz – Auswirkungen des neuen Datenschutzgesetzes auf die Erarbeitung von Rechtsgrundlagen, Bern 2024, Kap. 3.2.3, 21, abrufbar

4.2.3.2 Nationale Sicherheit

Artikel 3 Absatz 2 ist das Ergebnis eines Konsenses zwischen den Staaten, die den Bereich der nationalen Sicherheit vom Anwendungsbereich der KI-Konvention ausnehmen wollten, und den Staaten, die eine vollständige Einbeziehung befürworteten.

Der Text sieht nun vor, dass die Vertragsparteien die KI-Konvention auf Tätigkeiten innerhalb des Lebenszyklus von KI-Systemen für den Schutz ihrer nationalen Sicherheitsinteressen anwenden können, aber nicht müssen, unabhängig davon, welche Stelle diese Tätigkeiten ausführt. In jedem Fall müssen die betreffenden Tätigkeiten im Einklang mit dem geltenden Völkerrecht stehen, da die nationale Sicherheit in den Anwendungsbereich vieler Übereinkommen fällt. Der erläuternde Bericht nennt als Beispiele die EMRK und die beiden Pakte der Vereinten Nationen[45] und stellt darüber hinaus klar, dass diese Tätigkeiten im Einklang mit den demokratischen Institutionen und Verfahren der Vertragsparteien erfolgen müssen.[46]

Nationale Sicherheit gilt als ein «im Wesentlichen umstrittener» Begriff, d. h. als ein Konzept, über dessen Inhalt auf internationaler Ebene kein Konsens erzielt werden kann.[47] Diese Situation eröffnet den Staaten einen grossen Handlungsspielraum und führt in der Rechtswissenschaft zu einer Vielfalt von Definitionsvorschlägen.[48]

Intern hat der Bundesrat im Rahmen eines breiten Begriffsverständnisses der nationalen Sicherheit folgende Klassifikation entwickelt: 1) Polizeiliche Gefahrenabwehr, Staatsschutz und Strafverfolgung; 2) Vorbeugung, Vorsorge und Bewältigung von natur- und zivilisationsbedingten Katastrophen und Notlagen; 3) Abhalten und Abwehr eines militärischen Angriffs; 4) Wahrung der Interessen der Schweiz im Ausland und Beiträge zum internationalen Krisenmanagement.[49]

---

unter: www.bj.admin.ch > Datenschutz > Informationen für Bundesorgane > Gesetzgebungsleitfaden Datenschutz (abgerufen am 27. August 2024).

45 Internationaler Pakt über wirtschaftliche, soziale und kulturelle Rechte (UNO-Pakt I, SR **0.103.1**) und Internationaler Pakt über bürgerliche und politische Rechte (UNO-Pakt II, SR **0.103.2**).

46 Erläuternder Bericht zur KI-Konvention (Fn. 20), N 32.

47 BSK BK-OLIVER DIGGELMANN/TILMANN ALTWICKER, Art. 57, N 9 und angegebene Referenz.

48 FRANCK ELONG MBOULÉ, Le régime juridique des biens des organisations internationales, Zürich 2022, 294 ff., 296.

49 Bericht des Bundesrates an die Bundesversammlung über die Sicherheitspolitik der Schweiz vom 23. Juni 2010, BBl **2010** 5133 5158 f.

Für die Parlamentarische Versammlung des Europarats besteht die nationale Sicherheit in der Abwehr jeder sichtbaren und tatsächlichen Bedrohung der demokratischen Ordnung des Staates und der Gesellschaft.[50] Diese Definition wird jedoch von den einzelnen europäischen Staaten nicht geteilt, da sie die Bedürfnisse der Nachrichtendienste bei der Bekämpfung von verdeckten oder spekulativen Bedrohungen nicht berücksichtigt.[51]

Die KI-Konvention soll diese Frage nicht klären, aber der erläuternde Bericht stellt klar, dass routinemässige Tätigkeiten zur Aufrechterhaltung der öffentlichen Ordnung im Rahmen von Prävention, Aufdeckung, Untersuchung und Verfolgung von Straftaten, einschliesslich von Bedrohungen der öffentlichen Sicherheit, nicht betroffen sind, sofern die nationale Sicherheit der Vertragsparteien nicht auf dem Spiel steht.[52] Nach Auslegung des BJ unterscheiden sich diese regulären Tätigkeiten von den geheimen Tätigkeiten der Nachrichtendienste, die grundsätzlich unter die Ausnahmebestimmung der nationalen Sicherheit fallen.

Zu beachten ist, dass die Ausnahme nicht für Aktivitäten von KI-Systemen mit doppeltem Verwendungszweck gilt, sofern auch ein anderer Zweck als die Wahrung der nationalen Sicherheit verfolgt wird.[53]

In den Verhandlungen zur KI-Konvention hat sich die Schweizer Delegation dafür ausgesprochen, den Bereich der nationalen Sicherheit vollumfänglich in den Geltungsbereich des Konventionstextes einzubeziehen. Es gibt keinen Grund, im Rahmen der vorliegenden Analyse von dieser Position abzuweichen. Zudem gelten im schweizerischen Recht auch in diesem Bereich die Grundrechte, wobei Einschränkungen unter den üblichen Voraussetzungen von Art. 36 BV, insbesondere bei Vorliegen eines überwiegenden öffentlichen Interesses, möglich sind. In der vorliegenden Analyse wird daher von der Hypothese ausgegangen, dass die Tätigkeiten von KI-Systemen im Zusammenhang mit dem Schutz der nationalen Sicherheit, insbesondere den Tätigkeiten des NDB, in den Anwendungsbereich der Konvention fallen.

4.2.3.3 Forschung und Entwicklung

Forschungs- und Entwicklungstätigkeiten sind unter bestimmten Voraussetzungen vom Anwendungsbereich der Konvention ausgenommen. Voraussetzung ist, dass die betreffenden KI-Systeme noch nicht zur Nutzung freigegeben sind und dass die Erprobung oder vergleichbare Aktivitäten so ausgelegt sind, dass sie Menschenrechte, Demokratie und Rechtsstaatlichkeit nicht beeinträchtigen (Art. 3 Abs. 3).

---

[50] Empfehlung 1402 (1999) der parlamentarischen Versammlung des Europarats, Kontrolle der internen Sicherheitsdienste in den Mitgliedstaaten des Europarats, Punkt A. 2.

[51] CR Cst.-OLIVIER BLEICKER, Art. 57 N 34 und Referenz in Fussnote 86.

[52] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 32 *in fine*.

[53] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 32.

Gemäss dem erläuternden Bericht sollten diese Tätigkeiten in jedem Fall den geltenden Menschenrechten und dem innerstaatlichen Recht sowie den anerkannten ethischen und professionellen Standards der wissenschaftlichen Forschung entsprechen. KI-Systeme, die nach der Forschungs- und Entwicklungsphase zur Nutzung freigegeben werden, sollten grundsätzlich mit der Konvention vereinbar sein, auch in Bezug auf ihre Konzeption und Entwicklung.[54]

Zu beachten ist, dass diese Ausnahme vom Anwendungsbereich nicht die Verpflichtungen aus Artikel 13 (Sichere Innovation) und Artikel 25 Absatz 2 (Internationale Zusammenarbeit) betrifft.

4.2.3.4 Nationale Verteidigung

Der Bereich der nationalen Verteidigung ist vom Anwendungsbereich der Konvention ausgenommen (Art. 3 Abs. 4). Dies entspricht dem in Artikel 1 Buchstabe d der Satzung des Europarats verankerten Grundsatz, dass Fragen der nationalen Verteidigung nicht in die Zuständigkeit der Organisation fallen.

Dies bedeutet aber keinesfalls, dass Tätigkeiten, die innerhalb des Lebenszyklus von KI-Systemen im Zusammenhang mit der nationalen Verteidigung durchgeführt werden, nicht unter das Völkerrecht fallen.[55] Wie bei der nationalen Sicherheit gelten in der Schweiz auch in diesem Bereich die Grundrechte. Daher sollten keine Ausschlüsse vorgesehen werden, sondern allenfalls Ausnahmen oder Anpassungen, wo dies im Interesse der nationalen Verteidigung der Schweiz gerechtfertigt sein könnte.

4.2.3.5 Zusammenfassung

Die Hauptelemente des Anwendungsbereichs der Konvention gemäss Artikel 3 lassen sich wie folgt zusammenfassen:

- Öffentlicher / privater Sektor (Art. 3 Abs. 1):

  Die Konvention gilt für beide Sektoren. Im Privatsektor haben die Vertragsparteien jedoch einen grösseren Handlungsspielraum, sofern sie die Risiken und Auswirkungen, die während des Lebenszyklus von KI-Systemen auftreten, in einer Weise angehen, die mit dem Ziel und dem Zweck des Übereinkommens vereinbar ist.

  Bei einer Ratifizierung würde die Schweiz die Grundsätze und Verpflichtungen der Kapitel II bis VI der Konvention auch auf den Privatsektor anwenden. Rein administrative oder freiwillige Massnahmen scheinen angesichts unserer Rechtsordnung nicht ausreichend. Zu berücksichtigen wären Unterscheidungen zwischen den beiden Sektoren, die den Besonderheiten des schweizerischen Rechts Rechnung tragen, insbesondere im Hinblick auf die grundsätzlich nur vertikale Wirkung der

[54] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 33 f.

[55] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 36.

Grundrechte und unter Berücksichtigung der Tatsache, dass die rechtlichen Anforderungen im öffentlichen Sektor grundsätzlich höher sind.

- Nationale Sicherheit (Art. 3 Abs. 2):

  Die Staaten sind nicht zur Anwendung der Konvention im Bereich der nationalen Sicherheit verpflichtet. Über den Geltungsbereich dieses Begriffs besteht kein internationaler Konsens. Er bezieht sich vor allem auf geheime nachrichtendienstliche Tätigkeiten. Für die Zwecke der vorliegenden Analyse geht das BJ davon aus, dass die Bestimmungen der Konvention auch in diesem Bereich Anwendung finden.

- Forschung und Entwicklung (Art. 3 Abs. 3):

  Die Konvention gilt nicht für Forschungs- und Entwicklungstätigkeiten, es sei denn, die Versuche werden in einer Weise durchgeführt, die die Menschenrechte, die Demokratie und die Rechtsstaatlichkeit beeinträchtigen könnte. Artikel 13 (Sichere Innovation) und Artikel 25 Absatz 2 (Internationale Zusammenarbeit) der KI-Konvention bleiben von dieser Ausnahme unberührt.

- Nationale Verteidigung (Art. 3 Abs. 4):

  Fragen der nationalen Verteidigung fallen nicht in die Zuständigkeit des Europarats und sind daher nicht Gegenstand der Konvention.

4.3 Pflichten und Grundsätze

4.3.1 Kapitel II: Allgemeine Verpflichtungen

4.3.1.1 Artikel 4 – Schutz der Menschenrechte

Nach Artikel 4 stellt jede Vertragspartei durch die Einführung oder Beibehaltung von Massnahmen sicher, dass die Tätigkeiten innerhalb des Lebenszyklus von KI-Systemen im Einklang mit den Verpflichtungen zum Schutz der Menschenrechte gemäss anwendbarem Völkerrecht und innerstaatlichem Recht stehen.

Die Vertragsparteien müssen sicherstellen, dass der bestehende rechtliche Rahmen einen ebenso wirksamen wie effizienten Schutz der Menschenrechte gewährleistet, sowohl im spezifischen Kontext von KI-Systemen als auch ausserhalb solcher Systeme.

Das eigentliche Ziel der KI-Konvention besteht darin, zu gewährleisten, dass die Tätigkeiten innerhalb des Lebenszyklus von KI-Systemen vollumfänglich mit der Achtung der Menschenrechte vereinbar sind. Zu diesem Zweck sind spezifische Verpflichtungen vorgesehen, insbesondere die Verpflichtung, einen Rahmen für das Risiko- und Folgenmanagement von KI-Systemen hinsichtlich Menschenrechte, Demokratie und Rechtsstaatlichkeit zu schaffen (vgl. Ziff. 4.3.4). Mit der Umsetzung dieser Verpflichtungen kommen die Vertragsparteien indirekt der allgemeinen Verpflichtung zum Schutz der Menschenrechte nach Artikel 4 der KI-Konvention nach.

Im schweizerischen Recht sind, wie oben erwähnt (vgl. Ziff. 3), beim Einsatz von KI-Systemen bereits verschiedene völkerrechtliche Instrumente zum Schutz der Menschenrechte an-

wendbar, insbesondere die EMRK. Im innerstaatlichen Recht gelten die Grundrechtsgarantien auch im spezifischen Kontext von KI-Systemen. Für den öffentlichen Sektor gilt insbesondere die in Artikel 36 BV verankerte Regelung, wonach Grundrechtseinschränkungen grundsätzlich vier kumulative Voraussetzungen (gesetzliche Grundlage, öffentliches Interesse, Verhältnismässigkeit und Wahrung des Kerngehalts des Grundrechts) erfüllen müssen. Dabei spielt der Ursprung des staatlichen Handelns (Mensch oder KI-System) für den Grundrechtsschutz keine Rolle.

Für den Privatsektor ist dabei zu beachten, dass nach Artikel 35 Absatz 1 BV die Grundrechte in der gesamten Rechtsordnung zur Geltung kommen. Nach Artikel 35 Absatz 3 BV sorgen die Behörden dafür, dass die Grundrechte, soweit sie sich dazu eignen, auch unter Privaten wirksam werden. Diese Bestimmung bestätigt, dass die Grundrechte ihre Wirkung zwischen Privaten grundsätzlich nicht unmittelbar entfalten können, sondern dass es zu ihrer Geltung eines besonderen Eingreifens der Behörden bedarf.[56] Dies gilt sowohl für den Erlass von Rechtsnormen als auch für deren Auslegung und Anwendung. Der einzige Fall, in dem die BV eine direkte horizontale Wirkung anerkennt, ist Artikel 8 Absatz 3 Satz 3 BV, der den Grundsatz der Lohngleichheit zwischen Mann und Frau garantiert.

Somit ist für die Umsetzung der in Artikel 4 der Konvention enthaltenen Verpflichtung zum Schutz der Menschenrechte in innerstaatliches Recht zu berücksichtigen, dass die Grundrechte im Schweizer Recht zwischen Privaten in der Regel nur eine indirekte horizontale Wirkung entfalten.

Aus dem Gesagten ergibt sich, dass der schweizerische Rechtsrahmen für den Schutz der Grundrechte beim Einsatz von KI-Systemen, insbesondere im öffentlichen Sektor, bereits zur Anwendung kommt. Die Konvention verlangt, dass dieser Rahmen ausreichend und wirksam ist, um den besonderen Herausforderungen von KI-Systemen gerecht zu werden. Das eigentliche Ziel aller in der KI-Konvention enthaltenen Regeln und allgemeinen Grundsätze ist die Stärkung des Grundrechtsschutzes. Bei der Prüfung der Wirksamkeit des Schutzes im schweizerischen Recht und des allfälligen Handlungsbedarfs wird deshalb auf die nachfolgenden Ausführungen Bezug genommen.

4.3.1.2 Artikel 5 – Integrität demokratischer Prozesse und Achtung der Rechtsstaatlichkeit

Nach Artikel 5 Absatz 1 stellt jede Vertragspartei durch die Einführung oder Beibehaltung von Massnahmen sicher, dass KI-Systeme nicht eingesetzt werden, um die Integrität, Unabhängigkeit und Effektivität demokratischer Institutionen und Prozesse zu untergraben. Dies schliesst die Gewaltenteilung, die Achtung der Unabhängigkeit der Justiz und den Zugang zur Justiz ein. Nach Artikel 5 Absatz 2 ergreift oder beibehält jede Vertragspartei Massnahmen zum Schutz und zur Aufrechterhaltung ihrer demokratischen Prozesse während der Tätigkei-

[56] GIORGIO MALINVERNI/MICHEL HOTTELIER/MAYA HERTIG RANDALL/ALEXANDRE FLÜCKIGER, Droit constitutionnel suisse – Volume II: Les droits fondamentaux, Bern 2021, N 135.

ten innerhalb des Lebenszyklus von KI-Systemen, einschliesslich des gleichberechtigten Zugangs und der gleichberechtigten Teilhabe einer Person am öffentlichen Diskurs und der Möglichkeit zur freien Meinungsbildung.

Allerdings werden die Begriffe «demokratische Institutionen und Prozesse» und «Rechtsstaatlichkeit» im Konventionstext nicht definiert und bleiben daher vage. In der Analyse wird diese Bestimmung definiert als Verpflichtung, Massnahmen zu ergreifen oder beizubehalten, um Demokratie und Rechtsstaatlichkeit als strukturelle Prinzipien demokratischer Länder zu schützen.

KI-Systeme können sehr unterschiedliche Auswirkungen auf demokratische Institutionen, Prozesse und die Rechtsstaatlichkeit haben. Wie im erläuternden Bericht ausgeführt, können sich die Vertragsparteien bei der Erfüllung ihrer Verpflichtungen nach Artikel 5 auf bestimmte von KI ausgehende Risiken konzentrieren, z. B. auf das Risiko einer Beeinträchtigung des politischen Pluralismus.[57] Die Konvention verweist weder auf ein bestimmtes zu berücksichtigendes Risiko, noch gibt sie eine zu ergreifende Massnahme vor. Die Bestimmung ist zwar weit gefasst, hat aber den Vorteil, dass sie zukunftssicher ist, da sie auch bis anhin noch nicht bekannte Risiken umfassen kann.

Nachfolgend werden Beispiele für negative Auswirkungen von KI-Systemen auf die Integrität demokratischer Prozesse und die Wahrung der Rechtsstaatlichkeit anhand ausgewählter Bestimmungen des schweizerischen Rechts aufgezeigt.

- Desinformation und Falschinformationen

  KI-Systeme können das Phänomen der Desinformation und der Verbreitung von Falschinformationen, deren Ziel es ist, die Öffentlichkeit zu falschen Annahmen zu verleiten, verstärken. Desinformation kann die öffentliche Meinung lenken und Wahlen und Abstimmungen beeinflussen. Falschinformationen können verschiedene Formen annehmen, z. B. Behauptungen, unbegründete Meinungsäusserungen oder Hassreden gegen gesellschaftliche Gruppen oder Minderheiten. Automatisierte Software («Bots»), die menschliches Verhalten in sozialen Medien nachahmen, indem sie Informationen posten, Beiträge «liken» oder sich an reale Personen wenden, können beispielsweise durch die Verbreitung von Hassreden die öffentliche Meinung polarisieren. KI-Systeme können auch dazu eingesetzt werden, Stimmen und Bilder zu imitieren, um alternative Realitäten bzw. «Deepfakes» zu schaffen und einzusetzen.[58]

---

[57] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 46.

[58] YVES-MARIE DOUBLET, Désinformation et campagnes électorales, Bericht zuhanden des Europarats, Juni 2019, 5, 11 und 16; siehe auch MURAT KARABOGA/NULA FREI et al., Deepfakes und manipulierte Realitäten. Technologiefolgenabschätzung und Handlungsempfehlungen für die Schweiz, Ta-Swiss, TA 81/2024, Zollikon:vdf.

Dabei verschwimmt die Unterscheidung zwischen Fiktion und Realität zunehmend. Darüber hinaus können erfundene und irreführende Informationen über Chatbots verbreitet werden, was zur Desinformation beiträgt und je nach Kontext zu einem Problem für die Demokratie werden kann.[59]

Die schweizerische Gesetzgebung enthält bereits einschlägige Normen, die unter bestimmten Umständen bei Desinformation und Fehlinformation zur Anwendung kommen:

- Bei Wahlen und Abstimmungen schützt Artikel 34 Absatz 2 BV die freie Willensbildung und die unverfälschte Stimmabgabe der Stimmbürgerinnen und Stimmbürger. Das BPR enthält spezifische Bestimmungen für Wahlen und Abstimmungen auf Bundesebene; so kann gemäss Artikel 77 Absatz 1 Buchstabe b BPR gegen Unregelmässigkeiten bei Abstimmungen Beschwerde erhoben werden.

  Nach Rechtsprechung des BGer darf ein Abstimmungsergebnis, das den freien Willen der Stimmberechtigten nicht zuverlässig und unverfälscht zum Ausdruck bringt, nicht anerkannt werden. Konnte der freie Wille nicht zuverlässig und unverfälscht zum Ausdruck gebracht werden, ist die Abstimmung zu wiederholen. Diese Bestimmungen sollen in erster Linie eine unzulässige Einflussnahme des Staates verhindern. Aber auch Private können Fehlinformationen verbreiten und die freie Willensbildung der Stimmberechtigten beeinflussen. Dies ist beispielsweise dann der Fall, wenn in einem so späten Zeitpunkt mit offensichtlich unwahren und irreführenden Angaben in den Abstimmungskampf eingegriffen wird, dass es den Stimmberechtigten nach den Umständen nicht mehr möglich ist, sich aus anderen Quellen ein zuverlässiges Bild von den tatsächlichen Verhältnissen zu machen. In Anbetracht der Meinungsäusserungsfreiheit wird eine derartige Beeinträchtigung der Stimmberechtigten nicht leichthin angenommen werden. Die Aufhebung einer Abstimmung fällt daher nur unter grösster Zurückhaltung und bei ganz schwerwiegenden Verstössen in Betracht.[60]

  Weiter anerkennt das BGer die Richtigstellung offensichtlich falscher oder irreführender Informationen als triftigen Grund für die Rechtfertigung von behördlichem Eingreifen in den Abstimmungskampf. Unter bestimmten Umständen besteht sogar eine Pflicht der Behörden, zur Sicherstellung des

---

[59] Vgl. die von Algorithmwatch und Forensics durchgeführte Studie, in der die Antworten eines Chatbots im Zusammenhang mit den Wahlen in der Schweiz, Bayern und Hessen im Oktober 2023 getestet wurden, abrufbar unter: https://algorithmwatch.ch > Publikationen > KI-Chatbot liefert falsche Antworten auf Fragen zu demokratischen Wahlen (abgerufen am 27. August 2024). Siehe auch die kürzliche Studie vom August 2024 im Zusammenhang mit den Regionalwahlen in Deutschland, abrufbar unter: https://algorithmwatch.org/de > Blog > Chatbots bringen noch immer viele Falschinformationen in Umlauf (abgerufen am 28. August 2024).

[60] BGE **135** I 292, E. 4.1.

bundesrechtlichen Anspruchs auf eine freie Willensbildung und unverfälschte Stimmabgabe zu intervenieren. Dabei steht den Behörden jedoch ein grosser Ermessensspielraum zu. Eine Interventionspflicht, deren Verletzung zur Aufhebung der Abstimmung führen kann, wird daher grundsätzlich nur anzunehmen sein, wenn die Einflussnahme privater Akteure die Willensbildung der Stimmberechtigten in ganz schwerwiegender Art beeinträchtigt oder geradezu verunmöglicht.[61]

Die obigen Überlegungen gelten auch für den Einsatz von KI-Systemen zur Desinformation. Die Verbreitung eines verzerrten Bildes kann in bestimmten Kontexten beispielsweise zur politischen Manipulation eingesetzt werden.[62]

Eine schwerwiegende Täuschung über den Gegenstand einer Abstimmung in den sozialen Medien durch Privatpersonen wirft erstens die Frage auf, ob die Behörden zum Eingreifen und zur Korrektur bzw. Berichtung der betreffenden Informationen verpflichtet sind.[63] Zweitens ist im Hinblick auf eine mögliche Annullierung der Abstimmung daran zu erinnern, dass die Rechtsprechung nicht den Nachweis eines Einflusses der Unregelmässigkeiten auf das Ergebnis der Abstimmung verlangt; es genügt vielmehr, dass ein solcher Einfluss möglich erscheint. In Ermangelung einer quantifizierten Feststellung der Auswirkungen einer Unregelmässigkeit im Verfahren wird ihr Einfluss auf das Ergebnis anhand der Gesamtumstände und grundsätzlich mit voller Kognition beurteilt. [64] Diese Erleichterung könnte sich als besonders nützlich erweisen, wenn Desinformation online über KI-Systeme verbreitet wird.

Die genannten Bestimmungen zu den politischen Rechten greifen jedoch erst nach einer Verletzung, und erlauben keine präventive Regulierung der Strukturen der Informationsproduktion und -verbreitung an sich.

- In der öffentlichen Kommunikation und Meinungsbildung nehmen Online-Plattformen heutzutage eine immer bedeutendere Stellung ein. Plattformen tragen damit einerseits zu einer besseren Verwirklichung der Meinungsäusserungs- und Informationsfreiheit bei, andererseits beeinflussen sie auch die öffentliche

---

[61] Urteil des BGer, 1C_472/2010, E. 4.3; NADJA BRAUN BINDER/MANUELA KÄLIN, Rechtliche Aspekte der politischen Meinungsbildung, in: Urs Bieri et al. (Hrsg.), Digitalisierung der Schweizer Demokratie. Technologische Revolution trifft auf traditionelles Meinungsbildungssystem, Ta-Swiss, TA 75/2021, Zürich:vdf, 125 ff., 129.

[62] Für eine Darstellung der verschiedenen möglichen Gefahren vgl. TOM LEBRUN, La liberté d'expression aux Etats-Unis et au Canada face au risque de propagande générée par l'intelligence artificielle, in: Céline Castets-Renard/Jessica Eynard (Hrsg.), Un droit de l'intelligence artificielle. Entre règles sectorielles et régime général. Perspectives comparées, Bruxelles, 2023, 409 ff.

[63] NADJA BRAUN BINDER/MANUELA KÄLIN, Rechtliche Aspekte der politischen Meinungsbildung (Fn. 61), 145.

[64] BGE **135** I 292, E. 4.4.

Debatte. Zudem hat die niederschwellige Möglichkeit zur Produktion und Diffusion von Inhalten verschiedene negative Auswirkungen zur Folge.

Neben den allgemeinen Rechtsregeln bestehen bislang keine spezifischen gesetzlichen Regelungen oder eine ausführliche Rechtspraxis bezüglich der Verantwortlichkeit von Intermediären für fremde widerrechtliche Inhalte auf ihren Plattformen. Auch bestehen keine bzw. nur punktuelle Vorschriften, welche die Rechte der Nutzerinnen und Nutzer gegenüber den Intermediären stärken oder die Plattformen zu mehr Transparenz verpflichten.

Aus diesem Grund hat der Bundesrat das UVEK (BAKOM) damit beauftragt, eine Vernehmlassungsvorlage für die Regulierung von grossen Online-Plattformen auszuarbeiten. Diese soll gewisse Sorgfaltspflichten für Plattformen enthalten. So sollen diese beispielsweise Melde- und Abhilfeverfahren für rechtswidrige Hassrede einrichten sowie eine Kontaktstelle und einen Rechtsvertreter in der Schweiz benennen müssen. Ebenfalls sollen die Rechte von Nutzenden gestärkt werden, z. B. indem Plattformen Massnahmen gegen Nutzende begründen und ihnen ein Beschwerdeverfahren zur Verfügung stellen müssen.

Verschiedene dieser Aspekte werden nicht zuletzt auch algorithmische Systeme betreffen, welche von Plattformen eingesetzt werden (z. B. zur Moderation oder zum Ranking von Inhalten). Entsprechend wird auch diese Regulierung zum Schutz demokratischer Prozesse (in diesem Fall der öffentlichen Meinungsbildung) und dem Rechtsstaat beitragen.

- Die Identifizierung von KI-generierten Inhalten ist ebenfalls ein Mittel zur Bekämpfung von Online-Desinformation. Eine solche Identifizierung schränkt die Verbreitung von Desinformation zwar nicht per se ein, ermöglicht den Empfängern aber eine bessere Beurteilung und Einordnung der erhaltenen Informationen. In diesem Sinne ist die Umsetzung der in Artikel 8 der KI-Konvention verankerten Transparenzpflicht im Zusammenhang mit dem Schutz der Integrität demokratischer Prozesse und der Rechtsstaatlichkeit von besonderer Relevanz. In der Analyse wird nachfolgend darauf eingegangen (vgl. Ziff. 4.3.2.3). Mit einer Kennzeichnung könnten die Nutzerinnen und Nutzer darüber aufgeklärt werden, dass ein Account oder eine Publikation nicht von einem Menschen stammt.

- Auch das Strafrecht bietet einen rechtlichen Rahmen, der unter bestimmten Voraussetzungen einschlägig sein und bestimmten strafrechtlich relevanten Verhaltensweisen Grenzen setzen kann. Die Bestimmungen des StGB sind grundsätzlich technologieneutral und gelten unabhängig vom Tatmittel des Täters. Das Strafrecht bietet daher einen allgemeinen Rahmen, der bestimmte Verbrechen und Vergehen gegen den öffentlichen Frieden (Art. 258 ff. StGB) und Vergehen gegen den Volkswillen (Art. 279 ff. StGB) erfasst. So ist beispielsweise Artikel 261bis StGB über Diskriminierung und Aufruf zu Hass unter bestimmten Voraussetzungen auf Hassreden in sozialen Netzwerken anwendbar. Gleiches gilt für Artikel 173 ff. StGB über die Ehrverletzungsdelikte oder Artikel 179^{decies} StGB über Identitätsmissbrauch (vgl. zu diesen Bestimmungen Ziff. 6.6). Zivilrechtlich kann auch der Persönlichkeitsschutz (Art. 28 ff. ZGB)

zur Anwendung kommen (vgl. zu den Überlegungen im Zusammenhang mit Deepfakes Ziff. 6.3.3).

Bei der Anwendung dieser Normen ist die in Artikel 16 Absatz 2 BV verankerte Meinungsfreiheit zu beachten. Diese gilt auch für falsche und irreführende Behauptungen.[65] Einschränkungen bleiben möglich, müssen aber den Anforderungen von Artikel 36 BV genügen.[66]

- Targeting zu politischen Zwecken

Im Zusammenhang mit den Risiken für die freie Meinungsbildung der Bürgerinnen und Bürger ist auch die Frage des Targetings zu politischen Zwecken durch den Einsatz von KI-Systemen und die Aspekte des Datenschutzes zu nennen. Mithilfe von Algorithmen lässt sich das Verhalten der Bürgerinnen und Bürger beeinflussen, z. B. durch das gezielte Versenden von Informationen an Gruppen, von denen ein Profil erstellt wurde, in der Annahme, dass deren Mitglieder besonders sensibel auf eine bestimmte Art von Nachrichten reagieren.

a) Rechtlicher Rahmen gemäss DSG

Im schweizerischen Recht ist in diesem Zusammenhang insbesondere der durch das neue, am 1. September 2023 in Kraft getretene DSG geschaffene rechtliche Rahmen zu beachten.[67]

Die Bearbeitung von Daten über die politische Einstellung der Bürgerinnen und Bürger stellt eine Bearbeitung von besonders schützenswerten Personendaten dar (Art. 5 Abs. 1 Bst. c DSG). Diese Daten geniessen somit einen qualifizierten Schutz. Wenn Daten aus verschiedenen Quellen mittels automatisierter Analysemethoden miteinander verknüpft werden, um politische Meinungen und Tendenzen der Stimmbürgerinnen und Stimmbürger zu ermitteln, handelt es sich um eine Profiling-Tätigkeit (Art. 5 Abs. 1 Bst. f DSG). Je nach Umständen kann es sich dabei sogar um ein Profiling mit hohem Risiko handeln (Art. 5 Abs. 1 Bst. g DSG).

Auch hier sind die allgemeinen Grundsätze von Treu und Glauben, Transparenz, Verhältnismässigkeit, Zweckbindung, Richtigkeit und Datensicherheit zu beachten. Wenn

---

65 FLORENT THOUVENIN/STEPHANIE VOLZ/MARK EISENEGGER/DANIEL VOGLER/MARIELA JAFFÉ, Governance von Desinformation in digitalisierten Öffentlichkeiten, Jusletter 5. Februar 2024, 9 N 24, und die angegebenen Ref.

66 RAPHAELA CUENI, Falsche und irreführende Informationen im Verfassungsrecht der Schweiz, ex ante 1/2019, 3 ff., insbesondere 9 ff.

67 Vgl. zu diesem Thema den Leitfaden vom 15. Dezember 2022 der Datenschutzbehörden des Bundes und der Kantone zur digitalen Bearbeitung von Personendaten im Rahmen von Wahlen und Abstimmungen in der Schweiz, abrufbar unter: https://www.edoeb.admin.ch > Datenschutz > Internet & Technologie > Leitfaden zu Wahlen und Abstimmungen (abgerufen am 27. August 2024); siehe auch MICHAEL MONTAVON, Big Data, un outil d'influence en période électorale, 2023, www.swissprivacy.law/258.

die geplante Bearbeitung ein hohes Risiko für die Persönlichkeit oder die Grundrechte der betroffenen Person mit sich bringen kann, muss der für die Bearbeitung Verantwortliche vorgängig eine Datenschutz-Folgenabschätzung durchführen (Art. 22 Abs. 1 DSG). Ob ein hohes Risiko besteht, insbesondere bei der Verwendung neuer Technologien wie KI, hängt von der Art, dem Umfang, den Umständen und dem Zweck der Bearbeitung ab. Ein solches Risiko besteht insbesondere bei der umfangreichen Bearbeitung von besonders schützenswerten Personendaten (Art. 22 Abs. 2 Bst. b DSG), namentlich beim umfassenden Profiling zur Ermittlung von politischen Meinungen und Tendenzen der Stimmberechtigten.

b) Politische Werbung

In der Schweiz ist politische Werbung in Radio und Fernsehen gemäss Artikel 10 RTVG und Artikel 17 Absatz 3 RTVV verboten. Der Anwendungsbereich dieser Bestimmungen ist jedoch beschränkt und andere Medien, auch die sozialen Medien, sind nicht erfasst.[68]

In der EU bestehen seit Anfang April 2024 spezifische Regeln für politische Werbung.[69] Diese umfassen auch spezifische Pflichten für politische Werbung online. Das BAKOM prüft im Rahmen der Vernehmlassungsvorlage für die Regulierung von Kommunikationsplattformen, ob ähnliche Pflichten wie in der EU auch in der Schweiz sinnvoll wären.

- Risiken für die Ausübung der Grundrechte, insbesondere die Kommunikationsfreiheiten

Der Schutz von Demokratie und Rechtsstaatlichkeit geht mit dem Schutz der Grundrechte einher, insbesondere der Kommunikationsfreiheiten (u. a. Meinungs- und Informationsfreiheit [Art. 16 BV], Medienfreiheit [Art. 17 BV], Versammlungsfreiheit [Art. 22 BV]), die für diese konstitutiv sind. KI-Systeme können insbesondere zu Überwachungszwecken eingesetzt werden, z. B. zur Gesichtserkennung im öffentlichen Raum; damit können sie potenziell insbesondere die Meinungsfreiheit, die Versammlungsfreiheit und die Privatsphäre einer grossen Anzahl von Personen beeinträchtigen.

Im bereits erwähnten Urteil vom 4. Juli 2023 Glukhin gegen Russland[70] verurteilte der EGMR Russland wegen Verletzung von Artikel 8 EMRK (Recht auf Achtung des Privatlebens) und Artikel 10 EMRK (Recht auf Freiheit der Meinungsäusserung) auf-

---

[68] NADJA BRAUN BINDER/MANUELA KÄLIN, Rechtliche Aspekte der politischen Meinungsbildung (n. 61), 134 ff.

[69] Verordnung (EU) 2024/900 des Europäischen Parlamentes und des Rates vom 13. März 2024 über die Transparenz und das Targeting politischer Werbung, PE/90/2023/REV/1; JO L, 2024/900, 20. März 2024.

[70] EGMR, Glukhin gegen Russland (Fn. 6).

grund des Einsatzes von Gesichtserkennungstechnologie zur Identifizierung und Lokalisierung von Herrn Glukhin, nachdem dieser allein in der Moskauer Metro demonstriert hatte. Der Gerichtshof kam zum Schluss, dass die Bearbeitung der Personendaten des Beschwerdeführers im Zusammenhang mit seiner friedlichen Demonstration, die weder die öffentliche Ordnung noch die öffentliche Sicherheit bedroht hatte, einen Eingriff von besonderer Schwere darstellte. Der Einsatz von Gesichtserkennungstechnologie in seinem Fall war mit den Idealen und Werten einer demokratischen und rechtsstaatlichen Gesellschaft unvereinbar. Im schweizerischen Recht finden sich die Artikel 8 und 10 EMRK in den Artikeln 13 und 16 Absatz 2 BV wieder. Die von der EMRK abgeleiteten Grundsätze sind für die Schweiz verbindlich.

Der in der Schweiz geltende gesetzliche Rahmen zum Schutz der Grundrechte erweist sich somit auch im Hinblick auf die Umsetzung von Artikel 5 der KI-Konvention als relevant.

Es bleibt somit festzuhalten, dass das Schutzobjekt von Artikel 5, nämlich die Integrität des demokratischen Prozesses und die Achtung der Rechtsstaatlichkeit, in der Schweiz von einem breiten Spektrum von Normen erfasst wird. Diese reichen vom Grundrechtsschutz über den Persönlichkeitsschutz und den Datenschutz bis hin zu den Rechten im Bereich der Ausübung der politischen Rechte. Die geplante Regulierung der grossen Kommunikationsplattformen soll zudem die Rechte der Nutzerinnen und Nutzer in der Schweiz und die Transparenz seitens dieser Plattformen stärken und damit auch zu einer gut funktionierenden Demokratie beitragen.

Aufgrund der ersten oben aufgeführten Analysen drängen sich keine weiteren Gesetzesänderungen auf. Da sich die Gefahren für die Demokratie besonders schnell entwickeln und potenziell wichtige Folgen zeitigen, ist die Situation wie auch die Entwicklung der Rechtsprechung und ihre Tragweite im digitalen Kontext weiterhin genau zu beobachten, um bei Bedarf rechtzeitig reagieren zu können. Der Bericht des Bundesrates in Erfüllung des Postulats 22.3006 SiK-N «Auslegeordnung zur Bedrohung der Schweiz durch Desinformationskampagnen»[71] ist diesbezüglich hilfreich.

Weiter ist hervorzuheben, dass die KI-Konvention auch andere Bestimmungen enthält, die für den Schutz von Demokratie und Rechtsstaatlichkeit relevant sein können, insbesondere das Transparenzprinzip (Art. 8) und der rechtliche Rahmen für das Risiko- und Folgenmanagement von KI-Systemen (Art. 16). Die Umsetzung dieser allgemein ausgerichteten Bestimmungen würde ebenfalls eine Vervollständigung und Verstärkung des geltenden rechtlichen Rahmens im Bereich des Schutzes der Demokratie und des Rechtsstaates ermöglichen.

[71] Beeinflussungsaktivität und Desinformation. Bericht des Bundesrates in Erfüllung des Postulats 22.3006 SiK-N, 19. Juni 2024, abrufbar unter: www.admin.ch > Dokumentation > Medienmitteilungen > Beeinflussungsaktivität und Desinformation: Bundesrat betont Resilienz und stärkt Analyse und Koordination (abgerufen am 27. August 2024).

4.3.2 Kapitel III: Grundsätze in Bezug auf Tätigkeiten im Lebenszyklus von Systemen künstlicher Intelligenz

4.3.2.1 Artikel 6 – Allgemeiner Ansatz

Artikel 6 sieht vor, dass die Vertragsparteien die in Kapitel III der KI-Konvention enthaltenen Grundsätze bei der Umsetzung an ihre innerstaatliche Rechtsordnung und ihre anderen Verpflichtungen aus der Konvention anpassen.

Laut dem erläuternden Bericht zur Konvention ist dieser Artikel insofern von besonderer Bedeutung, da jede Vertragspartei über ein detailliertes Rechtssystem zum Schutz der Menschenrechte mit eigenen Regeln, Grundsätzen und Gepflogenheiten verfügt, die den Geltungsbereich, den Inhalt der wesentlichen Rechte und Einschränkungen, mögliche Abweichungen oder Ausnahmen von diesen Rechten sowie die Funktionsweise der anwendbaren Überwachungs- und Durchsetzungsmechanismen betreffen. [72]

Zusammenfassend lässt sich sagen, dass diese Bestimmung den bereits oben (vgl. Ziff. 4.2.1) dargelegten völkerrechtlichen Grundsatz widerspiegelt, wonach die Staaten in der Umsetzung ihrer völkerrechtlichen Verpflichtungen frei sind. Im schweizerischen Recht sieht Artikel 164 BV vor, dass dort, wo Bundeskompetenz besteht, alle wichtigen rechtsetzenden Bestimmungen in der Form eines Bundesgesetzes zu erlassen sind.

4.3.2.2 Artikel 7 – Menschenwürde und individuelle Autonomie

Gemäss Artikel 7 beschliessen die Vertragsparteien die Einführung oder Beibehaltung von Massnahmen zur Wahrung der Menschenwürde und der individuellen Autonomie in Bezug auf Tätigkeiten innerhalb des Lebenszyklus von KI-Systemen.

Diese Bestimmung unterstreicht die Bedeutung der Menschenwürde und der individuellen Autonomie im Rahmen einer menschenzentrierten Regulierung. So dürfen Tätigkeiten innerhalb des Lebenszyklus von KI-Systemen nicht zur Entmenschlichung des Einzelnen führen.[73]

Individuelle Autonomie ist ein wichtiger Aspekt der Menschenwürde; sie betrifft die Selbstbestimmungsfähigkeit des Individuums, d. h. seine Fähigkeit, Wahlmöglichkeiten zu nutzen und Entscheidungen ohne unzulässige Beeinflussung oder Zwang zu treffen. In Bezug auf KI erfordert die individuelle Autonomie, dass der Mensch die Kontrolle über den Einsatz und die Auswirkungen von KI-Technologien auf sein Leben hat.[74] Dieses Prinzip ist daher eng mit dem Grundsatz von Transparenz und Aufsicht verbunden (Art. 8 der Konvention, vgl. Ziff. 4.3.2.3).

---

[72] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 51.

[73] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 53.

[74] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 55.

Im schweizerischen Recht wird die Menschenwürde auf Verfassungsstufe durch Artikel 7 BV garantiert. Die Rechtsprechung erkennt dieser Bestimmung die Bedeutung eines eigenständigen Grundrechts zu, das per Definition einen absoluten Schutz all jener Aspekte der Menschenwürde gewährleistet, die von den anderen Grundrechten nicht erfasst werden.[75] In der Praxis wird die Menschenwürde jedoch vor allem zur Unterstützung eines anderen Grundrechts herangezogen und war offenbar noch nie allein ausschlaggebend für ein Urteil einer Justizbehörde des Bundes.[76] Sie hat daher eher den Charakter eines Auffanggrundrechts oder eines letzten Rechtsmittels. Dies liegt an der Unbestimmtheit ihres Schutzbereichs, denn es geht um nichts anderes als um das Menschsein selbst.[77] Der Schutzbereich der Menschenwürde deckt sich mit ihrem unantastbaren Kern (Art. 36 Abs. 4 BV). Jeder Eingriff des Staates stellt somit eine Verletzung dieses Grundrechtes dar.[78] Selbstbestimmung und Schutz der individuellen Autonomie sind sowohl durch Artikel 10 Absatz 2 BV (Recht auf persönliche Freiheit) als auch durch Artikel 13 BV (Schutz der Privatsphäre) abgedeckt.

Angesichts der mit KI verbundenen technologischen Herausforderungen, insbesondere der Gefahr der Entmenschlichung des Einzelnen[79], dürfte die Rolle des Garanten der Menschenwürde noch an Bedeutung gewinnen. Bestimmte Anwendungen, wie etwa der Einsatz von KI-Tools durch Behörden im Rahmen einer permanenten Überwachung des Einzelnen zum Zwecke des «Social Scoring» oder der Einsatz von Emotionserkennung in bestimmten Situationen, können per se einen nicht zu rechtfertigenden Eingriff in die Menschenwürde darstellen. Solche Anwendungen von KI sind daher bereits verfassungsrechtlich verboten.

In diesem Zusammenhang sind auch die vom Bund im November 2020 verabschiedeten KI-Leitlinien relevant. Diese unverbindlichen Leitlinien sollen als allgemeine Orientierungshilfe für die Entwicklung von KI in der Bundesverwaltung dienen und eine einheitliche Praxis sicherstellen. Die erste Leitlinie stellt den Menschen in den Mittelpunkt. Betont wird darin die besondere Bedeutung des Grundrechtsschutzes beim Einsatz von KI.[80]

---

[75] CR Cst.-JACQUES DUBEY, Art. 7, N 25, und die angegebenen Referenzen.

[76] CR Cst.-JACQUES DUBEY, Art. 7, N 26; BSK BV-EVA MARIA BELSER/EVA MOLINARI, Art. 7 N 39 f.; *contra* SGK BV-RAINER J. SCHWEIZER/CHRISTOPH SPENLÉ, Art. 7 N 22.

[77] CR Cst.-JACQUES DUBEY, Art. 7, N 27.

[78] CR Cst.-JACQUES DUBEY Art. 7 N 56 ff., insbesondere N 59; BSK BV-EVA MARIA BELSER/EVA MOLINARI, Art. 7 N 63.

[79] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 53.

[80] Leitlinien «Künstliche Intelligenz» für den Bund, Orientierungsrahmen für den Umgang mit künstlicher Intelligenz in der Bundesverwaltung, 2020, abrufbar unter: www.sbfi.admin.ch > BFI-Politik > Bildungs-, Forschungs- und Innovationspolitik 2021-2024 > Transversale Themen > Digitalisierung im BFI-Bereich > Künstliche Intelligenz (abgerufen am 26. August 2024). Die Leitlinien sollen alle zwei Jahre evaluiert werden. Die nächste Evaluierung ist für dieses Jahr im Rahmen der Bestandsaufnahme der Regulierungsansätze im Bereich der KI vorgesehen. Die Evaluation im Jahr 2022 hat gezeigt, dass die KI-Leitlinien in der Bundesverwaltung gut bekannt sind und von den Mitarbeitenden, die mit KI arbeiten (sei es beim Einsatz von KI, auf regulatorischer Ebene usw.), berücksichtigt und angewendet werden. Es wird zwar darauf hingewiesen, dass die Leitlinien und ihre konkrete Anwendung innerhalb der Bundesverwaltung weiter diskutiert werden müssen, die Evaluation hat

Es stellt sich die Frage nach der Tragweite von Artikel 7 der Konvention bei der Umsetzung dieser Bestimmung im privaten Sektor (vgl. Art. 3 Abs. 1 Bst. b der Konvention, Ziff. 4.2.3.1). In der schweizerischen Rechtsordnung bestehen bereits Regeln, welche die Garantie der Menschenwürde und der individuellen Autonomie zwischen Privaten verwirklichen. Dies gilt insbesondere für Artikel 28 ZGB zum Schutz der Persönlichkeit (vgl. zur Entwicklung in diesem Zusammenhang Ziff. 6.3.3).

Angesichts der obigen Ausführungen ergibt sich, dass die Verfassung das Recht auf Menschenwürde und individuelle Autonomie gegenüber dem Staat bereits schützt. Entsprechende Regelungen bestehen auch im privaten Sektor. Grundsätzlich eignet sich das schweizerische Zivilrecht mit seinen offenen Generalklauseln gut zur Bewältigung der anstehenden Probleme. Dies gilt beispielsweise für Artikel 28 ZGB zum Schutz der Persönlichkeit. Die Frage, ob diese Regeln im Einzelfall präzisiert werden müssen, kann im Falle einer Ratifikation der Konvention vertieft werden.

4.3.2.3 Artikel 8 – Transparenz und Aufsicht

Artikel 8 verpflichtet die Vertragsparteien zur Einführung oder Beibehaltung von Massnahmen, um die auf den jeweiligen Kontext und die spezifischen Risiken zugeschnittenen Transparenz- und Kontrollanforderungen für Tätigkeiten innerhalb des Lebenszyklus von KI-Systemen, einschliesslich der Identifizierung von KI-generierten Inhalten, sicherzustellen.

Diese Bestimmung deckt zwei Aspekte ab:

- Der erste Aspekt betrifft die Transparenz von KI-Systemen. Transparenz soll durch die Verfügbarkeit angemessener Informationen über das KI-System gewährleistet werden, wie z. B. Zweck(e), bekannte Beschränkungen, Annahmen, technische Entscheidungen bei der Konzeption, Merkmale, Einzelheiten der zugrunde liegenden Modelle oder Algorithmen, Lernmethoden und Qualitätssicherungsverfahren. Darüber hinaus kann Transparenz bedeuten, dass die betroffenen Personen oder die Öffentlichkeit gegebenenfalls über die Einzelheiten der zum Training des Systems verwendeten Daten und den Schutz von Personendaten sowie über den Zweck des Systems und die Art und Weise, in der es konzipiert und eingesetzt wird, informiert werden müssen.[81]

  Das Prinzip der Transparenz zielt auch darauf ab, dass KI-Systeme für die relevanten KI-Akteure verständlich und zugänglich sein müssen. Dieser Grundsatz umfasst die Erklärbarkeit und die Interpretierbarkeit von KI-Systemen. *Erklärbarkeit* bezieht sich

jedoch keinen Bedarf für eine Anpassung oder Aktualisierung der Leitlinien ergeben, siehe BAKOM-Bericht, Monitoring der Leitlinien «Künstliche Intelligenz für den Bund», Evaluation der Anwendung und Aktualität der Leitlinien, 9. Dezember 2022, abrufbar unter: www.cnai.swiss > Ressourcen > KI-Leitlinien (abgerufen am 26. August 2024).

[81] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 58.

auf die Fähigkeit, vorbehaltlich der technischen Machbarkeit, hinreichend verständliche Erklärungen darüber zu liefern, warum ein KI-System Informationen, Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen generiert. Dies ist besonders in sensiblen Bereichen wie dem Gesundheitswesen, dem Finanzwesen und der Strafverfolgung wichtig, wo es unerlässlich ist, die Logik zu verstehen, auf der die mithilfe eines KI-Systems getroffenen Entscheidungen beruhen.[82]

Interpretierbarkeit bezieht sich auf die Fähigkeit zu verstehen, wie ein KI-System Vorhersagen macht oder Entscheidungen trifft. Es geht darum, die interne Funktionsweise, die Logik und die Entscheidungsprozesse von KI-Systemen für menschliche Nutzerinnen und Nutzer, einschliesslich Entwicklern, Interessenvertretern und Endnutzern, sowie für Betroffene verständlich und zugänglich zu machen.[83]

Unter dem Gesichtspunkt der Erklärbarkeit und Interpretierbarkeit stellen sich bei KI-Systemen besondere Herausforderungen, die je nach Art des KI-Systems mehr oder weniger stark zum Tragen kommen:[84] Ein regelbasierter Algorithmus bedeutet, dass Menschen Arbeitsschritte programmieren, mit denen die Maschine ein Ergebnis generiert. Dabei führt ein *Input* immer zum gleichen *Output*. Das Ergebnis lässt sich aus den einzelnen Schritten erklären. Diese Algorithmen können jedoch auch eine gewisse Komplexität aufweisen, was die Begründung der auf ihnen basierenden Entscheidungen erschweren kann.

Ganz anders verhält es sich mit datenbasierten Algorithmen des maschinellen Lernens («machine learning»). Diese Algorithmen analysieren riesige Datensätze, um Korrelationen zwischen den Daten zu finden und so selbst eine verallgemeinerbare Regel zu definieren. Solche Algorithmen arbeiten mit Wahrscheinlichkeiten. Das bedeutet, dass die gleiche Eingabe nicht immer zum gleichen Ergebnis führt. Ausserdem ist das Zustandekommen der Ergebnisse nicht immer nachvollziehbar. Da die Ergebnisse in diesem Fall nicht auf kausalen Zusammenhängen, sondern auf Korrelationen zwischen verschiedenen Daten beruhen, kann es im Einzelfall schwieriger sein, die für das Ergebnis des Systems massgeblichen Sachdaten und Rechtsgrundlagen zu erklären. Die Erklärbarkeit und Interpretierbarkeit solcher Systeme ist daher nur sehr schwer zu gewährleisten (Black-Box-Effekt).

- Der zweite in Artikel 8 der Konvention angesprochene Aspekt betrifft die Notwendigkeit von Kontrollmechanismen zur Überwachung, Bewertung und Steuerung der Tätig-

---

[82] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 60.

[83] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 61.

[84] Zur Begründungspflicht der Behörden vgl. NADJA BRAUN BINDER/MATTHIAS SPIELKAMP/CATHERINE EGLI et al., Einsatz Künstlicher Intelligenz (Fn. 39), 37; NADJA BRAUN BINDER/NINA LAUKENMANN/LILIANE OBRECHT, KI in der Verwaltung (Fn. 28), 10 N 23; NADJA BRAUN BINDER/LILIANE OBRECHT, Die Begründung von Verfügungen beim Einsatz algorithmischer Systeme, SJZ 2024, 707 ff.

> keiten innerhalb des Lebenszyklus von KI-Systemen. Gemäss dem erläuternden Bericht kann die Kontrolle aus einer Vielzahl von Rahmenwerken und Prozessen bestehen. Genannt werden beispielsweise Rahmen für das Risiko- und Folgenmanagement, ethische Leitlinien, Zertifizierungsverfahren, Instrumente zur Erkennung und Abschwächung von Verzerrungen, die Rolle der zuständigen Behörde wie z. B. die sektoriellen Aufsichtsbehörden, Datenschutzbehörden usw.[85] In diesem Zusammenhang soll die Verpflichtung zur Transparenz auch sicherstellen, dass KI-Systeme so konzipiert, entwickelt und eingesetzt werden, dass menschliche Kontrollmechanismen während des gesamten Lebenszyklus von KI-Systemen wirksam eingesetzt werden können.[86]

Die Bestimmung bezieht sich auch auf die Transparenz und Aufsicht bei der Identifizierung von KI-generierten Inhalten. Ziel ist es, das Risiko der Täuschung zu verringern und die Unterscheidung zwischen authentischen, von Menschen erzeugten Inhalten und KI-generierten Inhalten zu ermöglichen. Solche Massnahmen könnten Techniken wie Kennzeichnungen oder Wasserzeichen umfassen, bei denen in der Regel eine erkennbare Signatur im KI-generierten Ergebnis eingebettet wird, unter Vorbehalt der Verfügbarkeit dieser Technologien und ihrer nachgewiesenen Wirksamkeit, des allgemein anerkannten Stands der Technik und der Besonderheiten der verschiedenen Inhaltstypen.[87]

Im schweizerischen Recht kann Transparenz im öffentlichen Sektor und im staatlichen Handeln als Element der Rechtsstaatlichkeit verstanden werden (Art. 5 BV);[88] eine weitere relevante verfassungsrechtliche Grundlage ist der Informationsauftrag des Bundesrates (Art. 180 Abs. 2 BV).[89]

Auf Gesetzesstufe soll das BGÖ die Kontrolle der Verwaltung durch die Bevölkerung verbessern, indem allen interessierten Personen Einsicht in die Dokumente der Bundesbehörden gewährt wird. Voraussetzung ist, dass die Dokumente vorhanden sind und der Zugang nicht aufgrund der im BGÖ vorgesehenen Ausnahmen ausgeschlossen ist. Die aktive Informationspflicht des Bundesrats ist hingegen insbesondere in Artikel 10 RVOG geregelt.

Im Übrigen ist zu beachten, dass die Transparenz staatlichen Handelns auch durch das Legalitätsprinzip gewährleistet werden kann (vgl. die Ausführungen zu diesem Prinzip in Ziff. 4.2.3.1.2).

---

[85] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 63.

[86] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 65.

[87] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 59.

[88] ROLF H. WEBER/SIMON HENSELER, Regulierung von Algorithmen in der EU und in der Schweiz, Zeitschrift für Europarecht 2020, 28 ff., 36.

[89] DOMINIQUE HÄNNI, Vers un principe d'intégrité de l'administration publique – La prévention de la corruption en droit administratif, Genf/Zürich/Basel 2019, 305 N 794.

Das Öffentlichkeitsprinzip wird in Bezug auf staatliches Handeln in einzelnen Bereichen teilweise konkretisiert. So müssen die Gerichte den betroffenen Parteien Akteneinsicht gewähren, damit diese ihr rechtliches Gehör wahrnehmen können (Art. 29 Abs. 2 BV). Zudem müssen die Behörden ihre Entscheide begründen. Für die Ausführungen zu diesen Aspekten und zu den Anforderungen der KI-Konvention in Bezug auf Rechtsmittel wird auf die nachfolgenden Ausführungen zu den Artikeln 14 und 15 der KI-Konvention verwiesen (vgl. Ziff. 4.3.3).

Auch das DSG konkretisiert das Öffentlichkeitsprinzips und betont die Transparenz bei der Bearbeitung von Personendaten. Nach Artikel 6 Absatz 3 DSG dürfen Personendaten nur für einen bestimmten und für die betroffene Person erkennbaren Zweck beschafft werden; sie dürfen nur so bearbeitet werden, dass es mit diesem Zweck vereinbar ist. Nach Artikel 12 DSG sind sowohl die Verantwortlichen als auch die Auftragsbearbeiter zum Führen eines Verzeichnisses über ihre Bearbeitungstätigkeiten verpflichtet (für weitere Einzelheiten siehe Ziff. 4.3.2.5).

Zudem sieht die Informationspflicht nach Artikel 19 DSG vor, dass die betroffene Person grundsätzlich angemessen über die Beschaffung von Personendaten informiert wird. Für automatisierte, von KI-Systemen getroffene Einzelentscheidungen sieht Artikel 21 Absatz 1 DSG vor, dass der Verantwortliche, d. h. der Betreiber des Systems, die betroffene Person über jede Entscheidung informieren muss, die ausschliesslich auf einer automatisierten Bearbeitung beruht und die für sie mit einer Rechtsfolge verbunden ist oder sie erheblich beeinträchtigt. Artikel 21 Absatz 4 DSG sieht insbesondere vor, dass eine automatisierte Einzelentscheidung eines Bundesorgans als solche zu kennzeichnen ist. Die betroffene Person hat zudem grundsätzlich das Recht, ihren Standpunkt darzulegen und zu verlangen, dass die automatisierte Einzelentscheidung durch eine natürliche Person überprüft wird. Artikel 25 Absatz 2 Buchstabe f DSG sieht ein entsprechendes Auskunftsrecht vor. Danach ist der betroffenen Person zumindest gegebenenfalls Auskunft darüber zu erteilen, ob eine automatisierte Einzelentscheidung vorliegt und auf welcher Logik diese Entscheidung beruht. Dies gilt allerdings nur für ausschliesslich automatisierte Entscheidungen und nicht, wenn KI-Systeme als – wenn auch sehr wesentliche – Entscheidungshilfen eingesetzt werden.[90] Schliesslich sieht Artikel 56 DSG vor, dass der EDÖB ein Register der Bearbeitungstätigkeiten der Bundesorgane führt und dass dieses Register veröffentlicht wird.

[90] Der EuGH scheint den Begriff der automatisierten Einzelentscheidung weiter auszulegen. So hat er kürzlich entschieden, dass bereits das «Scoring» eines Unternehmens, das einen Kreditwürdigkeitsprüfung durchführt, eine solche Entscheidung darstellt, wenn sich der Empfänger bei seiner Entscheidung über einen Kreditantrag der betroffenen Person massgeblich auf dieses Ergebnis stützt (EuGH, Rechtssache C-634/21 vom 7. Dezember 2023, SCHUFA Holding [Scoring]). Die Auswirkungen dieser Rechtsprechung in der Schweiz sollten weiterverfolgt werden. In jedem Fall werden die Besonderheiten des Schweizer Rechts zu berücksichtigen sein.

Überdies sehen einige technische Regelwerke punktuell Transparenzanforderungen für die Verwendung von Algorithmen vor.[91]

Im privaten Sektor sollten die Grundsätze der Transparenz und der Aufsicht auch im Verhältnis zwischen Privaten gelten, wo eine direkte oder indirekte horizontale Wirkung der Grundrechte besteht oder in Zukunft erkannt wird (vgl. Ziff. 4.2.3.1). Das DSG gilt diesbezüglich auch für Private und bietet Schutz in Bezug auf die Transparenz der Bearbeitung (vgl. insbesondere Art. 6 Abs. 3, Art. 12, 19, 21 und Art. 25 Abs. 2 Bst. f DSG). Zur allfälligen Pflicht, im Streitfall vor Gericht die Funktionsweise eines in einer Rechtsbeziehung eingesetzten KI-Systems zu erklären, siehe Ziff. 6.3.1.

Zusammenfassend lässt sich feststellen, dass das Prinzip der Transparenz und der Aufsicht in der schweizerischen Rechtsordnung bereits punktuell verwirklicht ist. Die Analyse kommt jedoch zum Schluss, dass der geltende Rechtsrahmen zumindest für den öffentlichen Sektor nicht ausreicht, um die Anforderungen von Artikel 8 der Konvention zu erfüllen. Auf Ebene der Verwaltung besteht insbesondere die Möglichkeit, ein öffentliches Register aller von den öffentlichen Stellen eingesetzten KI-Systeme mit Meldepflicht einzurichten. Um dem Black-Box-Effekt entgegenzuwirken, könnten Anforderungen wie eine detaillierte Dokumentation sowie regelmässige Audits und Tests der KI-Systeme vorgesehen werden. Ziel ist es, die Kontrolle über die Systeme zu behalten, indem sichergestellt wird, dass die Funktionsweise der Systeme nachvollziehbar bleibt oder zumindest, dass die bei der generativen KI inhärenten Halluzinationen[92] erkannt und korrigiert werden können.

Sowohl im öffentlichen als auch im privaten Sektor stellt sich die Frage nach einer allfälligen Ausdehnung der Informationspflichten bzw. der zu erteilenden Auskünfte bei der Ausübung des Auskunftsrechts nach DSG auf teilautomatisierte Entscheide.

Hinsichtlich der Kennzeichnung von KI-generierten Inhalten enthält das schweizerische Recht grundsätzlich keine Kennzeichnungspflicht. Die Einführung einer solchen Massnahme könnte angezeigt sein, um KI-generierte Inhalte erkennen zu können.

Dabei ist zu beachten, dass die getroffenen Massnahmen den entgegenstehenden Interessen Rechnung tragen müssen, die mit dem Interesse an der Offenlegung bestimmter Informationen kollidieren können, wie z. B. der Schutz von Daten Dritter, Geschäftsgeheimnisse oder Interessen der Strafverfolgung.

---

[91] Vgl. Artikel 24 Absatz 1 der Verordnung des EDI über den Strahlenschutz bei medizinischen Teilchenbeschleunigeranlagen (SR **814.501.513**): «Die Lieferantin oder der Lieferant des Bestrahlungsplanungssystems hat in der technischen Beschreibung genaue Angaben über die zur Berechnung der Dosisverteilungen verwendeten Algorithmen zu machen […]».

[92] Halluzinationen äussern sich in Antworten, die in keiner Weise der Realität entsprechen, vgl. Merkblätter vom 26. April 2024 zur Sensibilisierung betreffend grossen KI-Sprachmodellen (large language models, LLM) in der Bundesverwaltung, 2, abrufbar unter: www.cnai.swiss > Dienstleistungen > Weitere Dienstleistungen > Merkblätter zu KI in der Bundesverwaltung (abgerufen am 26. August 2024).

4.3.2.4 Artikel 9 – Rechenschaftspflicht und Verantwortung

Artikel 9 verpflichtet die Staaten zur Einführung oder Beibehaltung von Massnahmen zur Sicherstellung der Rechenschaftspflicht und der Verantwortung für nachteilige Auswirkungen, die sich aus Tätigkeiten im Lebenszyklus von KI-Systemen ergeben.

Gemäss dem erläuternden Bericht zur Konvention bedeutet diese Verpflichtung, dass Personen, Organisationen oder Stellen, die für Tätigkeiten während des Lebenszyklus von KI-Systemen verantwortlich sind, für nachteilige Auswirkungen dieser Tätigkeiten auf die Menschenrechte, die Demokratie und die Rechtsstaatlichkeit zur Rechenschaft gezogen werden müssen. Die Bestimmung fordert die Schaffung von Rahmenbedingungen und Mechanismen zur Durchsetzung dieser Verpflichtung.[93]

Artikel 9 fordert darüber hinaus die Schaffung klarer Verantwortungslinien, um Handlungen und Entscheidungen bis zu bestimmten Personen oder Institutionen zurückverfolgen zu können. Es muss sichergestellt werden, dass nachteilige Auswirkungen ermittelt und die Verantwortlichkeiten angemessen zugewiesen werden können.[94] Dies kann neue Rahmenbedingungen und Mechanismen umfassen, aber auch gerichtliche und administrative Massnahmen, zivil-, straf- und andere Haftungsregelungen sowie im öffentlichen Sektor Verwaltungs- und andere Verfahren, die eine Anfechtung von Entscheidungen ermöglichen, oder spezifische Verantwortlichkeiten und Pflichten für Betreiber, die an den Tätigkeiten im Lebenszyklus von KI-Systemen beteiligt sind.[95]

Diese Verpflichtung ist untrennbar mit dem Grundsatz der Transparenz und Aufsicht (Art. 8) verbunden, da über die Transparenz indirekt erklärt werden kann, wie KI-Systeme funktionieren und wie sie Entscheidungen generieren. Sie steht auch im Zusammenhang mit der Verpflichtung, einen Rahmen für das Risiko- und Folgenmanagement (Art. 16) zur proaktiven Vermeidung und Minderung von Risiken und Folgen zu schaffen.[96]

Für den privaten Sektor wird das in Artikel 9 umschriebene Verantwortlichkeitsprinzip zwischen Privaten dort umzusetzen sein, wo eine direkte oder indirekte horizontale Wirkung der Grundrechte besteht oder in Zukunft erkannt wird (vgl. Ziff. 4.2.3.1). Für Überlegungen zum Zivil- und Strafrecht, insbesondere zur Haftung, bzw. Verantwortlichkeit, wird auf die nachfolgenden Ausführungen verwiesen (vgl. Ziff. 6.3 und 6.6).

Im schweizerischen Recht sind die Haftungsregeln so formuliert, dass sie unabhängig davon gelten, ob eine Verletzung durch ein KI-System verursacht wurde oder nicht. Das DSG

[93] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 66.

[94] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 68.

[95] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 66.

[96] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 69 f.

sieht zudem, unabhängig von KI, Massnahmen zur Rechenschaftspflicht («Accountability») vor (vgl. Ziff. 4.3.2.6).

Im öffentlichen Sektor gelten die Staatshaftungsnormen und die Vorschriften, die den Rahmen für die Tätigkeit seiner Organe bilden, unabhängig davon, ob bei dieser Tätigkeit ein KI-System eingesetzt wird oder nicht. Es muss jedoch sichergestellt werden, dass der aus Regeln, Rechtsnormen und anderen Mechanismen bestehende Rahmen eine wirkungsvolle Zuweisung der Verantwortungen erlaubt. Die Analyse zeigt, dass die geplanten Massnahmen zur Umsetzung von Artikel 8 (Transparenz und Aufsicht), Artikel 14 (Rechtsmittel), Artikel 15 (Verfahrensgarantien) und 16 (Rahmen für das Risiko- und Folgenmanagement), für welche Handlungsbedarf festgestellt wurde, dazu beitragen werden, die Wirksamkeit der bestehenden Normen zu erhöhen.

Bezüglich der Überlegungen zur zivilrechtlichen Haftung und strafrechtlichen Verantwortlichkeit wird auf die nachfolgenden Ausführungen verwiesen (vgl. Ziff. 6.3 und 6.6).

4.3.2.5 Artikel 10 – Gleichstellung und Nichtdiskriminierung

Artikel 10 Absatz 1 sieht vor, dass jede Vertragspartei Massnahmen ergreift oder beibehält, um die Achtung der Gleichheit, einschliesslich der Gleichstellung der Geschlechter, und der Nichtdiskriminierung bei Tätigkeiten innerhalb des Lebenszyklus von KI-Systemen in Übereinstimmung mit dem anwendbaren Völkerrecht und dem innerstaatlichen Recht zu gewährleisten. Diese Formulierung bezieht sich insbesondere auf alle für die Vertragsparteien relevanten völkerrechtlichen und innerstaatlichen Rechtsinstrumente. Diese bieten zusammen eine solide Rechtsgrundlage und Orientierungshilfen, die es jedem Staat ermöglichen, Massnahmen zur Gewährleistung der Gleichstellung aller und zur Nichtdiskriminierung bei Tätigkeiten innerhalb des Lebenszyklus von KI-Systemen zu ergreifen.[97]

Nach Absatz 2 verpflichtet sich jede Vertragspartei dazu, Massnahmen zur Beseitigung von Ungleichheiten zu ergreifen oder beizubehalten, um im Einklang mit ihren innerstaatlichen und völkerrechtlichen Verpflichtungen auf dem Gebiet der Menschenrechte unparteiische, faire und gleichberechtigte Ergebnisse der Tätigkeit innerhalb des Lebenszyklus von KI-Systemen zu erzielen. Gemäss dem erläuternden Bericht soll sich der hier geforderte Ansatz nicht auf die Forderung beschränken, dass eine Person «ohne objektive und angemessene Rechtfertigung» aufgrund eines oder mehrerer ihrer geschützten Merkmale nicht weniger günstig behandelt werden darf. Vielmehr verpflichten sich die Vertragsparteien zur Einführung neuer oder zur Beibehaltung bestehender Massnahmen zur Überwindung struktureller und historisch bedingter Ungleichheiten.[98]

[97] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 71.

[98] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 77.

Der Einsatz von KI stellt eine grosse Herausforderung für die Wahrung des Grundsatzes der Gleichstellung und Nichtdiskriminierung dar.[99] Da Algorithmen lediglich das Ergebnis von Berechnungen darstellen, die von Menschen festgelegt werden und auf Daten basieren, die von Menschen, Maschinen oder einer Kombination aus beiden gesammelt wurden, können sie menschliche Verzerrungen widerspiegeln und verarbeiten, die in ihre Programmierung und Datenverarbeitung eingeflossen sind.[100] Diese meist unbeabsichtigten Verzerrungen[101] können sich aus den verwendeten Daten (mangelnde Repräsentativität, veraltete Daten, unzureichende Datenverarbeitung), aus dem Algorithmus selbst (welche Parameter werden im Modell berücksichtigt, welche nicht[102]) oder aus der Art und Weise seiner Anwendung (welche Bedürfnisse soll das System erfüllen und wie wird es in der Praxis eingesetzt) ergeben.[103] Diese Risiken können durch Rückkopplungsschleifen und mangelnde Transparenz der Prozesse noch verstärkt werden[104] (zur Frage der Transparenz vgl. Ziff. 4.3.2.3).

Das Arbeitsrecht, insbesondere die Rekrutierungsphase, wird diesbezüglich häufig als sensibler Bereich benannt. Die in diesem Bereich verwendeten KI-Systeme sind in der Tat sehr vielfältig und reichen von der «einfachen» Auswahl von Lebensläufen bis hin zu Programmen, welche die Mimik eines Bewerbers während eines Videointerviews analysieren oder die berufliche Leistung auf der Grundlage eines Lebenslaufs vorhersagen.[105] Ein bekanntes Beispiel in diesem Zusammenhang ist Amazon. Im Jahr 2018 berichtete die Nachrichtenagentur Reuters, dass das Unternehmen eine auf maschinellem Lernen beruhende Bewerbungssoftware entwickelt habe, um die besten Lebensläufe ausfindig zu machen. Das Programm benachteiligte jedoch systematisch die Lebensläufe von Frauen, da es die geschlechtsspezifischen Unterschiede bei den Einstellungen in den letzten zehn Jahren widerspiegelte. Auch der Verzicht auf das Ankreuzen des Geschlechts änderte nichts daran, da das Programm aus anderen Abgleichen auf das Geschlecht der Person schliessen konnte.[106]

Der Einsatz von KI kann in allen Bereichen zu Ungleichheiten und Diskriminierungen führen. Beispiele finden sich insbesondere in folgenden Bereichen: Migration; Zugang zu Gütern und

---

99 Bericht «Herausforderungen der künstlichen Intelligenz» (Fn. 1), 38 f.

100 Europarat, Commission pour l'égalité de genre (GEC) und Comité directeur sur l'anti-discrimination, la diversité et l'inclusion (CDADI), Etude sur l'impact des systèmes d'intelligence artificielle, leur potentiel de promotion de l'égalité, y compris l'égalité de genre, et les risques qu'ils peuvent entraîner en matière de non-discrimination, erstellt durch IVANA BARTOLETTI/RAPHAËLE XENIDIS, Strassburg 2023, 17, abrufbar unter: https://rm.coe.int/prems-107623-fra-2530-etude-sur-l-impact-de-ai-a5-web/1680ac99e2 (abgerufen am 15. August 2024).

101 GRÉGOIRE LOISEAU, Intelligence artificielle et droit des personnes, in: Alexandra Bensamoun/Grégoire Loiseau (Hrsg.), Droit et intelligence artificielle, Paris 2022, 35 ff., 49.

102 Die Proxy-Diskriminierung ist im Zusammenhang mit KI besonders relevant. Proxys (oder Stellvertreter) sind scheinbar harmlose Kriterien, die jedoch stark mit sensiblen Kriterien korrelieren können. Beispielsweise bedeutet die Angabe «30 Jahre Berufserfahrung», dass die betreffende Person mindestens 45 Jahre alt sein muss, und die Postleitzahl lässt Annahmen über den sozioökonomischen Status oder die Herkunft einer Person zu, vgl. NADJA BRAUN BINDER/NINA LAUKENMANN/LILIANE OBRECHT, KI in der Verwaltung (Fn. 28), 12 N 30; AlgorithmWatch/CH, Positionspapier: Schutz vor algorithmischer Diskriminierung, September 2023, 6, abrufbar unter: https://algorithmwatch.ch > Positionen > Diskriminierende Algorithmen: So gelingt der Schutz (abgerufen am 15. August 2024).

103 Erläuternder Bericht zur KI-Konvention (Fn. 20), N 75 und 76.

104 Bericht «Herausforderungen der künstlichen Intelligenz» (Fn. 1), 38 f. Vgl. für eine umfassende Übersicht der Problematik: FLORENT THOUVENIN/STEPHANIE VOLZ/SORAYA WEINER/CHRISTOPH HEITZ, Diskriminierung beim Einsatz von Künstlicher Intelligenz (KI), Jusletter IT 4. Juli 2024.

105 AVI ASHER-SCHAPIRO, AI is taking over job hiring, but can it be racist?, Reuters 2021, abrufbar unter: https://www.reuters.com/article/business/healthcare-pharmaceuticals/ai-is-taking-over-job-hiring-but-can-it-be-racist-idUSL5N2NF5ZC/ (abgerufen am 15. August 2024).

106 JEFFREY DASTIN, Insight – Amazon scraps secret AI recruiting tool that showed bias against women, Reuters 2018, in englischer Sprache abrufbar unter: https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/ (abgerufen am 15 August 2024).

Dienstleistungen, Banken und Versicherungen; Zugang zu Sozialleistungen[107]; Risikobewertung in den Bereichen Sicherheit, Verbrechensprävention, Aufrechterhaltung der öffentlichen Ordnung und Justizwesen; Zugang zu öffentlichen und administrativen Dienstleistungen; Bildung; Medien und Suchmaschinen; Gesundheitswesen.[108] Insbesondere im letztgenannten Bereich zeigen Beispiele, dass Systeme manchmal auf der Grundlage von nicht repräsentativen Daten trainiert werden. Dies kann zu Fehldiagnosen und zu Diskriminierung von Frauen[109], Minderheiten und Menschen *of Color*[110] oder älteren Menschen[111] führen.

Es ist wichtig anzumerken, dass ein gut konzipiertes KI-System auch dazu beitragen kann, diskriminierten oder benachteiligten Menschen zu helfen. KI-Systeme können beispielsweise eingesetzt werden, um den Zugang zu Informationen oder zu vorhandenen Gütern und Dienstleistungen zu verbessern. Der Einsatz maschineller Übersetzungssysteme für Regional- oder Minderheitensprachen könnte den Zugang zur Grundversorgung verbessern. KI könnte auch zur Förderung der Gleichstellung im Polizeisektor eingesetzt werden, z. B. zur Prävention geschlechtsspezifischer Gewalt (siehe die Software VioGen in Spanien). Im Gesundheitswesen könnte KI eingesetzt werden, um den Zugang zur Gesundheitsversorgung in benachteiligten Gebieten und die Diagnostik für traditionell unterrepräsentierte Gruppen zu verbessern.[112] KI-Systeme können auch menschliche Fehler (Rückkopplungsfehler, unvollständige Informationen) korrigieren. In solchen Fällen kann das KI-System zur Aufdeckung von Ungleichbehandlungen beitragen, um diese zu korrigieren und ihnen entgegenzuwirken.[113]

Im schweizerische Recht gibt es keine KI-spezifischen Bestimmungen zur Rechtsgleichheit und zur Nichtdiskriminierung. Es ist daher zu prüfen, ob der bestehende Rechtsrahmen, der auch für den Einsatz von KI gilt, die Herausforderungen dieser Technologie ausreichend abdeckt.

Die internationalen Verpflichtungen der Schweiz in den Bereichen Gleichstellung und Nichtdiskriminierung ergeben sich insbesondere aus den beiden UNO-Pakten (Art. 2 Abs. 2 UNO-Pakt I und Art. 2 Abs. 1 UNO-Pakt II, Art. 3 beider UNO-Pakte und Art. 26 UNO-Pakt II) sowie aus der EMRK (Art. 14). Diese Bestimmungen haben jedoch aus verschiedenen Gründen

---

[107] Hier ist der Skandal bezüglich der Familienzulagen in der Niederlanden zu nennen. Bei der Entwicklung des algorithmischen Systems zur Feststellung, ob Anträge für Familienzulagen als fehlerhaft und potenziell betrügerisch einzustufen sind, wurden Kriterien aus dem Bereich des *Racial Profiling* einbezogen. Dies führte dazu, dass Zehntausende von Eltern und Erziehungsberechtigten, zumeist aus einkommensschwachen Familien, von den niederländischen Steuerbehörden fälschlicherweise des Betrugs bezichtigt wurden, vgl. https://www.amnesty.org/fr/latest/news/2021/10/xenophobic-machines-dutch-child-benefit-scandal/ (abgerufen am 15. August 2024).

[108] Conseil de l'Europe, Etude sur l'impact des systèmes d'intelligence artificielle, leur potentiel de promotion de l'égalité, y compris l'égalité de genre, et les risques qu'ils peuvent entraîner en matière de non-discrimination (Fn. 100), 23 ff. (abgerufen am 15. August 2024); Algorithm-Watch/CH, Positionspapier: Schutz vor algorithmischer Diskriminierung, September 2023 (Fn. 102) (abgerufen am 15. August 2024).

[109] ANNE-SOPHIE MORAND, A WEIRD AI system?, Jusletter 20. September 2021, 9 N 21 ff.

[110] ANNE-SOPHIE MORAND, A WEIRD AI system? (Fn. 109), 7 N 16 ff.

[111] Vgl. JUSTYNA STYPINSKA, AI ageism: a critical roadmap for studying age discrimination and exclusion in digitalized societies, AI & SOCIETY 2023, 665 ff.

[112] Europarat, Etude sur l'impact des systèmes d'intelligence artificielle, leur potentiel de promotion de l'égalité, y compris l'égalité de genre, et les risques qu'ils peuvent entraîner en matière de non-discrimination (Fn. 100), 89.

[113] Universität Zürich, Digital Society Initiative, Positionspapier, Ein Rechtsrahmen für Künstliche Intelligenz, November 2021, 4, abrufbar unter: https://www.dsi.uzh.ch > Forschung > Strategy Lab > 1. DSI Strategy Lab (abgerufen am 19. August 2024); FLORENT THOUVENIN/STEPHANIE VOLZ/SORAYA WEINER/CHRISTOPH HEITZ, Diskriminierung (Fn. 104), 17.

keine eigenständige Bedeutung und können in der Schweiz nur zusammen mit anderen Bestimmungen der Texte geltend gemacht werden.[114] Die Schweiz hat zudem mehrere UNO-Übereinkommen ratifiziert, welche die Rechtsgleichheit in bestimmten Bereichen garantieren, wie das Übereinkommen zur Beseitigung jeder Form von Rassendiskriminierung[115] und das Übereinkommen zur Beseitigung jeder Form von Diskriminierung der Frau[116]. Die meisten Bestimmungen dieser Texte sind vor dem Bundesgericht einklagbar.[117] Zu erwähnen ist auch das Übereinkommen über die Rechte von Menschen mit Behinderungen[118], dessen Artikel 5 direkt einklagbar ist.[119]

Auf Verfassungsstufe ist der Grundsatz der Rechtsgleichheit in Artikel 8 Absatz 1 und das Diskriminierungsverbot in Artikel 8 Absatz 2 verankert. Absatz 3 dieser Bestimmung garantiert die rechtliche und tatsächliche Gleichstellung von Mann und Frau, und Absatz 4 enthält eine Verpflichtung zur Beseitigung von Benachteiligungen von Menschen mit Behinderungen. Der Grundsatz der Rechtsgleichheit gilt für alle Beziehungen, die der Staat eingehen kann. Er steht in engem Zusammenhang mit dem Gebot der Gerechtigkeit und Billigkeit. Mit Ausnahme der Lohngleichheit von Mann und Frau in Absatz 3 haben das Gebot der Rechtsgleichheit und das Diskriminierungsverbot keine unmittelbare Wirkung gegenüber Privaten. Allerdings verpflichtet Artikel 35 Absatz 1 BV den Staat, dafür zu sorgen, dass die Grundrechte in der gesamten Rechtsordnung zur Geltung kommen und, soweit sie sich dazu eignen, auch unter Privaten wirksam werden (Art. 35 Abs. 3 BV, vgl. Ziff. 4.3.1.1). Andere Verfassungsbestimmungen wiederholen das Gebot der Rechtsgleichheit in spezifischen Bereichen. So garantiert etwa Artikel 27 BV den freien Wettbewerb, oder Artikel 15 BV verbietet es dem Staat, für oder gegen eine Religion oder Weltanschauung Partei zu ergreifen (Grundsatz der religiösen Neutralität des Staates).[120]

Der schweizerische Gesetzgeber hat den Grundsatz der Rechtsgleichheit in mehreren Gesetzestexten konkretisiert. Diese gelten, wie bereits erwähnt, auch für KI-Systeme:

- GIG: Das GIG bezweckt die rechtliche und tatsächliche Gleichstellung von Frau und Mann in Arbeitsverhältnissen im öffentlichen und privaten Sektor. Konkret schreibt es

---

[114] GIORGIO MALINVERNI et. al., Droit constitutionnel suisse – Volume II (Fn. 56), N 1103 ff.; für Art. 14 EMRK vgl. BGE **123** II 472, E. 4c.

[115] SR **0.104**.

[116] SR **0.108**.

[117] Zur ersten dieser beiden Übereinkommen siehe BGE **123** IV 202, E. 2; BGE **123** II 472, 479, E. 4d; für das zweite, BGE **125** I 21, E. 4a; BGE **145** I 308, E. 3.1.

[118] SR **0.109**.

[119] Urteil des BGer 8C_633/2021, E. 4.2.

[120] GIORGIO MALINVERNI et al., Droit constitutionnel suisse – Volume II (Fn. 56), N 1100, 1113 f.

vor, dass Arbeitnehmerinnen und Arbeitnehmer aufgrund ihres Geschlechts weder direkt noch indirekt benachteiligt werden dürfen, namentlich nicht unter Berufung auf den Zivilstand, auf die familiäre Situation oder, bei Arbeitnehmerinnen, auf eine Schwangerschaft (Art. 3 Abs. 1 und 2). Dieses Diskriminierungsverbot gilt insbesondere für die Einstellung, Aufgabenzuteilung, Entlöhnung und Entlassung. Weil Diskriminierungen schwierig zu beweisen sind, führt das Gesetz auch eine Beweislasterleichterung ein (Art. 6). Das Gesetz sieht auch ein besonderes Klagerecht für Organisationen vor, wenn eine erhebliche Anzahl von Arbeitsverhältnissen betroffen ist (Art. 7).

- BehiG und dazugehörende Ausführungsverordnungen: Dieses Gesetz soll die Benachteiligung von Menschen mit Behinderungen verhindern, verringern und beseitigen. Es gilt insbesondere für verschiedene Arten von öffentlich zugänglichen Bauten oder für den Zugang zu Bildung. Es richtet sich in erster Linie an den Staat, unter Vorbehalt der in Artikel 3 genannten Bauten.

- Zu erwähnen ist auch Artikel 261bis StGB, der im Rahmen der Ratifizierung des Übereinkommens zur Beseitigung jeder Form von Rassendiskriminierung[121] eingeführt wurde, um diese völkerrechtlichen Verpflichtungen in nationales Recht umzusetzen. Dieser Artikel bezweckt in erster Linie den Schutz der Menschenwürde derjenigen Personen, gegen die sich rassistische Herabsetzung richtet. In zweiter Linie schützt er den öffentlichen Frieden, der durch Handlungen bedroht ist, die geeignet sind, Menschen gegeneinander aufzubringen.[122] Mit Ausnahme von Absatz 5 (Verweigerung einer der Allgemeinheit angebotenen Leistung aus Gründen der Rasse, Ethnie oder Religion) muss die strafrechtlich relevante Handlung in der Öffentlichkeit stattfinden. Der Begriff der Öffentlichkeit wird vom BGer relativ weit gefasst: Darunter fallen alle Äusserungen oder Verhaltensweisen, die nicht dem privaten Bereich zugeordnet werden können, also z. B. im Familien- oder Freundeskreis oder in einem Umfeld persönlicher Beziehungen oder besonderen Vertrauens.[123]

- Schliesslich ist auf das Bundesgesetz über genetische Untersuchungen beim Menschen[124] hinzuweisen, das in Artikel 4 festhält, dass niemand wegen seines Erbguts

[121] SR **0.104**.

[122] CR CP II-MIRIAM MAZOU, Art. 261bis N 3 und die zitierten Referenzen.

[123] BGE **130** IV 111, E. 5.2.1

[124] SR **810.12**.

diskriminiert werden darf, oder auf Artikel 2 des Freizügigkeitsabkommens[125], der die Diskriminierung aufgrund der Staatsangehörigkeit verbietet.

In den nicht von diesen Texten erfassten Bereichen können sich diskriminierte Personen je nach Umständen direkt auf die Verfassung oder auf allgemeinere Schutzbestimmungen wie Artikel 6 des ArG berufen.[126] Im schweizerischen Recht hat sich der Gesetzgeber dafür entschieden, den Grundsatz der Rechtsgleichheit und Nichtdiskriminierung unter Privaten je nach Art der Diskriminierung oder in bestimmten Bereichen umzusetzen. Im Privatrecht bieten bestimmte Normen indirekten Schutz vor Diskriminierung. So kann man sich beispielsweise bei einer die Persönlichkeit tangierenden Diskriminierung auf Artikel 28 ZGB[127] und Artikel 328 OR[128] berufen, um die Unterlassung der Diskriminierung oder eine Entschädigung zu verlangen. Es besteht auch ein Schutz vor diskriminierenden Kündigungen, die in die Kategorie der missbräuchlichen Kündigungen fallen, wenn diese wegen einer Eigenschaft, die der anderen Partei kraft ihrer Persönlichkeit zusteht, oder wegen der Ausübung eines verfassungsmässigen Rechts ausgesprochen wurde (Art. 336 Abs. 1 Bst. a und b OR).

Neben den oben genannten Bestimmungen ist im Hinblick auf den Schutz natürlicher Personen auch das DSG zu erwähnen. Dieses Gesetz zielt zwar nicht primär auf die Bekämpfung von Diskriminierungen und Ungleichheiten ab, enthält aber dennoch zahlreiche Bestimmungen, die einen indirekten Schutz im privaten und öffentlichen Sektor bieten können. Die Bearbeitung von Personendaten innerhalb des Lebenszyklus eines KI-Systems muss unter Einhaltung der im Text festgelegten Regeln erfolgen. Im Hinblick auf die hier behandelte Thematik erscheinen folgende Punkte relevant:

- *Besonders schützenswerte Personendaten:* Das DSG sieht einen erhöhten Schutz vor, wenn besonders schützenswerte Personendaten bearbeitet werden, sowie bei Profiling und Profiling mit hohem Risiko (Art. 5 Bst. c, f und g), was einen Schutz vor Diskriminierung bieten kann. Besonders schützenswerte Daten sind insbesondere Daten über religiöse, weltanschauliche, politische oder gewerkschaftliche Ansichten oder

---

[125] Abkommen zwischen der Schweizerischen Eidgenossenschaft einerseits und der Europäischen Gemeinschaft und ihren Mitgliedstaaten andererseits über die Freizügigkeit vom 21. Juni 1999, SR **0.142.112.681**.

[126] Siehe insbesondere SECO, Kommentar, Art. 2 ArGV 3, 302-4, abrufbar unter: https://www.seco.admin.ch > Arbeit > Arbeitsgesetz und Verordnungen > Kommentar zum Arbeitsgesetz und seinen Verordnungen > Kommentar zur ArGV 3 (abgerufen am 19. August 2024).

[127] Vgl. zum Beispiel: FLORENT THOUVENIN/STEPHANIE VOLZ, White Paper: Diskriminierung, Juni 2024, 7, abrufbar unter: https://www.itsl.uzh.ch > Wissenstransfer > Publikationen > Positionspapiere > Nachvollziehbare Algorithmen: ein Rechtsrahmen für den Einsatz von künstlicher Intelligenz (abgerufen am 19. August 2024); ELEONOR KLEBER, La discrimination multiple – Etude de droit international, suisse et européen, Zürich 2015, 210 und die zitierten Referenzen.

[128] CR CO I-LEMPEN, Art. 328 N 1; BSK OR I-PORTMANN/RUDOLPH, Art. 328 N 45 f. und Art. 320 N 2a für diese vorvertragliche Verhandlungsphase.

Tätigkeiten, Daten über die Gesundheit, die Intimsphäre oder die Zugehörigkeit zu einer Rasse oder Ethnie, genetische Daten sowie Daten über Massnahmen der sozialen Hilfe.

- *Erfordernis einer gesetzlichen Grundlage:* Bundesorgane dürfen Personendaten grundsätzlich nur bearbeiten, wenn eine gesetzliche Grundlage dies erlaubt. Der Entscheid über den Einsatz von KI läge somit bei Vorliegen eines Gesetzes im materiellen Sinn beim Bundesrat, und bei einer formell gesetzlichen Grundlage allenfalls beim Parlament, sofern nicht bereits eine genügende gesetzliche Grundlage besteht (vgl. Ziff. 4.2.3.1). Dieses Erfordernis stellt auch eine Form des indirekten Schutzes vor Ungleichbehandlung und Diskriminierung dar, da der Staat bei der Ausarbeitung von Gesetzen an den Grundsatz der Rechtsgleichheit gebunden ist. Es sei daran erinnert, dass das geltende Recht bei der Ausarbeitung von Gesetzesentwürfen auch eine spezifische Verpflichtung des Gesetzgebers vorsieht, in bestimmten Fällen zusätzlich zur Regulierungsfolgenabschätzung (RFA) eine Analyse der Auswirkungen des Gesetzes auf die Gleichstellung von Frauen und Männern durchzuführen.

- *Grundsätze der Verhältnismässigkeit, der Richtigkeit und des Datenschutzes durch Technik und datenschutzfreundliche Voreinstellungen* (Art. 6 Abs. 2 und Art. 7 DSG)*:* Die Umsetzung dieser Grundsätze kann dazu beitragen, die Beschaffung und Bearbeitung von Daten auf ein Minimum zu beschränken, ihre Richtigkeit und Sicherheit zu gewährleisten und sie sogar zu anonymisieren, wodurch das Risiko einer Diskriminierung aufgrund dieser Daten verringert wird. Dies kann auch dazu beitragen, die Datenbearbeitung transparent zu machen und die betroffenen Personen in die Lage zu versetzen, ihre Rechte geltend zu machen und so mögliche Diskriminierungen festzustellen.

- *Pflicht zur Durchführung einer Folgenabschätzung*: Wenn die geplante Bearbeitung voraussichtlich ein hohes Risiko für die Persönlichkeit oder die Grundrechte mit sich bringt, muss der Verantwortliche eine Datenschutz-Folgenabschätzung durchführen (Art. 22 DSG). Bei der Risikobeurteilung muss er die möglichen Schäden oder Folgen für die betroffene Person berücksichtigen, die mit einer möglichen Verletzung ihrer Persönlichkeit und ihres Rechts auf informationelle Selbstbestimmung verbunden sind. Zu den möglichen Schäden gehören z. B. die Verweigerung von Krediten oder die Verweigerung einer Anstellung.[129] Diese Bestimmungen können bei der Aufdeckung von Diskriminierung helfen. Die Folgenabschätzung enthält eine Beschreibung der geplanten Bearbeitung, eine Bewertung der Risiken für die Persönlichkeit oder die Grundrechte der betroffenen Person sowie die geplanten Massnahmen zum Schutz ihrer Persönlichkeit und ihrer Grundrechte. Zu beachten ist, dass es im privaten Sektor Ausnahmen von der Pflicht zur Durchführung einer Folgenabschätzung gibt (Art. 22 Abs. 4 und 5 DSG).

[129] CR LPD-PHILIPPE GILLIÉRON, Art. 22 N 24; DAVID ROSENTHAL/SAMIRA STUDER/ALEXANDRE LOMBARD (für die Übersetzung), La nouvelle loi sur la protection des données, Jusletter 16. November 2020, 58, N 149.

- *Informationspflicht im Allgemeinen und Informationspflicht bei einer automatisierten Einzelentscheidung und Auskunftsrecht (Art. 19, 21 und 25 Abs. 2 Bst. f DSG)*: Der für die Bearbeitung Verantwortliche muss die betroffene Person über die Beschaffung ihrer Daten informieren, sofern keine Ausnahme vorliegt. Darüber hinaus muss er die betroffene Person grundsätzlich über jede Entscheidung informieren, die ausschliesslich aufgrund einer automatisierten Bearbeitung von Personendaten getroffen wird und die für die betroffene Person rechtliche Folgen nach sich zieht oder sie erheblich beeinträchtigt. Die betroffene Person hat grundsätzlich das Recht, ihren Standpunkt darzulegen und zu verlangen, dass die Entscheidung von einer natürlichen Person überprüft wird. Diese Verpflichtungen gelten nicht, wenn die Entscheidung im Zusammenhang mit dem Abschluss oder der unmittelbaren Erfüllung eines Vertrags steht und dem Antrag der betroffenen Person stattgegeben wird, oder wenn diese ausdrücklich in eine automatisierte Entscheidung eingewilligt hat. Im Rahmen des Auskunftsrechts erhält die betroffene Person die erforderlichen Informationen, damit sie ihre Rechte gemäss DSG wahrnehmen kann und die Transparenz der Datenbearbeitung gewährleistet ist. Gegebenenfalls muss sie über das Vorliegen einer automatisierten Einzelentscheidung und über die Logik, auf der diese Entscheidung beruht, informiert werden (Art. 25 Abs. 2 DSG). Diese Bestimmungen können zur Aufdeckung von Diskriminierungen führen. Das Vorgesagte gilt jedoch nur für den Fall einer ausschliesslich automatisierten Entscheidung und nicht, wenn KI-Systeme als – wenn auch sehr wesentliche – Entscheidungshilfe eingesetzt werden.[130]

- *Dokumentations- und Protokollierungspflichten*: Gemäss Artikel 12 DSG und Artikel 3, 4 und 5 DSV ist der Verantwortliche verpflichtet, ein Verzeichnis der Bearbeitungstätigkeiten zu führen und unter bestimmten Voraussetzungen zu protokollieren. Zudem muss er technische und organisatorische Massnahmen treffen, um die Vertraulichkeit, die Verfügbarkeit, die Integrität und die Nachvollziehbarkeit der Daten zu gewährleisten. Diese Massnahmen können somit eine Kontrolle der durchgeführten Bearbeitung[131] und die Aufdeckung von Unregelmässigkeiten, die zu Diskriminierungen führen könnten, ermöglichen.

- *Grundsatz der Richtigkeit:* Gemäss Artikel 6 Absatz 5 DSG muss, wer Personendaten bearbeitet, sich über deren Richtigkeit vergewissern. Es sind alle angemessenen

[130] Der EuGH scheint den Begriff der automatisierten Einzelentscheidung weiter auszulegen. So hat er kürzlich entschieden, dass bereits das «Scoring» eines Unternehmens, das einen Kreditwürdigkeitsprüfung durchführt, eine solche Entscheidung darstellt, wenn sich der Empfänger bei seiner Entscheidung über einen Kreditantrag der betroffenen Person massgeblich auf dieses Ergebnis stützt (EuGH, Rechtssache C-634/21 vom 7. Dezember 2023, SCHUFA Holding [Scoring]). Die Auswirkungen dieser Rechtsprechung in der Schweiz sollten weiterverfolgt werden. In jedem Fall werden die Besonderheiten des Schweizer Rechts zu berücksichtigen sein.

[131] YVES POULLET, L'IA un défi pour nos législations vie privée, in: Astrid Epiney/Sophia Rovelli (Hrsg.): Künstliche Intelligenz und Datenschutz / L'intelligence artificielle et la protection des données, Freiburg 2021, 1 ff., 40, im Zusammenhang mit dem in der DSGVO vorgesehenen Grundsatz der Verantwortlichkeit.

Massnahmen zu treffen, damit Daten, die im Hinblick auf den Zweck ihrer Beschaffung oder Bearbeitung unrichtig oder unvollständig sind, berichtigt, gelöscht oder vernichtet werden.

- *Grundsatz von Treu und Glauben:* Einige Autoren vertreten die Ansicht, dass eine Bearbeitung von Personendaten, die zu einer Diskriminierung führt, als Verletzung des datenschutzrechtlichen Grundsatzes von Treu und Glauben betrachtet werden könnte (Art. 6 Abs. 2 DSG). Für eine solche Auslegung spricht, dass der Zweck des DSG der Schutz der Grundrechte und der Persönlichkeit der betroffenen Personen ist (Art. 1 DSG).[132]

Ungeachtet dessen vertreten verschiedene Autorinnen und Autoren die Auffassung, dass das schweizerische Diskriminierungsverbot heute – unabhängig davon, ob KI zur Anwendung gelangt oder nicht – bisweilen lückenhaft ist.[133] Auch das Schweizerische Kompetenzzentrum für Menschenrechte (SKMR) hat seinerzeit im Rahmen einer Studie, die der Bundesrat auf das Postulat Naef 12.3543 hin in Auftrag gegeben hatte, Empfehlungen zur Verstärkung der Gesetzgebung in diesem Bereich formuliert.[134] Der Bundesrat ist jedoch den meisten dieser Empfehlungen nicht gefolgt und hat auch auf die Einführung eines privatrechtlichen Diskriminierungsverbots in Ergänzung zu den Artikeln 27 ff. ZGB und auf eine Beweislasterleichterung im Diskriminierungsfall verzichtet, und auch den Anwendungsbereich von Artikel 261bis StGB nicht erweitert.[135] Der Bundesrat hat einen bereichsspezifischen Ansatz bevorzugt.

Die Überlegungen zu algorithmischer Diskriminierung kommen meist zu dem Schluss, dass der Einsatz von KI die bereits festgestellten Schwächen noch verschärft. So wird die Erarbei-

[132] In diesem Sinn: FLORENT THOUVENIN/STEPHANIE VOLZ, White Paper: Diskriminierung (Fn. 127), 7.

[133] Insbesondere CR Cst.-VINCENT MARTENET, Art. 8, N 141 ff. und die zitierten Ref.; PREVITALI ADRIANO/MICHAEL MONTAVON, L'interdiction des discriminations, in: Oliver Diggelmann/Maya Hertig Randall/Benjamin Schindler (Hrsg.), Verfassungsrecht der Schweiz Bd. II / Droit constitutionnel suisse Vol. II, 2020, 1453 ff.; SAMANTHA BESSON, L'égalité horizontale: l'égalité de traitement entre particuliers, Freiburg 1999, 1352 ff.

[134] SCHWEIZERISCHES KOMPETENZZENTRUM FÜR MENSCHENRECHTE (SKMR), Der Zugang zur Justiz in Diskriminierungsfällen, Synthesebericht, Bern 2015, abrufbar unter: www.bj.admin.ch > Gesellschaft > Gleichstellung der Geschlechter und Schutz vor Diskriminierung > Schutz vor Diskriminierung > Frühere Berichte und Studien > Postulat Naef 12.3543 (abgerufen am 19. August 2024).

[135] Bericht des Bundesrates vom 25. Mai 2016 in Erfüllung des Postulats Naef 12.3543, Recht auf Schutz vor Diskriminierung, insbesondere Ziff. 4.2.1, 4.3.1, 4.2.3 und 4.2.6, abrufbar unter: www.bj.admin.ch > Gesellschaft > Gleichstellung der Geschlechter und Schutz vor Diskriminierung > Schutz vor Diskriminierung > Frühere Berichte und Studien > Postulat Naef 12.3543 (abgerufen am 27. August 2024).

tung eines allgemeinen Gleichbehandlungsgesetzes bzw. eines gesetzlichen Diskriminierungsverbots[136], die Einführung eines Grundsatzes im DSG, der diskriminierende Datenbearbeitungen verbietet,[137] die Einführung eines Grundsatzes der *Non-Discrimination by Design*[138] oder auch eine Umkehr bzw. Erleichterung der Beweislast gefordert.[139]

Zusammenfassend lässt sich sagen, dass die Problematik der Rechtsgleichheit und des Diskriminierungsverbots beim Einsatz von KI in einigen Aspekten durch die geltende Gesetzgebung abgedeckt wird.

Es stellt sich jedoch die Frage, ob der bestehende Rechtsrahmen gegebenenfalls ergänzt werden muss. Es ist darauf hinzuweisen, dass die im Zusammenhang mit KI geäusserte Kritik und die Forderungen hinsichtlich der Eignung des geltenden Rechts zur Lösung der anstehenden Probleme nicht neu sind. Sie entsprechen vielmehr den allgemeinen und bereits bekannten Forderungen.

Es ist jedoch zu erwarten, dass KI die bereits identifizierten Probleme verschärfen, beschleunigen und ihnen eine kollektive Dimension verleihen wird. Beim Diskriminierungsschutz sollte grundsätzlich nicht danach unterschieden werden, ob die Diskriminierung durch ein KI-System verursacht wird oder nicht. Ist der politische Wille vorhanden, wird jede allgemeine Stärkung des Diskriminierungsverbots, d. h. unabhängig von der verwendeten Technologie, auch im Bereich der KI von Vorteil sein.

Unabhängig von dieser Frage erscheint ein Eingreifen des Gesetzgebers notwendig, um die Transparenz und die Folgenabschätzung im Zusammenhang mit möglichen Diskriminierungen zu verbessern, da es nach wie vor schwierig ist, Verzerrungen im Zusammenhang mit KI-Systemen zu erkennen und nachzuweisen. Denkbar sind dabei folgende Massnahmen:

- Pflicht zur Berücksichtigung von Gleichstellungs- und Nichtdiskriminierungsaspekten bei der Risiko- und Folgenabschätzung von KI-Systemen (vgl. Ziff. 4.3.4, Art. 16 der KI- Konvention).

- Informationspflicht bei automatisierten Einzelentscheidungen: Derzeit besteht diese Pflicht nur bei Entscheidungen, die ausschliesslich auf einer automatisierten Bearbeitung von Personendaten beruhen (vgl. Art. 21 und 25 DSG). In vielen Fällen werden KI-Systeme jedoch nur als Unterstützung oder Hilfe bei der Entscheidungsfindung eingesetzt und fallen nicht unter diese Bestimmung. Eine Ausweitung der

[136] FLORENT THOUVENIN/STEPHANIE VOLZ, White Paper: Diskriminierung (Fn. 127), 6; AlgorithmWatch/CH, Papier de position: Protection contre la discrimination algorithmique (Fn. 102), 8.

[137] FLORENT THOUVENIN/STEPHANIE VOLZ, White Paper: Diskriminierung (Fn. 127), 6.

[138] FLORENT THOUVENIN/STEPHANIE VOLZ/SORAYA/WEINER/CHRISTOPH HEITZ, Diskriminierung (Fn. 104), 26 f.

[139] In diesem Sinn: FLORENT THOUVENIN/STEPHANIE VOLZ, White Paper: Diskriminierung (Fn. 127), 6 f.

Informationspflicht auf teilautomatisierte Entscheidungen könnte daher von Vorteil sein.

- Meldung von KI-Systemen, die im öffentlichen Sektor eingesetzt werden: Derzeit führt das CNAI ein Register der im öffentlichen Sektor eingesetzten KI-Systeme. Eine solche Meldung erfolgt derzeit auf freiwilliger Basis. Die Meldung könnte obligatorisch werden.

- Bekämpfung des Black-Box-Effekts: Detaillierte Dokumentation, regelmässige Audits und Tests von KI-Systemen. Hier geht es darum, die Kontrolle über die Systeme zu behalten, indem sichergestellt wird, dass man weiterhin versteht, wie sie funktionieren, oder dass man sogenannte Halluzinationen erkennen kann. Die getroffenen Massnahmen müssen jedoch auch entgegenstehende Interessen berücksichtigen, die mit dem Interesse an der Offenlegung bestimmter Informationen kollidieren können, z. B. der Schutz von Daten Dritter, Geschäftsgeheimnisse oder Interessen der Strafverfolgung.

4.3.2.6 Artikel 11 – Privatsphäre und Schutz personenbezogener Daten

Artikel 11 der Konvention sieht vor, dass jede Vertragspartei Massnahmen ergreift oder beibehält, um sicherzustellen, dass das Recht des Einzelnen auf Schutz der Privatsphäre und seiner personenbezogenen Daten gewährleistet ist, unter anderem durch geltende nationale und internationale Gesetze, Normen und Rahmenvorschriften. Der Artikel sieht auch vor, dass wirksame Garantien und Schutzmassnahmen in Übereinstimmung mit den geltenden nationalen und internationalen rechtlichen Verpflichtungen eingeführt werden müssen.

Der erläuternde Bericht stellt fest, dass der Schutz der Privatsphäre und der Schutz personenbezogener Daten ein gemeinsamer Grundsatz ist, der für die Einhaltung fast aller anderen Grundsätze des Übereinkommens erforderlich ist. Die Beschaffung und Bearbeitung von Personendaten sind im KI-Bereich allgegenwärtig, und die eingesetzten Technologien und maschinengestützten Systeme wirken sich unmittelbar auf das Leben der Menschen aus. Da KI-Systeme in erster Linie datenbasiert sind, könnten die Tätigkeiten während des gesamten Lebenszyklus dieser Systeme ohne angemessene Garantien ernsthafte Risiken mit sich bringen.[140]

Herausforderungen bestehen in allen Phasen des Lebenszyklus von KI-Systemen, insbesondere bei generativer KI. Je nachdem, ob man sich in der Konzeptions-, Trainings-, Nutzungs- oder Feedbackphase des Systems befindet, ergeben sich unterschiedliche Problematiken.[141]

---

[140] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 79.

[141] SAMUEL KLAUS, KI trifft Datenschutz – Risiken und Lösungsansätze, in: Astrid Epiney/Sophia Rovelli (Hrsg.): Künstliche Intelligenz und Datenschutz (Fn. 131), 81 ff., 83 f.; DAVID ROSENTHAL, Datenschutz und KI: Worauf in der Praxis zu achten ist, Jusletter 26. April 2022, 3; vgl. auch

Es ist zu prüfen, ob der bestehende regulatorische Rahmen im nationalen Recht, der auch auf den Bereich der KI anwendbar ist, die Risiken dieser Technologie ausreichend abdeckt.

Auf Verfassungsstufe werden das Grundrecht auf Privatleben und Datenschutz durch Artikel 10 Absatz 2 BV (persönliche Freiheit) sowie spezifischer durch Artikel 13 BV (in geringerem Masse durch Art. 7) geschützt.[142] Artikel 13 sieht einerseits vor, dass jede Person Anspruch auf Achtung ihres Privat- und Familienlebens, ihrer Wohnung, sowie ihres Brief-, Post- und Fernmeldeverkehr hat (Abs. 1), und andererseits, dass jede Person Anspruch auf Schutz vor Missbrauch ihrer persönlichen Daten hat (Abs. 2). In diesem zweiten Absatz ist das Recht auf informationelle Selbstbestimmung verankert, d. h. das Recht der betroffenen Person, selbst bestimmen zu können, ob und zu welchem Zweck Informationen über sie bearbeitet werden dürfen.[143]

Auf Gesetzesstufe wird Artikel 13 BV für natürliche Personen hauptsächlich durch das DSG und seine Ausführungsverordnungen, namentlich die DSV und die VDSC, gewährleistet. Diese Bestimmungen gelten sowohl für den privaten als auch für den öffentlichen (Bundes-) Sektor. Das Datenschutzgesetz wurde kürzlich revidiert; die Änderungen traten am 1. September 2023 in Kraft. Dieser rechtliche Rahmen wird durch zahlreiche sektorspezifische Bundesgesetze ergänzt. Hervorzuheben sind hier das Fernmeldegesetz[144], welches das Fernmeldegeheimnis konkretisiert, und die in Artikel 1 des Bundesgesetzes betreffend die Überwachung des Post- und Fernmeldeverkehrs[145] aufgeführten Gesetze, welche die Voraussetzungen festlegen, unter denen eine Überwachung durchgeführt werden darf. Ebenfalls relevant sind weitere Bestimmungen wie die Artikel 28 ff. ZGB, die insbesondere das Recht auf Privatleben und das Recht am eigenen Bild schützen[146], oder die Artikel 179 ff. StGB (strafbare Handlungen gegen den Geheim- oder Privatbereich).

European Data Protection Board, Report of the work undertaken by the ChatGPT Taskforce, 23. Mai 2024, 4 N 14, abrufbar unter: https://www.edpb.europa.eu/our-work-tools/our-documents/other/report-work-undertaken-chatgpt-taskforce_en (abgerufen am 19. August 2024).

142 CR LPD-BERTIL COTTIER, Art. 1 N 19, BSK DSG/BGÖ - MATTHIAS R. SCHÖNBÄCHLER/URS MAURER-LAMBROU/SIMON KUNZ, Art. 1, N 5.

143 BGE **145** IV 42, E., 4.2; JULIEN FRANÇAIS, in: Petit commentaire LPD (Fn. 44), Art. 1 N 12.

144 SR **784.10**.

145 SR **780.1**.

146 BSK ZGB I-ANDREAS MEILI, Art. 28 N 17.

Das DSG und seine Verordnungen sind auf KI-Systeme anwendbar, wenn diese – was häufig der Fall ist – Personendaten verwenden.[147] Die für die Datenbearbeitung Verantwortlichen und in bestimmten Fällen auch die Auftragsbearbeiter müssen daher die zahlreichen in diesen Texten vorgesehenen Verpflichtungen einhalten. Dazu gehören insbesondere die Einhaltung der allgemeinen Grundsätze (Rechtmässigkeit, Erkennbarkeit, Zweckbindung, Richtigkeit, Verhältnismässigkeit, Treu und Glauben, Datenschutz durch Technik und durch datenschutzfreundliche Voreinstellungen, Sicherheit; vgl. Art. 6, 7 und 8 DSG), die Meldepflicht für automatisierte Bearbeitungen und das Führen von Verzeichnissen (Art. 12 DSG und Art. 31 DSV), die Dokumentations- und Protokollierungspflichten (Art. 3 bis 5 DSV), die Informationspflichten (Art. 19 DSG), insbesondere bei automatisierten Entscheidungen (Art. 21 DSG), die Informationspflichten bei Ausübung des Auskunftsrechts (Art. 25 DSG), die Pflicht zur Durchführung einer Folgenabschätzung (Art. 22). Diese verschiedenen Elemente werden in Art. 10 der Konvention (Gleichstellung und Nichtdiskriminierung) näher erläutert (vgl. Ziff. 4.3.2.5).

Für die Bearbeitung von Personendaten im öffentlichen Bereich gilt zudem der Grundsatz der gesetzlichen Grundlage (Ziff. 4.2.3.1.2). Das DSG hat diesen Grundsatz konkretisiert. So bedarf der staatliche Einsatz eines Gesichtserkennungsalgorithmus, der biometrische Personendaten (besonders schützenswerte Personendaten) bearbeitet oder Vorhersagen aufgrund von Profiling macht, eines Gesetzes im formellen Sinn (Art. 34 Abs. 2 Bst. a und b DSG). Zudem kann auch der blosse Einsatz von KI unter diese Voraussetzung fallen, wenn der Zweck oder die Art und Weise der Bearbeitung von Personendaten voraussichtlich schwerwiegende Auswirkungen auf die Grundrechte der betroffenen Person hat (Art. 34 Abs. 2 Bst. c DSG).[148] Das DSG verfolgt damit einen «risikobasierten» Ansatz.

Bei ungerechtfertigter Verletzung der vorgenannten Pflichten kann die betroffene Person von der für die Datenbearbeitung verantwortlichen Person insbesondere die Einstellung der Datenbearbeitung, die Berichtigung unrichtiger Daten oder die Löschung der Daten verlangen (Art. 30 und 41 DSG). Sie kann sich auch an den Eidgenössischen Datenschutz- und Öffentlichkeitsbeauftragten (EDÖB) wenden, der Datenbearbeitungen, die gegen das DSG verstossen, verbieten kann. Der EDÖB kann auch von Amtes wegen tätig werden.

Allerdings ist an dieser Stelle anzumerken, dass die Entwicklung von KI-Systemen sehr oft im Widerspruch zu den allgemeinen Grundsätzen des DSG steht. Eine umfassende Beschaffung von Personendaten aus einer Vielzahl von Quellen für das Training von KI-Systemen ist

---

[147] FLORENT THOUVENIN/STEPHANIE VOLZ, White Paper: Datenschutz, Juni 2024, 3, abrufbar unter: https://www.itsl.uzh.ch > Wissenstransfer > Publikationen > Positionspapiere > Nachvollziehbare Algorithmen: ein Rechtsrahmen für den Einsatz von künstlicher Intelligenz (abgerufen am 19. August 2024).

[148] BJ, Totalrevision des Bundesgesetzes über den Datenschutz (DSG) – Übersicht zu den wichtigsten Änderungen (Fn. 38), 10 ff., Ziff. 2.2.1; siehe auch BJ, Gesetzgebungsleitfaden Datenschutz (Fn. 44).

schlecht vereinbar mit den Grundsätzen der Datenminimierung, der Erkennbarkeit, der Richtigkeit oder der Zweckbindung.[149] Es stellt sich die Frage, ob der private Datenbearbeiter in diesem Fall ein überwiegendes Interesse geltend machen kann, insbesondere den Rechtfertigungsgrund der Bearbeitung zu nicht personenbezogenen Zwecken, namentlich für Forschung, Planung oder Statistik (vgl. Art. 31 Abs. 2 Bst. e DSG). Dies ist im Einzelfall und insbesondere in Abhängigkeit von den jeweiligen Interessen und der Art des KI-Systems zu beurteilen.[150] Im öffentlichen Sektor wird die Anwendbarkeit von Art. 39 DSG ebenfalls im Einzelfall beurteilt.[151]

Es sei noch darauf hingewiesen, dass im Rahmen der Totalrevision des DSG das StGB um eine Bestimmung zum Identitätsmissbrauch (Art. 179^{decies} StGB) erweitert wurde, welche die zahlreichen Bestimmungen zum Schutz der Privatsphäre und der Daten ergänzt. Diese Bestimmung wird im Zusammenhang mit KI und «Deepfakes» sicherlich an Bedeutung gewinnen (vgl. Ziff. 6.6.1).

Schliesslich sehen die Leitlinien des Bundes vor, dass die vom Bund eingesetzten KI-Technologien so zu gestalten sind, dass die Privatsphäre gemäss den Vorschriften geschützt und die Datenschutzbestimmungen jederzeit eingehalten werden.[152]

Zusammenfassend lässt sich sagen, dass KI zahlreiche Herausforderungen für den Schutz der Privatsphäre und den Datenschutz mit sich bringt, die jedoch durch den bestehenden Rechtsrahmen relativ gut abgedeckt sind.

Gleichwohl ist ein Tätigwerden des Gesetzgebers unter dem Gesichtspunkt der Transparenz nicht auszuschliessen, um eine wirksamere Umsetzung der bestehenden Garantien zu ermöglichen:

- Informationspflicht bei automatisierten Einzelentscheidungen: Derzeit besteht diese Pflicht nur für Entscheidungen, die ausschliesslich auf einer automatisierten Datenbearbeitung beruhen. In vielen Fällen werden KI-Systeme jedoch nur zur Unterstützung oder als Hilfe bei der Entscheidungsfindung eingesetzt und fallen nicht unter diese Vorschrift.

[149] In diesem Sinne im Zusammenhang mit Gesichtserkennungssoftware, vgl. NADJA BRAUN BINDER/NINA LAUKENMANN/LILIANE OBRECHT, KI in der Verwaltung (Fn. 28), 16 N 39.

[150] DAVID ROSENTHAL, Datenschutz beim Einsatz generativer künstlicher Intelligenz, Jusletter 6. November 2023, 12 N 34 ff.; ROLF H. WEBER, Künstliche Intelligenz und Datenschutz, Jusletter IT 4. Juli 2024, 7 N 19 ff.

[151] Die Problematik der Weiterverwendung von Daten wird im Rahmen der Arbeiten zur Erfüllung der Motion 22.3890 «Rahmengesetz für die Sekundärnutzung von Daten» vom 22. August 2022 untersucht.

[152] Leitlinien «Künstliche Intelligenz» für den Bund (Fn. 80), 3.

- Meldung der eingesetzten KI-Systeme: Das CNAI führt aktuell ein Register der im öffentlichen Sektor eingesetzten KI-Systeme. Eine Meldung ist derzeit freiwillig, könnte aber verpflichtend werden.

- Bekämpfung des Black-Box-Effekts: detaillierte Dokumentation, regelmässige Audits und Tests von KI-Systemen. Hier geht es darum, die Kontrolle über die Systeme zu behalten, indem sichergestellt wird, dass man weiterhin versteht, wie sie funktionieren, oder dass man sogenannte Halluzinationen erkennen kann. Wie bereits erwähnt, müssen die getroffenen Massnahmen den entgegenstehenden Interessen Rechnung tragen, die mit dem Interesse an der Offenlegung bestimmter Informationen kollidieren können, wie z. B. der Schutz der Daten Dritter, Geschäftsgeheimnisse oder Interessen der Strafverfolgung.

- Zudem muss der Gesetzgeber die Koordination zwischen einer allfälligen Regulierung der KI und dem DSG regeln, insbesondere im Hinblick auf die Pflicht zur Durchführung einer Folgenabschätzung.

4.3.2.7 Artikel 12 – Zuverlässigkeit

Artikel 12 sieht vor, dass jede Vertragspartei geeignete Massnahmen zur Förderung der Zuverlässigkeit von KI-Systemen und des Vertrauens in deren Ergebnisse trifft. Dies kann angemessene Qualitäts- und Sicherheitsanforderungen für den gesamten Lebenszyklus von KI-Systemen umfassen.

Gemäss dem erläuternden Bericht unterstreicht diese Bestimmung die wichtige Rolle, die insbesondere Normen, technische Spezifikationen oder Absicherungstechniken bei der Überwachung und Überprüfung der Zuverlässigkeit von KI-Systemen und bei der transparenten Dokumentation und Kommunikation von Nachweisen für diesen Prozess spielen können. Der Schwerpunkt liegt dabei auf spezifischen Schlüsselelementen in der Funktionsweise von KI-Systemen wie Zuverlässigkeit, Robustheit, Sicherheit, Genauigkeit und Leistung sowie auf funktionalen Voraussetzungen wie Qualität, Datenintegrität und -sicherheit sowie Cybersicherheit. Ziel dieser technischen Standards ist die Schaffung eines begründeten Vertrauens der Öffentlichkeit. Soweit deren Entwicklungsprozess transparent und inklusiv ist, können diese Normen dazu beitragen, eine gegenseitig verstandene und erweiterbare Sicherheit und Konformität im KI-Bereich zu schaffen.[153]

In diesem Rahmen sollen Massnahmen getroffen werden, die darauf abzielen, KI-Systeme wie jedes andere Softwaresystem technisch sicher und zuverlässig zu machen. Sicherheit und Zuverlässigkeit sollten als Grundvoraussetzungen für den gesamten Lebenszyklus von KI-Systemen und nicht als blosse Eigenschaften betrachtet werden. Je nach Einzelfall können die Massnahmen die Bereitstellung klarer Informationen darüber umfassen, wie die KI-Akteure diese Standards praktisch umgesetzt haben. Dies bedeutet, dass die Verantwortlichkeit von Anfang bis Ende mittels ausreichender Transparenz der Prozesse und Dokumentationsprotokolle sicherzustellen ist. Zwischen dieser Bestimmung und Artikel 8 (Transparenz

[153] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 84 ff.

und Aufsicht) sowie Artikel 9 (Rechenschaftspflicht und Verantwortung) besteht ein klarer Zusammenhang. Technische Normen spielen eine wichtige Rolle bei der Risikobewertung und -minderung, wenn die bestehenden Vorschriften keine ausreichenden Vorgaben enthalten. Dies gilt insbesondere für den Rahmen für das Risiko- und Folgenmanagement.[154]

Die schweizerische Datenschutzgesetzgebung sieht relativ weitgehende Verpflichtungen bezüglich der Sicherheit von Personendaten vor. Die Einhaltung der in Artikel 6 DSG festgelegten Grundsätze, die Grundsätze des Datenschutzes durch Technik und durch datenschutzfreundliche Voreinstellungen (Art. 7 DSG) sowie die Vorschriften über die Datensicherheit (Art. 8 DSG, ergänzt durch Art. 1–6 DSV), die Zertifizierung (Art. 13 DSG) und die Meldung von Verletzungen der Datensicherheit (Art. 24 DSG) gewährleisten einen kohärenten Schutz. Was insbesondere die Zertifizierung betrifft, sieht die VDSZ (Art. 6) vor, dass der EDÖB Richtlinien über die Mindestanforderungen für ein Managementsystem erlässt, wobei insbesondere folgende technische Normen zu berücksichtigen sind: SN EN ISO 9001, SN EN ISO/IEC 27001 und SN EN ISO/IEC 2770.

Das am 1. Januar 2024 in Kraft getretene Informationssicherheitsgesetz vom 18. November 2020[155] (ISG) und die dazugehörenden Verordnungen[156] enthalten ebenfalls Regeln zur Gewährleistung der Sicherheit der Bearbeitung von Informationen im Zuständigkeitsbereich des Bundes und der Sicherheit seiner Informatikmittel (Art. 1 Abs. 1 ISG). Dieses Gesetz legt den Schwerpunkt auf die Vereinheitlichung der Massnahmen. In diesem Rahmen kann es auch zum Schutz von KI-Systemen beitragen, die ohne Personendaten betrieben werden können.[157]

Die KI-Leitlinien des Bundes halten zudem fest, dass KI-Systeme sicher, robust und resilient konzipiert sein müssen, damit sie sich positiv auf Mensch und Umwelt auswirken und nicht missbraucht oder missbräuchlich eingesetzt werden können. Es sind geeignete Massnahmen zu treffen, um schwerwiegende Fehlentscheidungen zu vermeiden.[158]

In diesem Zusammenhang stellt sich die Frage nach der Bedeutung harmonisierter Normen bei der Konformitätsbewertung eines KI-Systems mit dem rechtlichen Rahmen. Im Folgenden wird erläutert, dass die KI-Verordnung der EU unter bestimmten Bedingungen, bei Einhaltung

---

[154] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 87 ff.

[155] SR **128**.

[156] Informationssicherheitsverordnung vom 8. November 2023 (SR **128.1**); Verordnung vom 8. November 2023 über das Betriebssicherheitsverfahren (SR **128.41**); Verordnung vom 19. Oktober 2016 über die Identitätsverwaltungs-Systeme und Verzeichnisdienste des Bundes (SR **172.010.59**).

[157] Universität Zürich, Digital Society Initiative, Positionspapier, Ein Rechtsrahmen für Künstliche Intelligenz (Fn. 113), 6, das zum damaligen Zeitpunkt die Schaffung eines allgemeinen Gesetzes zur Informatiksicherheit forderte.

[158] Leitlinien «Künstliche Intelligenz» für den Bund (Fn. 80), 5.

der harmonisierten Normen von der Vermutung der Konformität mit der Verordnung ausgeht (vgl. Ziff. 5.2.11).

Zusammenfassend lässt sich sagen, dass die Schweiz bereits gewisse Standards im Bereich des Datenschutzes und der Informationssicherheit anwendet, welche die Zuverlässigkeit und die funktionalen Voraussetzungen eines KI-Systems wie Qualität, Integrität, Datensicherheit und Cybersicherheit verbessern können.

Im Falle einer Ratifizierung der KI-Konvention muss die Schweiz ihre Kontakte und ihre Zusammenarbeit mit den Normierungsgremien zur Entwicklung von Normen und Standards im Bereich der KI fortsetzen.

Gegenwärtig sieht das schweizerische Recht keinen Mechanismus vor, der es erlaubt, von der Einhaltung bestimmter standardisierter KI-Normen auf die Konformität mit dem geltenden Rechtsrahmen zu schliessen. Im Falle einer schweizerischen Regulierung von KI würde sich unweigerlich die Frage nach der rechtlichen Tragweite von Normen stellen.

4.3.2.8 Artikel 13 – Sichere Innovation

Gemäss Artikel 13 wird jede Vertragspartei aufgefordert, zur Förderung von Innovation und zur Vermeidung negativer Auswirkungen auf Menschenrechte, Demokratie und Rechtsstaatlichkeit die Schaffung kontrollierter Umgebungen für die Entwicklung, Erprobung und die Testung von KI-Systemen unter Aufsicht der zuständigen Behörden zu ermöglichen.

Es sollen Anreize geschaffen werden, damit bereits in den frühen Phasen der Konzeption von KI-Systemen Fragen der System- und Datenqualität, Sicherheit und Absicherung sowie Menschenrechtsbelange berücksichtigt werden. Dies ist umso wichtiger, als bestimmte Risiken erst nach Tests erkannt und/oder nur in der Konzeptionsphase wirksam angegangen werden können.[159]

Die Konvention überlässt die Art und Weise der Umsetzung dieser Bestimmung den Vertragsparteien. Laut dem erläuternden Bericht kann es sich dabei insbesondere um «regulatorische Sandboxen» («regulatory sandboxes», bzw. Reallabore), informelle regulatorische Leitlinien oder Garantien der Verfahrensaussetzung («no-action letter»)[160] handeln, um zu

[159] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 90 ff.

[160] Gemäss dem erläuternden Bericht sollen Garantien der Verfahrensaussetzung klarstellen, wie die Regulierungsbehörden mit der Gestaltung, der Entwicklung oder dem Einsatz von KI-Systemen in neuen Kontexten umgehen werden (vgl. erläuternder Bericht KI-Konvention [Fn. 20], N 92). Sie kann auch in einer behördlichen Zusage bestehen, keine Aufsichtsmassnahmen zu ergreifen, solange sich die Unternehmen an die getroffenen Vereinbarungen oder an die von den Behörden gegebenen Auslegungshilfen zu den anwendbaren Gesetzen halten. Dies gibt den Betroffenen Rechts- und Planungssicherheit, insbesondere in Fällen, in denen unklar ist, wie ein Geschäftsmodell rechtlich zu behandeln ist. Garantien der Verfahrensaussetzung bedeutet für Unternehmen, dass während der Testphase in der Sandbox keine Rechtsanwendung erfolgt (vgl. STEPHANIE VOLZ, KI Sandboxen für die Schweiz?, SZW 2022, 51 ff., 63 f.).

klären, wie die Regulierungsbehörden die verschiedenen Lebenszyklen der KI angehen werden.[161]

Eine Sandbox beinhaltet eine sehr breite Palette an Instrumenten, deren Merkmale sich sehr oft vermischen. Die Analyse folgt hier der Kategorisierung des vom SECO im Jahr 2022 erteilten «Prüfauftrags zu Regulatory Sandboxes». Demnach umfasst eine Sandbox im weiteren Sinne zwei Hauptinstrumente: «Pilotprojekte» mit experimenteller Regulierung (im Folgenden: Pilotprojekte) und eine Sandbox im engeren Sinne.[162] In beiden Fällen wird durch die Behörden ein Rahmen festgelegt und überwacht.

Gemäss der obgenannten Studie dienen Pilotprojekte dazu, neue Technologien oder Verfahren für einen begrenzten Zeitraum unter realen Bedingungen zu testen, wobei in bestimmten Aspekten von bestehenden Vorschriften abgewichen werden kann. Ihre Ergebnisse zeigen, inwieweit die Vorschriften überarbeitet werden müssen.[163] Eine Sandbox im engeren Sinne ermöglicht es Unternehmen, Verfahren, Produkte und Dienstleistungen am Markt zu erproben, indem sie zeitlich befristet von bestimmten Vorschriften ausgenommen wird. Verstösse gegen solche Regeln werden korrigiert, aber nicht sanktioniert. Dies erlaubt Unternehmen herauszufinden, ob ihre innovativen Geschäftsmodelle innerhalb des bestehenden rechtlichen Rahmens bestehen können oder nicht.[164] Sandbox-Tests können sowohl unter realen Bedingungen als auch in einer simulierten und/oder kontrollierten Umgebung durchgeführt werden.[165]

Gemäss SECO-Studie sind die Anforderungen an die gesetzliche Grundlage bei der Entwicklung einer Sandbox weniger streng. Während Pilotprojekte oft ein Gesetz im formellen Sinn als Grundlage benötigen, genügt für eine Sandbox im engeren Sinn je nach Fall eine Verordnung oder sogar nur eine Änderung der behördlichen Praxis.[166]

---

[161] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 92.

[162] YVES SCHNEIDER/PETER HETTICH/PATRICK ZENHÄUSERN/ANNICK PIETZONKA/GUIDO SAURER, Prüfauftrag zu Regulatory Sandboxes, in SECO (Hrsg.), Grundlagen für die Wirtschaftspolitik, Juni 2022, 6.

[163] Vergleiche auch BJ, Gesetzgebungsleitfaden 4. Auflage 2019, mit Aktualisierung 2023, N 1044 ff., abrufbar unter: www.bj.admin.ch > Staat & Bürger > Legistik > Legistische Hauptinstrumente (abgerufen am 25. Juli 2024).

[164] YVES SCHNEIDER/PETER HETTICH/PATRICK ZENHÄUSERN u. a., Prüfauftrag zu Regulatory Sandboxes (Fn. 162), N 2.1.2 f.

[165] Vgl. auch Art. 3 Nr. 55 KI-Verordnung.

[166] YVES SCHNEIDER/PETER HETTICH/PATRICK ZENHÄUSERN u. a., Prüfauftrag zu Regulatory Sandboxes (Fn. 162), N 2.2.2 und 2.2.3.

Eine Alternative zur Sandbox im weiteren Sinne sind sogenannte «risikobasierte Regulierungen». In diesem Fall wird die Regulierung proportional zu den von den Unternehmen eingegangenen Risiken angepasst. Die eingeführten regulatorischen Vereinfachungen sind dauerhaft. Unternehmen mit relativ geringen Risiken für Kunden und Wirtschaft profitieren so von vereinfachten Vorschriften.[167] Ein weiteres wichtiges Instrument, das ebenfalls in diesen Rahmen der sicheren Innovation passt und eine Ergänzung oder sogar eine Alternative zur Sandbox im engeren Sinne darstellt, sind die «Innovation Hubs». Das sind Plattformen, auf denen sich Behörden, Unternehmen, Universitäten und Investoren über technologische Trends und branchenspezifisches Wissen austauschen können.[168]

Die schweizerische Gesetzgebung enthält Bestimmungen, die Pilotprojekte in vielen sektoriellen Bereichen zulassen. Die meisten dieser Bestimmungen sind so neutral gehalten, dass sie auch auf KI anwendbar sind.

Im Bereich des Datenschutzes sieht Art. 35 DSG vor, dass der Bundesrat die automatisierte Bearbeitung von besonders schützenswerten Personendaten oder andere Datenbearbeitungen im Sinne von Artikel 34 Absatz 2 Buchstaben b und c bereits vor Inkrafttreten eines formellen Gesetzes bewilligen kann, wenn: 1. die Aufgaben, für welche die Bearbeitung erforderlich ist, in einem bestehenden Gesetz formell geregelt sind; 2. ausreichende Massnahmen getroffen werden, um den Eingriff in die Grundrechte der betroffenen Personen auf ein Minimum zu beschränken; und 3. für die praktische Umsetzung der Datenbearbeitung, insbesondere aus technischen Gründen, eine Testphase vor dem Inkrafttreten unerlässlich ist. Diese Bedingungen müssen kumulativ erfüllt sein.[169] Das Genehmigungsverfahren durch den EDÖB und die Modalitäten sind in der DSV (Art. 32 ff.) detailliert beschrieben. Pilotversuche können dazu dienen, die Funktionsweise neuer Technologien, mit denen Personendaten bearbeitet werden, zu testen[170]; dies gilt also möglicherweise auch für KI-Systeme.

Ein weiteres Beispiel für ein Pilotprojekt, das für KI zur Anwendung gelangen könnte, ist Artikel 15 des teilweise am 1. Januar 2024 in Kraft getretenen Bundesgesetzes vom 17. März

---

[167] Dies ist beispielsweise in der Schweiz bei der Fintech-Regulierung der Fall, die es ermöglicht, Unternehmen, die Publikumseinlagen von bis zu einer Million Franken verwalten, von der Bewilligungspflicht zu befreien (vgl. YVES SCHNEIDER/PATRICK ZENHÄUSERN/PETER HETTICH/ROGER KÜTTEL, Welche Regulierungen begünstigen Innovation, Die Volkswirtschaft – Plattform für Wirtschaftspolitik, 23. Juni 2022, abrufbar unter: https://dievolkswirtschaft.ch/de/2022/06/welche-regulierungen-beguenstigen-innovationen/ [abgerufen am 25. Juli 2024]).

[168] YVES SCHNEIDER/PETER HETTICH/PATRICK ZENHÄUSERN U. A., Prüfauftrag zu Regulatory Sandboxes (Fn. 162), N 2.2.4. Die Schweiz verfügt über zahlreiche Innovation Hubs, die einen Informations- und Beratungsaustausch zwischen den verschiedenen KI-Akteuren ermöglichen. Auf Bundesebene wäre zum Beispiel Innosuisse zu nennen, der Schweizerische Innovationspark oder das KI-Zentrum der EPFL. Weitere wichtige Zentren gibt es auf kantonaler Ebene.

[169] BSK DSG/BGÖ-ANDREAS STÖCKLI/CHRISTOPH GRÜNINGER, Art. 35 N 8; CR DSG-ASTRID EPINEY/SAMAH POSSE, Art. 35 N 2.

[170] MONIQUE COSSALI SAUVAIN, in: Petit commentaire LPD (Fn. 44), Art. 35 N 20.

2023 über den Einsatz elektronischer Mittel zur Erfüllung von Behördenaufgaben[171] (EMBAG). Dieses Gesetz bezweckt die Förderung der elektronischen Abwicklung von Geschäftsprozessen im Bund (Grundsatz «digital first»). Diese Prozesse umfassen die Interaktion der Behörden aller Staatsebenen untereinander sowie der Behörden zu Unternehmen und zur Bevölkerung. Im EMBAG werden im Wesentlichen die allgemeinen Rahmenbedingungen für verschiedene Tätigkeiten festgelegt, namentlich die Entwicklung des Einsatzes von E-Government auf Bundesebene, die Zusammenarbeitsformen des Bundes mit anderen Gemeinwesen und Organisationen im Bereich E-Government sowie für die elektronischen Behördenleistungen des Bundes.[172] Artikel 15 EMBAG sieht auch die Möglichkeit der Durchführung von Pilotversuchen vor. Wenn Letztere Bearbeitungen beinhalten, die in den Anwendungsbereich von Artikel 35 DSG fallen, dürfte von den dort festgelegten Voraussetzungen nicht abgewichen werden können.[173]

Weitere Beispiele finden sich im sektoriellen Recht. So ermächtigt Artikel 106 Absatz 5 SVG den Bundesrat, Pilotprojekte im Bereich des automatisierten Fahrens zu bewilligen. Mehr als zehn Projekte wurden bereits auf dieser Grundlage realisiert.[174] Auch im Bereich der Stromversorgung können unter bestimmten Voraussetzungen Pilotprojekte zur Entwicklung innovativer Technologien, Geschäftsmodelle oder Produkte im Energiesektor bewilligt werden.[175] Pilotversuche werden auch in der Krankenversicherung[176] zur Dämpfung des Kostenwachstums im Gesundheitswesen, zur Stärkung der Qualität oder zur Förderung der Digitalisierung, in der Invalidenversicherung zur Erleichterung der Eingliederung[177] oder in der Berufsbildung[178] durchgeführt.

Die obigen Ausführungen lassen den Schluss zu, dass das Bundesrecht Bestimmungen enthält, die gegebenenfalls die Durchführung von Pilotprojekten im KI-Bereich ermöglichen. Das Spektrum der Innovationszentren in der Schweiz ist ebenfalls sehr reichhaltig. Es fällt jedoch auf, dass keine oder wenige Sandboxen im engeren Sinne, die es Unternehmen ermöglichen würden, innovative Systeme unter realen oder simulierten Bedingungen zu testen, entwickelt werden. Eine solche Sandbox wäre aber gemäss einer Umfrage im Rahmen der vom SECO in Auftrag gegebenen Studie zur Prüfung regulatorischer Sandboxen gerade im Zusammen-

---

171 SR **172.019**.

172 Pressemitteilung vom 22. November 2023, Bundesrat setzt Bundesgesetz über den Einsatz elektronischer Mittel zur Erfüllung von Behördenaufgaben in Kraft, abrufbar unter: www.admin.ch > Dokumentation > Medienmitteilungen > Bundesrat setzt Bundesgesetz über den Einsatz elektronischer Mittel zur Erfüllung von Behördenaufgaben in Kraft (abgerufen am 27. August 2024).

173 MONIQUE COSSALI SAUVAIN, Petit commentaire LPD (Fn. 44), Art. 35 N 11 ff.

174 Vgl. die auf der Website des ASTRA veröffentlichte Liste, «Intelligente Mobilität in der Schweiz: Übersicht über abgeschlossene Pilotversuche», Stand 3. Februar 2023, abrufbar unter: www.astra.admin.ch > Themen > Intelligente Mobilität > Pilotversuche (abgerufen am 25. Juli 2024).

175 Art. 23*a* des Stromversorgungsgesetzes (SR **734.7**).

176 Art. 59*b* des Bundesgesetzes über die Krankenversicherung (SR **832.10**).

177 Art. 68^{quater} des Bundesgesetzes über die Invalidenversicherung (SR **831.20**).

178 Art. 4 des Bundesgesetzes über die Berufsbildung (SR **412.10**).

hang mit KI und Datenschutz sinnvoll. Datenschutzbestimmungen bzw. Unsicherheiten bezüglich ihrer konkreten Anwendung würden die Entwicklung von potenziell gesellschaftlich nützlichen Anwendungen behindern.[179]

Vor diesem Hintergrund sollte im Rahmen einer möglichen Umsetzung der KI-Konvention näher geprüft werden, ob die bestehenden Regelungen zu Pilotprojekten bereits eine ausreichende Grundlage für die Entwicklung einer (ggf. sektorspezifischen) Sandbox für KI bieten. Gegebenenfalls ist eine Einrichtung solcher Instrumente zu prüfen. Dabei ist zu beachten, dass einige Autoren die Sandbox als «technisch-organisatorische Massnahme» zur Begrenzung von Grundrechtseingriffen im Rahmen von Pilotprojekten unter Federführung des Bundes ansehen.[180]

Eine Inspirationsquelle in diesem Zusammenhang könnte die «*Innovation Sandbox*»[181] sein, die 2022 vom Kanton Zürich ins Leben gerufen wurde. Dieses Projekt schafft eine Testumgebung, in der Akteure KI-Projekte innerhalb eines klar definierten Rahmens durchführen können. Unterschiedliche Organisationen wie Start-ups, KMUs, Grossunternehmen oder Forschungsinstitute erhalten durch diese Sandbox Zugang zu regulatorischem Know-how und neuen Datenquellen. Im Gegenzug werden alle Erkenntnisse und Ergebnisse öffentlich zugänglich gemacht. Im Gegensatz zu vielen ausländischen Ansätzen geht die «Innovation Sandbox» für KI insofern einen Schritt weiter, als einige der eingereichten Projekte nicht nur geprüft, sondern auch in die Praxis umgesetzt werden. Ziel ist es, verantwortungsvolle Innovation unter Berücksichtigung rechtlicher und ethischer Kriterien zu fördern und die Verbreitung von KI in Verwaltung, Wirtschaft und Forschung zu unterstützen.

Der schweizerische rechtliche Rahmen sieht sektorspezifische Möglichkeiten für Pilotprojekte im KI-Bereich vor. Auch die Innovationzentren sind hier gut vertreten. Die Schaffung von kontrollierten Umgebungen, in denen Unternehmen ihre Produkte und Systeme unter realen Bedingungen testen können (Sandbox im engeren Sinne), scheint hingegen wenig oder gar nicht entwickelt zu werden. Der diesbezügliche Bedarf wird im Rahmen der Umsetzung der Konvention genauer zu ermitteln sein.

[179] YVES SCHNEIDER/PETER HETTICH/PATRICK ZENHÄUSERN et al., Prüfauftrag zu Regulatory Sandboxes (Fn. 162), N 3.1. Ebenso: STEPHANIE VOLZ, KI Sandboxen für die Schweiz? (Fn. 160), 67 f., die sich für die Einführung einer Experimentierklausel im DSG ausspricht.

[180] SANDRA HUSI-STÄMPFLI/ANNE-SOPHIE MORAND, Datenschutzrecht, Zürich 2024, 180 N 336.

[181] Innovation Sandbox for Artificial Intelligence (AI), abrufbar unter: https://www.zh.ch > Wirtschaft & Arbeit > Wirtschaftsstandort > Innovation-Sandbox für KI (abgerufen am 25. Juli 2024).

4.3.3 Kapitel IV: Rechtsmittel

4.3.3.1 Allgemeines

Kapitel IV ist den Rechtsmitteln gewidmet. Artikel 14 enthält Bestimmungen über Massnahmen, die zu ergreifen sind, um zugängliche und wirksame Rechtsmittel gegen durch KI-Systeme entstandene Menschenrechtsverletzungen zu gewährleisten. Artikel 15 behandelt Verfahrensgarantien im Allgemeinen.

4.3.3.2 Artikel 14 – Rechtsmittel

Artikel 14 Absatz 1 – Zugängliche und wirksame Beschwerde

Gemäss Artikel 14 Absatz 1 ergreift jede Vertragspartei im Einklang mit ihrer innerstaatlichen Rechtsordnung Massnahmen oder behält diese bei, um die Verfügbarkeit zugänglicher und wirksamer Rechtsmittel gegen Menschenrechtsverletzungen sicherzustellen, die sich aus Tätigkeiten im Lebenszyklus von KI-Systemen ergeben.

Diese Bestimmung verlangt keine Schaffung neuer Rechtsmittel. Sie fordert die Vertragsparteien jedoch auf, die bestehenden Rechtsmittel bei Menschenrechtsverletzungen durch Massnahmen zu stärken, die den von KI-Systemen ausgehenden Gefahren Rechnung tragen. KI-Systeme sind technisch sehr komplex und ihre Funktionsweise ist oft undurchsichtig. Dies erschwert es Personen, die vom Einsatz von KI-Systemen betroffen sein können, ihre Rechte geltend zu machen. Insbesondere ist es für betroffene Personen schwierig, Zugang zu relevanten Informationen zu erhalten und diese zu verstehen.[182]

Aus diesem Grund betont Artikel 14 der Konvention die Notwendigkeit sowohl zugänglicher als auch wirksamer Beschwerde. Diese Terminologie lehnt sich insbesondere an Artikel 2 UNO-Pakt II und Artikel 13 EMRK an. Um wirksam zu sein, muss die betreffende Beschwerde in der jeweiligen Situation unmittelbar Abhilfe schaffen können. Um zugänglich zu sein, muss die Beschwerde mit ausreichenden Verfahrensgarantien ausgestattet sein, damit sie für die betroffene Person von Nutzen ist.[183] Bei der Zugänglichkeit ist die Möglichkeit des praktischen Zugangs zu den notwendigen Informationen und zu sachkundiger Beratung zu berücksichtigen. Bezüglich der rechtlichen Zugänglichkeit muss die Beschwerde insbesondere hinreichende Erfolgsaussichten bieten.[184]

Es stellt sich die Frage nach der Anwendbarkeit von Artikel 14 im Verhältnis zwischen Privaten. Artikel 14 der Konvention ist auf Rechtsmittel *gegen Menschenrechtsverletzungen* anwendbar, die sich aus Tätigkeiten während des Lebenszyklus von KI-Systemen ergeben. Die Verpflichtungen aus Artikel 14 der KI-Konvention sollten dort Anwendung finden, wo eine direkte oder indirekte horizontale Wirkung der Grundrechte zwischen Privaten besteht oder in

---

[182] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 98 f.

[183] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 98.

[184] OLIVIER BIGLER, in: Luc Gonin/Olivier Bigler (Hrsg.), Commentaire de la Convention européenne des droits de l'homme (CEDH), Bern 2018, Art. 13 N 21 f., und die zitierten Referenzen.

Zukunft erkannt wird (vgl. die Überlegungen in Ziff. 4.2.3.1). In diesem beschränkten Rahmen bieten die zivilrechtlichen Rechtsmittel bereits gewisse Garantien (vgl. insbesondere die Erwägungen in Ziff. 6.3).

Artikel 14 Absatz 2 – Massnahmen zur Verstärkung der Tragweite von Artikel 14 Absatz 1

Die Konvention enthält bereits in Kapitel III, namentlich mit Artikel 8 (Transparenz und Aufsicht) und Artikel 9 (Rechenschaftspflicht und Verantwortung), einschlägige Verpflichtungen in Bezug auf die Rechtsmittel. Artikel 14 Absatz 2 geht darüber hinaus, indem er die Verpflichtung einführt, drei spezifische Massnahmen zur Verstärkung der Tragweite von Artikel 14 Absatz 1 zu ergreifen oder beizubehalten:

- Buchstabe a sieht eine Dokumentationspflicht vor. Diese beinhaltet einschlägige Informationen über KI-Systeme, die die Menschenrechte erheblich beeinträchtigen können, und über ihre entsprechende Nutzung. Diese Informationen sind den zum Zugang zu diesen Informationen befugten Stellen zu liefern sowie – sofern angemessen und anwendbar – den betroffenen Personen zur Verfügung zu stellen oder mitzuteilen.

  Diese Bestimmung soll Transparenz beim Einsatz von KI-Systemen gewährleisten. Letztlich besteht der Zweck dieses Artikels jedoch darin, betroffenen Personen die Möglichkeit zu geben, Entscheidungen, die mit Hilfe von KI-Systemen vorbereitet oder getroffen wurden, und gegebenenfalls sogar den Einsatz des Systems an sich anzufechten (vgl. Art. 14 Abs. 2 Bst. b und die Erläuterungen dazu weiter unten).

  Im Schweizer Recht ergeben sich die Dokumentations- und Informationspflichten gegenüber den betroffenen Personen aus dem Anspruch auf rechtliches Gehör gemäss Artikel 29 Absatz 2 BV. Zudem gewährt das BGÖ unter bestimmten Voraussetzungen Zugang zu amtlichen Dokumenten der Bundesverwaltung.

  Im Datenschutzrecht besteht zudem gemäss DSG die Pflicht zur Führung eines Verzeichnisses der Bearbeitungstätigkeiten (vgl. Art. 12 DSG). Im Zusammenhang mit der Information der betroffenen Personen sind auch die Artikel 19, 21 und 25 DSG über die Informationspflicht und das Auskunftsrecht zu erwähnen (vgl. Ziff. 4.3.2.5 und 4.3.2.6). Angesichts der weitergehenden Verpflichtungen nach Artikel 14 Absatz 2 Buchstabe a erscheinen die geltenden Bestimmungen jedoch nicht ausreichend.

Zusammenfassend kann festgestellt werden, dass das schweizerische Recht bereits Bestimmungen enthält, an die die Verpflichtung nach Artikel 14 Absatz 2 Buchstabe a der Konvention geknüpft werden kann. Der bestehende Rechtsrahmen scheint jedoch nicht ausreichend zu sein. Es bleibt jedoch darauf hinzuweisen, dass die Begriffe «einschlägig», «erheblich» und «sofern anwendbar» den Vertragsparteien einen Ermessensspielraum bei der Umsetzung einräumen. Insbesondere fallen nur Systeme, die «erhebliche» Auswirkungen auf die Menschenrechte haben können, unter diese Bestimmung, die somit keine Anforderung für alle in Artikel 3 der Konvention definierten KI-Systeme darstellt. Es wird Aufgabe des Gesetzgebers sein, die entsprechende Schwelle festzulegen.

Darüber hinaus könnte die Pflicht, relevante Informationen zur Verfügung zu stellen oder zugänglich zu machen, durch andere überwiegende Interessen eingeschränkt werden, beispielsweise im Zusammenhang mit dem Schutz von Daten Dritter oder von Geschäftsgeheimnissen.

Artikel 14 Absatz 2 Buchstabe a erfordert daher einer Präzisierung des anwendbaren Rechtsrahmens, so dass ein Tätigwerden des Gesetzgebers in diesem Sinne angezeigt erscheint. Zudem werden die Informationen in der Regel nicht bei den Behörden, sondern bei den Entwicklern der KI-Systeme vorhanden sein.

Das Führen von Registern relevanter Informationen zu KI-Systemen ist in der Bestimmung nicht vorgesehen, könnte aber eine mögliche Form der Umsetzung dieser Verpflichtung sein.

- Buchstabe b schreibt vor, dass die in Artikel 14 Absatz 2 Buchstabe a genannten Informationen ausreichen müssen, damit die betroffene Person die Entscheidungen, die aufgrund der Verwendung des Systems getroffen wurden oder massgeblich auf der Verwendung des Systems beruhen, oder, falls erforderlich und angemessen, den Einsatz des Systems anfechten kann. Der letzte Fall kann Situationen abdecken, in denen die Verwendung eines bestimmten Systems hypothetisch verboten wäre. Denkbar sind auch Situationen, in denen der Behörde vorgeworfen wird, ein KI-System ohne angemessene Rechtsgrundlage eingesetzt zu haben, und in denen dann dieser Einsatz des KI-Systems als solcher angefochten wird.

  KI-Systeme stellen besondere Herausforderungen bezüglich Begründungspflicht.[185] Diese sind je nach Art des eingesetzten KI-Systems unterschiedlich. Wie oben erläutert (vgl. *ad* Art. 8, Ziff. 4.3.2.3), sind regelbasierte Algorithmen hinsichtlich Erklärbarkeit und Interpretierbarkeit ihrer Ergebnisse weniger problematisch als Algorithmen des maschinellen Lernens («machine learning»). Bei Letzteren erweist sich die Erfüllung der Begründungspflicht als besonders schwierig.

  In jedem Fall sollte sich die Begründung nicht auf die allgemeine Funktionsweise eines KI-Systems beziehen, sondern auf die *Logik, auf der die Entscheidung im Einzel-*

[185] Vgl. zu den Herausforderungen im öffentlichen Sektor NADJA BRAUN BINDER/MATTHIAS SPIELKAMP/CATHERINE EGLI et al., Einsatz Künstlicher Intelligenz (Fn. 39), 37; NADJA BRAUN BINDER/LILIANE OBRECHT, White Paper: Transparenz durch Begründung von Verfügungen, Juni 2024, 3, abrufbar unter: https://www.itsl.uzh.ch > Wissenstransfer > Publikationen > Positionspapiere > Nachvollziehbare Algorithmen: ein Rechtsrahmen für den Einsatz von künstlicher Intelligenz (abgerufen am 27. August 2024); NADJA BRAUN BINDER/NINA LAUKENMANN/LILIANE OBRECHT, KI in der Verwaltung (Fn. 28), 10 N 23; NADJA BRAUN BINDER/LILIANE OBRECHT, Die Begründung von Verfügungen (Fn. 84), 707 ff.

fall beruht. Inhaltlich müssen die übermittelten Informationen dem Kontext entsprechen, hinreichend klar und vor allem für die betroffene Person nützlich sein, damit sie sich in einem Rechtsmittel wirksam darauf berufen kann.[186]

Im schweizerischen Recht ergibt sich das Recht auf Begründung von Verwaltungs- und Gerichtsentscheiden aus dem Anspruch auf rechtliches Gehör nach Artikel 29 Absatz 2 BV. Die Begründung eines Entscheides muss sich sowohl auf den zugrunde gelegten Sachverhalt als auch auf die rechtlichen Erwägungen stützen. Das Bundesgericht betont, dass es entscheidend ist, dass aus dem Entscheid klar hervorgeht, welche Tatsachen festgestellt und welche rechtlichen Folgerungen aus den massgebenden Tatsachen gezogen werden.[187] Der Anspruch auf rechtliches Gehör wird durch das anwendbare Verfahrensrecht konkretisiert. Für das Verwaltungsverfahren auf Bundesebene sind insbesondere die Bestimmungen des VwVG, namentlich Artikel 35 VwVG, massgebend. Die Begründung muss angemessen sein und variiert je nach konkretem Fall. Je komplexer der Sachverhalt, desto dichter muss die Begründung sein. Hat die Behörde einen grossen Ermessensspielraum, ist der Entscheid ebenfalls ausführlicher zu begründen. In Bereichen, in denen viele gleichartige Entscheidungen getroffen werden, darf die Begründung aus Gründen der Verfahrensökonomie weniger detailliert ausfallen. Dabei ist zu beachten, dass eine ausreichende Begründung staatlichen Handelns auch der Rechtssicherheit dient, indem sie staatliches Handeln vorhersehbar macht und eine einheitliche Praxis fördert. Setzt die Behörde KI-Systeme ein, um behördliche oder gerichtliche Entscheidungen zu treffen, sind daher die Garantien zu beachten, die sich aus dem Anspruch auf rechtliches Gehör ergeben.

In diesem Zusammenhang ist auch das DSG zu erwähnen, das entsprechende Bestimmungen für automatisierte Einzelentscheidungen kennt (vgl. Art. 21 und 25 Abs. 2 Bst. f DSG). Eine automatisierte Einzelentscheidung liegt vor, wenn eine Datenbearbeitung automatisch, ohne Dazutun einer natürlichen Person erfolgt, und für die betroffene Person eine rechtliche Folge nach sich zieht oder sie erheblich beeinträchtigt (automatisierte Einzelentscheidung). Es kann sich dabei um eine Entscheidung einer privaten Stelle oder eines öffentlichen (Bundes-)Organs handeln, die Entscheidung muss jedoch einen gewissen Komplexitätsgrad aufweisen. Einfache Entscheidungen, wie sie etwa bei einem Geldbezug am Bancomaten getroffen werden, fallen nicht darunter.[188] Im öffentlichen Recht entfaltet eine Verfügung im Sinne von Artikel 5 VwVG, die ausschliesslich auf einer automatisierten Bearbeitung von Personendaten beruht, Rechtswirkungen gegenüber der betroffenen Person und ist daher als automatisierte

---

[186] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 99.

[187] BGE **142** II 154, E. 4.2; JACQUES DUBEY, Droits fondamentaux, Volume II: Liberté, garanties de l'Etat de droit, droits sociaux et politiques, Basel 2018, 815.

[188] Botschaft vom 15. September 2017 zum Bundesgesetz über die Totalrevision des Bundesgesetzes über den Datenschutz und die Änderung weiterer Erlasse zum Datenschutz, BBl **2017** 6941 7056 s.

Einzelentscheidung im Sinne von Art. 21 DSG zu qualifizieren (vgl. z. B. Art. 38 Abs. 2 ZG).[189]

Im Rahmen des Auskunftsrechts sieht Artikel 25 Absatz 2 Buchstabe f DSG vor, dass die betroffene Person zumindest über das Vorliegen einer automatisierten Einzelentscheidung informiert werden muss, falls eine solche vorliegt, sowie über die Logik, auf der diese Entscheidung beruht. Gemäss Botschaft des Bundesrates sind nicht die verwendeten Algorithmen, die häufig unter das Geschäftsgeheimnis fallen, sondern die Grundannahmen der Algorithmus-Logik, auf der die automatisierte Einzelentscheidung beruht, offenzulegen.[190]

Zu beachten ist jedoch, dass diese Bestimmungen nur anwendbar sind, wenn die Entscheidung auf einer *ausschliesslich* automatisierten Personendatenbearbeitung beruht (vgl. Art. 21 Abs. 1 DSG). Artikel 14 Absatz 2 Buchstabe b KI-Konvention ist demgegenüber weiter gefasst, da er auch Fälle erfasst, in denen die Entscheidung «erheblich» auf dem Einsatz eines KI-Systems beruht. Erfasst werden also auch Situationen, in denen KI-Systeme bei einer Entscheidung als Unterstützung eingesetzt werden, wobei diese Entscheidung dann aber letztlich von einer natürlichen Person getroffen wird. Im Gegensatz zu Artikel 21 DSG beschränkt sich Artikel 14 Absatz 2 Buchstabe b zudem nicht auf Fälle von automatisierten Entscheidungen, bei denen personenbezogene Daten verarbeitet werden. Denn auch wenn bei den meisten KI-Systemen personenbezogene Daten bearbeitet werden, ist dies doch nicht immer der Fall.[191]

Im Übrigen ist daran zu erinnern, dass die Bestimmungen des DSG auf die Bearbeitung von Personendaten im Rahmen von Gerichtsverfahren oder bundesrechtlich geregelte Verfahren nicht anwendbar sind. Auf erstinstanzliche Verwaltungsverfahren hingegen ist das DSG anwendbar (Art. 1 Abs. 3 DSG).

Zusammenfassend lässt sich sagen, dass das schweizerische Recht bereits verschiedene einschlägige Bestimmungen zur Begründungspflicht im Zusammenhang mit automatisierten Einzelentscheidungen enthält, insbesondere im Rahmen des DSG. Diese decken jedoch nicht alle Konstellationen ab, d. h. insbesondere nicht Entscheidungen, die *mithilfe* von KI-Systemen getroffen werden. Ein Tätigwerden des Gesetzgebers erscheint notwendig, um auch diesen Fall zu erfassen und die Anforderungen an die Begründungspflicht in diesem Zusammenhang zu definieren.

Bei erstinstanzlichen Verwaltungsentscheidungen ist selbstverständlich der Anspruch des rechtlichen Gehörs zu beachten. Es sollte jedoch weiter geprüft werden,

[189] LISA JACCOUD/SÉBASTIEN FANTI/ALEXANDRE STAGER, in: Petit commentaire LPD (Fn. 44), Art. 21 N 22.

[190] Botschaft Totalrevision DSG (Fn. 188), BBl **2017** 6941 7067.

[191] NADJA BRAUN BINDER/MATTHIAS SPIELKAMP/CATHERINE EGLI u. a., Einsatz Künstlicher Intelligenz (Fn. 39), 37, Fussnote 246.

ob der Rechtsrahmen in Bezug auf die Art und Weise, wie die Einhaltung dieser Anforderungen im Rahmen automatisierter Verwaltungsverfahren gewährleistet wird, einer Präzisierung bedarf.

Für gerichtliche Entscheidungen scheint kein Handlungsbedarf zu bestehen, da es nach dem derzeitigen Stand der Verfahrensvorschriften keine Rechtsgrundlage für solche Entscheidungen gibt. Sollten solche Rechtsgrundlagen geschaffen werden, wäre ein allfälliger Anpassungsbedarf vertieft zu prüfen.

- Artikel 14 Absatz 2 Buchstabe c sieht vor, dass den Betroffenen eine wirksame Beschwerde bei den zuständigen Behörden zur Verfügung stehen muss. Gemäss dem erläuternden Bericht kann dies die in Artikel 26 der Konvention vorgesehenen Kontrollmechanismen einschliessen (vgl. Ziff. 4.4.5).[192]

Gemäss Analyse des BJ verlangt die Bestimmung keine Schaffung neuer Rechtsmittel. Vielmehr verlangt sie von den Vertragsparteien die Gewährleistung der Wirksamkeit der bestehenden Rechtsmittel bei Menschenrechtsverletzungen im Zusammenhang mit Tätigkeiten im Rahmen des Lebenszyklus von KI-Systemen. Diesem Zweck dienen die in Artikel 14 Absatz 2 Buchstaben a und b der Konvention vorgesehenen Massnahmen. Hinsichtlich eines allfälligen Handlungsbedarfs im schweizerischen Recht kann daher auf die obigen Ausführungen verwiesen werden.

4.3.3.3 Artikel 15 – Verfahrensgarantien

Artikel 15 Absatz 1 – Wirksame Verfahrensgarantien, Schutzmassnahmen und Verfahrensrechte

Nach Artikel 15 Absatz 1 stellt jede Vertragspartei sicher, dass in Fällen, in denen sich ein KI-System erheblich auf die Wahrnehmung der Menschenrechte auswirkt, den betroffenen Personen im Einklang mit dem geltenden Völker- und innerstaatlichen Recht wirksame Verfahrensgarantien, Schutzmassnahmen und Verfahrensrechte zur Verfügung stehen.

Die Bestimmung beschränkt sich auf KI-Systeme, die «erhebliche» Auswirkungen auf die Ausübung der Menschenrechte haben, und gilt nicht für alle KI-Systeme im Sinne von Artikel 3 der Konvention.

Hinsichtlich der Anwendung der Bestimmung auf den Privatsektor wird auf die obigen Ausführungen zu Artikel 14 verwiesen (vgl. Ziff. 4.3.3.2). Das Erfordernis einer «erheblichen» Auswirkung wird die erfassten privatrechtlichen Sachverhalte weiter einschränken.

[192] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 100.

Diese Bestimmung betrifft die Verfahrensgarantien im Allgemeinen. Sie soll sicherstellen, dass die üblichen Verfahrensgarantien unabhängig vom Einsatz von KI gleich wirksam sind.

Im schweizerischen Recht ist dies insbesondere das Recht auf gleiche und gerechte Behandlung im Verfahren, das durch die Artikel 29–32 BV sowie durch Artikel 6 EMRK gewährleistet wird. Hervorzuheben ist insbesondere Folgendes:

- Eine Person hat das aus dem *Anspruch auf rechtliches Gehör* abgeleitete Recht, der Behörde *ihren Standpunkt darzulegen*, bevor diese eine sie betreffende Entscheidung fällt. Sie hat das Recht, sich an der Untersuchung zum Fall zu beteiligen, Tatsachen vorzubringen, Beweise anzubieten und sich zu den vorgebrachten Beweisen zu äussern.

  Das schweizerische Recht sieht in Artikel 21 Absätze 1 und 2 DSG eine Informationspflicht bei automatisierten Einzelentscheidungen sowie die Möglichkeit vor, dass die betroffene Person auf Wunsch ihren Standpunkt darlegen kann. Artikel 21 Absätze 3 und 4 DSG nennen mögliche Einschränkungen. So gelten die Rechte nach Artikel 21 Absatz 2 DSG für Bundesorgane nicht, wenn die betroffene Person nach Artikel 30 Absatz 2 VwVG nicht angehört werden muss.

  Wie bereits erwähnt, deckt der Anwendungsbereich der Bestimmungen des DSG jedoch nicht alle in Artikel 15 Absatz 1 der Konvention genannten Fälle ab, insbesondere nicht den Fall von teilautomatisierten Entscheidungen. Darüber hinaus bestimmt Artikel 1 Absatz 3 DSG, dass die Bearbeitung von Personendaten im Rahmen von Gerichtsverfahren oder von Verfahren, die durch Bundesrecht geregelt sind, sowie die Rechte der betroffenen Personen unter Ausschluss des DSG durch das anwendbare Verfahrensrecht geregelt werden. Diese Gesetze sehen jedoch ihrerseits Garantien des rechtlichen Gehörs vor.

- Eine wichtige Rolle spielt auch das Recht auf *Akteneinsicht*. Parteirechte laufen ins Leere, wenn die Verfahrensbeteiligten keine Kenntnis von den Unterlagen haben, die der Entscheidung der Behörde zugrunde liegen.

  Bei automatisierten Verfahren sollte die betroffene Person Zugang zu den spezifischen Informationen eines Falles haben, die von einem KI-System verarbeitet werden (*Input*), sowie zu dem daraus resultierenden Ergebnis (*Output*). Darüber hinaus sollte sie wissen können, wie diese Informationen maschinell verarbeitet werden. Unter Umständen sind diese Inhalte für die betroffene Person jedoch nur schwer verständlich. Trotzdem hat das Bundesgericht, wenn auch in einem anderen Zusammenhang, darauf hingewiesen, dass potenzielle faktische Schwierigkeiten, die sich aus der Menge

der einzusehenden Dokumente ergeben, dem Recht auf Akteneinsicht keinen Abbruch tun.[193] Bei KI-Systemen könnte jedoch eine Grenze erreicht sein, da die betroffene Person möglicherweise nicht in der Lage ist, aus den zur Verfügung gestellten Informationen Schlussfolgerungen zu ziehen.[194]

- Im Rahmen des Verwaltungsverfahren sind insbesondere die *Pflicht der Behörden zur Ermittlung des Sachverhalts von Amtes wegen* (Untersuchungsgrundsatz, Art. 12 VwVG) und die *Mitwirkungspflicht der Parteien* (Art. 13 VwVG) zu beachten.

  Bei voll- und teilautomatisierten Verfahren ist insbesondere sicherzustellen, dass alle relevanten Tatsachen berücksichtigt werden konnten.[195] Dies setzt einerseits die Richtigkeit und Vollständigkeit der Daten voraus, andererseits dass alle Daten, deren Relevanz für den spezifischen Kontext der Entscheidungsfindung zweifelhaft erscheint, unberücksichtigt bleiben.[196]

  Sachdaten können auch auf andere Weise beschafft werden, z. B. indem diese von der oder dem Beschwerdeführenden selber eingefordert werden. Die für die automatisierte Bearbeitung notwendige standardisierte Datenerfassung kann jedoch dazu führen, dass die oder der Beschwerdeführende nicht alle Angaben machen kann, die aus ihrer bzw. seiner Sicht für das Verfahren erforderlich wären. Wenn in einem Formular in bestimmten Feldern nur numerische Werte zulässig sind, können zusätzliche Erläuterungen, die für eine korrekte Beurteilung des Falles erforderlich wären, unter Umständen nicht gegeben werden.[197]

---

[193] BGE **144** II 427, 436, E. 3.2.3.

[194] In diesem Sinn: REGINA WEDER, Verfahrensgrundrechtliche Anforderungen an automatisierte Verwaltungsverfahren, in: Monika Simmler (Hrsg.), Smart Criminal Justice. Der Einsatz von Algorithmen in der Polizeiarbeit und Strafrechtspflege, Basel 2021, 237 ff., 251 f.

[195] NADJA BRAUN BINDER/NINA LAUKENMANN/LILIANE OBRECHT, KI in der Verwaltung (Fn. 28), 10 N 25 f.; vgl. auch NADJA BRAUN BINDER, Der Untersuchungsgrundsatz als Herausforderung vollautomatisierter Verfahren, zsis 2/2020, N 28.

[196] Europarat, L'administration et vous – Un manuel. Principes de droit administratif concernant les relations entre l'Administration et les personnes, Strassburg 2024, 18, abrufbar unter: https://www.coe.int/fr/web/cdcj/-/publication-of-the-new-handbook-the-administration-and-you-that-takes-into-account-the-increasing-use-of-ai (abgerufen am 26. August 2024).

[197] NADJA BRAUN BINDER/MATTHIAS SPIELKAMP/CATHERINE EGLI et al., Einsatz Künstlicher Intelligenz (Fn. 39), 38; NADJA BRAUN BINDER/NINA LAUKENMANN/LILIANE OBRECHT, KI in der Verwaltung (Fn. 28), 10 N 25 f.; NADJA BRAUN BINDER, Der Untersuchungsgrundsatz (Fn. 185), N 28.

- Die *Rechtsweggarantie* (Art. 29*a* BV) garantiert jeder Person das Recht auf Beurteilung durch eine richterliche Behörde. Daraus folgt, dass ein vollautomatisierter Entscheid im Einzelfall nicht von einer einzigen Instanz gefällt werden darf.[198]

Aus den obigen Ausführungen ergibt sich, dass die bestehenden Verfahrensgarantien im Zusammenhang mit KI-Systemen bereits vollumfänglich anwendbar sind. Sie werden in den Verwaltungs-, Zivil- und Strafprozessordnungen konkretisiert. Zudem bietet Artikel 21 DSG im erstinstanzlichen Verwaltungsverfahren bereits einen gewissen Schutz bei vollautomatisierten Einzelentscheidungen.

Abgesehen von den obigen Überlegungen zur Begründungspflicht (vgl. Ziff. 4.3.3.2) erscheinen die geltenden Regelungen ausreichend. Die Rechtsprechung wird die Konturen dieser Garantien im Zusammenhang mit KI-Systemen schärfen können. Bei Problemen oder Lücken können diejenigen Fälle geprüft werden, in denen der Einsatz von KI die Betroffenen daran hindern würde, ihre Rechte nach geltendem Recht gerichtlich geltend zu machen.

In jedem Fall sind mögliche entgegenstehende Interessen zu berücksichtigen, die den Anwendungsbereich dieser Garantien einschränken können, insbesondere das private Interesse des Entwicklers des Algorithmus, der sich auf das Geschäftsgeheimnis berufen könnte, oder ein anderes öffentliches Interesse, wie das Interesse an der Strafverfolgung. Im Verwaltungsverfahren ist es beispielsweise gemäss Artikel 27 VwVG bereits heute möglich, private Interessen, die dem Recht auf Akteneinsicht entgegenstehen, bei der Behandlung entsprechender Gesuche zu berücksichtigen.

Artikel 15 Absatz 2 – Information über Interaktion mit einem KI-System

Gemäss Artikel 15 Absatz 2 bemüht sich jede Vertragspartei sicherzustellen, dass Personen, die mit KI-Systemen interagieren, in einer dem Kontext angemessenen Weise darüber informiert werden, dass sie mit solchen Systemen und nicht mit einem Menschen interagieren. Damit soll insbesondere die Gefahr der Manipulation und Täuschung ausgeschlossen werden.[199]

Im schweizerischen Recht besteht für automatisierte Einzelentscheidungen eine entsprechende Verpflichtung im DSG. So sieht Artikel 21 Absatz 1 DSG vor, dass der Verantwortliche die betroffene Person über jede automatisierte Einzelentscheidung informieren muss. Ergeht die automatisierte Einzelentscheidung durch ein Bundesorgan, muss es die Entscheidung entsprechend kennzeichnen (Art. 21 Abs. 4 DSG). Artikel 25 Absatz 2 Buchstabe f DSG

[198] MICHAEL MONTAVON, Cyberadministration et protection des données – Etude théorique et pratique de la transition numérique en Suisse du point de vue de l'Etat, des citoyen-ne-s et des autorités de contrôle, Freiburg 2021, 667.

[199] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 104.

sieht im Rahmen des Auskunftsrechts vor, dass die betroffene Person über das Vorliegen einer automatisierten Einzelentscheidung informiert wird.

Wie bereits erwähnt deckt Artikel 21 DSG nicht alle Fälle ab, weil er nur in Fällen von automatisierten Einzelentscheidungen zur Anwendung gelangt. Artikel 15 Absatz 2 der Konvention würde hingegen eine Informationspflicht auch im Fall einer Interaktion mit KI-gestützten Chatbots auf den Websites der öffentlichen Verwaltung verlangen.[200]

Zusammenfassend kann festgestellt werden, dass die Umsetzung von Artikel 15 Absatz 2 in innerstaatliches Recht ein Tätigwerden des Gesetzgebers erfordert, um das im DSG bereits vorgesehene Auskunftsrecht zu erweitern. Eine solche Information ermöglicht es der betroffenen Person, ihre Verfahrensrechte wahrzunehmen.

Die Bestimmung lässt dem Gesetzgeber jedoch einen Spielraum, da die Formulierung «in einer dem Kontext angemessenen Weise» (französisch «en fonction du contexte») auch Situationen erfassen kann, in denen es offensichtlich ist, dass die Interaktion mit einer Maschine erfolgt und daher keine Information erforderlich ist. Nicht erfasst sind jedoch z. B. Situationen, in denen der eigentliche Zweck der Nutzung des KI-Systems durch die Information vereitelt würde, beispielsweise im Bereich der Strafverfolgung.

4.3.4 Kapitel V: Bewertung und Minderung von Risiken und nachteiligen Auswirkungen

Dieses Kapitel besteht aus einem einzigen Artikel (Art. 16) über den Rahmen für das Risiko- und Folgenmanagement.

Artikel 16 Absatz 1 sieht vor, dass jede Vertragspartei unter Berücksichtigung der Grundsätze in Kapitel III Massnahmen trifft oder aufrechterhält, um die von KI-Systemen ausgehenden Risiken zu ermitteln, zu bewerten, zu vermeiden und zu mindern, wobei die tatsächlichen und potenziellen Auswirkungen auf die Menschenrechte, die Demokratie und die Rechtsstaatlichkeit zu berücksichtigen sind.

Diese Bestimmung verpflichtet zur Schaffung eines Rahmens für das Risiko- und Folgenmanagement. Damit sollen *ex ante*, d. h. vor dem Einsatz eines KI-Systems, und gegebenenfalls iterativ, während des gesamten Lebenszyklus des KI-Systems, die relevanten Risiken des KI-Systems für Menschenrechte, Demokratie und Rechtsstaatlichkeit anhand einer Methodik mit konkreten Kriterien bewertet werden. Dies ist ein wichtiges Instrument, um die Einhaltung der Anforderungen der Konvention sicherzustellen, insbesondere der in Kapitel III niedergelegten Grundsätze.

Artikel 16 stellt sicher, dass die Vertragsparteien bei der Bestimmung, der Analyse und der Bewertung der Risiken und Folgen von KI-Systemen einen gemeinsamen Ansatz verfolgen. Gleichzeitig geht er davon aus, dass die Vertragsparteien besser in der Lage sind, die ent-

[200] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 104.

sprechenden Regulierungsentscheidungen zu treffen, und lässt ihnen daher einen Ermessensspielraum bei der Umsetzung. Aus diesem Grund sieht Artikel 16 Absatz 2 vor, dass die zu treffenden Massnahmen gegebenenfalls abgestuft und differenziert sein müssen und den Kontext des geplanten Einsatzes von KI-Systemen sowie die Schwere und Wahrscheinlichkeit möglicher Folgen angemessen berücksichtigen müssen (vgl. Art. 16 Abs. 2 Bst. a und b).

Artikel 16 Absatz 2 enthält weitere Details zu den zu ergreifenden oder beizubehaltenden Massnahmen. Buchstabe c schreibt vor, dass im Rahmen des Risiko- und Folgenmanagements die Standpunkte der relevanten beteiligten Parteien, insbesondere der Personen, deren Rechte betroffen sein könnten, zu berücksichtigen sind. Andere Beteiligte können beispielsweise externe Fachexperten oder Vertreter der Zivilgesellschaft sein.[201]

Buchstabe d sieht den iterativen Charakter der Massnahmen in dem Sinne vor, dass sie während des gesamten Lebenszyklus der KI-Systeme wiederholt werden müssen. Buchstabe e legt fest, dass die Massnahmen auch eine Überwachung der Risiken und negativen Auswirkungen auf Menschenrechte, Demokratie und Rechtsstaatlichkeit umfassen müssen. Diese Merkmale ermöglichen es, die Auswirkungen von KI-Systemen auch *ex post*, d. h. nach deren Einsatz, zu identifizieren und zu bewerten. Buchstabe f enthält eine Verpflichtung zur Dokumentation der Risiken, der tatsächlichen und potenziellen Auswirkungen und des Risikomanagementansatzes. Buchstabe g stellt klar, dass der Rahmen für das Risikomanagement gegebenenfalls eine Vorabprüfung von KI-Systemen vor ihrer Bereitstellung zur erstmaligen Verwendung sowie bei wesentlichen Änderungen vorsehen muss.

Artikel 16 Absatz 3 verpflichtet die Vertragsparteien, Massnahmen zum angemessenen Umgang mit nachgewiesenen nachteiligen Auswirkungen von KI-Systemen zu ergreifen. Dadurch sollen solche Auswirkungen mithilfe geeigneter Massnahmen behoben werden können. Diese Phase ist zu dokumentieren und in den Rahmen für das Risiko- und Folgenmanagement gemäss Artikel 16 Absatz 2 zu integrieren.

Schliesslich sieht Artikel 16 Absatz 4 vor, dass eine Vertragspartei, die den Einsatz eines KI-Systems als mit der Achtung der Menschenrechte, der Demokratie und der Rechtsstaatlichkeit unvereinbar hält, prüft, ob ein Moratorium, ein Verbot oder andere geeignete Massnahmen erforderlich sind. Angesichts ihres einschneidenden Charakters sollten Massnahmen wie ein Moratorium oder ein Verbot nur dann erwogen werden, wenn eine Vertragspartei der Auffassung ist, dass ein bestimmter Einsatz von KI-Systemen ein unannehmbares Risiko darstellt und eine sorgfältige Prüfung ergibt, dass keine anderen Massnahmen zur Minderung dieses Risikos zur Verfügung stehen. Diese Massnahmen sollten auch geeignete Überprüfungsverfahren umfassen, damit sie aufgehoben werden können, sobald die betreffenden Risiken hinreichend gemindert wurden oder geeignete Massnahmen der Risikominderung verfügbar geworden sind.

Der Europarat wird eine nicht bindende Methodik entwickeln, um einen möglichen Weg für die Umsetzung der Verpflichtungen aus Artikel 16 aufzuzeigen. So soll den Vertragsparteien

[201] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 108.

ein Modell zur Verfügung gestellt werden, das sie bei der Entwicklung ihres eigenen rechtlichen Rahmens für das Risiko- und Folgenmanagement unterstützt.

Der Rahmen für das Risiko- und Folgenmanagement von KI-Systemen nach Artikel 16 der KI-Konvention mag der Datenschutz-Folgenabschätzung (DSFA) nach Artikel 22 DSG ähnlich erscheinen, es handelt sich jedoch um zwei unterschiedliche Instrumente.

Gemäss Artikel 22 Absatz 1 DSG hat der für die Bearbeitung Verantwortliche vorgängig eine DSFA durchzuführen, wenn die geplante Bearbeitung voraussichtlich ein hohes Risiko für die Persönlichkeit oder die Grundrechte der betroffenen Person mit sich bringt. Gemäss der Botschaft zum DSG gilt die Pflicht zur Durchführung einer DSFA unter bestimmten Voraussetzungen sowohl für private Datenbearbeiter als auch für Bundesorgane, weshalb in der Bestimmung nicht nur von einem hohen Risiko für die Persönlichkeit der betroffenen Person, sondern auch für deren Grundrechte die Rede ist.[202]

Die DSFA ist ein weniger umfassendes Instrument als der Rahmen für das Risiko- und Folgenmanagement in Artikel 16 der KI-Konvention. Zusätzlich unterscheidet sie sich auch im Anwendungsbereich: Obwohl die meisten KI-Systeme auch Personendaten bearbeiten, ist dies nicht immer der Fall. Darüber hinaus sind die Risiken im Rahmen des Risiko- und Folgenmanagement weiter gefasst als in der DSFA.

Für die Datenbearbeitung durch Bundesorgane sieht Artikel 22 Absatz 1 DSG die Durchführung einer DSFA vor, wenn die geplante Bearbeitung voraussichtlich *ein hohes Risiko für die Grundrechte* mit sich bringt. Die vom DSG erfassten Grundrechte sind in erster Linie das Recht auf persönliche Freiheit (Art. 10 Abs. 2 BV), das Recht auf Schutz der Privatsphäre (Art. 13 Abs. 1 BV) und das Recht auf informationelle Selbstbestimmung im Sinne von Artikel 13 Absatz 2 BV, das auch in Artikel 8 EMRK verankert ist.[203] In der Praxis wird bei einer DSFA nach Artikel 22 DSG in der Regel abgeschätzt, welche negativen Folgen die Bearbeitung von Personendaten mit einer gewissen Wahrscheinlichkeit für die betroffene Person haben könnte, insbesondere physische Folgen (z. B. können falsche Daten zu einer falschen medizinischen Behandlung führen), materielle Folgen (z. B. die Verweigerung einer Arbeitsstelle, die betrügerische Verwendung einer Kreditkarte) oder immaterielle Folgen (z. B. kann das Wissen, beobachtet zu werden, bei einer Person zu einer Verhaltensänderung führen).[204] Artikel 16 der KI-Konvention zielt darauf ab, die Risiken und Auswirkungen von KI-Systemen

---

[202] Botschaft Totalrevision DSG (Fn. 188), BBl **2017** 6941 7059.

[203] JULIEN FRANÇAIS, in: Petit commentaire LPD (Fn. 44), Art. 1 N 11; CR LPD-BERTIL COTTIER, Art. 1 N 19; BSK DSG/BGÖ-MATTHIAS R. SCHÖNBÄCHLER/URS MAURER-LAMBROU/SIMON KUNZ, Art. 1 N 18; MARCO FREY, in: Bruno Baeriswy/Kurt Pärli/Dominika Blonski (Hrsg.), Datenschutzgesetz, Stämpfli Handkommentar, Art. 1 N 29; DAVID ROSENTHAL, in: David Rosenthal/Yvonne Jöhri (Hrsg.), Handkommentar zum Datenschutzrecht, Art. 1 N 3 f.

[204] In diesem Sinn: DAVID ROSENTHAL/SAMIRA STUDER/ALEXANDRE LOMBARD (für die Übersetzung), La nouvelle loi sur la protection des données (Fn. 129), 58, N 149; CR LPD-PHILIPPE GILLIÉRON, Art. 22 N 24.

auf alle Menschenrechte sowie auf Demokratie und Rechtsstaatlichkeit zu bewerten, unabhängig von der Bearbeitung personenbezogener Daten.

Für den Privatsektor hätte die Umsetzung von Artikel 16 zur Folge, dass die betroffenen privaten Akteure zumindest in jenen Bereichen eine Grundrechte-Folgenabschätzung vornehmen müssten, in denen eine direkte oder indirekte horizontale Wirkung der Grundrechte zwischen Privaten besteht oder in Zukunft erkannt wird (vgl. die Überlegungen in Ziff. 4.2.3.1). Dies bedeutet beispielsweise, dass die Datenschutzgesetzgebung, die Normen zum Persönlichkeitsschutz, wie sie in ZGB und OR verankert sind, als auch das GIG oder das BehiG usw. zu beachten sind. Die KI-Konvention führt aber nicht zu einer Erweiterung des Schutzbereichs der Grundrechte.

Nach der hier vorgenommenen Auslegung lässt Artikel 16 den Vertragsstaaten einen gewissen Gestaltungsspielraum, einschliesslich der Möglichkeit, im Privatsektor in bestimmten Fällen weniger weit zu gehen. So wäre es beispielsweise denkbar, die Pflicht zur Durchführung einer Grundrechte-Folgenabschätzung auf gewisse Anbieter grundlegender Dienste zu beschränken, wie dies in der KI-Verordnung der EU vorgesehen ist (vgl. Art. 27 der Verordnung, Ziff. 5.2.7.3.3).

Aus den obigen Ausführungen ergibt sich, dass das schweizerische Recht eine Pflicht zur Durchführung einer Folgenabschätzung bei der Bearbeitung von Personendaten (DFAS) kennt. Diese Pflicht reicht jedoch nicht aus, um den in Artikel 16 der KI-Konvention beschriebenen Rahmen des Risiko- und Folgenmanagements zu erfüllen. Der Gesetzgeber müsste deshalb handeln. Dabei sind jedoch die Besonderheiten des schweizerischen Systems zu berücksichtigen, insbesondere in Bezug auf die Wirkung der Grundrechte zwischen Privaten.

Im schweizerischen Recht regelt Artikel 36 BV die Voraussetzungen für die Einschränkung von Grundrechten. Ein Eingriff durch ein KI-System ohne gesetzliche Grundlage oder mit Auswirkungen auf den Kerngehalt der Grundrechte wäre rechtswidrig. Für den öffentlichen Sektor ergibt sich daher aus Artikel 16 Absatz 4 der Konvention kein Handlungsbedarf. Ein Verbot oder ein Moratorium für bestimmte Anwendungen von KI-Systemen könnte im privaten Sektor sinnvoll sein; oder auch im öffentlichen Sektor im Rahmen einer politischen Entscheidung zum aktiven Eingreifen zwecks Verbots bestimmter Praktiken. Diese Frage könnte sich zum Beispiel bei den Technologien zur Emotionserkennung stellen.

Bei der Einführung eines Rahmens für das Risiko- und Folgenmanagement von KI-Systemen in das schweizerische Recht sollte zudem die bereits bestehende Datenschutzfolgenabschätzung (DSFA) berücksichtigt werden.

In der KI-Verordnung ist zu beachten, dass, sofern Anbieter von KI-Systemen bereits eine Datenschutzfolgenabschätzung durchführen müssen, diese zusammen mit der Folgenabschätzung gemäss KI-Verordnung durchgeführt wird (vgl. Ziff. 5.2.7.3.3). Bei einer Ratifizierung der KI-Konvention müsste somit auch im Schweizer Recht ein Koordinationsmechanismus geschaffen werden.

4.3.5 Kapitel VI: Umsetzung des Übereinkommens

4.3.5.1 Allgemeines

Kapitel VI enthält Verpflichtungen der Vertragsparteien, die sich ausschliesslich auf die Phase der Umsetzung der Konvention in innerstaatliches Recht beziehen.

4.3.5.2 Artikel 17 – Nichtdiskriminierung

Diese Bestimmung legt fest, dass die Umsetzung der KI-Konvention durch die Vertragsparteien in Übereinstimmung mit ihren völkerrechtlichen Verpflichtungen auf dem Gebiet der Menschenrechte ohne Diskriminierung zu erfolgen hat.

Artikel 17 verbietet somit jegliche Diskriminierung bei der *Umsetzung* der Konvention durch die Vertragsparteien. Der Begriff der Diskriminierung ist identisch mit dem des geltenden Völkerrechts und umfasst ein breites Spektrum von Diskriminierungsgründen im Zusammenhang mit den persönlichen Merkmalen einer Person, ihrer Situation oder ihrer Zugehörigkeit zu einer Gruppe.[205]

In der schweizerischen Rechtsordnung verbietet Artikel 8 Absatz 2 BV bereits jegliche Diskriminierung. Insbesondere haben die rechtsanwendenden Behörden die gesetzlichen Bestimmungen im Lichte der Nichtdiskriminierung auszulegen.

4.3.5.3 Artikel 18 – Rechte von Menschen mit Behinderungen und von Kindern

Diese Bestimmung verpflichtet die Vertragsparteien, im Einklang mit ihrem innerstaatlichen Recht und den geltenden völkerrechtlichen Verpflichtungen die besondere Schutzbedürftigkeit im Zusammenhang mit der Achtung der Rechte von Menschen mit Behinderungen und der Rechte von Kindern gebührend zu berücksichtigen. Sie verweist daher auf die Bestimmungen und den Rechtsrahmen des Übereinkommens über die Rechte von Menschen mit Behinderungen[206] und des Übereinkommens über die Rechte des Kindes[207] sowie auf das geltende innerstaatliche Recht jeder Vertragspartei in Bezug auf die Rechte von Menschen mit Behinderungen und die Rechte von Kindern.

Ziel ist eine grösstmögliche Berücksichtigung der besonderen Bedürfnisse und der Schutzbedürftigkeit von Menschen mit Behinderungen und von Kindern im Hinblick auf die Achtung ihrer Rechte, einschliesslich der Vermittlung digitaler Kompetenz.[208]

---

[205] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 114.

[206] SR **0.109**.

[207] SR **0.107**.

[208] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 117.

Die Schweiz hat sowohl das Übereinkommen über die Rechte von Menschen mit Behinderungen als auch das Übereinkommen über die Rechte des Kindes ratifiziert und sich damit verpflichtet, die in diesen Übereinkommen enthaltenen Verpflichtungen in allen Lebensbereichen umzusetzen.

Im Falle einer Ratifizierung müsste die Schweiz bei der Umsetzung der Verpflichtungen aus den Übereinkommen die besonderen Bedürfnisse von Menschen mit Behinderungen und von Kindern berücksichtigen.

4.3.5.4 Artikel 19 – Öffentliche Konsultation

Diese Bestimmung fordert die Vertragsparteien auf, die Gesellschaft als Ganzes in die Diskussion über wichtige, durch den Einsatz von KI-Systemen aufgeworfene Fragen einzubeziehen, insbesondere hinsichtlich ihrer sozialen, wirtschaftlichen, rechtlichen, ethischen und ökologischen Auswirkungen.

Der Ausdruck «gegebenenfalls» überlässt es den Parteien, die Modalitäten des Einbezugs der Öffentlichkeit zu gestalten.[209] Die Bestimmung empfiehlt öffentliche Gespräche und Konsultationen verschiedener Interessenträger.

Die Schweiz verfügt bereits über Instrumente, die eine Beteiligung der Öffentlichkeit bei wichtigen Fragen im Zusammenhang mit KI-Systemen ermöglichen. So sehen insbesondere die Gesetzesentwürfe des Bundes in der Regel eine externe Vernehmlassung vor (vgl. VIG). Eine solche wäre auch für Gesetzesentwürfe im Bereich der KI vorzusehen. Das Vernehmlassungsverfahren bezweckt die Mitwirkung der Kantone, der politischen Parteien und der interessierten Kreise an der Meinungsbildung und Entscheidfindung des Bundes (Art. 2 VIG). Gemäss Artikel 4 Absatz 1 VIG kann jede Person oder Organisation an einem Vernehmlassungsverfahren teilnehmen und eine Stellungnahme einreichen.

Im spezifischen Kontext der KI-Systeme sind auch andere Mechanismen zur Beteiligung interessierter Kreise in der Schweiz zu nennen, insbesondere die «*Plateforme Tripartite Suisse* für digitale Gouvernanz und künstliche Intelligenz»[210]. Dabei handelt es sich um ein nationales Informationsnetzwerk für den Austausch über Themen im Zusammenhang mit der Digitalisierung. Diese Plateforme Tripartite wird vom BAKOM betreibt und steht allen Interessierten aus dem Privatsektor, der Zivilgesellschaft, der Wissenschaft und aus allen Verwaltungsebenen offen. Sie soll es ermöglichen, die Bedürfnisse, Anliegen und Erwartungen der verschiedenen Teilnehmer (Stakeholder) zu verstehen und diese bei der Entwicklung geeigneter Regeln im digitalen Bereich, einschliesslich KI, zu berücksichtigen.

---

[209] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 121.

[210] Vgl. www.bakom.admin.ch > Das BAKOM > Internationale Aktivitäten > Informationsgesellschaft International / WSIS > Die Plateforme Tripartite Suisse für digitale Gouvernanz und künstliche Intelligenz.

Aus den obigen Ausführungen geht hervor, dass die Schweiz bereits über angemessene Instrumente zur Erfüllung der Verpflichtungen aus Artikel 19 der KI-Konvention verfügt.

4.3.5.5 Artikel 20 – Digitale Kompetenzen und Fähigkeiten

Diese Bestimmung enthält die Verpflichtung der Vertragsparteien, angemessene digitale Kompetenzen und Fähigkeiten in allen Bevölkerungsgruppen zu fördern und zu unterstützen, einschliesslich spezifischer Spitzenkompetenz für Personen, die für die Identifizierung, Bewertung, Vermeidung und Minderung von Risiken im Zusammenhang mit KI-Systemen zuständig sind.

Die Begriffe der digitalen Kompetenzen und Fähigkeiten beziehen sich auf die Kompetenz, digitale Technologien, einschliesslich KI-Technologien, effizient zu nutzen, zu verstehen und mit ihnen zu interagieren. Dies sind die grundlegenden Fähigkeiten, die die breite Bevölkerung benötigt, um die Chancen und Risiken von KI-Systemen zu verstehen.[211]

Digitale Kompetenzen und Fähigkeiten sind besonders wichtig für diejenigen, die KI bei ihrer Arbeit nutzen. Dazu gehören beispielsweise Mitarbeitende, die ein KI-System in einem Entscheidungsfindungsprozess einsetzen, oder Personen, die für die Beschaffung von KI-Systemen in öffentlichen Verwaltungen zuständig sind. Die beteiligten Personen müssen in der Lage sein, die Herausforderungen zu verstehen, die KI-Systeme mit sich bringen. Solche Kompetenzen ermöglichen es beispielsweise, der Automatisierungs- oder Bestätigungsverzerrung entgegenzuwirken, die dazu führt, dass Menschen Maschinen und technologischen Artefakten mehr Vertrauen schenken als ihrem eigenen, möglicherweise widersprüchlichen Urteilsvermögen, und daher algorithmische Ergebnisse blind und ohne sie zu hinterfragen bestätigen.

Artikel 20 bezieht sich auf die spezifischen Fachkenntnisse der Personen, die gemäss Artikel 16 der Konvention am Rahmen für das Risiko- und Folgenmanagement beteiligt sind. Es versteht sich von selbst, dass diese Personen über die spezifischen Fähigkeiten verfügen müssen, die für die ordnungsgemässe Durchführung der in dieser Bestimmung vorgesehenen Risiko- und Folgenanalyse erforderlich sind.[212]

Der Bund könnte in seinen Kompetenzbereichen Massnahmen ergreifen, zum Beispiel bestimmte Anforderungen an die Ausbildung derjenigen Mitarbeitenden vorsehen, die zur Erfüllung ihrer Aufgaben KI-Systeme einsetzen.

4.3.5.6 Artikel 21 – Schutz der bestehenden Menschenrechte

Diese Bestimmung soll die Koexistenz der KI-Konvention mit anderen internationalen Menschenrechtsübereinkommen und -instrumenten sicherstellen. Sie besagt, dass keine Bestim-

[211] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 122.

[212] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 124.

mung des Übereinkommens so ausgelegt werden darf, dass sie eine einschränkende, abweichende oder nachteilige Wirkung auf die Menschenrechte oder andere damit zusammenhängende Rechte und Rechtspflichten hat, die durch das innerstaatliche Recht einer Vertragspartei oder durch eine andere einschlägige internationale Übereinkunft garantiert sind.

Jede Bezugnahme in der Konvention auf innerstaatliches Recht ist daher so auszulegen, dass sie sich auf Fälle beschränkt, in denen das innerstaatliche Recht ein höheres Schutzniveau für die Menschenrechte vorsieht als das einschlägige Völkerrecht.[213]

Hinsichtlich der Umsetzung in innerstaatliches Recht ist zu diesem Artikel nichts anzumerken.

4.3.5.7 Artikel 22 – Umfassenderer Schutz

Diese Bestimmung schützt die Bestimmungen des innerstaatlichen Rechts und bestehender oder künftiger verbindlicher völkerrechtlicher Übereinkommen, die einen zusätzlichen Schutz für Tätigkeiten im Lebenszyklus von KI-Systemen vorsehen, der über den durch die Konvention gewährleisteten Schutz hinausgeht. Dieser Schutz darf durch die Konvention nicht eingeschränkt werden.[214]

Hinsichtlich der Umsetzung in innerstaatliches Recht ist zu diesem Artikel nichts anzumerken.

4.4 Kapitel VII: Nachfolgemechanismus und Zusammenarbeit

4.4.1 Allgemeines

Dieses Kapitel enthält die Bestimmungen, die den in Artikel 1 Absatz 3 der Konvention vorgesehenen Nachfolgemechanismus bilden. Ziel dieses Mechanismus ist es, die wirksame Umsetzung der KI-Konvention durch die Vertragsparteien zu gewährleisten. Von den vorgesehenen Instrumenten sind die ersten drei auf internationaler Ebene (Art. 23–25) und das vierte auf nationaler Ebene (Art. 26) angesiedelt.

4.4.2 Artikel 23 – Konferenz der Vertragsparteien

Dieser Artikel sieht die Einrichtung eines Organs für die Konvention vor, namentlich die Konferenz der Vertragsparteien, die sich aus Vertretern der Vertragsparteien zusammensetzt. Die Einrichtung dieses Organs gewährleistet die gleichberechtigte Beteiligung aller Vertragsparteien am Entscheidungs- und Überwachungsprozess der Konvention und stärkt die Zusammenarbeit zwischen den Vertragsparteien, um eine angemessene und wirksame Implementierung der Konvention sicherzustellen.[215]

---

[213] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 126.

[214] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 127.

[215] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 130.

Dieses Organ wird seine eigene Geschäftsordnung verabschieden (Art. 23 Abs. 4). Es wird von der Generalsekretärin oder dem Generalsekretär des Europarats je nach Bedarf einberufen, in jedem Fall aber, wenn die Mehrheit der Vertragsparteien oder des Ministerkomitees die Einberufung verlangt (Art. 23 Abs. 3).

Die Konferenz der Vertragsparteien verfügt über die üblichen Überwachungsbefugnisse. Hervorzuheben ist insbesondere die Möglichkeit, Änderungen der Konvention durch Vorschläge gemäss Artikel 28 der Konvention vorzunehmen. Darüber hinaus hat die Vertragsparteienkonferenz eine allgemeine beratende Funktion in Bezug auf die Konvention. Sie kann spezifische Empfehlungen zu allen Fragen der Auslegung oder der Anwendung der Konvention abgeben, beispielsweise durch Vorschläge zur Auslegung verschiedener in der Konvention verwendeter Rechtsbegriffe. Diese Empfehlungen sind zwar nicht rechtsverbindlich, können aber als Ausdruck eines gemeinsamen Verständnisses der Vertragsparteien zu einer bestimmten Frage angesehen werden, das von den Vertragsparteien bei der Umsetzung der Konvention nach Treu und Glauben zu berücksichtigen ist.[216]

Im Falle eines Beitritts der Schweiz zur Konvention muss sie eine aus einer oder mehreren Personen bestehende Vertretung für die Konferenz der Vertragsparteien ernennen.

4.4.3 Artikel 24 – Berichterstattungspflicht

Zur Erleichterung der Zusammenarbeit und zur regelmässigen Unterrichtung über die Umsetzung der Konvention legt jede Vertragspartei der Konferenz der Vertragsparteien innerhalb von zwei Jahren nach dem Tag, an dem sie Vertragspartei geworden ist, und danach in regelmässigen Abständen einen Bericht vor, in dem die von ihr ergriffenen Massnahmen zur Umsetzung von Artikel 3 Absatz 1 Buchstaben a und b (vgl. Ziff. 4.2.3.1) im Einzelnen aufgeführt sind.

Die Konferenz der Vertragsparteien legt die Form und das Verfahren für die Vorlage des Berichts in Übereinstimmung mit ihrer Geschäftsordnung fest.

Wenn die Schweiz der Konvention beitritt, muss sie gemäss dieser Bestimmung und gemäss den von der Konferenz der Vertragsparteien festgelegten Modalitäten Bericht erstatten.

4.4.4 Artikel 25 – Internationale Zusammenarbeit

Diese Bestimmung legt die Verpflichtung der Vertragsparteien fest, bei der Umsetzung der Ziele der Konvention zusammenzuarbeiten.

Darüber hinaus werden die Vertragsparteien ermutigt, gegebenenfalls Nichtvertragsstaaten dabei zu unterstützen, im Einklang mit den Bestimmungen der Konvention zu handeln und dieser beizutreten (Abs. 1).

---

[216] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 132, Bst. c.

Darüber hinaus tauschen die Vertragsparteien relevante und nützliche Informationen über Aspekte von KI aus, die erhebliche positive oder negative Auswirkungen auf die Ausübung der Menschenrechte, das Funktionieren der Demokratie und die Achtung der Rechtsstaatlichkeit haben können, einschliesslich der im privaten Sektor festgestellten Risiken und Auswirkungen. Dabei werden sie ermutigt, gegebenenfalls relevante Interessengruppen und Staaten, die nicht Vertragsparteien der Konvention sind, einzubeziehen (Abs. 2).

Schliesslich werden die Vertragsparteien ermutigt, die Zusammenarbeit zu verstärken, gegebenenfalls auch mit relevanten Teilnehmergruppen (z. B. Vertreterinnen und Vertreter von Nichtregierungsorganisationen), um Risiken und nachteilige Auswirkungen auf Menschenrechte, Demokratie und Rechtsstaatlichkeit im Zusammenhang mit Tätigkeiten während des Lebenszyklus von KI-Systemen zu vermeiden und zu mindern.

Diese Bestimmung ist eher vage und lässt den Parteien relativ freie Hand bei der Art der geplanten Zusammenarbeit. Es handelt sich hier eher um einen Informationsaustausch als um eine operative Zusammenarbeit zwischen den Behörden.

4.4.5 Artikel 26 – Wirksame Aufsichtsmechanismen

Artikel 26 verpflichtet die Vertragsparteien, einen oder mehrere wirksame Mechanismen zur Aufsicht über die Einhaltung der Verpflichtungen aus der Konvention einzurichten oder zu bestimmen. Dies bedeutet, dass die Vertragsparteien die Mechanismen ihrer jeweiligen Rechtsordnungen überprüfen und gegebenenfalls deren Aufgaben neu definieren oder sogar gänzlich neue Strukturen schaffen müssen.

Gemäss dem erläuternden Bericht müssen diese Mechanismen funktionell von der Exekutive und der Legislative unabhängig sein. Dieser Begriff umfasst verschiedene Arten der funktionellen Unabhängigkeit. Dabei kann es sich um Kontrollfunktionen handeln, die in bestimmte Regierungsorgane integriert sind und die Entwicklung und den Einsatz von KI-Systemen beurteilen oder überwachen.

Die betreffenden Stellen müssen auch über die zur wirksamen Erfüllung ihrer Aufgaben erforderlichen Befugnisse, Fachkenntnisse, einschliesslich technischer Kenntnisse und Fähigkeiten, und sonstigen Ressourcen verfügen. Im Falle einer gemeinsamen Aufsicht durch mehrere Stellen verlangt das Übereinkommen eine wirksame Zusammenarbeit.[217]

Wirksame Aufsichtsmechanismen sollten auch im privaten Sektor vorgesehen werden, wo eine direkte oder indirekte horizontale Wirkung der Grundrechte zwischen Privaten besteht oder in Zukunft erkannt wird (vgl. Ziff. 4.2.3.1).

Diese Bestimmung verlangt ein Tätigwerden des Gesetzgebers, um entweder eine oder mehrere bestehende Behörden mit der Aufsicht über die Einhaltung der Konvention zu betrauen oder eine neue Behörde zu schaffen. Im schweizerischen Recht gibt es verschie-

[217] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 141 ff.

dene unabhängige Aufsichtsbehörden (z. B. EDÖB), aber keine, die den gesamten Anwendungsbereich der Konvention abdeckt. Der Gesetzgeber sollte die Organisation und die Eingriffsbefugnisse der Aufsichtsbehörde(n) im Zusammenhang mit der Konvention regeln. Wären mehrere Aufsichtsbehörden zuständig, sollte zudem die Koordination zwischen ihnen geregelt werden.

4.5 Kapitel VIII: Schlussbestimmungen

Die Bestimmungen der Artikel 27 bis 36 lehnen sich mit wenigen Ausnahmen im Wesentlichen an das Muster für die Schlussbestimmungen der im Rahmen des Europarats geschlossenen Übereinkommen, Zusatzprotokolle und Änderungsprotokolle an, das vom Ministerkomitee des Europarats anlässlich der 1291. Versammlung der Ministerdelegierten vom 5. Juli 2017 verabschiedet wurde.[218]

Besonders hervorzuheben ist Artikel 30 über die Unterzeichnung und das Inkrafttreten der Konvention. Nach Absatz 1 liegt das Übereinkommen für die Mitgliedstaaten des Europarats, für die Nichtmitgliedstaaten des Europarats, die an der Ausarbeitung der Konvention beteiligt waren (Argentinien, Australien, Kanada, Costa Rica, der Heilige Stuhl, Israel, Japan, Mexiko, Peru, die Vereinigten Staaten von Amerika und Uruguay) sowie für die Europäische Union zur Unterzeichnung auf.

Nach Artikel 30 Absatz 3 tritt die Konvention am ersten Tag des Monats in Kraft, der auf einen Zeitabschnitt von drei Monaten nach dem Tag folgt, an dem fünf Unterzeichnerstaaten, darunter mindestens drei Mitgliedstaaten des Europarats, erklärt haben, durch die Konvention gebunden zu sein (durch Ratifikation, Annahme oder Genehmigung, vgl. Art. 30 Abs. 2).

Nach Inkrafttreten der Konvention können andere Nichtmitgliedstaaten, die sich nicht an der Ausarbeitung der Konvention beteiligt haben, vom Ministerkomitee des Europarats eingeladen werden, dem Übereinkommen gemäss Artikel 31 beizutreten.

4.6 Zwischenfazit

Mit der vorliegenden Analyse der KI-Konvention konnte der gesetzgeberische Handlungsbedarf im schweizerischen Recht für den Fall einer Ratifikation der Konvention durch die Schweiz skizziert werden.

Unabhängig von der Frage der Ratifizierung ermöglichte diese Analyse auch eine Beurteilung des Schutzniveaus des schweizerischen Rechts im Hinblick auf die Herausforderungen, die KI-Systeme heute für Menschenrechte, Demokratie und Rechtsstaatlichkeit im Allgemeinen darstellen, da die KI-Konvention die grossen, international anerkannten Herausforderungen beleuchtet. Die KI-Konvention stellt somit eine Art Leitfaden für die Überprüfung des aktuellen Stands des schweizerischen Rechts in diesen Bereichen dar.

Es ist darauf hinzuweisen, dass die Analyse in Bezug auf bestimmte Aspekte vertieft werden muss, um die Gesamtheit der bestehenden einschlägigen Rechtsvorschriften einzubeziehen.

---

[218] Erläuternder Bericht zur KI-Konvention (Fn. 20), N 145.

Die vorliegende Analyse ist nur ein erster Schritt im Hinblick auf eine umfassendere Untersuchung, die gegebenenfalls im Rahmen der weiteren Arbeiten nach dem Grundsatzentscheid des Bundesrates zur Regulierung der KI durchgeführt wird.

Die Analyse hat folgendes Zwischenfazit hervorgebracht:

- Ziel der KI-Konvention ist der Schutz der Menschenrechte, einer funktionierenden Demokratie und der Rechtsstaatlichkeit. Die KI-Konvention richtet sich grundsätzlich an Staaten und enthält Normen, die nicht unmittelbar anwendbar sind.

- Die KI-Konvention schreibt keine genauen und konkreten Massnahmen zur Erreichung ihrer Ziele vor. Die Vertragsparteien haben daher einen grossen Ermessensspielraum bei der Umsetzung der Konvention.

  So legt die KI-Konvention beispielsweise keine allgemeinen Zulassungsvoraussetzungen für KI-Systeme fest, sondern bleibt auf einer allgemeineren Ebene. Zum Verhältnis zwischen der KI-Konvention und der KI-Verordnung siehe Kapitel Ziff. 5.3.4.

- Für den Handlungsbedarf des Gesetzgebers im Falle einer Ratifikation hat die Analyse drei Konstellationen ergeben:

  - Hinsichtlich einiger Bestimmungen der Konvention scheint das schweizerische Recht ein Schutzniveau zu bieten, das den Anforderungen der Konvention genügt, so dass ein Tätigwerden des Gesetzgebers derzeit nicht erforderlich erscheint. Dies gilt beispielsweise für Artikel 5 (Integrität demokratischer Prozesse und Achtung der Rechtsstaatlichkeit) und Artikel 19 (Öffentliche Konsultation).

  - Hinsichtlich anderer Bestimmungen der Konvention enthält das schweizerische Recht bereits einschlägige Bestimmungen, die jedoch im Vergleich zu den Verpflichtungen der Konvention nicht weit genug gehen. Zur Umsetzung der Konventionsbestimmungen wären daher Anpassungen notwendig. Zu erwähnen sind insbesondere Artikel 8 (Transparenz und Aufsicht), Artikel 13 (Sichere Innovation), Artikel 14 (Rechtsmittel) und Artikel 15 (Verfahrensgarantien).

    Die Umsetzung einzelner Schlüsselprinzipien, wie z. B. jenes in Artikel 8 (Transparenz und Aufsicht), ermöglicht es zudem, die Wirksamkeit des bereits bestehenden rechtlichen Rahmens des schweizerischen Rechts zu verbessern, z. B. in den Bereichen Verantwortlichkeit, Gleichstellung und Nichtdiskriminierung sowie Datenschutz.

  - Für andere Bestimmungen sieht das schweizerische Recht keine entsprechenden Normen vor. Dies gilt insbesondere für Artikel 16 (Rahmen für das Risiko- und Folgenmanagement) und Artikel 26 (Wirksame Aufsichtsmechanismen). Hier könnte der Gesetzgeber erwägen, neue Massnahmen zu ergreifen, da das schweizerische Recht nur punktuelle Aspekte regelt.

Zusammenfassend kann festgehalten werden, dass der geltende Rechtsrahmen zwar relevante Bestimmungen enthält, aber in vielen Fällen ergänzt werden müsste. Daraus lässt sich schliessen, dass im Falle einer Ratifizierung der KI-Konvention durch die Schweiz Anpassungen notwendig wären.

- Über Art und Umfang der zu treffenden Massnahmen hat gegebenenfalls der Gesetzgeber zu entscheiden. Bei der Ausarbeitung eines Gesetzgebungsvorhabens selbst ist nach den Grundsätzen der Gesetzgebungstechnik und unter Berücksichtigung der zu treffenden politischen Entscheidungen auch festzulegen, wie die neuen Normen in das geltende Recht integriert werden sollen. So ist zum Beispiel zu prüfen, ob durch die neuen Massnahmen ein oder mehrere bestehende Gesetze ergänzt werden sollen, oder ob ein neues Gesetz zu erlassen ist. Hinsichtlich der Regelungsdichte sind insbesondere die Vorgaben von Artikel 164 BV zu beachten.

  Es zeichnet sich jedoch bereits jetzt ab, dass aufgrund der möglicherweise betroffenen Rechtsbereiche voraussichtlich mehrere Gesetze angepasst werden müssen. Bereits jetzt ist ein erheblicher Koordinationsbedarf absehbar. Ein solcher besteht insbesondere mit dem DSG.

- Die Analyse hat auch gezeigt, dass die Verhältnismässigkeit ein wichtiger Aspekt bei der Umsetzung ist. Die Tragweite der Verpflichtungen aus der Konvention sollte an die Schwere möglicher Verletzungen angepasst werden. Auch hier ist die Frage, wie diese Abstufung konkret aussehen sollte, in erster Linie von gesetzgeberischer und politischer Natur.

- Hinsichtlich des Privatsektors hat die Analyse gezeigt, dass sich der Anwendungsbereich der Konvention auf Fälle beschränken würde, in denen eine direkte oder indirekte horizontale Wirkung der Grundrechte zwischen Privaten besteht oder in Zukunft erkannt wird (vgl. Ziff. 4.2.3.1). Die KI-Konvention lässt Raum für Besonderheiten des schweizerischen Rechts.

5 Verordnung der Europäischen Union zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz

5.1 Struktur des Kapitels und Methode

In diesem Kapitel wird zunächst der Inhalt der KI-Verordnung präsentiert (vgl. Ziff. 5.2).

Danach folgt eine rechtliche Würdigung der KI-Verordnung, in der einige Aspekte, die aus Sicht des Schweizer Rechts relevant sind, analysiert werden (vgl. Ziff. 5.3).

In diesem Rahmen wird auf die rechtlichen Auswirkungen der KI-Verordnung auf die Schweizer Akteure (vgl. Ziff. 5.3.1), das Verhältnis zwischen der KI-Verordnung und dem Abkommen zwischen der Schweizerischen Eidgenossenschaft und der Europäischen Union über die gegenseitige Anerkennung von Konformitätsbewertungen (MRA) (vgl. Ziff. 5.3.2) und das Verhältnis zum Angemessenheitsbeschluss der Europäischen Kommission im Bereich Datenschutz (vgl. Ziff. 5.3.3) eingegangen. Das Verhältnis zur KI-Konvention wird ebenfalls thematisiert (vgl. Ziff. 5.3.4). Anschliessend werden weitere ausgewählte Aspekte erörtert (vgl. Ziff. 5.3.5), bevor das Kapitel mit einem Zwischenfazit abgeschlossen wird (vgl. Ziff. 5.4).

Aus methodischer Sicht weicht die Präsentation von der vorstehenden Analyse der KI-Konvention des Europarats ab (vgl. Ziff. 4). Die KI-Verordnung ist eine EU-interne Regelung und ist für die Schweiz nicht bindend. Grundsätzlich hat diese Verordnung derzeit im Schweizer Recht keine Entsprechung. Eine Analyse, in der jede ihrer Bestimmungen in Bezug auf das Schweizer Recht untersucht wird, wie es bei der KI-Konvention gemacht wurde, erscheint daher derzeit nicht angebracht. Stattdessen soll der Inhalt der KI-Verordnung präsentiert und gewisse relevante Aspekte hervorgehoben werden, um so eine allgemeine Würdigung aus Sicht des Schweizer Rechts vornehmen zu können. Da die Verordnung in der EU erst vor Kurzem verabschiedet wurde und noch viele Fragen hinsichtlich ihrer Anwendung geklärt werden müssen, wird hier nur auf die wichtigen Herausforderungen eingegangen, die für die Schweiz derzeit relevant sind.

5.2 Inhalt der Verordnung

5.2.1 Ausgangslage

Am 21. April 2021 hat die Europäische Kommission ein Legislativpaket zu KI vorgestellt, das einen Vorschlag für eine Verordnung zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und einen Vorschlag für eine Verordnung über Maschinen und Geräte (neue Konformitätsverfahren, die u. a. das Internet der Dinge regeln) enthielt.[219]

Aus Sicht der Europäischen Kommission will die EU mit dieser Verordnung ihre Rolle als Pionierin bei der Regulierung der digitalen Wirtschaft und als Vorreiterin bei der Festlegung von

[219] Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates vom 21. April 2021 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Gesetz über künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union, COM/2021/206 final, abrufbar unter: https://eur-lex.europa.eu/legal-content/DE/ALL/?uri=CELEX%3A52021PC0206 (abgerufen am 26. August 2024); Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates über Maschinenprodukte, COM(2021) 202 final, abrufbar unter: https://eur-lex.europa.eu/legal-content/DE/ALL/?uri=CELEX%3A52021PC0202 (abgerufen am 26. August 2024).

Standards festigen, die Einfluss nehmen sollen auf die internationale Debatte zum Umgang mit KI.

Die KI-Verordnung stellt die erste horizontale und direkt verbindliche KI-Regulierung auf regionaler Ebene dar. Sie deckt insbesondere die Entwicklung, das Inverkehrbringen, die Inbetriebnahme und die Verwendung von KI-Systemen in der EU ab. Sie wurde vom Europäischen Parlament am 13. März 2024 und vom Rat am 21. Mai 2024 verabschiedet, am 12. Juni 2024 unterzeichnet und im Amtsblatt der EU am 12. Juli 2024 veröffentlicht. Sie ist am 1. August 2024 in Kraft getreten.[220]

Diese Verordnung ist für die Schweiz nicht bindend.

5.2.2 Ziele der Regelung

Gemäss Artikel 1 Absatz 1 besteht der Zweck der Verordnung darin, das Funktionieren des Binnenmarkts zu verbessern und die Einführung einer auf den Menschen ausgerichteten und vertrauenswürdigen KI zu fördern und gleichzeitig ein hohes Schutzniveau in Bezug auf Gesundheit, Sicherheit und die in der EU-Charta verankerten Grundrechte, einschliesslich Demokratie, Rechtsstaatlichkeit und Umweltschutz, vor schädlichen Auswirkungen von KI-Systemen in der EU zu gewährleisten und die Innovation zu unterstützen.

Das Ziel besteht also darin, einen einheitlichen Rechtsrahmen, insbesondere für die Entwicklung, das Inverkehrbringen, die Inbetriebnahme und die Verwendung von KI-Systemen in der EU, zu schaffen.[221] Die Verordnung soll gewährleisten, dass die in der EU in Verkehr gebrachten KI-Systeme sicher sind und den geltenden europäischen Normen entsprechen, namentlich im Bereich der Produktsicherheit.[222] Sie soll auch den grenzüberschreitenden Verkehr KI-gestützter Waren und Dienstleistungen ermöglichen und eine Fragmentierung des Markts verhindern.

Die Wendung «und gleichzeitig ein hohes Schutzniveau in Bezug auf Gesundheit, Sicherheit und die in der Charta verankerten Grundrechte [...] zu gewährleisten» (vgl. Art. 1 Abs. 1) zeigt, dass der europäische Gesetzgeber über diesen auf der Produktsicherheit basierenden Regulierungsansatz auch andere Interessen schützen wollte, insbesondere die Grundrechte der betroffenen Personen. Dies führt zu einer Verordnung mit hybriden Zielen (vgl. Ziff. 5.3.4).

5.2.3 Risikobasierter Ansatz

Die KI-Verordnung weist einen risikobasierten Ansatz auf. So nimmt sie eine Einstufung der KI-Systeme nach ihrem Risiko für die Gesundheit, die Sicherheit und die Grundrechte vor.

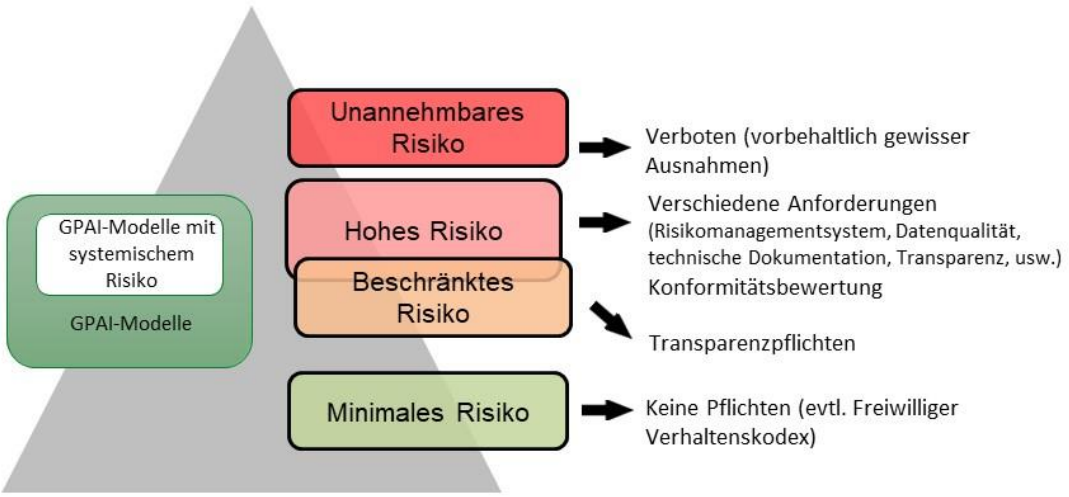
---

[220] Verordnung (EU) 2024/1689 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Fn. 3).

[221] DAVID ROSENTHAL, Der EU AI Act – Verordnung über künstliche Intelligenz, Jusletter 5. August 2024, 5 N 6.

[222] MARTINA ARIOLI, Risikomanagement nach der EU-Verordnung über Künstliche Intelligenz, Jusletter IT 4. Juli 2024, 4 N 8 f.

Gestützt auf diese Einstufung sieht die Verordnung verschiedene Pflichten für den Marktzugang vor.[223]



Quelle: BJ

Gemäss diesem Ansatz sind KI-Systeme, die unannehmbare Risiken aufweisen, bis auf einige Ausnahmen verboten (vgl. Ziff. 5.2.6).

KI-Systeme, die als hochriskant eingestuft werden, sind zulässig, müssen aber viele Anforderungen erfüllen, damit sie in der EU in Verkehr gebracht werden können (vgl. Ziff. 5.2.7). Insbesondere müssen sie einer Konformitätsbewertung unterzogen werden (vgl. Ziff. 5.2.7.4). In ihrer Folgenabschätzung der Verordnung[224] kam die Europäische Kommission zum Schluss, dass die Hochrisiko-KI-Anwendungen schätzungsweise nur 5 bis 15 Prozent aller Anwendungen auf dem Markt ausmachen.

KI-Systeme, die nur ein beschränktes Risiko aufweisen, unterliegen lediglich Transparenzpflichten (vgl. Ziff. 5.2.8). Wenn diese Systeme jedoch die Kriterien erfüllen, um als hochriskant eingestuft zu werden, gelten die für diese Systeme vorgesehenen Pflichten ebenfalls.

---

[223] Vgl. ANGELA MÜLLER, Der Artificial intelligence Act der EU: Ein risikobasierter Ansatz zur Regulierung von Künstlicher Intelligenz, Zeitschrift für Europarecht 1/2022, 1 ff., 7 ff. und 15 f. (im Zusammenhang mit einer früheren Version der KI-Verordnung).

[224] Commission staff working document, Impact assessment – Accompanying the Proposal for a Regulation of the European Parliament and of the Council, Laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative act, SWD(2021) 84 final, 68 (abrufbar unter: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021SC0084 (abgerufen am 26. August 2024).

Fällt ein KI-System unter keine der oben erwähnten Kategorien, so wird es als KI-System mit minimalem Risiko betrachtet. Gemäss der Europäischen Kommission weist die Mehrheit der KI-Systeme ein minimales Risiko auf. Die Verordnung sieht für diese Systeme daher kein staatliches Eingreifen vor. Sie erlaubt die freie Verwendung von Anwendungen wie Videospielen mit integrierten KI-Systemen oder KI-basierten Spam-Filtern. In diesem Fall werden die betroffenen Akteure angehalten, Verhaltenskodizes zu erstellen.

Die Verordnung regelt auch die KI-Modelle mit allgemeinem Verwendungszweck (Englisch: «general purpose AI model», im Folgenden: GPAI-Modelle), die im ursprünglichen Verordnungsvorschlag nicht enthalten waren. Die Verordnung sieht nun Pflichten für alle GPAI-Modelle und zusätzliche Pflichten für GPAI-Modelle vor, die systemische Risiken beinhalten (vgl. Ziff. 5.2.10).

5.2.4 Begriffsbestimmungen

In Artikel 3 der KI-Verordnung werden gewisse Begriffe definiert (insgesamt 68). Diese Analyse konzentriert sich auf die Begriffe «KI-System» und «KI-Modell mit allgemeinem Verwendungszweck».

5.2.4.1 System der künstlichen Intelligenz

Gemäss Artikel 3 Nummer 1 der KI-Verordnung ist ein KI-System «ein maschinengestütztes System, das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können».

Die Verordnung will zukunftssicher sein und sowohl aktuelle als auch künftige technologische Entwicklungen im KI-Bereich abdecken. So lehnt sich die Begriffsbestimmung der EU für das KI-System eng an die Arbeiten der internationalen Organisationen an, die im KI-Bereich tätig sind, wie die OECD, um die Rechtssicherheit zu gewährleisten sowie die internationale Konvergenz und die allgemeine Akzeptanz zu fördern.

Angesichts der Ähnlichkeiten dieser Definition des KI-Systems mit jener in der KI-Konvention, die sich ebenfalls an den Arbeiten der OECD orientiert, wird für weitere Erläuterungen auf die Ausführungen zur Begriffsbestimmung der KI-Konvention verwiesen (vgl. Ziff. 4.2.2.1).

Ausserdem wird die Europäische Kommission Leitlinien für die Anwendung der Definition eines KI-Systems erarbeiten (vgl. Art. 96 Abs. 1 Bst. f).

5.2.4.2 Modell der künstlichen Intelligenz mit allgemeinem Verwendungszweck

In der Verordnung wird zwischen KI-Modellen mit allgemeinem Verwendungszweck und KI-Systemen unterschieden.

Gemäss Artikel 3 Nummer 63 ist ein KI-Modell mit allgemeinem Verwendungszweck «ein KI-Modell – einschliesslich der Fälle, in denen ein solches KI-Modell mit einer grossen Datenmenge unter umfassender Selbstüberwachung trainiert wird –, das eine erhebliche allgemeine Verwendbarkeit aufweist und in der Lage ist, unabhängig von der Art und Weise seines Inverkehrbringens ein breites Spektrum unterschiedlicher Aufgaben kompetent zu

erfüllen, und das in eine Vielzahl nachgelagerter Systeme oder Anwendungen integriert werden kann, ausgenommen KI-Modelle, die vor ihrem Inverkehrbringen für Forschungs- und Entwicklungstätigkeiten oder die Konzipierung von Prototypen eingesetzt werden».

Die Begriffsbestimmung des KI-Modells mit allgemeinem Verwendungszweck beruht auf den wesentlichen funktionalen Merkmalen dieser Art von Modell, insbesondere auf der allgemeinen Verwendbarkeit und der Fähigkeit, ein breites Spektrum unterschiedlicher Aufgaben kompetent zu erfüllen. Bespiele dafür sind «GPT» und «Llama».[225]

Aus dieser Begriffsbestimmung geht hervor, dass KI-Modelle mit allgemeinem Verwendungszweck wesentliche Komponenten von KI-Systemen sind, aber für sich allein keine KI-Systeme darstellen. Damit KI-Modelle mit allgemeinem Verwendungszweck zu KI-Systemen werden, ist die Hinzufügung weiterer Komponenten, beispielsweise einer Nutzerschnittstelle, erforderlich. Daher sind diese Modelle in der Regel in KI-Systeme integriert und Teil davon.

Wenn ein KI-System auf einem KI-Modell mit allgemeinem Verwendungszweck beruht, spricht man von KI-System mit allgemeinem Verwendungszweck. Dabei handelt es sich um ein System, das in der Lage ist, einer Vielzahl von Zwecken sowohl für die direkte Verwendung als auch für die Integration in andere KI-Systeme zu dienen (vgl. Art. 3 Nr. 66).

5.2.5 Geltungsbereich

5.2.5.1 Umfang des Geltungsbereichs

Die KI-Verordnung hat einen weiten Geltungsbereich. Sie gilt sowohl für Private als auch für Behörden.[226]

Um den Geltungsbereich zu verstehen, ist zunächst zu beachten, dass in der Verordnung zwischen folgenden Rollen unterschieden wird:

- Anbieter: Dabei handelt es sich um jede natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, die ein KI-System oder ein KI-Modell mit allgemeinem Verwendungszweck entwickelt oder entwickeln lässt *und* es unter ihrem eigenen Namen oder ihrer Handelsmarke in Verkehr bringt oder das KI-System unter ihrem eigenen Namen oder ihrer Handelsmarke in Betrieb nimmt, sei es entgeltlich oder unentgeltlich. Er kann in der EU oder in einem Drittland niedergelassen sein (vgl. Art. 3 Nr. 3).

  Unter «Inverkehrbringen» ist die erstmalige Bereitstellung eines KI-Systems oder eines KI-Modells mit allgemeinem Verwendungszweck auf dem EU-Markt zu verstehen (vgl. Art. 3 Nr. 9).

[225] DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 3 N 5.

[226] DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 2 N 2; 6 N 11.

Mit «Inbetriebnahme» ist die Bereitstellung eines KI-Systems zum Erstgebrauch direkt an den Betreiber oder zum Eigengebrauch entsprechend seiner Zweckbestimmung gemeint (vgl. Art. 3 Nr. 11).

- Betreiber: Hier handelt es sich um jede natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, die ein KI-System in eigener Verantwortung verwendet, es sei denn, diese Verwendung erfolgt im Rahmen einer persönlichen und nicht beruflichen Tätigkeit (vgl. Art. 3 Nr. 4).

- Einführer: Jede in der EU ansässige oder niedergelassene natürliche oder juristische Person, die ein KI-System in Verkehr bringt, das den Namen oder die Handelsmarke einer in einem Drittland niedergelassenen natürlichen oder juristischen Person trägt (vgl. Art. 3 Nr. 6).

- Händler: Jede natürliche oder juristische Person in der Lieferkette, die ein KI-System auf dem EU-Markt bereitstellt, mit Ausnahme des Anbieters oder des Einführers (vgl. Art. 3 Nr. 7).

- Akteur: ein Anbieter, Produkthersteller, Betreiber, Bevollmächtigter, Einführer oder Händler (vgl. Art. 3 Nr. 8).

Die KI-Verordnung gilt in erster Linie für die in der EU oder einem Drittland niedergelassenen oder ansässigen Anbieter, die in der EU KI-Systeme in Verkehr bringen oder in Betrieb nehmen oder KI-Modelle mit allgemeinem Verwendungszweck in Verkehr bringen (vgl. Art. 2 Abs. 1 Bst. a).

Sie gilt auch für Betreiber von KI-Systemen, die ihren Sitz in der EU haben oder sich in der EU befinden (vgl. Art. 2 Abs. 1 Bst. b).

Weiter gilt die Verordnung für Anbieter und Betreiber von KI-Systemen, die ihren Sitz in einem Drittland haben oder sich in einem Drittland befinden, wenn die vom KI-System hervorgebrachte Ausgabe in der EU verwendet wird (Art. 2 Abs. 1 Bst. c). Mit dieser Bestimmung soll verhindert werden, dass Massnahmen zur Umgehung der Rechtsvorschriften umgesetzt werden könnten. Auf die Frage, welche Auswirkungen die KI-Verordnung auf die Schweizer Akteure hat, wird weiter unten eingegangen (vgl. Ziff. 5.3.1).

Die Verordnung gilt ebenfalls für Einführer und Händler von KI-Systemen (Art. 2 Abs. 1 Bst. d), Produkthersteller, die ein KI-System zusammen mit ihrem Produkt unter ihrem eigenen Namen oder ihrer Handelsmarke in Verkehr bringen oder in Betrieb nehmen (Art. 2 Abs. 1 Bst. e) und Bevollmächtigte von Anbietern, die nicht in der EU niedergelassen sind (Art. 2 Abs. 1 Bst. f).

Ferner gilt sie für betroffene Personen, die sich in der EU befinden (Art. 2 Abs. 1 Bst. g). Dies betrifft beispielsweise die Rechtsbehelfe gemäss den Artikeln 85 ff. der Verordnung.

5.2.5.2 Ausnahmen

Die Verordnung sieht mehrere Ausnahmen von ihrer Anwendung vor. Dies sind namentlich Folgende:

- KI-Systeme, die ausschliesslich für militärische Zwecke, Verteidigungszwecke oder Zwecke der nationalen Sicherheit in Verkehr gebracht, in Betrieb genommen oder, mit oder ohne Änderungen, verwendet werden, unabhängig von der Art der Einrichtung, die diese Tätigkeiten ausübt, sind vom Geltungsbereich der Verordnung ausgenommen (Art. 2 Abs. 3). Der Bereich der nationalen Sicherheit ist nicht Bestandteil des EU-Rechts, sondern liegt in der ausschliesslichen Verantwortung der Mitgliedstaaten (Art. 4 Abs. 2 EUV). Der Ausschluss des Militär- und Verteidigungsbereichs gründet sich auf Artikel 4 Absatz 2 EUV sowie auf die Besonderheiten der Verteidigungspolitik der Mitgliedstaaten und der gemeinsamen Verteidigungspolitik der EU gemäss Titel V Kapitel 2 EUV, die dem Völkerrecht unterliegen. Wird ein KI-System jedoch vorübergehend oder ständig ausserhalb der erwähnten Bereiche für andere Zwecke verwendet, etwa für zivile oder humanitäre Zwecke oder für Zwecke der Strafverfolgung oder öffentlichen Sicherheit, so fällt dieses System in den Geltungsbereich der Verordnung (vgl. E. 24 der Verordnung).

- Die Verordnung gilt weder für Behörden in Drittländern noch für internationale Organisationen, die gemäss Artikel 2 Absatz 1 in den Geltungsbereich der Verordnung fallen, soweit diese Behörden oder Organisationen KI-Systeme im Rahmen internationaler Übereinkünfte im Bereich der Strafverfolgung und justiziellen Zusammenarbeit mit der EU oder mit einem oder mehreren Mitgliedstaaten verwenden, sofern Garantien hinsichtlich des Schutzes der Grundrechte und der Grundfreiheiten von Personen gewährt werden (Art. 2 Abs. 4).

- Für KI-Systeme, die als Hochrisiko-KI-Systeme gemäss Artikel 6 Absatz 1 eingestuft sind und im Zusammenhang mit Produkten stehen, die unter die in Anhang I Abschnitt B aufgeführten Harmonisierungsrechtsvorschriften fallen (z. B. Zivilluftfahrt, land- und forstwirtschaftliche Fahrzeuge, Schiffsausrüstung), gelten nur Artikel 6 Absatz 1 und die Artikel 102–109 und 112 der Verordnung (vgl. Art. 2 Abs. 2).

- Die Verordnung gilt nicht für KI-Systeme oder KI-Modelle, einschliesslich ihrer Ausgabe, die für den alleinigen Zweck der wissenschaftlichen Forschung und Entwicklung entwickelt und in Betrieb genommen werden (Art. 2 Abs. 6). Forschungs-, Test- und Entwicklungstätigkeiten fallen nicht in den Geltungsbereich der Verordnung, müssen aber im Einklang mit dem geltenden EU-Recht durchgeführt werden. Tests unter Realbedingungen sind von diesem Ausschluss nicht betroffen (Art. 2 Abs. 8 und Art. 60).

- Die Verordnung gilt nicht für die Pflichten von Betreibern, die natürliche Personen sind und KI-Systeme im Rahmen einer ausschliesslich persönlichen und nicht beruflichen Tätigkeit verwenden (Art. 2 Abs. 10).

- Ferner gilt die Verordnung ebenfalls nicht für KI-Systeme, die unter freien und quelloffenen Lizenzen bereitgestellt werden, es sei denn, sie werden als Hochrisiko-KI-Systeme oder als ein KI-System, das unter Artikel 5 (Verbotene Praktiken) oder 50 (Transparenzpflichten unterliegende KI-Systeme) fällt, in Verkehr gebracht oder in Betrieb genommen (vgl. Art. 2 Abs. 12).

5.2.6 Verbotene Praktiken

Gemäss dem oben beschriebenen risikobasierten Ansatz (vgl. Ziff. 5.2.3) verbietet das Kapitel II der KI-Verordnung acht Praktiken, die als unannehmbares Risiko betrachtet werden (vgl. Art. 5 Abs. 1):[227]

- KI-Systeme, die Techniken der unterschwelligen Beeinflussung einsetzen: Verboten sind das Inverkehrbringen, die Inbetriebnahme oder die Verwendung von KI-Systemen, die Techniken der unterschwelligen Beeinflussung oder absichtlich manipulative oder täuschende Techniken mit dem Ziel oder der Wirkung einsetzen, das Verhalten einer Person oder einer Gruppe von Personen wesentlich zu verändern, indem ihre Fähigkeit, eine fundierte Entscheidung zu treffen, deutlich beeinträchtigt wird, wodurch sie veranlasst wird, eine Entscheidung zu treffen, die sie andernfalls nicht getroffen hätte, und zwar in einer Weise, die dieser Person, einer anderen Person oder einer Gruppe von Personen erheblichen Schaden zufügt oder mit hinreichender Wahrscheinlichkeit zufügen wird (Art. 5 Abs. 1 Bst. a).

- KI-Systeme, die eine Vulnerabilität oder Schutzbedürftigkeit ausnutzen: Verboten sind Systeme, die eine Vulnerabilität oder Schutzbedürftigkeit einer natürlichen Person oder einer Gruppe von Personen aufgrund ihres Alters, einer Behinderung oder der sozialen oder wirtschaftlichen Situation mit dem Ziel oder der Wirkung ausnutzen, das Verhalten dieser Person oder einer dieser Gruppe angehörenden Person in einer Weise wesentlich zu verändern, die dieser Person oder einer anderen Person erheblichen Schaden zufügt oder mit hinreichender Wahrscheinlichkeit zufügen wird (Art. 5 Abs. 1 Bst. b).

- Soziale Bewertung durch öffentliche und private Akteure: Verboten sind Systeme, die natürliche Personen oder Gruppen von Personen auf der Grundlage ihres sozialen Verhaltens oder bekannter, abgeleiteter oder vorhergesagter persönlicher Eigenschaften oder Persönlichkeitsmerkmale bewerten oder einstufen, *wenn* die soziale Bewertung zu einem oder beiden der folgenden Ergebnisse führt (Art. 5 Abs. 1 Bst. c):

  - Schlechterstellung oder Benachteiligung in sozialen Zusammenhängen, die in keinem Zusammenhang zu den Umständen stehen, unter denen die Daten ursprünglich erzeugt oder erhoben wurden;

  - Schlechterstellung oder Benachteiligung in einer Weise, die im Hinblick auf ihr soziales Verhalten oder dessen Tragweite ungerechtfertigt oder unverhältnismässig ist.

- «Predictive Policing»: Verboten sind KI-Systeme, um das Risiko, dass eine natürliche Person eine Straftat begeht, ausschliesslich auf der Grundlage ihres Profilings oder der Bewertung ihrer persönlichen Merkmale und Eigenschaften zu beurteilen oder vorherzusagen. Das Verbot gilt nicht, wenn das System dazu verwendet wird, eine durch Menschen durchgeführte Bewertung zu unterstützen (Art. 5 Abs. 1 Bst. d).

---

[227] Für kritische Erwägungen s. DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 18 N 42 ff.

- Massenhafte Sammlung von Bildern zum Zweck der Gesichtserkennung: Verboten sind KI-Systeme, die Datenbanken zur Gesichtserkennung durch das ungezielte Auslesen von Gesichtsbildern aus dem Internet oder von Überwachungsaufnahmen erstellen oder erweitern (Art. 5 Abs. 1 Bst. e).

- Erkennung von Emotionen: Verboten sind KI-Systeme, mit denen Emotionen einer Person am Arbeitsplatz und in Bildungseinrichtungen abgeleitet werden können, es sei denn, die Verwendung des KI-Systems erfolgt aus medizinischen Gründen oder Sicherheitsgründen (Art. 5 Abs. 1 Bst. f).

- Systeme zur biometrischen Kategorisierung: Systeme, mit denen Personen auf der Grundlage ihrer biometrischen Daten, wie des Gesichts oder der Fingerabdrücke, kategorisiert werden, um ihre politischen Einstellungen, ihre Gewerkschaftszugehörigkeit, ihre religiösen oder weltanschaulichen Überzeugungen, ihre Rasse, ihr Sexualleben oder ihre sexuelle Ausrichtung abzuleiten, sind verboten (Art. 5 Abs. 1 Bst. g).

- Biometrische Echtzeit-Fernidentifizierung in öffentlich zugänglichen Räumen zu Strafverfolgungszwecken: Ein solches Vorgehen ist grundsätzlich verboten (Art. 5 Abs. 1 Bst. h). Der Begriff «Echtzeit» bedeutet, dass die Erfassung biometrischer Daten, der Abgleich und die Identifizierung ohne erhebliche Verzögerung erfolgen (vgl. Art. 3 Nr. 42).

  Die Verwendung dieser Technik wird als besonders starker Eingriff in die Rechte und Freiheiten der Personen betrachtet, da sie die Privatsphäre der Bevölkerung beeinträchtigen, ein Gefühl der ständigen Überwachung wecken und indirekt von der Ausübung der Versammlungsfreiheit und anderer Grundrechte abhalten kann. Zudem können die technischen Ungenauigkeiten von KI-Systemen, die für die biometrische Fernidentifizierung natürlicher Personen bestimmt sind, zu verzerrten Ergebnissen führen und eine diskriminierende Wirkung haben (E. 32 der Verordnung).

  Die Verwendung dieser Systeme ist nur in drei Situationen zulässig (vgl. Art. 5 Abs. 1 Bst. h): i) gezielte Suche nach bestimmten Opfern von Entführung, Menschenhandel oder sexueller Ausbeutung sowie die Suche nach vermissten Personen; ii) Abwenden einer konkreten, erheblichen und unmittelbaren Gefahr für das Leben oder die Sicherheit natürlicher Personen oder einer tatsächlichen und bestehenden oder tatsächlichen und vorhersehbaren Gefahr eines Terroranschlags; iii) Aufspüren oder Identifizieren von Personen, die beschuldigt oder verdächtigt werden, eine in Anhang II aufgeführte Straftat begangen zu haben, die im betreffenden Mitgliedstaat mit einer Freiheitsstrafe oder einer Freiheit beschränkenden sichernden Massnahme im Höchstmass von mindestens vier Jahren bedroht ist.

  Artikel 5 Absätze 2 und 3 legt die Voraussetzungen fest, unter denen die biometrische Fernidentifizierung in diesen ausserordentlichen Situationen eingesetzt werden darf. Insbesondere muss jede Verwendung von einer Justizbehörde oder einer unabhängigen Verwaltungsbehörde genehmigt werden. Eine solche Genehmigung muss grundsätzlich vor der Verwendung des Systems zur Identifizierung einer Person oder mehrerer Personen eingeholt werden, ausser in begründeten dringenden Fällen. In diesen Fällen ist es möglich, das System ohne Genehmigung zu verwenden, sofern diese unverzüglich, spätestens jedoch innerhalb von 24 Stunden beantragt wird. Die Genehmigung wird nur gewährt, wenn die Verwendung für das Erreichen eines der in Artikel 5

Absatz 1 Buchstabe h genannten Ziele notwendig und verhältnismässig ist und insbesondere wenn sie auf das unbedingt erforderliche Mass beschränkt bleibt.

Zudem wird die Verwendung dieser Systeme – ausser in Ausnahmefällen – nur gestattet, wenn die Strafverfolgungsbehörde eine Folgenabschätzung im Hinblick auf die Grundrechte gemäss Artikel 27 der Verordnung durchgeführt und das System gemäss Artikel 49 der Verordnung in der Datenbank registriert hat (vgl. Art. 5 Abs. 2 *in fine*) (vgl. Ziff. 5.2.7.3.3).

Die nationale Marktüberwachungsbehörde (vgl. Ziff. 5.2.14) und die nationale Datenschutzbehörde müssen über jeden Einsatz eines biometrischen Echtzeit-Fernidentifizierungssystems informiert werden und der Europäischen Kommission jährlich einen Bericht über die Verwendung dieser Systeme vorlegen. Auf der Grundlage der Berichte der Mitgliedstaaten veröffentlicht die Europäische Kommission jährlich einen Gesamtbericht (Art. 5 Abs. 6 und 7).

Innerhalb der in den Artikeln 5 Absatz 1 Buchstabe h und 5 Absätze 2 und 3 aufgeführten Grenzen können die Mitgliedstaaten in ihrem nationalen Recht detaillierte Vorschriften, namentlich für die Beantragung und Erteilung der Genehmigungen, vorsehen (vgl. Art. 5 Abs. 5).

5.2.7 Hochrisiko-Systeme der künstlichen Intelligenz

5.2.7.1 Einstufung

Gemäss der KI-Verordnung dürfen Hochrisiko-KI-Systeme nur in Verkehr gebracht, in Betrieb genommen oder verwendet werden, wenn sie bestimmte Bedingungen erfüllen.

Die Verordnung sieht zwei Kategorien von Hochrisiko-KI-Systemen vor:

- KI-Systeme, die als Sicherheitsbauteil eines Produkts (z. B. Spielzeug, Medizinprodukte) verwendet werden, das unter die in Anhang I aufgeführten Harmonisierungsrechtsvorschriften der EU fällt, sowie KI-Systeme, die selber solche Produkte sind, sofern diese Produkte einer Konformitätsbewertung durch Dritte im Hinblick auf ihr Inverkehrbringen oder ihre Inbetriebnahme gemäss den Harmonisierungsrechtsvorschriften der EU unterzogen werden müssen (Art. 6 Abs. 1).

- Die in Anhang III genannten KI-Systeme (Art. 6 Abs. 2), also KI-Systeme, die in acht spezifischen Bereichen eingesetzt werden:[228]

[228] Für kritische Erwägungen s. DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 22 N 51 ff.

- Biometrie[229];
- kritische Infrastruktur;
- allgemeine und berufliche Bildung;
- Beschäftigung, Personalmanagement und Zugang zur Selbstständigkeit;
- Zugänglichkeit und Inanspruchnahme grundlegender privater und grundlegender öffentlicher Dienste und Leistungen;
- Strafverfolgung;
- Migration, Asyl und Grenzkontrolle;
- Rechtspflege und demokratische Prozesse.

Die Europäische Kommission ist befugt, diese Liste zu ändern (vgl. Art. 7).

Allerdings gilt ein KI-System, obwohl es in Anhang III genannt ist, in bestimmten Fällen, wenn es die Entscheidungsfindung nicht wesentlich beeinflusst oder die geschützten rechtlichen Interessen nicht beeinträchtigt, nicht als hochriskant (vgl. Art. 6 Abs. 3 und E. 53):

- Das KI-System ist dazu bestimmt, eine eng gefasste Verfahrensaufgabe durchzuführen. Dabei kann es sich beispielsweise um ein System handeln, das unstrukturierte Daten in strukturierte Daten umwandelt, oder ein KI-System, das zur Erkennung von Duplikaten unter einer grossen Zahl von Anwendungen eingesetzt wird (E. 53 der Verordnung).

- Das KI-System ist dazu bestimmt, das Ergebnis einer zuvor abgeschlossenen menschlichen Tätigkeit zu verbessern, beispielsweise wenn ein KI-System zum Ziel hat, die in zuvor verfassten Dokumenten verwendete Sprache zu verbessern (E. 53 der Verordnung).

- Das KI-System ist dazu bestimmt, Entscheidungsmuster oder Abweichungen von früheren Entscheidungsmustern zu erkennen, und ist nicht dazu gedacht, die zuvor abgeschlossene menschliche Bewertung ohne eine angemessene menschliche Überprüfung zu ersetzen oder zu beeinflussen. Das Risiko wird als geringer betrachtet, da die Verwendung des KI-Systems einer zuvor abgeschlossenen menschlichen Bewertung folgt, die das KI-System ohne angemessene menschliche Überprüfung nicht ersetzen oder beeinflussen soll (E. 53).

[229] Diese Kategorie umfasst keine KI-Systeme, die bestimmungsgemäss für die biometrische Verifizierung, wozu die Authentifizierung gehört, verwendet werden, deren einziger Zweck darin besteht zu bestätigen, dass eine bestimmte natürliche Person die Person ist, für die sie sich ausgibt, und die Identität einer natürlichen Person zu dem alleinigen Zweck zu bestätigen, Zugang zu einem Dienst zu erhalten, ein Gerät zu entriegeln oder Sicherheitszugang zu Räumlichkeiten zu erhalten. Biometrische Systeme, die ausschliesslich dazu bestimmt sind, Massnahmen zur Cybersicherheit und zum Schutz personenbezogener Daten durchführen zu können, sollten nicht als Hochrisiko-KI-Systeme gelten (vgl. Anhang III Nr. 1 Bst. a und E. 54).

- Das KI-System ist dazu bestimmt, eine Aufgabe auszuführen, die eine Bewertung, die für die Zwecke der in Anhang III aufgeführten Anwendungsfälle relevant ist, lediglich vorbereitet (E. 53).

Ungeachtet dieser vier Fälle gilt ein in Anhang III genanntes KI-System immer als hochriskant, wenn es ein Profiling natürlicher Personen vornimmt (vgl. Art. 6 Abs. 3 *in fine*).

Um die Nachvollziehbarkeit und Transparenz der Systeme zu gewährleisten, die gemäss den oben genannten Kriterien ausgenommen sind, muss ein Anbieter, der der Auffassung ist, dass von einem in Anhang III aufgeführten KI-System kein hohes Risiko ausgeht, seine Bewertung dokumentieren, bevor dieses System in Verkehr gebracht oder in Betrieb genommen wird. Zudem ist dieser Anbieter verpflichtet, das System in der EU-Datenbank zu registrieren. Er muss diese Dokumentation den zuständigen nationalen Behörden auf Anfrage zur Verfügung stellen (Art. 6 Abs. 4).

Die Europäische Kommission muss nach Konsultation des Europäischen Gremiums für Künstliche Intelligenz weitere Orientierungshilfen für die praktische Umsetzung von Artikel 6 der Verordnung bereitstellen (Ziff. 5.2.13). Sie erarbeitet Leitlinien für die Umsetzung und stellt eine umfassende Liste von Anwendungsfällen für KI-Systeme bereit, die hochriskant oder nicht hochriskant sind (vgl. Art. 6 Abs. 5).

5.2.7.2 Anforderungen an Hochrisiko-KI-Systeme

Hochrisiko-KI-Systeme müssen die in Abschnitt 2 des Kapitels III festgelegten Anforderungen, also die Artikel 8 bis 15, erfüllen. Auf die Frage, welche Pflichten im Zusammenhang mit der Einhaltung der Grundanforderungen für welche Akteure (Anbieter, Betreiber oder andere) gelten, wird in den Artikeln 16 ff. der Verordnung eingegangen (vgl. Ziff. 5.2.7.3).

Es müssen folgende Anforderungen eingehalten werden:

- Einrichtung eines Risikomanagementsystems (Art. 9): Das Risikomanagementsystem hat zum Ziel, die bekannten und vernünftigerweise vorhersehbaren Risiken zu ermitteln und zu mindern, die von einem Hochrisiko-KI-System für die Gesundheit, die Sicherheit und die Grundrechte ausgehen können, wenn es entsprechend seiner Zweckbestimmung verwendet wird (Art. 9 Abs. 2 Bst. a). Weiter dient es der Abschätzung und Bewertung der Risiken, die bei einer vernünftigerweise vorhersehbaren Fehlanwendung entstehen können (Art. 9 Abs. 2 Bst. b). Es handelt sich um einen kontinuierlichen iterativen Prozess, der während des gesamten Lebenszyklus eines Hochrisiko-KI-Systems geplant und durchgeführt wird. Das Risikomanagementsystem umfasst auch das Ergreifen geeigneter und gezielter Risikomanagementmassnahmen zur Bewältigung der ermittelten Risiken (Art. 9 Abs. 2 Bst. d).[230]

- Daten und Daten-Governance (Art. 10): Für Hochrisiko-KI-Systeme, in denen Techniken eingesetzt werden, mit denen KI-Modelle mit Daten trainiert werden, gelten im Bereich Daten und Daten-Governance besondere Bedingungen. Insbesondere müssen

[230] Für eine Vertiefung s. MARTINA ARIOLI, Risikomanagement (Fn. 222), 6 N 19 ff.

die Trainings-, Validierungs- und Testdatensätze den Anforderungen von Artikel 10 Absätze 2 bis 5 der Verordnung entsprechen.

Die verwendeten Daten müssen namentlich hochwertig sein, um sicherzustellen, dass das System bestimmungsgemäss funktioniert und nicht zur Ursache für Diskriminierung wird. Dies erfordert, dass alle verwendeten Daten im Hinblick auf die Zweckbestimmung des Systems relevant, hinreichend repräsentativ und so weit wie möglich fehlerfrei und vollständig sind. Alle Daten müssen zudem geeignete statistische Merkmale haben, insbesondere bezüglich der Personen oder Personengruppen, auf die das KI-System bestimmungsgemäss angewandt werden soll, und der Minderung möglicher Verzerrungen in allen Daten besondere Beachtung beimessen. Zudem müssen Massnahmen vorgesehen werden, mit denen sichergestellt wird, dass die verarbeiteten Daten geschützt und Gegenstand angemessener Sicherheitsvorkehrungen sind. Für personenbezogene Daten, einschliesslich sensibler Daten, sind ausserdem besondere Vorschriften vorgesehen.

- Technische Dokumentation (Art. 11): Um die Nachvollziehbarkeit der Hochrisiko-KI-Systeme, die Überprüfung der Einhaltung der Anforderungen der Verordnung sowie die Kontrolle ihres Betriebs und die Beobachtung nach dem Inverkehrbringen zu ermöglichen, braucht es umfassende Informationen darüber, wie diese Systeme entwickelt wurden und wie sie während ihrer gesamten Lebensdauer funktionieren. Artikel 11 sieht daher die Pflicht vor, eine technische Dokumentation in klarer und verständlicher Form bereitzustellen, mit der die Konformität des KI-Systems beurteilt werden kann und die zumindest die in Anhang IV genannten Angaben enthält. Dabei handelt es sich namentlich um allgemeine Merkmale, die Fähigkeiten und Grenzen des Systems, die Algorithmen, die Daten, die Trainingsmethoden, die Tests und die verwendeten Validierungsverfahren sowie die Dokumentation des Risikomanagementsystems.

- Aufzeichnungspflichten (Art. 12): Die Technik der Hochrisiko-KI-Systeme muss die automatische Aufzeichnung von Ereignissen (Protokollierung) während des gesamten Lebenszyklus des Systems ermöglichen. Eines der Ziele besteht darin, die Beobachtung nach dem Inverkehrbringen gemäss Artikel 72 zu erleichtern und den Betrieb des Systems gemäss Artikel 26 Absatz 5 der Verordnung zu überwachen (vgl. Art. 12 Abs. 2).

- Transparenz und Bereitstellung von Informationen für die Betreiber (Art. 13): Hochrisiko-KI-Systeme müssen so konzipiert und entwickelt werden, dass die Betreiber die Ausgaben eines Systems angemessen interpretieren und verwenden können. Sie müssen mit geeigneten Informationen in Form einer Betriebsanleitung bereitgestellt werden.

- Menschliche Aufsicht (Art. 14): Hochrisiko-KI-Systeme müssen so konzipiert und entwickelt werden, dass sie während der Dauer ihrer Verwendung von natürlichen Personen wirksam beaufsichtigt werden können. Diese wirksame Aufsicht dient der Verhinderung oder Minimierung der Risiken für Gesundheit, Sicherheit oder Grundrechte, die entstehen können, wenn ein Hochrisiko-KI-System im Einklang mit seiner Zweckbestimmung oder im Rahmen einer vernünftigerweise vorhersehbaren Fehlanwendung verwendet wird. Es müssen den Risiken angemessene Massnahmen ergriffen werden. Sie können vom Anbieter direkt in das System eingebaut werden oder von diesem bestimmt und

vom Betreiber umgesetzt werden (vgl. Art. 14 Abs. 3). Dabei kann es sich beispielsweise um die Möglichkeit handeln, den Betrieb eines KI-Systems mit einer Stopptaste zu unterbrechen (vgl. Art. 14 Abs. 4 Bst. e).

Die den in Anhang III Nummer 1 Buchstabe a genannten Systeme unterliegen einer erweiterten menschlichen Aufsicht. So darf der Betreiber keine Massnahmen oder Entscheidungen allein aufgrund des vom System hervorgebrachten Identifizierungsergebnisses treffen, solange die Identifizierung nicht von mindestens *zwei* natürlichen Personen, die über die notwendige Kompetenz, Ausbildung und Befugnis verfügen, getrennt überprüft und bestätigt wurde. Diese Pflicht gilt jedoch nicht für Hochrisiko-KI-Systeme, die in den Bereichen Strafverfolgung, Migration, Grenzkontrolle oder Asyl verwendet werden, wenn die Umsetzung dieser Anforderung nach EU-Recht oder nationalem Recht unverhältnismässig wäre (Art. 14 Abs. 5 *in fine*).

- Genauigkeit, Robustheit und Cybersicherheit (Art. 15): Es müssen technische Massnahmen ergriffen werden, um ein angemessenes Mass an Genauigkeit, Robustheit und Cybersicherheit der Hochrisiko-KI-Systeme zu gewährleisten.

5.2.7.3 Pflichten der Anbieter und anderer Beteiligter

5.2.7.3.1 Pflichten der Anbieter

Die Verordnung legt den Schwerpunkt auf die zentrale Rolle der Anbieter, denen sie die Verantwortung für die Konformität der Hochrisiko-KI-Systeme überträgt.[231]

Die meisten Pflichten im Zusammenhang mit Hochrisiko-KI-Systemen entfallen daher auf die Anbieter von solchen Systemen. Dieses Unterkapitel konzentriert sich auf die wichtigsten Pflichten:

- Gemäss Artikel 16 Buchstabe a stellen die Anbieter sicher, dass ihre Systeme die in Kapitel III Abschnitt 2 (Art. 8–15 der Verordnung, vgl. Ziff. 5.2.7.2) festgelegten Anforderungen erfüllen. Auf begründete Anfrage einer zuständigen nationalen Behörde haben sie nachzuweisen, dass das KI-System diese Anforderungen erfüllt (vgl. Art. 16 Bst. k).

- Die Anbieter müssen identifizierbar sein (Art. 16 Bst. b).

- Gemäss Artikel 16 Buchstabe c und Artikel 17 müssen die Anbieter von Hochrisiko-KI-Systemen über ein *Qualitätsmanagementsystem* verfügen, das das in Artikel 9 genannte Risikomanagementsystem und verschiedene andere Aspekte umfasst, wie ein Konzept zur Einhaltung der Regulierungsvorschriften, Systeme für die Aufzeichnung der Dokumentation, Ressourcenmanagement, einen internen Rechenschaftsrahmen

[231] Für eine Kritik zu diesem Ansatz vgl. ANGELA MÜLLER, Der Artificial intelligence Act der EU (Fn. 223), 18.

usw. Wenn in der sektorspezifischen Gesetzgebung bereits ein Qualitätsmanagementmechanismus vorgesehen ist, können diese Aspekte in das Qualitätsmanagementsystem gemäss dieser Gesetzgebung integriert werden (vgl. Art. 17 Abs. 3).

- Die Artikel 18 und 19 sehen Pflichten für die Aufbewahrung der Dokumentation und der automatisch erzeugten Protokolle vor.

- Artikel 20 sieht die Pflicht vor, bei Verdacht, dass ein in Verkehr gebrachtes KI-System nicht der Verordnung entspricht, Massnahmen zu ergreifen.

  Die Anbieter von Hochrisiko-KI-Systemen müssen auch über ein «System zur Beobachtung nach dem Inverkehrbringen» verfügen (vgl. Art. 3 Nrn. 25 und 72), damit den möglichen Risiken im Zusammenhang mit KI-Systemen, die nach dem Inverkehrbringen oder der Inbetriebnahme «dazulernen», effizient begegnet werden kann (E. 155).

- Gemäss Artikel 21 ist der Anbieter zudem verpflichtet, mit den zuständigen Behörden zusammenzuarbeiten.

- Die Anbieter müssen sicherstellen, dass das Hochrisiko-KI-System dem Konformitätsbewertungsverfahren gemäss Artikel 43 unterzogen wird, bevor es in Verkehr gebracht oder in Betrieb genommen wird (Art. 16 Bst. f). Weiter haben sie eine Konformitätserklärung gemäss Artikel 47 auszustellen, am System eine CE-Kennzeichnung («Conformité Européenne»-Kennzeichnung) anzubringen (Art. 48) und das System ordnungsgemäss zu registrieren (vgl. Art. 49).

  Für diese Aspekte wird auf die Ausführungen unten verwiesen (vgl. Ziff. 5.2.7.4).

Liegen besondere Umstände vor, so gehen die in den Artikeln 16 ff. vorgesehenen Pflichten auf den Händler, Einführer oder Betreiber oder auf sonstige Dritte über. Dies ist beispielsweise der Fall, wenn einer dieser Akteure die Zweckbestimmung eines KI-Systems so verändert, dass dieses zu einem Hochrisiko-KI-System wird (z. B. ein System in der Art von ChatGPT wird für die Analyse der Lebensläufe von Bewerberinnen und Bewerbern verwendet).[232] In diesen Fällen gilt der «ursprüngliche» Anbieter nicht mehr als Anbieter im Sinne der Verordnung. Für die anderen Fälle wird auf Artikel 25 der Verordnung verwiesen.

5.2.7.3.2 Pflichten der Einführer und der Händler

Die Einführer müssen sicherstellen, dass ein Hochrisiko-KI-System der Verordnung entspricht, bevor sie es in Verkehr bringen (vgl. Art. 23). Sie müssen beispielsweise überprüfen, ob der Anbieter das entsprechende Konformitätsbewertungsverfahren gemäss Artikel 43 durchgeführt und die technische Dokumentation gemäss Artikel 11 und Anhang IV erstellt hat. Weiter müssen sie ihren Namen und die Anschrift angeben, unter der sie kontaktiert werden können (vgl. Art. 23 Abs. 3).

---

[232] DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 10 N 25.

Die Händler ihrerseits müssen sicherstellen, dass das System mit der erforderlichen CE-Kennzeichnung versehen ist, ihm eine Kopie der EU-Konformitätserklärung und der Betriebsanleitung beigefügt ist und der Anbieter und gegebenenfalls der Einführer dieses Systems ihre in Artikel 16 Buchstaben b und c sowie in Artikel 23 Absatz 3 festgelegten jeweiligen Pflichten erfüllt haben (vgl. Art. 24).

5.2.7.3.3 Pflichten der Betreiber

Gemäss der Verordnung sind die Betreiber nach den Anbietern die zweite Gruppe von Akteuren, die im Zusammenhang mit den Hochrisiko-KI-Systemen am meisten Pflichten haben.

Die Verordnung legt zunächst die allgemeinen Pflichten fest. Dazu zählen insbesondere:

- Die Pflicht, geeignete technische und organisatorische Massnahmen zu treffen, um eine vorschriftsgemässe Verwendung der Hochrisiko-KI-Systeme sicherzustellen (Art. 26 Abs. 1).

- Die Betreiber müssen zudem gewährleisten, dass die menschliche Aufsicht natürlichen Personen, die über die erforderliche Kompetenz, Ausbildung und Befugnis verfügen, übertragen wird (Art. 26 Abs. 2).

- Wenn der Betreiber die Möglichkeit hat, die Eingabedaten zu kontrollieren, muss er sicherstellen, dass diese der Zweckbestimmung des Hochrisiko-KI-Systems entsprechen und ausreichend repräsentativ sind (Art. 26 Abs. 4).

- Die Betreiber sind verpflichtet, den Betrieb des Hochrisiko-KI-Systems zu überwachen. Sie haben eine Informationspflicht, wenn ein Risiko im Sinne von Artikel 79 vorliegt oder ein schwerwiegender Vorfall festgestellt wurde (Art. 26 Abs. 5).

- Sie bewahren die von ihrem Hochrisiko-KI-System automatisch erzeugten Protokolle auf, soweit diese Protokolle ihrer Kontrolle unterliegen (Art. 26 Abs. 6).

Anschliessend legt die Verordnung spezifische Pflichten je nach Betreiber fest, namentlich:

- Betreiber, die Arbeitgeber sind, informieren die Arbeitnehmervertreterinnen und -vertreter und die betroffenen Arbeitnehmenden darüber, dass sie der Verwendung eines Hochrisiko-KI-Systems unterliegen werden, und zwar vor seiner Inbetriebnahme oder Verwendung (Art. 26 Abs. 7).

- Betreiber, bei denen es sich um Behörden oder Organe, Einrichtungen oder sonstige Stellen der EU handelt, müssen den Registrierungspflichten gemäss Artikel 49 nachkommen (Art. 26 Abs. 8). Sie müssen insbesondere die Verwendung der in Anhang III aufgeführten Systeme registrieren, mit Ausnahme der unter Nummer 2 genannten (die auf nationaler Ebene registriert werden) (vgl. Art. 49 Abs. 3 und 5). Die Registrierung der in Anhang III Nummern 1, 6 und 7 genannten Hochrisiko-KI-Systeme erfolgt in den Bereichen Strafverfolgung, Migration, Asyl und Grenzkontrolle in einem nicht öffentlichen Teil der Datenbank, auf den nur die Europäische Kommission und bestimmte nationale Behörden Zugriff haben (Art. 49 Abs. 4).

- Betreiber eines Hochrisiko-KI-Systems zur nachträglichen[233] biometrischen Fernidentifizierung im Rahmen von Ermittlungen zur gezielten Suche einer Person, die der Begehung einer Straftat verdächtigt wird oder aufgrund einer solchen verurteilt wurde, müssen vorab oder unverzüglich, spätestens jedoch innerhalb von 48 Stunden bei einer Justizbehörde oder einer Verwaltungsbehörde eine Genehmigung für die Verwendung des Systems beantragen (vgl. Art. 26 Abs. 10).

  Weiter legt die Verordnung fest, dass diese Systeme in keinem Fall zu Strafverfolgungszwecken in nicht zielgerichteter Weise und ohne jeglichen Zusammenhang mit einer Straftat, einem Strafverfahren, einer tatsächlichen und bestehenden oder tatsächlichen und vorhersehbaren Gefahr einer Straftat oder der Suche nach einer bestimmten vermissten Person verwendet werden dürfen. Ausserdem muss sichergestellt werden, dass die Strafverfolgungsbehörden keine ausschliesslich auf der Grundlage der Ausgabe solcher Systeme beruhende Entscheidung treffen, aus der sich eine nachteilige Rechtsfolge für eine Person ergibt.

  Keine Genehmigung ist erforderlich, wenn das System zur erstmaligen Identifizierung einer potenziell verdächtigen Person auf der Grundlage objektiver und nachprüfbarer Tatsachen, die in unmittelbarem Zusammenhang mit der Straftat stehen, verwendet wird.

  Ferner müssen die Betreiber dieser Systeme den zuständigen Marktüberwachungsbehörden und den nationalen Datenschutzbehörden Jahresberichte über ihre Verwendung vorlegen, wovon die Offenlegung sensibler operativer Daten im Zusammenhang mit der Strafverfolgung ausgenommen ist (Art. 26 Abs. 10).

- Wenn der Betreiber ein in Anhang III aufgeführtes Hochrisiko-KI-System verwendet, um Entscheidungen betreffend natürliche Personen zu treffen oder solche Entscheidungen zu unterstützen, muss er die betroffenen Personen informieren (Art. 26 Abs. 11). Diese Pflicht gilt unbeschadet von Artikel 50 (Transparenzpflichten, vgl. Ziff. 5.2.8). Auf die zu Strafverfolgungszwecken verwendeten Systeme ist Artikel 13 der Richtlinie (EU) 2016/680[234] anwendbar.

  Zusätzlich hat die betroffene Person das Recht, eine Erläuterung im Zusammenhang mit automatisierten Einzelentscheidungen zu erhalten (vgl. Art. 86, Ziff. 5.2.15).

---

[233] Der Begriff «nachträglich» bedeutet, dass es eine zeitliche Verzögerung zwischen der Erfassung biometrischer Daten, dem Abgleich und der Identifizierung gibt (vgl. Art. 3 Nrn. 42 und 43).

[234] Richtlinie (EU) 2016/680 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten durch die zuständigen Behörden zum Zwecke der Verhütung, Ermittlung, Aufdeckung oder Verfolgung von Straftaten oder der Strafvollstreckung sowie zum freien Datenverkehr und zur Aufhebung des Rahmenbeschlusses 2008/977/JI des Rates, ABl. L 119/89, 4. Mai 2016.

- Schliesslich haben bestimmte Betreiber die Pflicht, eine Abschätzung der Folgen von Hochrisiko-KI-Systemen auf die Grundrechte durchzuführen.[235] Diese Pflicht ist in Artikel 27 geregelt:

  - Adressaten der Pflicht: Eine Folgenabschätzung durchführen müssen 1) Einrichtungen des öffentlichen Rechts oder private Einrichtungen, die öffentliche Dienste[236] erbringen, 2) (öffentliche oder private) Betreiber von Hochrisiko-KI-Systemen gemäss Anhang III Nummer 5 Buchstaben b und c[237].

  - Anwendungsbereich: Die Grundrechte-Folgenabschätzung muss vor der Inbetriebnahme der Hochrisiko-KI-Systeme gemäss Artikel 6 Absatz 2 der Verordnung, also der in Anhang III genannten Systeme, durchgeführt werden. Von dieser Pflicht ausgenommen sind KI-Systeme, die im Bereich kritischer Infrastruktur verwendet werden (vgl. Anhang III Nummer 2).

  - Inhalt: Die Grundrechte-Folgenabschätzung muss dem Betreiber ermöglichen, die spezifischen Risiken für die Grundrechte des Einzelnen, die wahrscheinlich berührt sein werden, zu ermitteln und die Massnahmen festzulegen, die zu ergreifen sind, wenn diese Risiken eintreten (vgl. Art. 27 Abs. 1 Bst. a–f). Um die Betreiber bei der Durchführung der Abschätzung zu unterstützen, erarbeitet das Büro für Künstliche Intelligenz ein Muster für einen Fragebogen, auch mithilfe eines automatisierten Instruments (Art. 27 Abs. 5).

    Wenn gewisse Aspekte der Abschätzung bereits durch die Datenschutz-Folgenabschätzung (Art. 35 DSGVO oder Art. 27 der Richtlinie (EU) 2016/680[238]) abgedeckt sind, kann die Grundrechte-Folgenabschätzung die Datenschutz-Folgenabschätzung ergänzen (vgl. 27 Abs. 4).

---

[235] Zum Unterschied zu den Risikomanagementsystemen (Art. 9) vgl. MARTINA ARIOLI, Risikomanagement (Fn. 222), 14 N 48 f.

[236] Es ist nicht klar, ob der Begriff «öffentliche Dienste» auch rein privatrechtliche Tätigkeiten umfasst, aber in Anbetracht der Erwägungen scheint dies der Fall zu sein (vgl. E. 96). Erwähnt sind die im öffentlichen Interesse erbrachten Dienste, namentlich in den Bereichen Bildung, Gesundheitsversorgung, Sozialdienste, Wohnungswesen und Justizverwaltung, vgl. DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 30 N 58.

[237] Hier handelt es sich um Betreiber von KI-Systemen, die bestimmungsgemäss für die Kreditwürdigkeitsprüfung und Bonitätsbewertung natürlicher Personen verwendet werden, mit Ausnahme von KI-Systemen, die zur Aufdeckung von Finanzbetrug verwendet werden (Anhang III Nr. 5 Bst. b); Betreiber von KI-Systemen, die bestimmungsgemäss für die Risikobewertung und Preisbildung in Bezug auf natürliche Personen im Fall von Lebens- und Krankenversicherungen verwendet werden (Anhang III Nr. 5 Bst. c).

[238] Richtlinie (EU) 2016/680 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten im Bereich der Strafverfolgung (Fn. 234).

- Überwachung: Wurde die Abschätzung durchgeführt, so teilt der Betreiber die Ergebnisse der Marktüberwachungsbehörde mit (vgl. Ziff. 5.2.14). In bestimmten Fällen sind Ausnahmen möglich (vgl. Art. 27 Abs. 3).

5.2.7.4 Konformitätsbewertung

5.2.7.4.1 Verfahren

Gemäss Artikel 16 Buchstabe f der Verordnung stellen die Anbieter von Hochrisiko-KI-Systemen sicher, dass diese Systeme dem Konformitätsbewertungsverfahren gemäss Artikel 43 unterzogen werden, bevor sie in Verkehr gebracht oder in Betrieb genommen werden.

Artikel 43 regelt die Konformitätsbewertung, wobei je nach System unterschiedliche Verfahren zur Anwendung gelangen:[239]

- Bei den Hochrisiko-KI-Systemen, die unter die Harmonisierungsrechtsvorschriften der EU fallen (vgl. Anhang I Abschnitt A), wird die Einhaltung der Anforderungen an Hochrisiko-KI-Systeme (Art. 8 ff., vgl. Ziff. 5.2.7.2) in das in diesen Vorschriften vorgesehene Konformitätsbewertungsverfahren einbezogen, um den Aufwand für die Akteure zu verringern und Doppelspurigkeiten zu vermeiden. Die Stellen, die gemäss diesen sektorspezifischen Vorschriften notifiziert wurden, sind berechtigt, die Konformität von Hochrisiko-KI-Systemen zu bewerten (vgl. Art. 43 Abs. 3).

  Die Anwendbarkeit der Anforderungen der KI-Verordnung muss die spezifische Logik, die Methodik oder die allgemeine Struktur der in den sektorspezifischen Rechtsvorschriften vorgesehenen Konformitätsbewertung unberührt lassen.

- Bei Hochrisiko-KI-Systemen, die nicht unter eine Harmonisierungsregelung (vgl. Anhang III) fallen, sieht die KI-Verordnung zwei Verfahrensarten für die Konformitätsbewertung vor: eine Konformitätsbewertung *auf der Grundlage einer internen Kontrolle* (gemäss Anhang VI) und in bestimmten Fällen eine Konformitätsbewertung *durch eine Drittpartei* (notifizierte Stelle[240]) (gemäss Anhang VII).

  Dieses Vorgehen soll dazu beitragen, bei den Anbietern, die zurzeit keinen Pflichten im Zusammenhang mit den Produkten unterstehen, eine Konformitätskultur zu schaffen

---

[239] Die Verfahren basieren auf dem neuen Rechtsrahmen («New Legislative Framework», NLF), s. dazu MARTINA ARIOLI, Risikomanagement (Fn. 222), 4 N 8 f.

[240] Damit KI-Systeme, falls vorgeschrieben, Konformitätsbewertungen durch Dritte unterzogen werden können, müssen die notifizierten Stellen gemäss den Rechtsvorschriften von den zuständigen nationalen Behörden notifiziert werden, sofern sie eine Reihe von Anforderungen erfüllen, insbesondere in Bezug auf Unabhängigkeit, Kompetenz, Nichtvorliegen von Interessenkonflikten und geeignete Anforderungen an die Cybersicherheit. Die zuständigen nationalen Behörden müssen die Notifizierung dieser Stellen der Europäischen Kommission und den anderen Mitgliedstaaten übermitteln (vgl. Art. 28 ff. der Verordnung).

und zu stärken, ohne dass von Anfang an langwierige Verfahren vorgeschrieben werden, die die Innovation unverhältnismässig hemmen könnten.

Daher muss die Konformitätsbewertung in der Regel vom Anbieter in eigener Verantwortung durchgeführt werden («self-assessment»). In diesem Sinn sieht Artikel 43 Absatz 2 der Verordnung vor, dass bei den in Anhang III Nummern 2 bis 8 aufgeführten Hochrisiko-KI-Systemen ein Konformitätsbewertungsverfahren auf der Grundlage einer internen Kontrolle durchgeführt werden muss, das keine Beteiligung einer notifizierten Stelle vorsieht.

Wenn erforderlich, namentlich zur Vermeidung und Minimierung von Risiken für die Sicherheit und die Grundrechte, ist die Europäische Kommission befugt, einzugreifen und diese Systeme der Konformitätsbewertung oder Teilen davon durch eine notifizierte Stelle zu unterstellen (Art. 43 Abs. 6).

Bei KI-Systemen, die bestimmungsgemäss zu biometrischen Zwecken verwendet werden (vgl. Anhang III Nr. 1), sieht Artikel 43 Absatz 1 der Verordnung einen detaillierteren Mechanismus vor.

In beiden Fällen misst die KI-Verordnung den harmonisierten Normen grosse Bedeutung bei (für diesen Punkt wird auf Ziff. 5.2.11 verwiesen).

In Artikel 46 der Verordnung sind die Voraussetzungen geregelt, unter denen die Marktüberwachungsbehörde eine Genehmigung erteilt, mit der Hochrisiko-KI-Systeme in Abweichung vom Konformitätsbewertungsverfahren in Verkehr gebracht oder in Betrieb genommen werden dürfen. Ausnahmen sind beispielsweise möglich aus aussergewöhnlichen Gründen der öffentlichen Sicherheit. Für Hochrisiko-KI-Systeme im Zusammenhang mit Produkten, die den in Anhang I Abschnitt A aufgeführten Harmonisierungsrechtsvorschriften der EU unterstehen, gelten nur die in diesen Vorschriften festgelegten Ausnahmen von den Konformitätsbewertungsverfahren (vgl. Art. 46 Abs. 7).

5.2.7.4.2 Konformitätserklärung, CE-Kennzeichnung und Registrierung

Wurde im Rahmen einer internen Kontrolle oder einer Bewertung durch eine notifizierte Stelle festgestellt, dass das Hochrisiko-KI-System den Anforderungen der Verordnung entspricht und wurde eine Konformitätsbescheinigung ausgestellt (vgl. Art. 44), so muss der Anbieter eine schriftliche EU-Konformitätserklärung erstellen, die die in Anhang V vorgesehenen Informationen enthält. Diese bestätigt, dass das Hochrisiko-KI-System die Anforderungen der Verordnung erfüllt. Mit dieser Erklärung übernimmt der Anbieter also die Verantwortung für die Erfüllung dieser Anforderungen (vgl. Art. 47 Abs. 4).

Wenn bereits in sektorspezifischen Rechtsvorschriften (Harmonisierungsrechtsvorschriften) eine EU-Konformitätserklärung vorgesehen ist, wird, um Doppelverfahren zu vermeiden, eine einzige Erklärung ausgestellt (Art. 47 Abs. 3).

Hochrisiko-KI-Systeme müssen mit einer CE-Kennzeichnung («Conformité Européenne»-Kennzeichnung) versehen werden, aus der ihre Konformität mit der Verordnung hervorgeht, damit sie frei im Binnenmarkt verkehren können (vgl. Art. 48). Wenn Hochrisiko-KI-Systeme

unter andere Rechtsvorschriften der EU fallen, in denen die CE-Kennzeichnung ebenfalls vorgesehen ist, bedeutet die CE-Kennzeichnung, dass die Hochrisiko-KI-Systeme auch die Anforderungen dieser anderen Rechtsvorschriften erfüllen.

Schliesslich müssen die Anbieter und bestimmte Betreiber vor dem Inverkehrbringen, der Inbetriebnahme oder der Verwendung eines in Anhang III (mit Ausnahme von Nr. 2) aufgeführten Hochrisiko-KI-Systems sich selber und ihre Systeme gemäss ihren diesbezüglichen Pflichten registrieren (vgl. Art. 49), und zwar in der von der Europäischen Kommission in Zusammenarbeit mit den Mitgliedstaaten geführten Datenbank (vgl. Art. 71) (die in Anhang III Nr. 2 aufgeführten Hochrisiko-KI-Systeme [Infrastruktur] werden hingegen auf nationaler Ebene registriert). Die bereitzustellenden Informationen sind in Anhang VIII aufgeführt.

5.2.8 Transparenzpflichten für Anbieter und Betreiber bestimmter KI-Systeme

In Kapitel IV sieht die Verordnung Transparenzpflichten für Anbieter und Betreiber in bestimmten Fällen vor (Art. 50 ff.). Diese Pflichten gelten für alle KI-Systeme, unabhängig davon, ob es sich um Hochrisiko-KI-Systeme handelt oder nicht. Bei Hochrisiko-KI-Systemen ergänzen diese Pflichten die anderen Transparenzpflichten, die für diese Systeme bereits vorgesehen sind (Art. 50 Abs. 6):

- KI-Systeme, die für die direkte Interaktion mit natürlichen Personen bestimmt sind: Die *Anbieter* müssen sicherstellen, dass die KI-Systeme so konzipiert und entwickelt werden, dass die betreffenden natürlichen Personen informiert werden, dass sie mit einem KI-System interagieren, beispielsweise einer Roboter-Software (Chatbot), es sei denn, dies ist aufgrund der Umstände und des Kontexts der Nutzung offensichtlich. Diese Pflicht gilt nicht für Systeme, die beispielsweise zur Aufdeckung von Straftaten verwendet werden (Art. 50 Abs. 1).

- KI-Systeme, einschliesslich KI-Systeme mit allgemeinem Verwendungszweck, die Audio-, Bild-, Video- oder Textinhalte erzeugen: Die *Anbieter* haben dafür zu sorgen, dass die Ausgaben des KI-Systems in einem maschinenlesbaren Format gekennzeichnet und als künstlich erzeugt oder manipuliert erkennbar sind. Die Verordnung sieht eine Ausnahme für KI-Systeme vor, die eine unterstützende Funktion für die Standardbearbeitung ausführen oder die vom Betreiber bereitgestellten Eingabedaten nicht wesentlich verändern, sowie für Systeme, die gesetzlich zur Aufdeckung von Straftaten zugelassen sind (Art. 50 Abs. 2).[241]

- Emotionserkennungssysteme oder Systeme zur biometrischen Kategorisierung: Die *Betreiber* müssen die von einem solchen System betroffenen natürlichen Personen über seinen Betrieb und die Verarbeitung personenbezogener Daten informieren. KI-Systeme, die rechtmässig zur Aufdeckung, Verhütung, Ermittlung und Verfolgung von Straftaten verwendet werden, sind von dieser Pflicht ausgenommen (Art. 50 Abs. 3).

- Deepfakes: Die *Betreiber* von KI-Systemen die Deepfakes erzeugen müssen offenlegen, dass der Inhalt künstlich erzeugt oder manipuliert wurde. Die Verwendung solcher

[241] Für praktische Erwägungen vgl. DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 34 N 69.

Techniken im gesetzlichen Rahmen zur Aufdeckung, Verhütung, Ermittlung und Verfolgung von Straftaten, ist hingegen erlaubt. Deepfakes, die in einem künstlerischen, kreativen, satirischen, fiktionalen oder analogen Kontext verwendet werden, dürfen in einer Weise offengelegt werden, die die Darstellung oder den Genuss des Werks nicht beeinträchtigt (Art. 50 Abs. 4).

- Bei Texten, die veröffentlicht werden, um die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse zu informieren, muss offengelegt werden, dass diese künstlich erzeugt oder manipuliert wurden, es sei denn, sie wurden einer menschlichen Überprüfung oder redaktionellen Kontrolle unterzogen und eine natürliche oder juristische Person trägt die redaktionelle Verantwortung für ihre Veröffentlichung (Art. 50 Abs. 4).

Die Verordnung sieht vor, dass die in Artikel 50 Absätze 1 bis 4 genannten Informationen den betreffenden natürlichen Personen spätestens zum Zeitpunkt der ersten Interaktion oder Aussetzung in klarer und eindeutiger Weise bereitgestellt werden müssen.

Das Büro für Künstliche Intelligenz (vgl. Ziff. 5.2.13) fördert die Ausarbeitung von Praxisleitfäden auf EU-Ebene, um die Umsetzung der Pflichten in Bezug auf die Feststellung und Kennzeichnung künstlich erzeugter oder manipulierter Inhalte zu erleichtern. Die Europäische Kommission kann solche Praxisleitfäden mittels Durchführungsrechtsakte genehmigen, oder einen Durchführungsrechtsakt erlassen, in dem gemeinsame Vorschriften für die Umsetzung dieser Pflichten festgelegt werden, wenn sie einen Praxisleitfaden für nicht angemessen hält (Art. 50 Abs. 7).

5.2.9 Andere Systeme der künstlichen Intelligenz

Wenn ein KI-System weder eine verbotene Praxis noch ein Hochrisiko-KI-System ist und auch nicht unter die Transparenzbestimmungen fällt (Art. 50), bestehen gemäss der KI-Verordnung grundsätzlich keine Pflichten für die Anbieter, Betreiber und anderen beteiligten Stellen.

Die Verordnung sieht jedoch vor, dass das Büro für Künstliche Intelligenz und die Mitgliedstaaten die Erarbeitung von Verhaltenskodizes fördern und erleichtern, mit denen die freiwillige Anwendung einiger oder aller der in Kapitel III Abschnitt 2 genannten Anforderungen (also die für die Hochrisiko-KI-Systeme geltenden Vorschriften) auf diese anderen Systeme unterstützt werden soll (Art. 95 Abs. 1). Die Verhaltenskodizes sollen auch andere Aspekte fördern, wie ethische Grundsätze, ein respektvoller Umgang mit den ökologischen Ressourcen, die KI-Kompetenz und die Verhinderung negativer Auswirkungen auf schutzbedürftige Personen oder Gruppen schutzbedürftiger Personen (Art. 95 Abs. 2).

Zudem sieht Artikel 4 der Verordnung eine Art allgemeine Pflicht für alle Anbieter und Betreiber von KI-Systemen vor, Massnahmen zu ergreifen, um nach besten Kräften sicherzustellen,

dass ihr Personal und andere Personen, die in ihrem Auftrag mit dem Betrieb und der Nutzung von KI-Systemen befasst sind, über ein ausreichendes Mass an KI-Kompetenz verfügen («AI literacy»).[242]

5.2.10 Modelle der künstlichen Intelligenz mit allgemeinem Verwendungszweck

5.2.10.1 Allgemeine Bemerkungen

Wie oben ausgeführt (vgl. Ziff. 5.2.4), unterscheidet die KI-Verordnung zwischen KI-Systemen und GPAI-Modellen. Das Kapitel V der Verordnung ist den GPAI-Modellen gewidmet.

Die GPAI-Modelle waren im Vorschlag für eine KI-Verordnung der Europäischen Kommission vom 4. April 2021 nicht enthalten. Das Aufkommen dieser Modelle mit der Lancierung von ChatGPT im November 2022 hat jedoch die Verhandlungen zwischen dem Europäischen Parlament und dem Rat wesentlich mitbestimmt. Der endgültige Regulierungsansatz, auf den sich die Mitgesetzgeber geeinigt haben, umfasst neue Pflichten für Anbieter von «einfachen» GPAI-Modellen und solchen mit systemischen Risiken.

5.2.10.2 Pflichten für alle Modelle der künstlichen Intelligenz mit allgemeinem Verwendungszweck

Die Anbieter[243] von GPAI-Modellen haben unabhängig von deren Risiko namentlich folgende Pflichten (vgl. Art. 53):

- Sie müssen die technische Dokumentation ihres Modells erstellen und regelmässig aktualisieren. Die Elemente, die die Dokumentation mindestens enthalten muss, sind in den Anhängen XI und XII beschrieben.

  Die Dokumentation muss insbesondere die Einzelheiten des Trainings- und Testverfahrens und der Ergebnisse der Bewertung enthalten. Auf Anfrage muss sie dem Büro für Künstliche Intelligenz und den zuständigen nationalen Behörden zur Verfügung gestellt werden (vgl. Art. 53 Abs. 1 Bst. a).

- Die Anbieter von GPAI-Modellen müssen auch Informationen und die Dokumentation vorbereiten und aktualisieren, die Anbietern von KI-Systemen zur Verfügung gestellt werden, die das GPAI-Modell in ihre Systeme integrieren wollen, um sie insbesondere dabei zu unterstützen, die Fähigkeiten und Grenzen des GPAI-Modells zu verstehen und die Verordnung einzuhalten (vgl. Art. 53 Abs. 1 Bst. b).

- Weiter sind die Anbieter von GPAI-Modellen verpflichtet, interne Vorschriften zur Einhaltung des Urheberrechts der EU einzuführen (Art. 53 Abs. 1 Bst. c).[244]

---

[242] DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 36 N 72.

[243] Für die GPAI-Modelle gibt es die Rolle des Betreibers nicht; vgl. DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 31 N 62.

[244] Zur Auslegung dieser Pflicht vgl. DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 33 N 65.

- Die Anbieter müssen ausserdem eine hinreichend detaillierte Zusammenfassung der für das Training des GPAI-Modells verwendeten Inhalte erstellen und veröffentlichen. Diese Zusammenfassung muss Parteien mit berechtigtem Interesse, einschliesslich der Inhaber von Urheberrechten, die Ausübung und Durchsetzung ihrer Rechte nach dem EU-Recht ermöglichen. Das Büro für Künstliche Intelligenz erarbeitet eine Vorlage, die den Anbietern helfen soll, die erforderliche Zusammenfassung in verbaler Form bereitzustellen (Art. 53 Abs. 1 Bst. d).

- Anbieter von GPAI-Modellen, die sich ausserhalb der EU befinden, aber in der EU GPAI-Modelle in Verkehr bringen, müssen einen Bevollmächtigten in der EU benennen (vgl. Art. 54).

GPAI-Modelle, die im Rahmen einer freien und quelloffenen Lizenz bereitgestellt werden, sind von den in Artikel 53 Absatz 1 Buchstaben a und b und Artikel 54 aufgeführten Pflichten ausgenommen. Diese Ausnahmeregelung gilt, wenn die Parameter der KI-Modelle, einschliesslich Gewichte, Informationen über die Modellarchitektur und Informationen über die Modellnutzung, öffentlich zugänglich sind. Sie gilt jedoch nicht für GPAI-Modelle, die ein systemisches Risiko beinhalten (vgl. Ziff. 5.2.10.3; Art. 53 Abs. 2 und 54 Abs. 6).

Anbieter von GPAI-Modellen können sich bis zur Veröffentlichung einer harmonisierten Norm auf Praxisleitfäden im Sinne von Artikel 56 stützen, um die Einhaltung ihrer Pflichten nachzuweisen. Die Einhaltung der harmonisierten Normen begründet für die Anbieter die Vermutung der Konformität (Art. 53 Abs. 4). Das Büro für Künstliche Intelligenz hat den Auftrag, die Schaffung von Praxisleitfäden unter Berücksichtigung internationaler Ansätze zu fördern und zu erleichtern.

5.2.10.3 Pflichten für Modelle der künstlichen Intelligenz mit allgemeinem Verwendungszweck mit systemischem Risiko

5.2.10.3.1 Begriff und Verfahren

Für GPAI-Modelle mit systemischem Risiko gelten gemäss dem Grundsatz, dass eine Steigerung der Fähigkeiten auch die Risiken erhöht, zusätzliche Pflichten.

Ein GPAI-Modell wird grundsätzlich als Modell mit systemischen Risiken angesehen, wenn es über «Fähigkeiten mit hohem Wirkungsgrad» verfügt (vgl. Art. 51). Als Kriterium verwendet die Verordnung die kumulierte Menge der für sein Training verwendeten Berechnungen, gemessen in Gleitkommaoperationen («floating point operations», FLOP) (vgl. E. 111 der Verordnung). Um zu bestimmen, ob es sich um ein GPAI-Modell mit systemischen Risiken handelt, wurde ein Grenzwert von 10^25 FLOP festgelegt (vgl. Art. 51 Abs. 2).

Dieser Grenzwert und dieses Kriterium können im Laufe der Zeit angepasst werden, um technologischen und industriellen Veränderungen, wie algorithmischen Verbesserungen oder erhöhter Hardware-Effizienz, Rechnung zu tragen. Das wissenschaftliche Gremium Sachverständiger (vgl. Ziff. 5.2.13) kann dem Büro für Künstliche Intelligenz eine qualifizierte Warnung übermitteln, wenn es Grund zur Annahme hat, dass ein GPAI-Modell ein systemisches Risiko auf EU-Ebene aufweist (Art. 90).

Die Vorschriften für die Einstufung eines GPAI-Modells als Modell mit systemischem Risiko sind in den Artikeln 51 und 52 der Verordnung enthalten. Wenn also ein KI-Modell mit allgemeinem Verwendungszweck die Bedingung nach Artikel 51 Absatz 1 Buchstabe a erfüllt (es

verfügt über Fähigkeiten mit hohem Wirkungsgrad), so teilt der betreffende Anbieter dies der Europäischen Kommission unverzüglich mit. Er kann gleichzeitig nachzuweisen versuchen, dass das Modell trotz allem keine systemischen Risiken aufweist (vgl. Art. 52 Abs. 2).

Die Europäische Kommission kann ein KI-Modell mit allgemeinem Verwendungszweck auch selber als Modell mit systemischen Risiken bezeichnen (Art. 52 Abs. 4). In diesem Fall kann der Anbieter eine Neubewertung beantragen.

Die Europäische Kommission veröffentlicht eine Liste der KI-Modelle mit allgemeinem Verwendungszweck, die ein systemisches Risiko aufweisen, und hält diese unter Beachtung der Vorschriften über geistiges Eigentum und der Geschäftsgeheimnisse auf dem neusten Stand (Art. 52 Abs. 6).

5.2.10.3.2 Zusätzliche Pflichten

Zusätzlich zu den Pflichten, die für alle GPAI-Modelle gelten (vgl. Art. 53 und 54, Ziff. 5.2.10.2), haben die Anbieter von GPAI-Modellen mit systemischen Risiken insbesondere folgende Pflichten (vgl. Art. 55):

- Angriffstests («adversarial testing»): Die Anbieter müssen die erforderlichen Bewertungen des Modells, insbesondere vor seinem ersten Inverkehrbringen, durchführen, wozu auch die Durchführung und Dokumentation von Angriffstests bei Modellen gehören, gegebenenfalls auch im Rahmen interner oder unabhängiger externer Tests, um systemische Risiken zu ermitteln und zu mindern.

- Minderung systemischer Risiken: Die Anbieter müssen mögliche systemische Risiken auf EU-Ebene, die sich aus der Entwicklung, dem Inverkehrbringen oder der Verwendung des Modells ergeben können, bewerten und mindern.

- Meldung von Vorfällen: Wenn trotz der Bemühungen, Risiken im Zusammenhang mit einem GPAI-Modell mit systemischen Risiken zu ermitteln und vorzubeugen, die Entwicklung oder Verwendung des Modells einen schwerwiegenden Vorfall verursacht, muss der Anbieter des Modells dem Vorfall unverzüglich nachgehen und dem Büro für Künstliche Intelligenz und den zuständigen nationalen Behörden alle einschlägigen Informationen und mögliche Korrekturmassnahmen mitteilen.

- Cybersicherheit: Die Anbieter müssen während des gesamten Lebenszyklus des Modells ein angemessenes Mass an Cybersicherheit für das Modell und seine physische Infrastruktur gewährleisten.

Die Anbieter können die Einhaltung dieser Pflicht über Praxisleitfäden nachweisen, deren Bereitstellung vom Büro für Künstliche Intelligenz bis zur Veröffentlichung einer harmonisierten Norm erleichtert wird. Die Einhaltung harmonisierter Normen begründet für die Anbieter die Vermutung der Konformität (Art. 55 Abs. 2; Ziff. 5.2.11).

5.2.11 Die harmonisierten Normen und ihre Rolle in der KI-Verordnung[245]

Die Normung spielt bei der Umsetzung der KI-Verordnung eine zentrale Rolle.[246] So können die Anbieter der betroffenen KI-Systeme und GPAI-Modelle durch die Einhaltung der harmonisierten Normen die Konformität mit den Anforderungen der Verordnung nachweisen.

Um zu verstehen, wie diese Normen mit der KI-Verordnung zusammenhängen, werden im Folgenden der Begriff «harmonisierte Normen» (vgl. Ziff. 5.2.11.1), das Normungssystem der EU (vgl. Ziff. 5.2.11.2) und der Mechanismus der Konformitätsvermutung (vgl. Ziff. 5.2.11.3) vorgestellt. Anschliessend werden die Normungsarbeiten im KI-Bereich erläutert (vgl. Ziff. 5.2.11.4).

5.2.11.1 Die harmonisierten Normen

Eine «harmonisierte Norm» ist eine europäische Norm, die auf der Grundlage eines Auftrags der Europäischen Kommission zur Durchführung von Harmonisierungsrechtsvorschriften der EU angenommen wurde (vgl. Art. 2 Abs. 1 Bst. c der Verordnung [EU] 1025/2012[247]). Diese technischen Normen sollen die europäische Rechtssetzung unterstützen. Dabei handelt es sich um nicht rechtsverbindliche Regeln, Leitlinien oder Merkmale, die von Fachpersonen in den europäischen Normungsorganisationen erarbeitet wurden und die fast alle Gebiete des modernen Wirtschafts- und Alltagslebens betreffen. Technische Normen regeln vielfältige materielle und immaterielle Gegenstände wie Produkte, Verfahren, Messmethoden, Prozesse und Dienstleistungen und kommen in nahezu allen Branchen und Fachgebieten zum Einsatz.[248]

5.2.11.2 Das Normungssystem der EU

Das System der europäischen Normungsorganisationen bilden das Europäische Komitee für Normung (CEN) und das Europäische Komitee für elektrotechnische Normung (CENELEC)

---

[245] Dieses Kapitel wurde auf der Grundlage von Beiträgen der DV verfasst.

[246] Vgl. E. 121 der KI-Verordnung; s. auch MARTINA ARIOLI, Risikomanagement (Fn. 222), 12 N 38 ff.

[247] Verordnung (EU) Nr. 1025/2012 des Europäischen Parlaments und des Rates vom 25. Oktober 2012 zur europäischen Normung, zur Änderung der Richtlinien 89/686/EWG und 93/15/EWG des Rates sowie der Richtlinien 94/9/EG, 94/25/EG, 95/16/EG, 97/23/EG, 98/34/EG, 2004/22/EG, 2007/23/EG, 2009/23/EG und 2009/105/EG des Europäischen Parlaments und des Rates und zur Aufhebung des Beschlusses 87/95/EWG des Rates und des Beschlusses Nr. 1673/2006/EG des Europäischen Parlaments und des Rates, ABl. L 316/12, 14. November 2012.

[248] SECO, Förderung der Normungsorganisationen im Bereich der Digitalisierung: Akteure und Erkenntnisse in ausgewählten Themenbereichen, Bericht an den Bundesrat, 16. August 2022, 6, abrufbar unter: www.admin.ch > Dokumentation > Medienmitteilungen > Publikationshinweis > Bericht Förderung der Normierungsorganisationen im Bereich der Digitalisierung: Akteure und Erkenntnisse in ausgewählten Themenbereichen (abgerufen am 26. August 2024); für eine Zusammenfassung der bestehenden Normen auf internationaler und europäischer Ebene s. MÉLANIE GORNET/WINSTON MAXWELL, Normes techniques et éthique de l'IA. CNIA 2023 – Conférence Nationale en Intelligence Artificielle, Strassburg 2023, abrufbar unter: https://hal.science/hal-04121843 (abgerufen am 26. August 2024).

zusammen mit dem Europäischen Institut für Telekommunikationsnormen (ETSI). Das CEN ist verantwortlich für die europäischen Normen (EN) in allen technischen Bereichen ausser der Elektrotechnik und der Telekommunikation. Das CENELEC ist zuständig für die europäische Normung im Bereich Elektrotechnik (ENEC-Kennzeichnung) und das ETSI im Bereich der Telekommunikation.

Aus dem 1985 von der EU eingeführten «New Approach»-Konzept auf dem Gebiet der Produktevorschriften ergibt sich eine Verknüpfung zwischen Gesetzgebung und Normung. Der Gesetzgeber beschränkt sich auf das Formulieren von grundlegenden Anforderungen und Schutzzielen, die technisch durch Normen konkretisiert werden. Dieses Vorgehen soll einerseits den EU-Rechtssetzungsmechanismus entlasten, andererseits die Anpassung an den Stand der Technik und des Wissens ermöglichen.

Diese harmonisierten Normen werden von den europäischen Normungsorganisationen (CEN, CENELEC, ETSI) aufgrund eines von der EU-Kommission erteilten Mandates erarbeitet und im Amtsblatt der EU veröffentlicht. Nach ihrer Veröffentlichung im Amtsblatt der EU begründet ihre Anwendung eine Konformitätsvermutung. Die Normung spielt daher bei der Umsetzung der KI-Verordnung eine zentrale Rolle.

5.2.11.3 Die Konformitätsvermutung

5.2.11.3.1 Im Allgemeinen

Bei Produkten, die in Übereinstimmung mit harmonisierten Normen hergestellt worden sind, wird davon ausgegangen, dass sie die entsprechenden wesentlichen Anforderungen der einschlägigen Rechtsvorschriften erfüllen. Dieses Verfahren wird auch «Konformitätsbewertungsverfahren» genannt und führt bei positivem Ergebnis zur Berechtigung, eine «Konformitätserklärung» zu unterzeichnen und auf dem Produkt das CE-Zeichen anzubringen.

Entsprechend der spezifischen EU-Richtlinie oder EU-Verordnung kann ein Inverkehrbringer das Konformitätsbewertungsverfahren im Rahmen der Selbstkontrolle durchführen oder er muss akkreditierte Prüfstellen («Certified Bodies») und Benannte Stellen («Notified Bodies») zuziehen. Im Bedarfsfall können auch andere Normen berücksichtigt werden; in diesem Fall muss aber der Nachweis geführt werden, dass auch mit diesen Normen die «grundlegenden Sicherheits- und Gesundheitsanforderungen» erfüllt werden. Bei anderen Lösungen muss der Stand der Technik und des Wissens nachgewiesen werden.

5.2.11.3.2 In der KI-Verordnung

Gemäss Artikel 40 Absatz 1 der KI-Verordnung wird bei Hochrisiko-KI-Systemen oder bei GPAI-Modellen, die mit harmonisierten Normen oder Teilen davon übereinstimmen, deren Fundstellen gemäss der Verordnung 1025/2012[249] im Amtsblatt der EU veröffentlicht wurden, eine Konformität mit den Anforderungen gemäss Kapitel III Abschnitt 2 (im Zusammenhang mit Hochrisiko-KI-Systemen; Art. 8 ff. der KI-Verordnung) oder gegebenenfalls mit den Pflich-

[249] Verordnung (EU) 1025/2012 zur europäischen Normung (Fn. 247).

ten gemäss Kapitel V Abschnitte 2 und 3 der Verordnung (im Zusammenhang mit GPAI-Modellen) vermutet, soweit diese Anforderungen oder Verpflichtungen von den Normen abgedeckt sind.

Mit dieser Bestimmung wird somit eine Vermutung der Konformität mit den Anforderungen der Verordnung eingeführt, wenn die harmonisierten Normen eingehalten werden.

Damit diese harmonisierten Normen erarbeitet werden, erteilt die Europäische Kommission den europäischen Normungsorganisationen einen Normungsauftrag (vgl. Ziff. 5.2.11.4).

5.2.11.4 Die Normungsarbeiten im KI-Bereich

Am 22. Mai 2023 erteilte die Europäische Kommission dem CEN und dem CENELEC den Auftrag[250], die nötigen technischen Normen für die Umsetzung der KI-Verordnung auszuarbeiten. Konkret hat die Europäische Kommission die beiden Komitees aufgefordert, bis am 30. April 2025 zehn neue europäische Normen im KI-Bereich zu erarbeiten. Die Normen betreffen die zehn folgenden Bereiche:

- Risikomanagement für KI-Systeme
- Governance und Qualität von Datensätzen, die zur Entwicklung von KI-Systemen verwendet werden
- Aufzeichnung durch Protokollierungsfunktionen von KI-Systemen
- Transparenz und Bereitstellung von Information an Nutzerinnen und Nutzer von KI-Systemen
- Menschliche Aufsicht über KI-Systeme
- Spezifikationen für die Genauigkeit von KI-Systemen
- Spezifikationen für die Robustheit von KI-Systemen
- Spezifikationen für die Cybersicherheit von KI-Systemen
- Qualitätsmanagementsysteme für Anbieter von KI-Systemen, einschliesslich Verfahren zur Beobachtung nach dem Inverkehrbringen
- Konformitätsbewertung für KI-Systeme

[250] Durchführungsbeschluss der Kommission vom 22. Mai 2023 über einen Normungsauftrag an das Europäische Komitee für Normung und das Europäische Komitee für elektrotechnische Normung zur Unterstützung der Unionspolitik im Bereich der künstlichen Intelligenz, C(2023) 3215 final, abrufbar unter: https://ec.europa.eu/transparency/documents-register/detail?ref=C(2023)3215&lang=de (abgerufen am 26. August 2024).

Die zehn oben erwähnten technischen Normen werden von Fachpersonen in einem gemischten technischen Komitee, dem «Joint CEN-CENELEC Technical Committee 21 'Artificial Intelligence'» (CEN-CLC/JTC 21), ausgearbeitet.

Ist dieses Komitee nicht in der Lage, die von der Europäischen Kommission verlangten Normen innerhalb der festgelegten Frist (bis am 30. April 2025) auszuarbeiten, so kann die Kommission gemäss Artikel 41 der Verordnung gemeinsame Spezifikationen erlassen. Dieser Mechanismus ermöglicht somit der Kommission, einen gewissen Druck auf die Normungsorganisationen auszuüben.

5.2.12 Massnahmen zur Innovationsförderung

Die KI-Verordnung versucht, die Innovation mit der Sicherheit der KI-Systeme in Einklang zu bringen, indem ein System von «KI-Reallaboren» («regulatory sandboxes») geschaffen wird, also rechtliche Rahmenwerke, die beschränkte Innovationstests unter gesetzlicher Kontrolle ermöglichen.[251]

Die Mitgliedstaaten sind verpflichtet, mindestens ein solches KI-Reallabor auf nationaler Ebene einzurichten, das zwei Jahre nach dem Inkrafttreten der Verordnung einsatzbereit sein sollte. Sie können dieser Verpflichtung auch durch Beteiligung an einem bestehenden Reallabor nachkommen, sofern diese Beteiligung die nationale Abdeckung der teilnehmenden Mitgliedstaaten in gleichwertigem Mass gewährleistet (vgl. Art. 57).

Das Ziel dieser Reallabore besteht darin, eine kontrollierte Umgebung bereitzustellen, um die Innovation zu fördern und die Entwicklung, das Training, das Testen und die Validierung innovativer KI-Systeme für einen begrenzten Zeitraum vor ihrem Inverkehrbringen oder ihrer Inbetriebnahme zu erleichtern (vgl. Art. 57 Abs. 5). Die zuständigen nationalen Behörden müssen das Büro für Künstliche Intelligenz und das Gremium für Künstliche Intelligenz (vgl. Ziff. 5.2.13) über die Einrichtung eines Reallabors unterrichten. Das Büro für Künstliche Intelligenz veröffentlicht eine Liste der geplanten und bestehenden Reallabore (Art. 57 Abs. 15).

Die zuständigen nationalen Behörden haben die Aufgabe, innerhalb der KI-Reallabore Anleitung, Aufsicht und Unterstützung bereitzustellen, um Risiken im Zusammenhang mit Grundrechten, Gesundheit und Sicherheit sowie Risikominderungsmassnahmen und deren Wirksamkeit hinsichtlich der Pflichten und Anforderungen der Verordnung zu ermitteln. Auf Anfrage der Anbieter oder zukünftigen Anbieter kann die zuständige Behörde einen Abschlussbericht vorlegen, in dem sie die im Reallabor durchgeführten Tätigkeiten und die entsprechenden Ergebnisse im Einzelnen darlegt. Die Anbieter können diese Unterlagen nutzen, um im Rahmen des Konformitätsbewertungsverfahrens die Einhaltung der Verordnung nachzuweisen (vgl. Art. 57 Abs. 7).

Die am KI-Reallabor beteiligten Anbieter und zukünftigen Anbieter bleiben nach dem Recht der EU und der Mitgliedstaaten für Schäden haftbar, die Dritten infolge der Erprobung im Re-

[251] Für eine Kritik zur Wirksamkeit dieser Massnahmen vgl. DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 4 N 5.

allabor entstehen. Wenn jedoch der Anbieter den Plan und die Beteiligungsbedingungen beachtet hat und der Anleitung der zuständigen nationalen Behörden in gutem Glauben gefolgt ist, werden von den Behörden keine Geldbussen für Verstösse gegen die KI-Verordnung verhängt (Art. 57 Abs. 12). Die Verordnung schafft somit einen starken Anreiz für die Beteiligung an einem KI-Reallabor.

Schliesslich sieht die Verordnung unter bestimmten Voraussetzungen auch die Möglichkeit vor, Hochrisiko-KI-Systeme unter Realbedingungen in und ausserhalb eines KI-Reallabors zu testen (Art. 57, 58 und 60 f. der Verordnung). Die Durchführung eines Tests unter Realbedingungen muss von der Marktüberwachungsbehörde genehmigt werden (vgl. Art. 60).

5.2.13 Governance

Mit der Verordnung wird ein mehrstufiger Governance-Rahmen geschaffen, der zum Ziel hat, die Anwendung der Verordnung auf nationaler Ebene zu koordinieren und zu unterstützen, auf EU-Ebene Kapazitäten aufzubauen und Interessenträger im KI-Bereich zu integrieren.

Auf EU-Ebene sieht die Verordnung folgende Governance-Organe vor:

- Büro für Künstliche Intelligenz («AI Office», Art. 64): Dieses innerhalb der Europäischen Kommission geschaffene Büro soll zur Umsetzung, Beobachtung und Überwachung von KI-Systemen und GPAI-Modellen sowie der KI-Governance beitragen (vgl. Art. 3 Nr. 47). Es stellt insbesondere Muster bereit, die bei der Anwendung der Verordnung helfen (vgl. z. B. Art. 27 Abs. 5, der vorsieht, dass das Büro für Künstliche Intelligenz ein Muster für einen Fragebogen, auch mithilfe eines automatisierten Instruments, ausarbeitet, das die Betreiber bei ihrer Pflicht, eine Folgenabschätzung durchzuführen, unterstützen soll).

- Europäisches Gremium für Künstliche Intelligenz («European AI Board», Art. 65): Dieses setzt sich aus Vertreterinnen und Vertretern der Mitgliedstaaten zusammen und hat die Aufgabe, die Europäische Kommission und die Mitgliedstaaten zu beraten und zu unterstützen, um die einheitliche und wirksame Anwendung der Verordnung zu erleichtern. Der Europäische Datenschutzbeauftragte nimmt als Beobachter teil.

- Beratungsforum («Advisory Forum», Art. 67): Dieses stellt technisches Fachwissen bereit und berät das Gremium für Künstliche Intelligenz und das Büro für Künstliche Intelligenz. Es hat die Aufgabe, die Teilnahme der Interessenträger an der Umsetzung und Anwendung der Verordnung sicherzustellen. Das Forum setzt sich unter anderem aus Interessenträger der Industrie, der Universitäten, der Zivilgesellschaft sowie aus der Agentur der EU für Grundrechte, der Agentur der EU für Cybersicherheit, dem CEN, dem CENELEC und dem ETSI zusammen.

- Wissenschaftliches Gremium unabhängiger Sachverständiger («Scientific Panel of Independent Experts», Art. 68): Dieses setzt sich aus Sachverständigen der Wissenschaft zusammen und soll die Umsetzung und Anwendung der Verordnung unterstützen, namentlich die Durchsetzungstätigkeiten des Büros für Künstliche Intelligenz in Bezug auf GPAI-Modelle (vgl. Ziff. 5.2.10). Die Mitgliedstaaten können das Gremium um Hilfe für ihre Durchsetzungstätigkeiten im Rahmen der Verordnung bitten (Art. 69).

Auf nationaler Ebene schaffen oder benennen die Mitgliedstaaten mindestens eine notifizierende Behörde und mindestens eine Marktüberwachungsbehörde als zuständige nationale Behörden (Art. 70):

- Notifizierende Behörde: Dies ist die nationale Behörde, die dafür zuständig ist, die erforderlichen Verfahren für die Bewertung, Benennung und Notifizierung von Konformitätsbewertungsstellen und für deren Überwachung einzurichten und durchzuführen, also derjenigen Stellen, die für die Konformitätsbewertung von Hochrisiko-KI-Systemen zuständig sind (vgl. Art. 3 Nrn. 19–22; Ziff. 5.2.7.4).

- Marktüberwachungsbehörde: Hier handelt es sich um die nationale Behörde, die die Tätigkeiten durchführt und die Massnahmen ergreift, die in der Verordnung (EU) 2019/1020 über Marktüberwachung und die Konformität von Produkten[252] vorgesehen sind (vgl. Art. 3 Nr. 26) (vgl. Ziff. 5.2.14).

  Jeder Mitgliedstaat muss eine Marktüberwachungsbehörde benennen, die als zentrale Anlaufstelle für die Verordnung fungiert (Art. 70 Abs. 2).

5.2.14 Überwachung und Durchsetzung

Den Mitgliedstaaten kommt bei der Anwendung und Durchsetzung der Verordnung eine Schlüsselrolle zu. Jeder Mitgliedstaat benennt eine oder mehrere zuständige nationale Behörden, die die Anwendung und Umsetzung der Vorschriften überwachen und Marktüberwachungstätigkeiten durchführen.

Für die Überwachung verweist die KI-Verordnung auf die Verordnung (EU) 2019/1020[253] (vgl. Art. 74). Die KI-Verordnung enthält jedoch wichtige Präzisierungen, insbesondere zur Organisation und Koordinierung der Marktüberwachungsbehörden. So beispielsweise:

- Bei Hochrisiko-KI-Systemen und damit in Zusammenhang stehenden Produkten, auf die die in Anhang I Abschnitt A aufgeführten Harmonisierungsrechtsvorschriften der EU Anwendung finden, gilt als Marktüberwachungsbehörde im Sinne der KI-Verordnung die in diesen Vorschriften für die Marktüberwachung benannte Behörde (Art. 74 Abs. 3).

- Bei Hochrisiko-KI-Systemen, die in Finanzinstituten verwendet werden, gilt gemäss dem Recht der EU im Bereich der Finanzdienstleistungen die für die Finanzaufsicht

[252] Verordnung (EU) 2019/1020 des Europäischen Parlaments und des Rates vom 20. Juni 2019 über Marktüberwachung und die Konformität von Produkten sowie zur Änderung der Richtlinie 2004/42/EG und der Verordnungen (EG) Nr. 765/2008 und (EU) Nr. 305/2011, ABl. L 169/1, 25. Juni 2019.

[253] Verordnung (EU) 2019/1020 über Marktüberwachung und die Konformität von Produkten (Fn. 252).

über diese Institute benannte nationale Behörde als Marktüberwachungsbehörde, sofern ein Zusammenhang mit der Erbringung dieser Finanzdienstleistungen besteht (Art. 74 Abs. 6).

– Für die in Anhang III Nummer 1 (Biometrie) genannten Hochrisiko-KI-Systeme, sofern diese Systeme für Strafverfolgungszwecke, Grenzmanagement sowie Justiz und Demokratie eingesetzt werden, und für die in Anhang III Nummern 6 bis 8 genannten Hochrisiko-KI-Systeme benennen die Mitgliedstaaten als Marktüberwachungsbehörden die für den Datenschutz zuständigen Aufsichtsbehörden (Art. 74 Abs. 8).[254]

Grundsätzlich gewähren die Anbieter den Marktüberwachungsbehörden uneingeschränkten Zugang zur Dokumentation sowie zu den für die Entwicklung von Hochrisiko-KI-Systemen verwendeten Trainings-, Validierungs- und Testdatensätzen (Art. 74 Abs. 12). Unter bestimmten Bedingungen können die Überwachungsbehörden Zugang zum Quellcode eines Hochrisiko-KI-Systems erhalten (vgl. Art. 74 Abs. 13).

In der KI-Verordnung werden die Interventionsbefugnisse und die Verfahren beschrieben, die die Überwachungsbehörden je nach Umständen einsetzen können (Art. 79 ff.) Wenn ein Hochrisiko-KI-System beispielsweise ein Risiko für die Gesundheit, die Sicherheit oder die Grundrechte von Personen darstellt, führt die Überwachungsbehörde selber eine Konformitätsbewertung unter Berücksichtigung aller in der Verordnung festgelegten Anforderungen durch. Stellt sie fest, dass die Voraussetzungen nicht erfüllt sind, fordert sie den betreffenden Akteur auf, alle geeigneten Korrekturmassnahmen zu ergreifen (Art. 79).

In einer Art Generalklausel ermächtigt die Verordnung die Marktüberwachungsbehörde zudem, für jedes KI-System, das zwar den Vorgaben der Verordnung entspricht, aber dennoch ein Risiko für die Gesundheit oder Sicherheit von Personen, für die Grundrechte oder für andere Aspekte des Schutzes öffentlicher Interessen darstellt, besondere Massnahmen anzuordnen (vgl. Art. 82 Abs. 1).[255]

Für die Organe, Einrichtungen und sonstigen Stellen der UE, die in den Geltungsbereich der Verordnung fallen, wird der Europäische Datenschutzbeauftragte als zuständige Marktüberwachungsbehörde benannt (Art. 74 Abs. 9 und E. 156).

---

[254] Vgl. DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 6 N 9, der der Ansicht ist, dass die Mitgliedstaaten, zusätzlich zu den in Art. 74 Abs. 8 der KI-Verordnung genannten Situationen, in ungefähr der Hälfte der Fälle die Datenschutzbehörden als Überwachungsbehörde bezeichnen werden. S. auch Erklärung des Europäischen Datenschutzausschusses vom 16. Juli 2024, in der die Datenschutzbehörden den Wunsch äussern, mit der Anwendung der KI-Verordnung in Bezug auf die Hochrisiko-KI-Systeme beauftragt zu werden. Die Erklärung ist abrufbar unter: https://www.edpb.europa.eu/news/news/2024/edpb-adopts-statement-dpas-role-ai-act-framework-eu-us-data-privacy-framework-faq_en (abgerufen am 26. August 2024).

[255] DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 30 N 59.

Bei GPAI-Modellen verfügt die Europäische Kommission über das Büro für Künstliche Intelligenz über ausschliessliche Befugnisse zur Beaufsichtigung und Durchsetzung der Verordnung (vgl. Art. 88).

Ferner überträgt die KI-Verordnung den für den Schutz der Grundrechte zuständigen Behörden gewisse Befugnisse. Gemäss Artikel 77 Absatz 1 sind die nationalen Behörden oder öffentlichen Stellen, die die Einhaltung des EU-Rechts zum Schutz der Grundrechte beaufsichtigen oder durchsetzen, befugt, sämtliche Dokumentation im Zusammenhang mit der Verwendung der in Anhang III genannten Hochrisiko-KI-Systeme anzufordern, sofern dies für die wirksame Ausübung ihrer Aufträge im Rahmen ihrer Befugnisse notwendig ist. Die Mitgliedstaaten erstellen eine Liste dieser Behörden oder Stellen (Art. 77 Abs. 2). Wenn dies zur Feststellung, ob gegen die Pflichten verstossen wurde, notwendig ist, kann die Behörde oder die Stelle die Durchführung eines Tests des Hochrisiko-KI-Systems beantragen (vgl. Art. 77 Abs. 3).

5.2.15 Individuelle Rechte

In einem separaten Abschnitt mit dem Titel «Rechtsbehelfe» (Art. 85 f.) enthält die KI-Verordnung zwei Bestimmungen, die individuelle Rechte betreffen:

- Recht auf Beschwerde: Jede natürliche oder juristische Person, die Grund zur Annahme hat, dass gegen die Bestimmungen der Verordnung verstossen wurde, kann bei der betreffenden Marktüberwachungsbehörde Beschwerde einreichen. Diese Beschwerden werden gemäss der Verordnung (EU) 2019/1020[256] und der Durchführung von Überwachungstätigkeiten berücksichtigt. Dieses Recht besteht unbeschadet anderer verwaltungsrechtlicher oder gerichtlicher Rechtsbehelfe (Art. 85).

- Recht auf Erläuterung der Entscheidungsfindung im Einzelfall: Jede betroffene Person hat unter folgenden Bedingungen das Recht, vom Betreiber eines KI-Systems eine klare und aussagekräftige Erläuterung zur Rolle des KI-Systems im Entscheidungsprozess und zu den wichtigsten Elementen der getroffenen Entscheidung zu erhalten (Art. 86 Abs. 1):

    - Es handelt sich um ein Hochrisiko-KI-System, das in Anhang III aufgeführt ist, mit Ausnahme der unter Nummer 2 genannten Systeme (kritische Infrastruktur).

    - Die Entscheidung wurde vom Betreiber auf der Grundlage der Ausgaben dieses Systems getroffen.

    - Die Entscheidung hat rechtliche Auswirkungen oder beeinträchtigt die Person in ähnlicher Art erheblich auf eine Weise, die ihrer Ansicht nach negative Auswirkungen auf ihre Gesundheit, ihre Sicherheit oder ihre Grundrechte hat.

[256] Verordnung (EU) 2019/1020 über Marktüberwachung und die Konformität von Produkten (Fn. 252).

Dieses Recht kann nicht geltend gemacht werden, wenn das EU-Recht oder das nationale Recht Ausnahmen oder Beschränkungen vorsieht. Schliesslich gilt Artikel 86 nur insoweit, als das Recht gemäss Absatz 1 nicht anderweitig im EU-Recht festgelegt ist.

5.2.16 Sanktionen

Die Verordnung sieht für Verstösse gegen die in der KI-Verordnung festgelegten Pflichten erhebliche finanzielle Sanktionen vor.

Insbesondere werden Verstösse gegen die verbotenen KI-Praktiken mit Geldbussen von bis zu 35 Millionen Euro oder, im Fall von Unternehmen, von bis zu 7 Prozent des gesamten weltweiten Jahresumsatzes des vorangegangenen Geschäftsjahrs geahndet, je nachdem, welcher Betrag höher ist (vgl. Art. 99 Abs. 3).

Verstösse gegen die Bestimmungen über die Akteure oder die notifizierten Stellen (vgl. z. B die Pflichten der Anbieter nach Art. 16) werden mit Geldbussen von bis zu 15 Millionen Euro oder, im Fall von Unternehmen, von bis zu 3 Prozent des gesamten weltweiten Jahresumsatzes des vorangegangenen Geschäftsjahres geahndet, je nachdem, welcher Betrag höher ist (vgl. Art. 99 Abs. 4). Gegen Anbieter von KI-Modellen mit allgemeinem Verwendungszweck können Bussen im gleichen Umfang verhängt werden (vgl. Art. 101).

Werden notifizierten Stellen oder zuständigen nationalen Behörden auf deren Auskunftsersuchen hin falsche, unvollständige oder irreführende Informationen bereitgestellt, so werden Geldbussen von bis zu 7,5 Millionen Euro oder, im Fall von Unternehmen, von bis zu 1 Prozent des gesamten weltweiten Jahresumsatzes des vorangegangenen Geschäftsjahres verhängt, je nachdem, welcher Betrag höher ist (vgl. Art. 99 Abs. 5).

Für KMU, einschliesslich Start-up-Unternehmen, belaufen sich die vorstehend genannten Bussen auf höchstens die in Artikel 99 Absätze 3 bis 5 genannten Prozentsätze oder Beträge (vgl. Art. 99 Abs. 6).

5.2.17 Inkrafttreten und Geltungsbeginn

Die KI-Verordnung gilt nach Ablauf von zwei Jahren nach ihrem Inkrafttreten am 1. August 2024 (Art. 113). Die Verordnung ist für ihre Adressatinnen (Anbieter, Betreiber, betroffene Personen usw.) direkt anwendbar, ohne dass sie vorgängig in das Recht der Mitgliedstaaten übertragen werden muss (Art. 288 AEUV).[257] Es gibt jedoch Ausnahmen von der Frist von zwei Jahren. So beispielsweise:

- Die Kapitel I (Allgemeine Bestimmungen) und II (Verbotene Praktiken) gelten bereits sechs Monate nach dem Inkrafttreten, also ab dem 2. Februar 2025.

- Die Vorschriften für die GPAI-Modelle gelten, mit Ausnahme von Artikel 101, bereits zwölf Monate nach dem Inkrafttreten, also ab dem 2. August 2025. Das Gleiche gilt für

[257] DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 2 N 2.

die Vorschriften für die Stellen, die Konformitätserklärungen ausstellen können, die Vorgaben für die Schaffung der in der Verordnung vorgesehenen Behörden und Stellen (z. B. das Büro für Künstliche Intelligenz) und die Bestimmungen über die Sanktionen.

- Artikel 6 Absatz 1 der Verordnung und die entsprechenden Pflichten (Hochrisiko-KI-Systeme im Rahmen der Harmonisierungsrechtsvorschriften der EU für Produkte) gilt erst 36 Monate nach dem Inkrafttreten, also ab dem 2. August 2027.

5.3 Würdigung

5.3.1 Rechtliche Auswirkungen auf die Schweizer Akteure

5.3.1.1 Betroffene Akteure

Wie oben erwähnt (vgl. Ziff. 5.2.5), gilt die KI-Verordnung für Anbieter, die sich in der EU *oder in einem Drittland* befinden oder dort niedergelassen sind und in der EU KI-Systeme in Verkehr bringen oder in Betrieb nehmen oder GPAI-Modelle in Verkehr bringen (Art. 2 Abs. 1 Bst. a). Das Inverkehrbringen oder die Inbetriebnahme betrifft gemäss den Begriffsbestimmungen (vgl. Art. 3 Nrn. 9 und 11) nur die erstmalige Bereitstellung beziehungsweise die erste Verwendung auf dem EU-Markt.

Gemäss Artikel 2 Absatz 1 Buchstabe c gilt die Verordnung auch für Anbieter und Betreiber von KI-Systemen, die ihren Sitz *in einem Drittland* haben oder sich *in einem Drittland* befinden, wenn die vom KI-System hervorgebrachte Ausgabe in der EU verwendet wird.

Mit diesen Bestimmungen soll verhindert werden, dass die Verordnung umgangen wird, indem KI-Systeme in einem Drittland entwickelt werden und das vom KI-System erzeugte Ergebnis schliesslich in der EU verwendet wird.

Ein Teil der Lehre weist jedoch darauf hin, dass die Auslegung von Artikel 2 Absatz 1 Buchstaben a und c nicht ganz unproblematisch ist: Da der Begriff «Anbieter» das Inverkehrbringen auf dem EU-Markt voraussetzt (vgl. Art. 3 Nrn. 3 und 9), bedeutet dies, dass ein Anbieter, der das KI-System nicht auf dem EU-Markt in Verkehr bringt, nicht als Anbieter im Sinne der Verordnung gilt, auch wenn sein KI-System letztlich den Weg in die EU findet oder dort Wirkung erzielt. Daher würde Artikel 2 Absatz 1 Buchstabe c, gemäss dem jeder Anbieter, der sich in einem Drittland befindet, der Verordnung untersteht, wenn die von seinem KI-System erzeugten Ergebnisse in der EU verwendet werden, den Begriff «Anbieter» *de facto* erweitern. In diesem Fall ist die Anwendung der Verordnung durch den blossen Umstand gegeben, dass das Ergebnis in der EU verwendet wird.[258] Es muss beobachtet werden, wie die beiden Bestimmungen künftig miteinander in Einklang gebracht werden. Es scheint auf jeden Fall, dass die Absicht des europäischen Gesetzgebers eher in Richtung einer umfassenden Anwendung der Verordnung geht.

Für die Betreiber, die sich in einem Drittland befinden, ist das Kriterium die Verwendung des Ergebnisses in der EU (vgl. Art. 2 Abs. 1 Bst. c). Als Beispiel kann der Fall eines Schweizer

[258] DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 9 N 24; s. auch MARTINA ARIOLI, Risikomanagement (Fn. 222), 15 N 50 f.

Unternehmens angeführt werden, das seinen Kunden in der EU Texte zustellt, die mit einem KI-System generiert wurden.[259]

In den Erwägungen der KI-Verordnung wird auch der Fall erwähnt, in dem ein in der EU niedergelassener Akteur bestimmte Dienstleistungen im Zusammenhang mit einer Tätigkeit, die von einem als hochriskant eingestuften KI-System ausgeübt werden soll, an einen in einem Drittland niedergelassenen Akteur vergibt. Unter diesen Umständen könnte das vom Akteur in einem Drittland betriebene KI-System Daten verarbeiten, die rechtmässig in der EU erhoben und aus der EU übermittelt wurden, und dem vertraglichen Akteur in der EU die aus dieser Verarbeitung resultierende Ausgabe dieses KI-Systems liefern. Die KI-Verordnung sollte daher auch für diesen Fall gelten (vgl. E. 22).

5.3.1.2 Auswirkungen

Wenn Schweizer Akteure von der Verordnung betroffen sind, müssen sie die darin vorgesehenen Pflichten für das betreffende KI-System oder GPAI-Modell einhalten.

Weist das KI-System ein minimales Risiko auf, so gibt es keine Pflichten, sondern höchstens eine freiwillige Anwendung der Verhaltenskodizes (vgl. Ziff. 5.2.9). Wenn das KI-System von einem der in Artikel 50 aufgeführten Fälle erfasst ist, müssen die entsprechenden Transparenzpflichten eingehalten werden (vgl. Ziff. 5.2.8).

Für Hochrisiko-KI-Systeme gelten die Pflichten von Kapitel III Abschnitt 2 (vgl. Ziff. 5.2.7). Insbesondere müssen Schweizer Anbieter von Hochrisiko-KI-Systemen eine Konformitätsbewertung ihrer Produkte durchführen. Für weitere Ausführungen zu diesem Punkt wird auf die Erläuterungen unten im Zusammenhang mit dem MRA verwiesen (vgl. Ziff. 5.3.2).

KI-Systeme mit unannehmbaren Risiken sind schlicht und einfach verboten (abgesehen von den unter Ziff. 5.2.6 genannten Ausnahmen). Bei GPAI-Modellen haben sich die Anbieter an die Pflichten gemäss den Artikeln 51 ff. der KI-Verordnung zu halten (vgl. Ziff. 5.2.10).

Weiter ist zu beachten, dass die in der Schweiz niedergelassenen Anbieter von Hochrisiko-KI-Systemen und GPAI-Modellen vor dem Inverkehrbringen dieser Systeme und Modelle in der EU schriftlich einen in der EU niedergelassenen Bevollmächtigten benennen müssen (vgl. Art. 22 und 54 der KI-Verordnung). Bei Hochrisiko-KI-Systemen muss der Bevollmächtigte insbesondere sicherstellen, dass das KI-System in der EU-Datenbank gemäss Artikel 49 registriert ist.

Auf begründete Anfrage der zuständigen Behörden beziehungsweise des Büros für Künstliche Intelligenz muss der Bevollmächtigte alle Informationen übermitteln, die erforderlich sind, um die Konformität eines Hochrisiko-KI-Systems oder eines GPAI-Modells nachzuweisen (vgl. Art. 22 Abs. 3 Bst. c und d und 54 Abs. 3 Bst. c und d der KI-Verordnung). Zudem muss der Bevollmächtigte mit diesen bei allen ergriffenen Massnahmen zusammenarbeiten. Wenn ein Bevollmächtigter eines in der Schweiz niedergelassenen Anbieters aufgefordert würde, Informationen bereitzustellen, die sich in der Schweiz befinden, und diese den zuständigen

[259] Für weitere Beispiele s. DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 14 N 32.

europäischen Behörden ohne Bewilligung übermitteln würde, obwohl dafür die Schweizer Behörden zuständig wären, würde sich die Frage eines allfälligen Verstosses gegen Artikel 271 StGB stellen. Diese Bestimmung stellt Handlungen, die ohne Bewilligung für einen fremden Staat vorgenommen werden, unter Strafe.

Die Verordnung sieht weiter vor, dass der Bevollmächtigte den Auftrag beendet, wenn er der Auffassung ist oder Grund zur Annahme hat, dass der Anbieter gegen seine Pflichten gemäss der Verordnung verstösst.

Anbieter von Hochrisiko-KI-Systemen geben zudem auf dem Hochrisiko-KI-System oder, falls dies nicht möglich ist, auf seiner Verpackung oder in der beigefügten Dokumentation ihren Namen, ihren eingetragenen Handelsnamen oder ihre eingetragene Handelsmarke und ihre Kontaktanschrift an (vgl. Art. 16 Bst. b).

5.3.2 Verhältnis zum Abkommen zwischen der Schweiz und der EU über die gegenseitige Anerkennung von Konformitätsbewertungen[260]

5.3.2.1 Funktionsweise des Abkommens

Mit dem Abkommen zwischen der Schweiz und der EU über die gegenseitige Anerkennung von Konformitätsbewertungen (MRA)[261] soll die gegenseitige Anerkennung von Konformitätsbewertungen für die Vermarktung von Industrieprodukten (z. B. Maschinen, Medizinprodukte, Aufzüge) in Sektoren sichergestellt werden, in denen im EU-Recht harmonisierte Vorschriften existieren und eine zwingende Konformitätsbewertung vorgesehen ist.[262] Das Abkommen ermöglicht einen erleichterten Zugang zum EU-Binnenmarkt für die Produktsektoren, die es abdeckt, und reduziert Zeit und Kosten für die Vermarktung der Produkte. Konkret leistet das MRA einen Beitrag zum Abbau der technischen Handelshemmnisse. Zum einen geschieht dies durch die Harmonisierung der Vorschriften der Schweiz und der EU und zum anderen durch die Durchführung einer einzigen Konformitätsbewertung für den Zutritt zum EU-Markt beziehungsweise zum Schweizer Markt, die auf der Grundlage der technischen Vorschriften der Schweiz oder der EU von einer durch das Abkommen anerkannten Konformitätsbewertungsstelle ausgestellt wird. Das Abkommen beruht auf der Gleichwertigkeit der Gesetzgebungen der Schweiz und der EU. Bei jeder substanziellen Revision der technischen Vorschriften der EU in einem durch das Abkommen abgedeckten Sektor, muss die schweizerische Gesetzgebung angepasst werden, um die Gleichwertigkeit mit den Rechtsvorschriften der EU zu erhalten, und das entsprechende Kapitel des Abkommens aktualisiert werden.

Seit einigen Jahren lehnt die EU jedoch aufgrund institutioneller Fragen in den Beziehungen zwischen der Schweiz und der EU eine Aktualisierung des MRA ab. 2017 hat die EU eine

---

[260] Dieses Kapitel wurde auf der Grundlage von Beiträgen des BAKOM verfasst.

[261] Abkommen zwischen der Schweizerischen Eidgenossenschaft und der Europäischen Gemeinschaft über die gegenseitige Anerkennung von Konformitätsbewertungen, SR **0.946.526.81**.

[262] Botschaft zur Genehmigung der sektoriellen Abkommen zwischen der Schweiz und der EG, BBl **1999** 6128, 6212.

neue Regulierung der Medizinprodukte erlassen, die im Mai 2021 in Kraft getreten ist. Die Schweiz hat eine Gesetzgebung verabschiedet, die mit jener der EU im Einklang steht. Doch seit dem 26. Mai 2021 weigert sich die EU, das MRA im Bereich der Medizinprodukte zu aktualisieren, weil bei den Verhandlungen zu den institutionellen Fragen keine Fortschritte erzielt werden. Der Schweiz wird deshalb bei der Konformitätsbewertung nicht mehr die gegenseitige Anerkennung gewährt und Schweizer Anbieter von Medizinprodukten sind beim Zugang zum EU-Markt mit vielfältigen Hindernissen konfrontiert, wie der Notwendigkeit, die Konformität der Produkte durch eine Konformitätsbewertungsstelle der EU bewerten zu lassen. In Zukunft könnte diese Blockade auch andere Sektoren des MRA betreffen, die in der EU einer umfassenden Revision unterzogen wurden oder derzeit unterzogen werden (Maschinen, Bauprodukte, Spielzeug). Dies ist jedoch abhängig vom Ausgang der laufenden Verhandlungen zwischen der Schweiz und der EU.

5.3.2.2 Auswirkungen der KI-Verordnung auf die vom Abkommen betroffenen Schweizer Akteure

Die KI-Verordnung regelt die KI-spezifischen Aspekte in der Mehrheit der vom MRA abgedeckten Produktsektoren. Sie gilt zusätzlich zu den sektorspezifischen Rechtsvorschriften für Produkte und sieht spezifische zusätzliche Anforderungen im Zusammenhang mit KI vor.

In der Verordnung werden «Hochrisiko-KI-Systeme» als KI-Systeme definiert, die als Sicherheitsbauteile von Produkten verwendet werden sollen oder selbst Produkte sind, die unter die in Anhang I der Verordnung aufgeführten Rechtsvorschriften fallen – so beispielsweise Maschinen oder Sicherheitsbauteile von Maschinen – und die einer Konformitätsbewertung durch eine als Dritte auftretende Stelle unterzogen werden (vgl. Art. 6 Abs. 1).[263]

Folgende Sektoren des MRA entsprechen den durch Anhang I Abschnitt A der KI-Verordnung abgedeckten Bereichen:

- Maschinen (Kapitel 1 MRA)
- Persönliche Schutzausrüstungen (Kapitel 2 MRA)
- Spielzeuge (Kapitel 3 MRA)
- Medizinprodukte (Kapitel 4 MRA)
- Gasverbrauchseinrichtungen und Heizkessel (Kapitel 5 MRA)

[263] In der Verordnung wird unterschieden zwischen Harmonisierungsrechtsvorschriften gemäss dem neuen Rechtsrahmen («New Legislative Framework», NFL; s. dazu MARTINA ARIOLI, Risikomanagement [Fn. 222], 4 N 8 f.), die in Anhang I Abschnitt A der Verordnung aufgeführt sind, und Rechtsvorschriften nach dem alten Konzept, die in Anhang I Abschnitt B der Verordnung genannt sind. Die KI-Verordnung gilt direkt für Bereiche, die von Anhang I Abschnitt A erfasst sind, während die Europäische Kommission für die Bereiche, die unter das alte Konzept fallen, den Anforderungen an Hochrisiko-KI-Systeme Rechnung tragen muss, wenn sie delegierte Rechtsakte oder Durchführungsrechtsakte für KI-Systeme erlässt, bei denen es sich um Sicherheitsbauteile gemäss KI-Verordnung handelt. Ein grosser Teil der vom MRA erfassten Bereiche fällt unter Abschnitt A und nur zwei im MRA geregelte Bereiche sowie die Zivilluftfahrt und die Interoperabilität des Eisenbahnsystems gemäss den bilateralen Abkommen über Land- und Luftverkehr betreffen Abschnitt B.

- Druckgeräte (Kapitel 6 MRA)
- Funkanlagen und Telekommunikationsendgeräte (Kapitel 7 MRA)
- Geräte und Schutzsysteme zur bestimmungsgemässen Verwendung in explosionsgefährdeten Bereichen (ATEX) (Kapitel 8 MRA)
- Aufzüge (Kapitel 17 MRA)
- Seilbahnen (Kapitel 19 MRA)

Folgende Sektoren des MRA entsprechen den durch Anhang I Abschnitt B der KI-Verordnung abgedeckten Bereichen:

- Kraftfahrzeuge (Kapitel 12 MRA)
- Land- und forstwirtschaftliche Zugmaschinen (Kapitel 13 MRA)

Die KI-Verordnung gilt somit für die Mehrheit der durch das MRA abgedeckten Sektoren, die somit drei Jahre nach dem Inkrafttreten der Verordnung (Datum, ab dem Art. 6 Abs. 1 der KI-Verordnung und die entsprechenden Pflichten gelten, vgl. Ziff. 5.2.17) den Bestimmungen für Hochrisiko-KI-Systeme unterstehen. In den Bereichen, die unter die in Anhang I der Verordnung aufgeführten Harmonisierungsrechtsvorschriften der EU fallen, wird daher die Bewertung der Einhaltung der Anforderungen an die Hochrisiko-KI-Systeme in die Konformitätsbewertungsverfahren gemäss diesen Vorschriften integriert (vgl. Ziff. 5.2.7.1 und 5.2.7.4.1). Die Pflichten der KI-Verordnung ergänzen somit die bestehenden sektorspezifischen Pflichten.

Für Schweizer Anbieter von KI-Produkten, die sowohl unter das MRA und die Harmonisierungsrechtsvorschriften der EU gemäss Anhang I der KI-Verordnung (Hochrisiko-KI-Systeme gemäss Art. 6 Abs. 1 KI-Verordnung) fallen, erfolgt die Konformitätsbewertung gemäss dem im MRA vorgesehenen Verfahren, muss aber auch die Bewertung der Konformität mit den Anforderungen an Hochrisiko-KI-Systeme gemäss der KI-Verordnung beinhalten. In der gegenwärtigen Situation ohne schweizerische KI-Gesetzgebung und solange die Bestimmungen der KI-Verordnung für die oben genannten Produkte nicht in das MRA aufgenommen wurden, muss die Konformitätsbewertung für die Einhaltung der Anforderungen der KI-Verordnung durch eine Konformitätsbewertungsstelle der EU und gemäss EU-Recht durchgeführt werden. Um ein solches Produkt oder einen solchen Produktbestandteil in der EU in Verkehr zu bringen, müssen Schweizer Akteure zudem in der EU einen Bevollmächtigten bezeichnen, der sicherstellen muss, dass das Konformitätsbewertungsverfahren durchgeführt wurde (vgl. Art. 22 und 54 KI-Verordnung). Ausserdem müssen die Einführer (die gemäss Definition der KI-Verordnung in der EU niedergelassen sind) ihre Kontaktdaten auf der Verpackung angeben (Art. 23). Somit kommt es zu neuen technischen Handelshemmnissen.

5.3.2.3 Mögliche Erweiterung des MRA

Zwei Schritte wären nötig, um im Rahmen des MRA diese technischen Handelshemmnisse beim EU-Marktzugang gänzlich zu beseitigen. Erstens muss die Schweiz ihre Produktevorschriften mit denjenigen der EU harmonisieren, im vorliegenden Fall also mit den relevanten Produktvorschriften der KI-Verordnung. Zweitens müsste das MRA entsprechend erweitert werden.

Die EU und die Schweiz könnten folglich beschliessen, das MRA zu erweitern und die Produktvorschriften der KI-Verordnung in das MRA zu integrieren. Dies würde die gegenseitige Anerkennung künftiger Konformitätsbewertungen von Produkten, die KI-Systeme beinhalten,

ermöglichen. Diese Option ist jedoch abhängig vom Ausgang der aktuellen Verhandlungen zwischen der Schweiz und der EU.

Selbst wenn die Schweiz eine KI-Gesetzgebung schaffen würde, die sich an jene der EU anlehnt, würden die in der Schweiz durchgeführten Konformitätsbewertungsverfahren für die KI-Aspekte nicht anerkannt werden, solange das MRA nicht aktualisiert wurde.

5.3.3 Verhältnis zum Angemessenheitsbeschluss der Europäischen Kommission im Bereich Datenschutz

Die Übermittlung von Personendaten aus einem EU-Staat in einen Drittstaat unterliegt, bis auf einige Ausnahmen, der Voraussetzung, dass die Daten im Empfängerstaat auf ähnliche Weise geschützt sind. Dieses Schutzniveau kann durch einen Vertrag, Ad-hoc-Garantien (z. B. Standardvertragsklauseln) oder durch einen Angemessenheitsbeschluss gewährleistet werden (s. Art. 45 DSGVO und Art. 36 Richtlinie (EU) 2016/680[264]).[265] Dieser Beschluss bestätigt, dass der Empfängerstaat über ein Datenschutzniveau verfügt, das als gleichwertig gilt. In diesem Rahmen nimmt die Europäische Kommission unter Berücksichtigung der Anforderungen des EU-Rechts eine eingehende Prüfung der rechtlichen Situation im betreffenden Staat vor, insbesondere der einschlägigen Rechtsvorschriften und deren Umsetzung.

Die Schweiz verfügt seit dem Jahr 2000 über einen Angemessenheitsbeschluss der Europäischen Kommission.[266] Dieser wurde im Januar 2024 nach einer mehrjährigen Überprüfung zusammen mit jenen von zehn anderen Ländern[267] bestätigt. Dieser Beschluss betrifft die Übermittlung von Daten im privaten und öffentlichen Sektor, mit Ausnahme der Übermittlung im Rahmen der Strafverfolgung. Er ist sehr wichtig, denn er erlaubt, Personendaten ohne zusätzliche Garantien zu übermitteln. Er erleichtert die Wirtschaftsbeziehungen.

Die Verwendung von KI-Systemen in der EU kann zum einen die Übermittlung von Personendaten in die Schweiz und zum anderen die Bearbeitung von in die Schweiz übermittelten Personendaten von EU-Staatsangehörigen beinhalten. Das im Schweizer Recht gebotene Schutzniveau könnte daher relevant sein.

---

[264] Richtlinie (EU) 2016/680 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten im Bereich der Strafverfolgung (Fn. 234).

[265] Die Liste der Länder mit einem Angemessenheitsbeschluss kann abgerufen werden unter: https://commission.europa.eu/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en (abgerufen am 26. August 2024).

[266] Entscheidung der Kommission vom 26. Juli 2000 gemäss der Richtlinie 95/46/EG des Europäischen Parlaments und des Rates über die Angemessenheit des Schutzes personenbezogener Daten in der Schweiz, ABl. L 215, 25. August 2000, 0001–0003.

[267] Bericht der Kommission an das Europäische Parlament und den Rat über die erste Überprüfung der Wirkungsweise der Angemessenheitsfeststellungen gemäss Artikel 25 Absatz 6 der Richtlinie 95/46/EG vom 15. Januar 2024.

5.3.4 Verhältnis zum Rahmenübereinkommen des Europarats über künstliche Intelligenz und Menschenrechte, Demokratie und Rechtsstaatlichkeit

Die KI-Konvention ist ein verbindlicher völkerrechtlicher Vertrag, dem die Mitgliedstaaten des Europarats, die Staaten, die nicht Mitglieder des Europarats sind, und die EU beitreten können. Die KI-Verordnung ist eine EU-interne Regelung. Die KI-Konvention richtet sich an die Staaten und enthält allgemeine Regeln und Prinzipien, die bei einer Ratifizierung grundsätzlich in innerstaatliches Recht umgesetzt werden müssen. Die KI-Verordnung ist hingegen direkt anwendbar und enthält ausführlichere Vorschriften für die KI-Systeme in der EU. Sie sieht Pflichten für die Akteure im KI-Bereich vor. Die Art und die Tragweite dieser Texte sind somit verschieden.

Das Ziel der KI-Konvention, die Menschenrechte, die Demokratie und die Rechtsstaatlichkeit zu schützen, erklärt sich aus dem Auftrag des Europarats, die Grundrechte zu schützen. Die KI-Verordnung ist in erster Linie eine Regulierung der KI-Produkte.[268] Ihr Zweck ist es, das Funktionieren des Binnenmarkts zu verbessern, indem ein einheitlicher Rechtsrahmen für KI-Systeme in der EU geschaffen wird, und gleichzeitig die Grundrechte zu schützen (vgl. Art. 1 Abs. 1 und E. 1 der Verordnung). Die Ziele sind also teilweise verschieden. Die KI-Konvention hat nicht zum Ziel, einheitliche Regeln für den Verkehr von KI-Produkten zu gewährleisten. In diesem Sinn unterscheidet sich die KI-Verordnung insbesondere durch ihre technischen Inhalte von der KI-Konvention. Es scheint daher, dass sich die beiden Texte gegenseitig ergänzen sollen.

Um dort, wo es möglich ist, einen Vergleich der KI-Konvention und der KI-Verordnung vornehmen zu können, enthält die Analyse eine Tabelle, die mögliche Korrelationen zwischen den Bestimmungen dieser beiden Texte aufzeigt (vgl. Anhang 1).

5.3.5 Weitere ausgewählte Elemente

In diesem Kapitel werden einige andere Merkmale der KI-Verordnung aus Sicht des Schweizer Rechts präsentiert, die sowohl im Zusammenhang mit der Frage der Auswirkungen der Verordnung auf die Schweizer Akteure als auch vor dem Hintergrund einer allfälligen Annäherung der schweizerischen Gesetzgebung an diese Verordnung wichtig erscheinen. Die Analyse erhebt keinen Anspruch auf Vollständigkeit:

- Wie bereits erwähnt, will die KI-Verordnung das Funktionieren des Binnenmarkts im Zusammenhang mit dem Verkehr von KI-Produkten in der EU verbessern und gleichzeitig die Grundrechte schützen (Art. 1 Abs. 1 der KI-Verordnung). Die Schweiz hat nicht den gleichen Bedarf, ihren Binnenmarkt zu vereinheitlichen.

- Die KI-Verordnung verweist häufig auf die europäischen Harmonisierungsrechtsvorschriften, die die im Binnenmarkt angebotenen Produkte regeln. Beispielsweise werden KI-Systeme als Hochrisiko-KI-Systeme eingestuft, wenn es sich um Produkte handelt, die unter die Harmonisierungsrechtsvorschriften der EU nach Anhang I fallen, oder um KI-Systeme, die als Sicherheitsbauteile dieser Produkte verwendet werden

[268] DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 5 N 7; MARTINA ARIOLI, Risikomanagement (Fn. 222), 4 N 8 f.

sollen (vgl. Art. 6 Abs. 1 der Verordnung). Zudem fügt sich die KI-Verordnung in die bestehenden Konformitätsbewertungsmechanismen ein und ergänzt diese mit Vorschriften (vgl. insb. Art. 43 der Verordnung).

Aus Sicht des Schweizer Rechts erhöht diese enge Verknüpfung mit den europäischen Harmonisierungsvorschriften die Komplexität der Verordnung.

- Die KI-Verordnung enthält zudem Mechanismen, die dem europäischen Gesetzgeber ein rasches Eingreifen ermöglichen, um den technologischen Entwicklungen Rechnung zu tragen. So sieht beispielsweise Artikel 7 Absatz 1 der Verordnung vor, dass die Europäische Kommission delegierte Rechtsakte zur Änderung von Anhang III durch Hinzufügen von Anwendungsfällen für Hochrisiko-KI-Systeme erlassen darf. Diese Befugnis hat zur Folge, dass sich der Inhalt der KI-Verordnung potenziell rasch verändern kann. Dies ist eine grosse Herausforderung für die Schweizer Akteure, die mögliche Entwicklungen genau beobachten und sich wenn möglich anpassen müssen. Im Fall einer allfälligen Annäherung der schweizerischen Gesetzgebung an die Rechtsvorschriften der EU müsste diesem Aspekt ebenfalls Rechnung getragen werden, um nötigenfalls rasch auf die Entwicklung des europäischen Rechts reagieren zu können.

- Die KI-Verordnung enthält transversale Vorschriften, die für alle KI-Systeme im privaten und öffentlichen Sektor gelten. In einem föderalistischen Staat wie der Schweiz wäre es aufgrund der Verteilung der Kompetenzen zwischen Bund und Kantonen nicht möglich, eine Bundesgesetzgebung mit einem so weiten Geltungsbereich im öffentlichen Recht zu haben. Die kantonalen Zuständigkeiten müssten gewahrt werden.

  Die KI-Verordnung sieht beispielsweise in Anhang III Nummer 3 Buchstabe c vor, dass «KI-Systeme, die bestimmungsgemäss zum Zweck der Bewertung des angemessenen Bildungsniveaus, das eine Person im Rahmen von oder innerhalb von Einrichtungen aller Ebenen der allgemeinen und beruflichen Bildung erhalten wird oder zu denen sie Zugang erhalten wird, verwendet werden sollen», als Hochrisiko-Systeme gelten. Da jedoch die Schweiz ein föderalistisches Bildungssystem hat, in dem die Kompetenzen der Kantone von besonderer Bedeutung sind, könnte eine solche Regel höchstens auf Bundesebene für die Bereiche vorgesehen werden, für die der Bund zuständig ist.

  Im Fall einer Annäherung der schweizerischen Gesetzgebung an die Rechtsvorschriften der EU im KI-Bereich müsste daher die Verteilung der Kompetenzen zwischen Bund und Kantonen gebührend berücksichtigt werden.

- Zudem müsste sich der Schweizer Gesetzgeber bewusst sein, dass die europäische Regelung die Erarbeitung von harmonisierten Normen, denen im KI-Bereich grosse Bedeutung zukommt, an die Normungsorganisationen delegiert (vgl. Ziff. 5.2.11).

- Die KI-Verordnung als solche enthält nur wenige individuelle Rechte (vgl. Art. 85 und 86 der Verordnung).[269] Sie ist jedoch nicht das einzige relevante Instrument in der EU im KI-Bereich. Andere europäische Rechtsvorschriften sind in diesem Bereich ebenfalls anwendbar.[270]

  Insbesondere im Zusammenhang mit dem Schutz der Grundrechte stellt die KI-Verordnung ein zusätzliches Element dar, das einen Beitrag leistet zur Umsetzung anderer grundlegender Rechtsvorschriften, die die Grundrechte schützen. So sind die Grundrechte in der EU und den Mitgliedstaaten beispielsweise bereits durch die geltende Gesetzgebung geschützt (z. B. im Bereich des Datenschutzes und der Nichtdiskriminierung). Die KI-Verordnung stellt sicher, dass die KI-Systeme ab der Konzeption diesen anderen Rechtsvorschriften entsprechen. Bei einem Verstoss erhalten die Behörden aufgrund der Anforderungen der KI-Verordnung zudem Zugang zu den Informationen, die nötig sind, um festzustellen, ob der Einsatz von KI rechtmässig erfolgt ist.[271] In welchem Umfang die Rechte des Einzelnen tatsächlich geschützt sind, hängt somit stark von den bestehenden Garantien in anderen materiellrechtlichen Vorschriften ab.

  Ein weiteres Beispiel ist die Haftung. In der KI-Verordnung sind beispielsweise die Fragen der zivilrechtlichen Haftung nicht geregelt. Um Ersatz für Schäden zu erhalten, die Personen durch die Verwendung von KI-Systemen erlitten haben, müssen sich geschädigte Personen auf andere materielle Rechtsvorschriften beziehen (s. z. B. Vorschlag für eine Richtlinie über die Haftung für fehlerhafte Produkte, vgl. Ziff. 6.3.2). Die KI-Verordnung schafft die Voraussetzungen, um die Umsetzung dieser materiellrechtlichen Vorschriften zu erleichtern.

  Eine allfällige Annäherung der Schweiz an die KI-Verordnung müsste dem Umstand Rechnung tragen, dass diese nicht alle Fragen im Zusammenhang mit KI abschliessend regelt.

Diese Erwägungen zeigen, dass die KI-Verordnung aus Sicht des Schweizer Rechts Besonderheiten aufweist, die in rechtlicher Hinsicht gewisse Fragen aufwerfen, sei es im Zusammenhang mit den Auswirkungen auf die Schweizer Akteure oder im Fall einer allfälligen Annäherung der schweizerischen Gesetzgebung an diese Verordnung.

---

[269] Für eine Kritik zu diesem Ansatz vgl. ANGELA MÜLLER, Der Artificial intelligence Act der EU (Fn. 223), 19 f.; MARTINA ARIOLI, Risikomanagement (Fn. 222), 16 N 54.

[270] Vgl. in diesem Sinn: DAVID ROSENTHAL, Der EU AI Act (Fn. 221), 5 N 9, der festhält, dass in der KI-Verordnung mehrmals erwähnt wird, dass diese zusätzlich zum bestehenden Recht gilt, vgl. z. B. Art. 2 Abs. 7 und E. 10.

[271] Vgl. Europäische Kommission, Künstliche Intelligenz – Fragen und Antworten, *Wie werden die Grundrechte durch die neuen Vorschriften geschützt*?, Dezember 2023, abrufbar unter: https://ec.europa.eu/commission/presscorner/detail/de/qanda_21_1683 (abgerufen am 26. August 2024).

5.4 Zwischenfazit

Die Untersuchung der KI-Verordnung hat ermöglicht, einen Überblick über ihren Inhalt zu erhalten und die wichtigsten Herausforderungen der Verordnung aus Sicht des Schweizer Rechts aufzuzeigen. Folgende Schlussfolgerungen können formuliert werden:

- Im Mittelpunkt der Verordnung stehen die Pflichten der Anbieter von Hochrisiko-KI-Systemen. Daher sind die Schweizer Akteure, die unter diese Kategorie fallen, von der KI-Verordnung am stärksten betroffen.

- Die KI-Verordnung ergänzt die harmonisierten Rechtsvorschriften der EU, indem sie darauf verweist und sie durch Vorschriften vervollständigt. Dies trägt zur Komplexität der Verordnung bei.

- Bestünde der politische Wille, sich an die KI-Verordnung anzunähern, müsste den jeweiligen Besonderheiten der schweizerischen und europäischen Rechtsordnung, beispielsweise der Verteilung der Kompetenzen zwischen Bund und Kantonen, Rechnung getragen werden.

- Die KI-Verordnung bezweckt, das Funktionieren des Binnenmarkts der EU im Zusammenhang mit den KI-Systemen zu harmonisieren und gleichzeitig die Grundrechte zu schützen. Sie enthält jedoch nur wenige individuelle Rechte. Der Schutz der individuellen Interessen im KI-Bereich ist grundsätzlich durch andere europäische Rechtsnormen, beispielsweise im Bereich des Datenschutzes und der Nichtdiskriminierung, gewährleistet. Die KI-Verordnung stärkt diese Rechte und deren Umsetzung im Rahmen von KI.

- In Bezug auf das Verhältnis zur KI-Konvention des Europarats hat die Analyse ergeben, dass die Art, die Tragweite und die Ziele der beiden Texte teilweise unterschiedlich sind. Es ist davon auszugehen, dass sie sich ergänzen.

- Würde die Schweiz eine der KI-Verordnung gleichwertige Gesetzgebung verabschieden und würden diese Gesetzgebung und die KI-Verordnung der EU in das MRA aufgenommen, so könnte die Schweiz von einer gegenseitigen Anerkennung der zukünftigen Konformitätsbewertungen von Produkten profitieren, die KI-Systeme beinhalten.

  Ohne gegenseitige Anerkennung müssen sich die betroffenen Schweizer Akteure, die Produkte auf den EU-Markt exportieren, für die Bewertung der Konformität der KI-Aspekte ihrer Produkte an die Anforderungen der KI-Verordnung halten.

Die KI-Verordnung wurde in der EU erst vor Kurzem verabschiedet. Viele Fragen im Zusammenhang mit ihrer Anwendung, sowohl aus rechtlicher als auch aus praktischer Sicht, müssen noch geklärt werden. Mit der Analyse wurden die wichtigsten Herausforderungen aufgezeigt, die sich zum jetzigen Zeitpunkt für die Schweiz abzeichnen. Die Analyse müsste vertieft werden, wenn die Politik eine Annäherung an diese Verordnung anstreben sollte.

6 Weitere ausgewählte Rechtsgebiete

6.1 Einleitung

Die KI-Konvention des Europarats und die KI-Verordnung der EU sind wichtige internationale Entwicklungen. Allerdings decken sie nicht alle für KI relevante Rechtsbereiche ab. In einigen Fällen verweisen sie lediglich auf Rechtsvorschriften aus anderen Bereichen. Dies ist beispielsweise der Fall beim Recht des geistigen Eigentums, beim Haftpflichtrecht und bei der strafrechtlichen Verantwortlichkeit.

Im Sinne eines etwas breiteren Überblicks über die Rechtslage in der Schweiz im Hinblick auf die Herausforderungen von KI werden in diesem Kapitel einige weitere ausgewählte Rechtsgebiete erörtert und der aktuelle Stand der Gesetzgebung skizziert.

Auch in diesen Bereichen sind weitere Vertiefungen erforderlich, um eine vollständige Übersicht zu erhalten.

6.2 Geistiges Eigentum[272]

KI verändert immer mehr die Art und Weise, wie heute Erfindungen oder Inhalte wie Texte, Musik oder Bilder entstehen. Das führt im Bereich des geistigen Eigentums dazu, dass etablierte Grundsätze allenfalls zu überdenken sind. Insbesondere im Urheberrecht und im Patentrecht tauchen Fragestellungen auf, welche zu Herausforderungen führen können. Es ist zu prüfen, inwieweit die veränderten Verhältnisse einer Anpassung des Urheber- und des Patentrechts bedürfen.

6.2.1 Urheberrecht und künstliche Intelligenz

6.2.1.1 Allgemeines

Das URG regelt primär den Schutz der Urheberinnen und Urheber von Werken der Literatur und Kunst (Art. 1 Abs. 1 Bst. a URG).

Artikel 2 Absatz 1 URG hält dabei fest, dass Werke «[…] unabhängig von ihrem Wert oder Zweck, geistige Schöpfungen der Literatur und Kunst [sind], die individuellen Charakter haben». Der Begriff der geistigen Schöpfung setzt eine Gedankenäusserung voraus. Urheberin oder Urheber kann nur eine natürliche Person sein (Art. 6 URG). Was nicht von einem Menschen geschaffen wurde, ist kein Werk.[273] Die geistige Schöpfung muss damit Ausdruck einer menschlichen Gedankenäusserung sein. Die Begriffe der Literatur und Kunst sind weit zu verstehen. Eine nicht abschliessende Aufzählung, was alles zu Literatur und Kunst gehören kann, findet sich in Artikel 2 Absatz 2 URG (z. B. literarische Sprachwerke, Werke der Musik, Werke der bildenden Kunst usw.). Geschützt soll aber in der Regel nur sein, was individuellen Charakter besitzt. Damit wird eine gewisse Einmaligkeit verlangt bzw. Merkmale, die eine

[272] Dieses Kapitel wurde vom IGE verfasst.

[273] BBl **1989** III 477, hier 521.

Schöpfung von anderen bestehenden oder möglichen Schöpfungen abheben.[274] Originalität im Sinne einer persönlichen Prägung durch die Urheberin oder den Urheber ist gemäss bundesgerichtlicher Rechtsprechung nicht erforderlich. Vorausgesetzt wird, dass der individuelle Charakter im Werk selbst zum Ausdruck kommt. Der individuelle Charakter hängt dabei vom Spielraum der Urheberin bzw. des Urhebers ab. Ist der Spielraum klein, wird der urheberrechtliche Schutz bereits gewährt, wenn nur ein geringer Grad an selbständiger Tätigkeit vorliegt.[275]

Die Urheberinnen und Urheber haben das ausschliessliche Recht zu bestimmen, ob, wann und wie ihre Werke verwendet werden (Art. 10 Abs. 1 URG). Unterschieden wird dabei üblicherweise zwischen Nutzungsrechten und Urheberpersönlichkeitsrechten. Die Nutzungsrechte haben in der Regel eine vermögensrechtliche Komponente; so erfolgt z. B. die Erlaubnis, ein Werk zu vervielfältigen, oft gegen Bezahlung einer Vergütung. Die Urheberpersönlichkeitsrechte weisen darüber hinaus eine ideelle Komponente auf und betonen damit die persönliche Verbundenheit zwischen Urheberin bzw. Urheber und Werk. Zu den Urheberpersönlichkeitsrechten zählt z. B. das Recht auf Anerkennung der Urheberschaft (Art. 9 URG).

Eine wichtige Eigenschaft des Urheberrechtsgesetzes ist seine sogenannte Technologieneutralität. Dies bedeutet, dass das Gesetz nicht danach unterscheidet, mit welchen technischen Verfahren ein Werk genutzt wird. Analoge und digitale Vervielfältigungen eines Werkes sind beispielsweise gleichermassen erfasst. Dies bietet den Vorteil, dass nicht jeglicher technischer Fortschritt eine Anpassung des Gesetzes notwendig macht. Dies bedeutet aber nicht, dass technologische Veränderungen nie Einfluss auf das Urheberrechtsgesetz haben. Die enorme technologische Entwicklung im digitalen Umfeld (man denke an das Internet) und dadurch entstehende Nutzungsformen führten mehrere Male zu einer Teilrevision des Urheberrechtsgesetzes.[276]

6.2.1.2 Fragestellungen und Regulierungsbedarf im Zusammenhang mit KI

Die Zunahme von Anwendungen generativer KI stellt das Urheberrecht erneut vor Herausforderungen und etablierte Grundsätze sind allenfalls zu hinterfragen. Im Zusammenhang mit generativen KI-Systemen sind folgende Bereiche urheberrechtlich genauer zu analysieren: erstens das Training von KI-Systemen, zweitens die Ergebnisse von KI-Systemen und drittens die erstellten Prompts.

---

[274] BBl **1989** III 477, hier 521.

[275] BGE **143** III 373, E. 2.1 S. 377.

[276] So zum Beispiel die Änderungen des Urheberrechtsgesetzes, welche 2008 und 2020 in Kraft traten.

Das Training von KI-Systemen bedarf einer grossen Menge[277] an Daten. Die Trainingsdatensätze setzen sich zum Beispiel aus im Internet verfügbaren Bildern und Bildbeschreibungen zusammen, welche in systematischen und automatisierten Verfahren heruntergeladen werden.[278] Die KI lernt anschliessend anhand der erstellten Trainingsdatensätze. Die für das Training von KI-Systemen verwendeten Inhalte können dabei urheberrechtlich geschützt sein. Es stellt sich deshalb die Frage, ob die Verwendung von urheberrechtlich geschützten Inhalten für das Training der KI-Systeme einer Einwilligung der Rechteinhaberinnen und Rechteinhaber (oder einer Schrankenregelung) bedarf. Die Antwort auf diese Frage ist von verschiedenen Faktoren abhängig. So zunächst davon, welches Recht auf das Training der KI anwendbar ist.[279] Mit Blick auf das vorliegend interessierende Schweizer Recht ist relevant, ob das Erstellen der Trainingsdatensets und/oder das anschliessende Training des KI-Systems eine urheberrechtlich relevante «Verwendung» von Werken darstellt. Ist die Verwendung urheberrechtlich relevant, greift sie in das Ausschliesslichkeitsrecht der Rechteinhaberinnen und Rechteinhaber ein und bedarf einer Erlaubnis. Artikel 10 Absatz 2 URG zählt verschiedene erlaubnispflichtige Werkverwendungen auf, so zum Beispiel die Vervielfältigung (Kopie). Eine relevante Werkverwendung liegt vor, wenn ein Werk nicht nur konsumiert, sondern vervielfältigt, an eine andere Person vermittelt und wahrnehmbar gemacht wird.[280] Sie ist damit oft eine Vorbereitungshandlung zum Werkgenuss[281] (dieser stellt selber keinen Eingriff in das Ausschliesslichkeitsrecht dar).

In der Lehre ist man sich nun uneinig darüber, ob nur jene Vervielfältigungen als urheberrechtlich relevant erfasst werden sollen, welche gezielt auf einen anschliessenden Werkgenuss ausgerichtet sind, oder auch solche, welche den Werkgenuss lediglich potenziell ermöglichen. Die herrschende Lehre scheint den letzteren Fall zu bejahen.[282] Folgt man diesem Verständnis, so sind Vervielfältigungen, welche für die Erstellung von Trainingsdatensätzen gemacht werden und allfällige Vervielfältigungen, welche für das Training der KI-Systeme benötigt werden, als urheberrechtlich relevant anzusehen (da sie potenziell den Werkgenuss ermöglichen). In die-

---

[277] Beispielhaft: ADITYA RAMESH/PRAFULLA DHARIWAL/ALEX NICHOL/CASEY CHU/MARK CHEN, Hierarchical Text-Conditional Image Generation with CLIP Latents, 2022, 23 (abrufbar unter: https://arxiv.org/abs/2204.061, Stand: 6.5.2024).

[278] PAULINA JO PESCH/RAINER BÖHME, Artpocalypse now? – Generative KI und die Vervielfältigung von Trainingsbildern, GRUR 2023/14, 997 ff., 998.

[279] Antwort auf diese Frage gibt das Internationale Privatrecht.

[280] DENIS BARRELET/WILLI EGLOFF, Das neue Urheberrecht. Kommentar zum Bundesgesetz über das Urheberrecht und verwandte Schutzrechte, Art. 10 N 8.

[281] MATHIS BERGER, Künstliche Intelligenz und Immaterialgüterrecht, Jusletter IT 4. Juli 2024, 4.

[282] SANDRA MARMY-BRÄNDLI/ISABELLE OEHRI, Das Training künstlicher Intelligenz, sic! 2023, 655 ff., 658 (mit weiteren Hinweisen).

sem Sinne findet sich in der Lehre die Auffassung, dass hier Eingriffe in das Vervielfältigungsrecht stattfinden.[283] Es gibt jedoch auch andere Auffassungen. Verschiedene Autorinnen und Autoren sehen in der Erstellung der Trainingsdatensets zwar eine urheberrechtlich relevante Vervielfältigung, beim Training des KI-Systems jedoch einen urheberrechtsfreien Werkgenuss.[284] Wiederum andere Autorinnen und Autoren wählen einen anderen Ansatz und stellen für die Frage, ob eine urheberrechtlich relevante Vervielfältigung vorliegt darauf ab, ob es zu einer wirtschaftlichen Verwertung des Werkes kommt (dies in Anlehnung an Überlegungen im US-amerikanischen Rechtsraum). Die Schweizer Gerichte hatten die urheberrechtliche Relevanz der beiden Sachverhalte, soweit ersichtlich, noch nicht zu beurteilen. Die Frage ist damit noch nicht geklärt. Sollten die Gerichte jedoch zum Schluss kommen, dass das Erstellen der Trainingsdatensätze und/oder das Training der KI-Systeme urheberrechtlich relevante Vervielfältigungen sind, so müssten hierfür Einwilligungen der Rechteinhaberinnen und Rechteinhaber eingeholt werden, soweit nicht eine Schrankenregelung greift.

Denkbar wären hier die Schranken des betriebsinternen Gebrauchs (Art. 19 Abs. 1 Bst. c URG), der vorübergehenden Vervielfältigungen (Art. 24*a* URG) oder zur Werkverwendung zum Zweck der wissenschaftlichen Forschung (Art. 24*d* URG). Bei jeder der genannten Schranken gibt es jedoch Elemente, welche, je nach konkretem Fall, vorliegend eine Anwendung verhindern könnten: bei der Schranke zum betriebsinternen Gebrauch beispielsweise das Kriterium des internen Informations- oder Dokumentationszwecks, bei der Schranke zur vorübergehenden Vervielfältigung das Kriterium der lediglich flüchtigen oder begleitenden Vervielfältigung und bei der Wissenschaftsschranke der Hauptzweck, der in der wissenschaftlichen Forschung liegen muss. Ob eine Schrankenregelung auf eine konkrete KI angewendet werden könnte, wäre deshalb im Einzelfall genauer zu prüfen.[285]

Unabhängig davon, wie die offenen Fragen beantwortet werden, kann sich daraus allfälliger gesetzgeberischer Handlungsbedarf ergeben: Sollte das Erstellen und/oder das Training von KI-Systemen urheberrechtlich relevante Vervielfältigungen beinhalten und, mangels einer bereits bestehenden Schrankenregelung, individuelle Einwilligungen der Rechteinhaberinnen und Rechteinhaber erfordern, so wäre sicherzustellen, dass die Entwicklung von KI-Systemen künftig in angemessener Weise möglich ist. Sollten diese Vorgänge jedoch als urheberrechtlich nicht relevant angesehen werden, so wäre wiederum zu klären, was dies für die Rechteinhaberinnen und Rechteinhaber der benötigten Werke bedeutet und welche allfälligen Massnahmen zu treffen wären.

Betreffend die Ergebnisse eines KI-Systems (Output) stellt sich einerseits die Frage, ob diese Urheberrechte verletzen können und andererseits, ob sie selber urheberrechtlichen Schutz geniessen. Mit Blick auf die Frage der Verletzung von Urheberrechten ist zu prüfen, ob das KI-Ergebnis ein Werk unverändert wiedergibt oder dann in einer Weise, in der das Original-

---

[283] SANDRA MARMY-BRÄNDLI/ISABELLE OEHRI, Das Training (Fn. 282), 657 f. (mit weiteren Hinweisen).

[284] Vgl. zur Auseinandersetzung hierzu SANDRA MARMY-BRÄNDLI/ISABELLE OEHRI, Das Training (Fn. 282), 658 f.

[285] Vgl. hierzu auch MATHIS BERGER, Künstliche Intelligenz und Immaterialgüterrecht (Fn. 281), 5 ff.

werk noch erkennbar ist. Wird ein solches KI-Ergebnis ohne Einwilligung (der Rechteinhaberin bzw. des Rechteinhabers oder einer Schrankenbestimmung) verwendet, stellt dies eine Urheberrechtsverletzung dar.

Ob das KI-Ergebnis selber urheberrechtlichen Schutz geniesst, hängt (neben der Frage der Individualität) davon ab, ob es sich dabei um eine «geistige Schöpfung» handelt.[286] Das Bundesgericht stellt an das Mass der geistigen Schöpfung keine hohen Anforderungen. Bereits beim Vorliegen eines geringen Grades von selbständiger geistiger Tätigkeit wird der urheberrechtliche Schutz gewährt.[287] Nicht urheberrechtlich geschützt sind demgegenüber Kreationen, welche beispielsweise von Tieren geschaffen werden oder reine Zufallsprodukte der Technik darstellen. Kein Hinderungsgrund für die geistige Schöpfung ist jedoch, wenn sich ein Mensch für die Schaffung eines Werkes bewusst bestimmter Zufallsprinzipien (z. B. Action Painting) oder der Technik (z. B. Computer) bedient.[288] Wo der menschliche Wille über das Ergebnis entscheidet, liegt eine geistige Schöpfung vor.[289] Es ist deshalb davon auszugehen, dass es sich beim Ergebnis des KI-Systems dann um eine geistige Schöpfung handelt, wenn die Nutzerin bzw. der Nutzer trotz des Einsatzes der KI noch auf wesentliche Elemente des Ergebnisses Einfluss nimmt und zwar selbst dann, wenn das KI-System ein gewisses Mass an Zufall in das Ergebnis einbringt. Nicht als geistige Schöpfung und damit auch urheberrechtlich nicht geschützt wäre demgegenüber ein Ergebnis, welches von einem KI-System eigenständig oder unter lediglich geringer Einflussnahme der Nutzerin oder des Nutzers entstand.[290]

Betreffend den urheberrechtlichen Schutz der von der Nutzerin bzw. vom Nutzer eingegebenen Prompts kann auf die vorangehenden Ausführungen verwiesen werden. Ein von einem Menschen geschriebener Prompt erfüllt die Kriterien der geistigen Schöpfung und der Literatur. Wesentlich für den urheberrechtlichen Schutz wird deshalb sein, ob der Prompt genügenden «individuellen Charakter» aufweist. Dies ist jeweils im Einzelfall zu beurteilen.

Generative KI-Systeme werden häufig mit urheberrechtlich geschützten Werken trainiert. Gleichzeitig nutzen immer mehr Personen KI-Systeme, um selber Inhalte wie Bilder, Texte oder Musik zu erstellen. Dies wirft urheberrechtliche Fragen auf, welche zurzeit noch nicht alle verbindlich beantwortet wurden. Sind beispielsweise die Ergebnisse von KI urheberrechtlich geschützt und ist das Training von KI urheberrechtlich relevant oder nicht? Es ist

[286] Vgl. hierfür vorangehend Ziff. 6.2.1.1.

[287] BGE **59** II 401, S. 405.

[288] DENIS BARRELET/WILLI EGLOFF, Das neue Urheberrecht. Kommentar (Fn. 280), Art. 2 N 8.

[289] BBl **1989** III 477, hier 521.

[290] In der Schweiz liegt diesbezüglich noch kein verbindlicher Entscheid vor. Im Ausland zeichnet sich jedoch eine entsprechende Tendenz im Immaterialgüterrecht ab, so z. B. in den USA im Urheberrecht (U.S. Copyright Office, Robert J. Kasunic, Zarya of the Dawn [Registration # VAu001480196]) oder in Deutschland im Patentrecht (Beschluss des Bundespatentgerichtes X ZB 5/22 vom 11. Juni 2024).

davon auszugehen, dass insbesondere die Antwort auf letztere Frage möglichen regulatorischen Bedarf aufzeigen wird. Sollte das Training von KI urheberrechtlich relevant sein, ist zu prüfen, wie die (Weiter-)Entwicklung von KI sichergestellt werden kann. Sollte das Training von KI hingegen nicht urheberrechtlich relevant sein, so ist zu prüfen, inwieweit den Interessen der Rechteinhaberinnen und Rechteinhabern von Werken Rechnung zu tragen ist.

6.2.2 KI und Patentrecht

6.2.2.1 Allgemeines

Das internationale Patentrecht ist in zahlreichen Übereinkommen geregelt.[291] Für die Schweiz ist das Europäische Patentübereinkommen[292] massgebend.[293] Auf nationaler Ebene regeln das Patentgesetz (PatG) und die Patentverordnung (PatV)[294] das Patentrecht.

Das Patentrecht schützt technische Lösungen für technische Probleme, sofern diese Lösungen neu, erfinderisch und gewerblich anwendbar sind (Art. 1 PatG). Abstrakte Ideen ohne konkrete technische Schritte, insbesondere mathematische Methoden, Algorithmen oder Lernverfahren, sind nicht patentierbar.

Das Patentrecht gilt für alle Gebiete der Technik.[295] Es ist insbesondere auf computerimplementierte Erfindungen[296] und auf KI-gestützte Erfindungen anwendbar. Die Weltorganisation

[291] Die Schweiz hat insbesondere die Pariser Übereinkunft zum Schutz des gewerblichen Eigentums (SR **0.232.04**), den Vertrag über die internationale Zusammenarbeit auf dem Gebiet des Patentwesens (SR **0.232.141.1**), den Patentrechtsvertrag (SR **0.232.141.2**) und das Abkommen über handelsbezogene Aspekte der Rechte an geistigem Eigentum (Anhang 1C des Abkommens zur Errichtung der Welthandelsorganisation, SR **0.632.20**) ratifiziert.

[292] SR **0.232.142.2**.

[293] Das Europäische Patentamt (EPA) ist eine von der EU unabhängige internationale Organisation, die ein zentralisiertes Patentprüfungsverfahren für 39 Länder anbietet. Die meisten Patentanmeldungen, die die Schweiz betreffen, werden direkt beim EPA eingereicht. Sobald das EPA ein Patent erteilt, ist es auch in der Schweiz gültig.

[294] SR **232.14** und SR **232.141**.

[295] Das PatG kennt einige Ausnahmen von den Artikel 1*a*, 1*b* und 2 PatG. Dazu gehören beispielsweise der menschliche Körper und seine Bestandteile, Gensequenzen, chirurgische, therapeutische oder diagnostische Verfahren, Pflanzensorten und Tierrassen. Nicht patentierbar ist ferner eine Erfindung, deren Verwendung die Menschenwürde verletzen oder die Würde der Kreatur missachten, oder in anderer Weise gegen die öffentliche Ordnung oder die guten Sitten verstossen würde (z. B. Folterwerkzeuge).

[296] Computerprogramme sind als solche nicht patentierbar. Sie können jedoch Teil einer Erfindung sein, wenn diese eine Wirkung hat, welche über das blosse Funktionieren des Programmes hinausgeht.

für geistiges Eigentum (WIPO) verzeichnete zwischen 2016 und 2022 einen Anstieg der Patentanmeldungen im Bereich der KI um 718 Prozent.[297] Aufgrund des Anstiegs der Anmeldungen in diesem Bereich hat das Europäische Patentamt (EPA) seine Prüfungsrichtlinien für KI-gestützte Erfindungen seit 2018 präzisiert.[298]

KI wird zunehmend als Werkzeug für Erfindungen im Bereich Forschung und Entwicklung eingesetzt. Dabei ermöglicht KI eine Beschleunigung, Ausweitung und Diversifizierung von Forschung und Entwicklung.

6.2.2.2 Herausforderungen

Generative KI könnte zu einer drastischen Zunahme der Menge an wissenschaftlicher und technischer Literatur führen. Diese spiegelt den Stand der Technik wider, woran Neuheit und Erfindungshöhe einer Patentanmeldung gemessen werden. Die Erlangung eines Patents könnte dadurch erschwert werden.[299]

Der Einsatz von KI in Forschung und Entwicklung dürfte die Beurteilung des Kriteriums der erfinderischen Tätigkeit für eine Erfindung verändern. Nicht jede technische Lösung oder Verbesserung einer technischen Lösung ist patentierbar. Die in der Patentanmeldung offenbarte Lösung muss erfinderisch sein, d. h. sie darf für Experten nicht naheliegend sein. KI kann komplexe Lösungen analysieren und generieren, die die Grenzen des herkömmlichen menschlichen Verständnisses überschreiten. Dies könnte den Bereich innovativer Ideen erweitern, indem Kombinationen gefunden werden, die für menschliche Fachleute nicht offensichtlich sind, und somit eine Weiterentwicklung des Kriteriums der erfinderischen Tätigkeit erzwingen.

Die Offenlegung von Trainingsdaten einer KI-basierten Erfindung wird diskutiert. Ein Patent wird erteilt, wenn die Erfindung in der Patentschrift so beschrieben wird, dass sie von einer Fachperson nachvollzogen werden kann. Bei KI-gestützten Erfindungen spielen Trainingsdaten eine wesentliche Rolle für die Reproduzierbarkeit der Erfindung. Es wird daher diskutiert, wie und in welchem Umfang diese Trainingsdaten offengelegt werden sollen. Dabei sind Verpflichtungen oder Interessen wie Datenschutz und Schutz von Geschäftsgeheimnissen oder

---

[297] https://www.wipo.int > Understand and Learn > IP in > Frontier Technologies > AI Inventions factsheet, 1 (abgerufen am 15. Mail 2024).

[298] Richtlinien für die Prüfung im Europäischen Patentamt, Teil G, 3.3.1, «Künstliche Intelligenz und maschinelles Lernen», Version 03.2024, (abrufbar unter: https://www.epo.org/de/legal/guidelines-epc/2024/g_ii_3_3_1.html, abgerufen am 15. Mai 2024); siehe auch https://www.epo.org/de/news-events/in-focus/ict/artificial-intelligence (abgerufen am 15. Mai 2024).

[299] Siehe insbesondere All Prior Art, ein Zufallsgenerator für wissenschaftliche und technische Texte, https://allpriorart.com (abgerufen am 15. Mai 2024). Das US-amerikanische-Patentamt hat übrigens am 30. April 2024 eine öffentliche Konsultation zu diesem Thema durchgeführt, Request for Comments Regarding the Impact of the Proliferation of Artificial Intelligence on Prior Art, the Knowledge of a Person Having Ordinary Skill in the Art, and Determinations of Patentability Made in View of the Foregoing, abrufbar unter: https://www.federalregister.gov/documents/2024/04/30/2024-08969/request-for-comments-regarding-the-impact-of-the-proliferation-of-artificial-intelligence-on-prior (abgerufen am 15. Mai 2024).

Geheimhaltungspflichten zu berücksichtigen. Das Europäische Patentamt hat diesen Punkt bei der letzten Aktualisierung seiner Prüfungsrichtlinien klargestellt: «*Hängt die technische Wirkung von bestimmten Merkmalen des verwendeten Trainingsdatensatzes ab, sind die zur Reproduktion dieser technischen Wirkung erforderlichen Merkmale zu offenbaren, es sei denn, der Fachmann kann sie ohne unzumutbaren Aufwand unter Verwendung seines allgemeinen Fachwissens ermitteln. Der konkrete Trainingsdatensatz selbst muss jedoch in der Regel nicht offenbart werden.*»[300] Die letzte Herausforderung betrifft das Erfordernis, in jeder Patentanmeldung eine natürliche Person als Erfinder anzugeben. Diese Frage wird durch das *Artificial Inventor Project*[301] beleuchtet, das versucht, zwei Patente im Namen von *DABUS*, einem System künstlicher Intelligenz, anzumelden, ohne dabei eine natürliche Person als Erfinder anzugeben. Dieses Vorgehen stellt diese Anforderung des Patentrechts in Frage und wirft grundsätzliche Thematiken betreffend den Zweck des Patentrechts auf.

6.2.2.3 Regulatorischer Bedarf

Ein gesetzgeberischer Handlungsbedarf scheint nicht zu bestehen. Der spektakuläre Anstieg der Anmeldungen von KI-basierten Erfindungen zwischen 2016 und 2022 belegt dies. Die Bedingungen für den Schutz dieser Erfindungen sind klar und stabil, und es werden Patente für diese Art von Erfindungen erteilt. Das System bietet zudem eine gewisse Flexibilität. Begriffe wie «Stand der Technik», «Neuheit», «erfinderische Tätigkeit» und «Erfordernis der Offenbarung der Erfindung» können von den Ämtern für geistiges Eigentum und den Gerichten bis zu einem gewissen Grad entsprechend den technologischen Entwicklungen ausgelegt werden.

Hinsichtlich des Erfordernisses einer natürlichen Person als Erfinder erscheint die durch das *Artificial Inventor Project* ausgelöste Diskussion verfrüht. Der Mensch spielt nach wie vor eine entscheidende Rolle in Forschung und Entwicklung, und es ist nach wie vor möglich, zumindest eine natürliche Person als Erfinder zu benennen.

- Der Einsatz von KI in Forschung und Entwicklung könnte zu einer Weiterentwicklung der Konzepte von Neuheit und erfinderischer Tätigkeit führen.
- Es sollte eine klare und stabile Praxis für die Offenbarung von Trainingsdaten für KI-basierte Erfindungen entwickelt werden.
- Die weltweite Entwicklung der Rechtsprechung hinsichtlich des Erfordernisses der Nennung einer natürlichen Person als Erfinder sollte aufmerksam verfolgt werden.

---

[300] Vgl. Fn. 298.

[301] The Artificial Inventor Project, https://artificialinventor.com/ (abgerufen am 15. Mai 2024). Dies ist ein Test des Patentsystems im echten Leben. Der Anmelder behauptet, ein KI-System geschaffen zu haben, das selbstständig Erfindungen generieren kann. Weltweit wurden Patentanträge gestellt. Der Fall ging bis zum Obersten Gerichtshof der USA und zum Obersten Gerichtshof Grossbritanniens. Beide sprachen sich gegen die Zulassung eines künstlichen Erfinders aus. Das EPA wies die Anmeldung zurück. In der Schweiz wies das EPA das Eintragungsgesuch ab. Eine Beschwerde ist beim Bundesverwaltungsgericht hängig.

- Gesetzgeberischer Handlungsbedarf scheint nicht zu bestehen. Das Patentrechtssystem ist flexibel genug, um sich den technologischen Entwicklungen ohne Änderungen anzupassen. Der spektakuläre Anstieg der Patentanmeldungen für KI-basierte Erfindungen seit 2016 zeigt, dass das System gut funktioniert.

6.3 Ausservertragliche Haftung

6.3.1 Allgemeines ausservertragliches Haftpflichtrecht

6.3.1.1 Vorschlag für eine EU-Richtlinie über KI-Haftung

Der Vorschlag der Europäischen Kommission vom 28. September 2022 für eine Richtlinie zur Anpassung der Vorschriften über ausservertragliche zivilrechtliche Haftung an künstliche Intelligenz (RL-Vorschlag KI-Haftung)[302] betrifft die *verschuldensabhängige ausservertragliche Haftung* und will den Herausforderungen, die KI für die Anwendung bestehender Haftungsvorschriften darstellt, mithilfe von Offenlegungsvorschriften und widerlegbaren Vermutungen begegnen. Es soll den besonderen Merkmalen bestimmter KI-Systeme, wie z. B. Undurchsichtigkeit, autonomes Verhalten und Komplexität, die es erschweren oder verunmöglichen, ein kausales Handeln oder Unterlassen von verantwortlichen Personen zu beweisen, mit Beweiserleichterungen begegnet werden. Der Vorschlag knüpft an die Begriffsbestimmungen und Kategorien des *KI-Gesetzes*[303] an und sieht für *Hochrisiko-KI-Systeme* besondere Rechtsfolgen vor, welche auch zur Durchsetzung der Anforderungen des KI-Gesetzes beitragen sollen.[304] Mit der widerlegbaren Vermutung wurde gemäss Begründung des RL-Vorschlags das am wenigsten einschneidende Instrument gewählt.[305] Der in der Entschliessung des Europäischen Parlaments vom 20. Oktober 2020[306] mit Empfehlungen an die Kommission für eine Regelung der zivilrechtlichen Haftung beim Einsatz künstlicher Intelligenz noch vorgesehene Ansatz einer Gefährdungshaftung[307] für Betreiber von Systemen mit hohen Risiken wurde nicht weiterverfolgt. Die Anwendung der Richtlinie soll fünf Jahre nach Ablauf der

[302] Vorschlag für eine Richtlinie des Parlaments und des Rates zur Anpassung der Vorschriften über aussenvertragliche zivilrechtliche Haftung an künstliche Intelligenz, COM/2022/496 final, abrufbar unter: https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52022PC0496 (abgerufen am 28. August 2024).

[303] S. oben: Ziff. 5.

[304] Vorschlag Richtlinie über KI-Haftung (Fn. 302), Begründung, 3.

[305] Vorschlag Richtlinie über KI-Haftung (Fn. 302), Begründung, 7.

[306] Abl. C 404/107, abrufbar unter: https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52020IP0276&from=PL (abgerufen am 28. August 2024).

[307] Vgl. GERHARD WAGNER, Haftung für Künstliche Intelligenz – Eine Gesetzesinitiative des Europäischen Parlaments, ZEuP 2021, 545 ff., 556 ff.

Umsetzungsfrist geprüft werden, wobei dann allenfalls die Notwendigkeit einer verschuldensunabhängigen Gefährdungshaftung gekoppelt mit einer Pflichtversicherung zu prüfen wäre.[308] Im Einzelnen:

- **Anwendungsbereich:** Der RL-Vorschlag enthält Vorschriften über die verschuldensabhängige ausservertragliche zivilrechtliche Haftung für durch ein KI-System verursachte Schäden. Die Richtlinie gilt nicht für die Haftung aus Vertrag. Sie gilt auch nicht für die strafrechtliche Haftung (Art. 1(1)+(2) RL-Vorschlag KI-Haftung).

- **Offenlegung von Beweismitteln über Hochrisiko-KI-Systeme:** Der RL-Vorschlag schafft für diejenigen, die Schadensersatz fordern, eine Möglichkeit, *Informationen über Hochrisiko-KI-Systeme* zu erhalten, die gemäss *KI-Gesetz* aufzuzeichnen / zu dokumentieren sind (Art. 3 RL-Vorschlag KI-Haftung).[309] Die Gerichte wären auf Antrag eines potenziellen Klägers befugt, von Anbietern und Nutzern von Hochrisiko-KI-Systemen, die im Verdacht stehen, einen Schaden verursacht zu haben, die Offenlegung von Beweismitteln anzuordnen. Der potentielle Kläger muss seinen Schadenersatz plausibel machen und belegen, dass eigene Anstrengungen nicht zum Ziel geführt haben. In einer Verhältnismässigkeitsabwägung sind weiter die Interessen aller Parteien und auch Dritter, namentlich auch Geschäftsgeheimnisse, zu beachten. Kommt ein Beklagter der Anordnung nicht nach, wird ein *Sorgfaltspflichtverstoss vermutet* (Art. 3[5] RL-Vorschlag KI-Haftung).

- **Kausalitätsvermutung:** Unter bestimmten Bedingungen soll sowohl für Hochrisiko-KI-Systeme als auch für andere KI-Systeme eine *Kausalitätsvermutung* vorgesehen werden. Diese soll den Nachweis, dass eine bestimmte Eingabe, für die die potenziell haftbare Person verantwortlich ist, ein bestimmtes Ergebnis des KI-Systems verursacht hat, erleichtern. Vermutet wird ein ursächlicher Zusammenhang zwischen dem Verschulden des Beklagten und dem vom KI-System hervorgebrachten Ergebnis oder der Tatsache, dass kein Ergebnis hervorgebracht wurde. Der Vorschlag bewirkt keine Umkehr der Beweislast, es wird eine widerlegbare Vermutung geschaffen.[310] Auch entbindet sie nicht vom Nachweis des Verschuldens, das heisst im Normalfall der Verletzung einer Sorgfaltspflicht. Es wird Bezug genommen auf im Unionsrecht oder im nationalen Recht festgelegte Sorgfaltspflichten, deren unmittelbarer Zweck darin besteht, den eingetretenen Schaden zu verhindern (Art. 4[1][a] RL-Vorschlag KI-Haftung), es besteht ein direkter Bezug zum KI-Gesetz. Die Kausalitätsvermutung gilt unter folgenden Voraussetzungen:

    - Für Anbieter von Hochrisiko-KI-Systemen, wenn nachgewiesen werden kann, dass gegen Anforderungen des KI-Gesetzes in Bezug auf Training, Transpa-

[308] Vorschlag Richtlinie über KI-Haftung (Fn. 302), Begründung, 7.

[309] Vorschlag Richtlinie über KI-Haftung (Fn. 302), Erw. 17.

[310] Vorschlag Richtlinie über KI-Haftung (Fn. 302), Begründung S. 7 (Verhältnismässigkeit).

renz, Beaufsichtigung, Genauigkeit, Robustheit und Cybersicherheit verstossen wurde bzw. erforderliche Korrekturmassnahmen nicht unverzüglich ergriffen wurden (Art. 4[2] RL-Vorschlag KI-Haftung).

- Für Nutzer von Hochrisiko-KI-Systemen, wenn nachgewiesen werden kann, dass gegen Anforderungen des KI-Gesetzes in Bezug auf Verwendung, Überwachung oder Zweckbestimmung der Eingabedaten verstossen wurde (Art. 4[3] RL- Vorschlag KI-Haftung).

- Bei Nicht-Hochrisiko-KI-Systemen gilt die Vermutung, wenn es nach Auffassung des Gerichtes für den Kläger übermässig schwierig ist, den ursächlichen Zusammenhang zwischen Schaden und Verschulden zu beweisen.

Der Richtlinienvorschlag der EU-Kommission befindet sich aktuell (Stand 31.08.2024) im Rechtsausschuss des Europäischen Parlaments.[311] Dieser hat laut Presseberichten eine Zusatzstudie zur Notwendigkeit dieser zusätzlichen Richtlinie – neben dem KI-Gesetz und der revidierten Richtlinie über die Haftung für fehlerhafte Produkte[312], welche beide im März 2024 vom Parlament verabschiedet wurden – in Auftrag gegeben.[313] Die Zukunft des Richtlinienvorschlags KI-Haftung scheint damit ungewiss.

6.3.1.2 Schweizer Recht

Allgemein ist das Schweizer Zivilrecht und namentlich auch das Haftpflichtrecht mit seinen offenen Generalklauseln in der Lage, technische Entwicklungen aufzufangen und stellt den Gerichten ein Instrumentarium zur Verfügung, um im Einzelfall zu gerechten Lösungen zu gelangen.[314] Der Bundesrat hat sich in der Vergangenheit deshalb verschiedentlich gegen Rechtsänderungen aufgrund von technischen Innovationen ausgesprochen, da diese die Gefahr mit sich bringen, innert Kürze zu veralten.[315] Eine zu stark auf technische Einzelheiten

---

311 2022/0303(COD): Adaptation des règles en matière de responsabilité civile extracontractuelle au domaine de l'intelligence artificielle (nicht auf Deutsch verfügbar), https://oeil.secure.europarl.europa.eu/oeil/popups/ficheprocedure.do?reference=2022%2F0303(COD)&l=fr (abgerufen am 26. August 2024).

312 S. dazu sogleich Ziff. 6.3.2.

313 https://www.euractiv.com/section/digital/news/picking-up-the-ai-liability-directive-after-the-tech-policy-spree/ (abgerufen am 26. August 2024)

314 Vgl. Bericht «Herausforderungen der künstlichen Intelligenz» (Fn. 1), 36 f.

315 Vgl. zu den Grundsätzen der Politik des Bundes im Umgang mit neuen Technologien Bericht «Herausforderungen der künstlichen Intelligenz» (Fn. 1), 34 ff.; Bericht des Bundesrates «Rechtliche Grundlagen für Distributed Ledger-Technologie und Blockchain in der Schweiz» vom 14. Dezember 2018, 13 ff., abrufbar unter: https://www.newsd.admin.ch/newsd/message/attachments/55150.pdf (abgerufen am 31. Juli 2024); vgl. auch Bericht des Bundesrates «Die zivilrechtliche Verantwortlichkeit von Providern» vom 11. Dezember 2015, 99 ff. abrufbar unter:

fokussierte Gesetzgebung läuft Gefahr, Regelungslücken gerade erst zu schaffen. Dennoch ist es im Zuge der Digitalisierung angezeigt, die bestehenden Rechtsgrundlagen regelmässig auf ihre Komptabilität mit neueren technischen Entwicklungen zu prüfen sowie allenfalls bestehende Lücken zu schliessen bzw. Rechtssicherheit zu schaffen.

Die Schweiz kennt keine den diskutierten EU-Regeln vergleichbaren spezifischen haftpflichtrechtlichen Regeln zur künstlichen Intelligenz. Nach der haftpflichtrechtlichen Generalklausel von Artikel 41 Absatz 1 OR ist haftbar und schadenersatzpflichtig, wer einem andern mit Absicht oder aus Fahrlässigkeit widerrechtlich Schaden zufügt. Notwendige Voraussetzungen der Haftung sind daher Schaden, Widerrechtlichkeit, Kausalität und Verschulden. Nach Artikel 8 des Zivilgesetzbuches (ZGB)[316] muss derjenige, der aus einer Tatsache Rechte ableitet, das Vorhandensein dieser Tatsache beweisen, gesetzliche Ausnahmen vorbehalten. Nach Artikel 55 Absatz 1 der Zivilprozessordnung (ZPO)[317] trifft diese Partei die sogenannte Behauptungs- und Substantiierungslast. Sie muss dem Gericht einen Beweisantrag stellen, in dem die Beweise bestimmt oder klar bestimmbar sind (Art. 152 Abs. 1 ZPO).[318] Reine Ausforschungsanträge («fishing expeditions») sind nicht zulässig. Die Gegenpartei ist zur Mitwirkung verpflichtet und muss dabei namentlich auch genau bezeichnete Urkunden herausgeben (Art. 160 Abs. 1 Bst. b ZPO). Ein Verweigerungsrecht, um sich selbst vor zivilrechtlicher Verantwortlichkeit zu schützen, besteht nicht.[319] Verweigert eine Partei die Mitwirkung unberechtigterweise, so berücksichtigt dies das Gericht bei der Beweiswürdigung (Art. 164 ZPO). Es ergibt sich daraus jedoch nicht ohne weiteres eine Tatsachenvermutung zu Ungunsten der verweigernden Partei.[320] Unter Umständen kann zur Abklärung von Prozesschancen auch eine vorsorgliche Beweisabnahme verlangt werden (Art. 158 Abs. 1 Bst. b ZPO).[321]

In gewissen Fällen sind auch Beweiserleichterungen denkbar. Gemäss Rechtsprechung und Lehre reicht eine überwiegende Wahrscheinlichkeit[322] aus, wenn «ein strikter Beweis nicht nur im Einzelfall, sondern der Natur der Sache nach nicht möglich oder nicht zumutbar ist und

---

www.bj.admin.ch > Publikationen & Service > Berichte, Gutachten und Verfügungen > Bericht und Gutachten > Die Zivilrechtliche Verantwortlichkeit von Providern (abgerufen am 31. Juli 2024).

316 SR **210**.

317 SR **272**.

318 BSK ZPO-GUYAN, Art. 152 N 3 f.

319 Botschaft ZPO vom 28. Juni 2006, BBl **2006** 7221 7317.

320 BGE **140** III 264, E. 2.3.

321 BSK ZPO-GUYAN, Art. 158 N 5.

322 «Nach dem Beweismass der überwiegenden Wahrscheinlichkeit [...] gilt ein Beweis als erbracht, wenn für die Richtigkeit der Sachbehauptung nach objektiven Gesichtspunkten derart gewichtige Gründe sprechen, dass andere denkbare Möglichkeiten vernünftigerweise nicht massgeblich in Betracht fallen»; BGE **132** III 715, E. 3.1. m.w.H.

insofern eine ‹Beweisnot› besteht.»[323] Hierzu gehört auch der Nachweis des Kausalzusammenhangs im Haftpflichtrecht.[324]

Diese Regeln gelten auch für Verfahren, in denen die Verhandlungsmaxime im engeren Sinne gilt. Im vereinfachten Verfahren, das bei Streitigkeiten mit einem Streitwert bis zu 30 000 Franken zur Anwendung kommt (Art. 243 Abs. 1 ZPO), ist die Rolle des Gerichts aktiver, da ihm eine verstärkte Fragepflicht obliegt (Art. 247 Abs. 1 ZPO).

6.3.1.3 Würdigung

Da sich die Anforderungen des Vorschlags einer Richtlinie über KI-Haftung an den Vorgaben des KI-Gesetzes ausrichten, ist eine isolierte Einführung der Richtlinie über KI-Haftung nicht denkbar. Sollte das KI-Gesetz für die Schweizer Rechtsordnung übernommen werden, wären die begleitenden Bestimmungen des Vorschlags der Richtlinie über KI-Haftung der Schweizer Rechtsordnung aber nicht komplett fremd. Die Offenlegung einer genau bezeichenbaren, bei der Gegenpartei vorhandenen Dokumentation ist unter Umständen mit den vorhandenen zivilprozessualen Instrumenten bereits nach geltendem Recht erreichbar. Die Rechtsfolgen bei Verweigerung der Mitwirkung bzw. Offenlegung von Beweismitteln ist ebenfalls vergleichbar, die im Schweizer Recht vorgesehene Berücksichtigung bei der Beweiswürdigung geht aber weniger weit als die im Richtlinienvorschlag vorgesehene Vermutung der Sorgfaltspflichtverletzung. Die Kausalitätsvermutungen des Vorschlags einer Richtlinie über KI-Haftung lassen sich aus dem geltenden Schweizer Recht jedoch nicht herleiten, auch wenn Beweiserleichterungen in solchen Fällen vorgesehen sind. Eine Übernahme der Richtlinie über KI-Haftung könnte damit die gerichtliche Durchsetzung von zivilrechtlichen Ansprüchen eventuell erleichtern und auch dazu beitragen, die Vorgaben der KI-Konvention des Europarats (namentlich Art. 8, 9 und 14 ; vgl. Ziff. 4.3) im privaten Sektor besser umzusetzen. Es bleiben aber die zusätzlichen Analysen und Ergebnisse der Diskussionen in der EU abzuwarten.

Anzumerken bleibt, dass der Anwendungsbereich dieser allgemeinen ausservertraglichen haftpflichtrechtlichen Regel wohl stark beschränkt wäre.[325] Für zahlreiche mögliche Einsatz-

---

[323] BGE **132** III 715, E. 3.1. m.w.H.

[324] BGE **132** III 715, E. 3.2. m.w.H. «Glaubhaft gemacht ist dabei eine Tatsache schon dann, wenn für deren Vorhandensein gewisse Elemente sprechen, selbst wenn das Gericht noch mit der Möglichkeit rechnet, dass sie sich nicht verwirklicht haben könnte»; BGE **132** III 715, E. 3.1. m.w.H.

[325] S. ISABELLE WILDHABER, KI und Haftung: Lösungsansätze für die Schweiz, Jusletter IT 4. Juli 2024, Rz. 52 ff.; s. auch schon: ISABELLE WILDHABER, Eine Einführung in die ausservertragliche Haftung für Künstliche Intelligenz (KI), in: HAVE (Hrsg.), Haftpflichtprozess 2021, Zürich 2021, 1 ff., 45; BERNHARD A. KOCH/PASCAL PICHONNAZ, Der Entwurf einer neuen EU-Produkthaftungsrichtlinie aus schweizerischer Sicht, SJZ 2023, 627 ff., 637.

bereiche von KI-Anwendungen bestehen nämlich bereits nach geltendem Recht verschuldensunabhängige Gefährdungshaftungen, welche potentiell geschädigte Personen besser zu schützen vermögen, namentlich:

- die Haftpflicht des Motorfahrzeughalters nach Artikel 58 SVG,
- die Haftpflicht des Halters eines Luftfahrzeugs nach Artikel 64 Luftfahrtgesetz (LFG),[326]
- die Herstellerhaftung nach dem Bundesgesetz vom 18. Juni 1993 über die Produktehaftpflicht (PrHG)[327] (s. dazu das folgende Kapitel Ziff. 6.3.2).

In wichtigen praktischen Anwendungsgebieten der künstlichen Intelligenz (selbstfahrende Autos, Drohnen) bestehen damit schon heute Haftungsregeln, die den Schutz von Geschädigten sicherstellen. Durch die Haftpflicht der Halter der Fahrzeuge – verbunden mit einer Versicherungspflicht (Art. 63 SVG; Art. 70 LFG) – können Haftungslücken ausgeschlossen werden.

6.3.2 Produktehaftpflichtrecht

Die Produktehaftpflicht sieht eine verschuldensunabhängige Haftung der Herstellerinnen und Hersteller für durch fehlerhafte Produkte verursachte Sach- und Personenschäden vor. Geregelt werden somit nur die Folgeschäden von fehlerhaften Produkten und nicht die Schäden am Produkt selbst.

6.3.2.1 Revidierte EU-Richtlinie über die Haftung für fehlerhafte Produkte

Die am 12. März 2024 vom Parlament verabschiedete Richtlinie[328] über die Haftung für fehlerhafte Produkte wird die bisherige Richtlinie 85/374/EWG[329] vom 25. Juli 1985 zur Angleichung der Rechts- und Verwaltungsvorschriften der Mitgliedstaaten über die Haftung für fehlerhafte Produkte («Produkthaftungsrichtlinie») ersetzen. Sie geht zurück auf den Vorschlag

---

[326] SR **748.0**.

[327] SR **221.112.944**.

[328] Legislative Entschliessung des Europäischen Parlaments vom 12. März 2024 zu dem Vorschlag für eine Richtlinie des Europäischen Parlaments und des Rates über die Haftung für fehlerhafte Produkte, abrufbar unter: https://www.europarl.europa.eu/doceo/document/TA-9-2024-0132_DE.html (abgerufen am 26. August 2024).

[329] Richtlinie 85/374/EWG des Rates vom 25. Juli 1985 zur Angleichung der Rechts- und Verwaltungsvorschriften der Mitgliedstaaten über die Haftung für fehlerhafte Produkte, ABl. L 210 vom 7. August 1985, 29–33, abrufbar unter: https://eur-lex.europa.eu/legal-content/DE/ALL/?uri=celex%3A31985L0374 (abgerufen am 26. August 2024).

der Europäischen Kommission vom 28. September 2022 für eine Richtlinie[330] über die Haftung für fehlerhafte Produkte. Die revidierte Richtlinie sieht umfangreiche Neuerungen vor, um *Software einschliesslich KI-Systeme* der verschuldensunabhängigen Herstellerhaftung zu unterstellen. Obwohl die revidierte Produkthaftungsrichtlinie als Teil des «KI-Pakets» vorgeschlagen wurde, geht es im Kern nicht um KI-Anwendungen, sondern um eine allgemeine Modernisierung der Herstellerhaftung für fehlerhafte Produkte im Zuge der fortschreitenden Digitalisierung. Damit soll der Entwicklung Rechnung getragen werden, dass die Hersteller die Kontrolle oder Einflussnahme auf ihre Produkte heute mit Inverkehrbringen nicht zwangsläufig verlieren.[331]

Sowohl der *Kreis der Haftenden* (Softwarehersteller, Online-Plattformen und Betreiber von KI-Systemen) als auch der *Fehlerbegriff* (fehlerhafte Upgrades und Updates, mangelhafte Cybersicherheit) werden erweitert. Auch soll die Herstellerhaftung neben Sach- und Personenschäden auf Schäden wegen dem Verlust oder der Verfälschung von Daten ausgedehnt werden. Wie in der vorgeschlagenen Richtlinie über KI-Haftung wird Beweisschwierigkeiten mit *Offenlegungspflichten* und *Kausalitätsvermutungen* begegnet. Weiter gibt es in Bezug auf Update- und Aktualisierungspflichten Schnittstellen mit zwei weiteren zivilrechtlichen EU-Richtlinien:

- *Richtlinie (EU) 2019/771*[332] vom 20. Mai 2019 über bestimmte vertragsrechtliche Aspekte des Warenkaufs (*Warenkauf-Richtlinie [EU] 2019/771*) und
- *Richtlinie (EU) 2019/770*[333] des Europäischen Parlaments und des Rates vom 20. Mai 2019 über bestimmte vertragsrechtliche Aspekte der Bereitstellung digitaler Inhalte und digitaler Dienstleistungen (*Digitale-Inhalte-und-Dienste-Richtlinie [EU] 2019/770).*

Die Bestätigung der am 12. März 2024 durch das Parlament verabschiedeten Richtlinie durch den Europäischen Rat und die Publikation im Amtsblatt stehen zur Zeit (Stand 31.08.2024)

---

[330] Vorschlag für eine Richtlinie des Europäischen Parlaments und des Rates über die Haftung für fehlerhafte Produkte, COM/2022/495 final, 28. September 2022, abrufbar unter: https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52022PC0495 (abgerufen am 26. August 2024).

[331] BERNHARD A. KOCH/PASCAL PICHONNAZ, Der Entwurf einer neuen EU-Produkthaftungsrichtlinie (Fn. 325), 630.

[332] Richtlinie (EU) 2019/771 vom 20. Mai 2019 über bestimmte vertragsrechtliche Aspekte des Warenkaufs, zur Änderung der Verordnung (EU) 2017/2394 und der Richtlinie 2009/22/EG sowie zur Aufhebung der Richtlinie 1999/44/EG, ABl. L 136 vom 22. Mai 2019, 28 ff.

[333] ABl. L 136 vom 22.5.2019, 1 ff.

noch aus. Nach dem Inkrafttreten haben die Mitgliedstaaten zwei Jahre Zeit, um die Richtlinie umzusetzen.[334]

6.3.2.2 Schweizer Recht

Die Schweiz hat das PrHG im Rahmen des Swisslex-Pakets eingeführt, es entspricht weitgehend der bisherigen Produkthaftungsrichtlinie 85/374/EWG. [335] Gestützt auf die Regeln des PrHG können Herstellerinnen und Hersteller für fehlerhafte Produkte haftbar gemacht werden. In der Lehre ist seit Jahren umstritten, ob auch «reine» Software, das heisst Software, die nicht mit einem körperlichen Produkt verbunden ist, als Produkt im Sinne des PrHG gelten kann.[336] Das Bundesgericht hat sich zu dieser Frage noch nicht geäussert. Ein Grossteil der jüngeren Lehre spricht sich für eine entsprechend weite, dem Wortlaut widersprechende Auslegung von Artikel 3 Absatz 1 PrHG aus, wünscht sich jedoch eine Klarstellung durch den Gesetzgeber.[337]

6.3.2.3 Würdigung

Die Unterstellung von Software unter die Herstellerhaftung des PrHG, wie sie im Vorschlag der EU-Richtlinie vorgesehen ist, wird von der Schweizer Lehre seit längerem gefordert.[338] Nicht enthalten im geltenden PrHG sind auch die weiteren Neuerungen im Vorschlag für die revidierte EU-Richtlinie, welche der zunehmenden Digitalisierung der Produkte Rechnung tragen.[339]

In seinem Bericht «Modernisierungsbedarf des Gewährleistungsrechts beim Kauf» vom 16. Juni 2023 hat der Bundesrat in Bezug auf die *Gewährleistung des Verkäufers für Sachmängel* festgehalten, dass das geltende Recht, welches auf den Einmalaustausch von Ware gegen Geld ausgerichtet ist, der Interessenlage bei digitalen Produkten oder Produkten mit digitalen Komponenten nicht gerecht wird.[340] Er hat sich deshalb für die Einführung einer Update- oder Aktualisierungspflicht für digitale Produkte und Produkte mit digitalen Komponenten nach dem Vorbild der Digitale-Inhalte-und-Dienste-Richtlinie (EU) 2019/770 und der Warenkauf-Richtlinie (EU) 2019/771 ausgesprochen. Mit den gleichlautenden Motionen 23.4316

---

[334] Art. 22 der Richtlinie über die Haftung für fehlerhafte Produkte, abrufbar unter: https://www.europarl.europa.eu/doceo/document/TA-9-2024-0132_DE.html (abgerufen am 26. August 2024).

[335] S. Botschaft über das Folgeprogramm nach der Ablehnung des EWR-Abkommens vom 24. Februar 1993, BBl **1993** I 805 884.

[336] S. BSK OR-FELLMANN, Art. 3 PrHG N 10 m.w.H.

[337] S. nur FELLMANN, Haftpflichtrecht im Zeichen der Digitalisierung, HAVE 2021, 105 ff.; ISABELLE WILDHABER, Eine Einführung (Fn. 325), 26; BERNHARD A. KOCH/PASCAL PICHONNAZ, Der Entwurf einer neuen EU-Produkthaftungsrichtlinie (Fn. 325), 638, je mit weiteren Hinweisen; für eine Unterstellung von KI-Produzenten unter die Herstellerhaftung de lege lata namentlich auch: ARIANE MORIN, L'opposabilité de la LRFP au fournisseur de l'intelligence artificielle, in: Alexandre Ri-Cha/Damiano Canapa (Hrsg.), Aspects juridiques de l'intelligence artificielle, Bern 2024, 117 ff., 120 f.

[338] S. die Nachweise oben in Fn. 337 sowie ISABELLE WILDHABER, KI und Haftung (Fn. 325), Rz. 26 ff.

[339] S. zum Anpassungsbedarf des PrHG im Einzelnen: BERNHARD A. KOCH/PASCAL PICHONNAZ, Der Entwurf einer neuen EU-Produkthaftungsrichtlinie (Fn. 325), 637 ff.

[340] Bericht des Bundesrates «Modernisierungsbedarf des Gewährleistungsrechts beim Kauf» vom 16. Juni 2023, Ziff. 4.1 ff., abrufbar unter: https://www.newsd.admin.ch/newsd/message/attachments/79587.pdf (abgerufen am 31. Juli 2024).

und 23.4345 der Rechtskommissionen beider Räte wurde der Bundesrat in der Wintersession 2023 beauftragt, eine entsprechende Vorlage auszuarbeiten.[341]

Der Bundesrat und das Parlament haben sich bislang nicht zu einer allfälligen Revision des PrHG geäussert.

6.3.3 Schutz der Persönlichkeit

Im Zusammenhang mit KI-Anwendungen sind schliesslich auch die Mittel des allgemeinen zivilrechtlichen Persönlichkeitsschutzes zu thematisieren. Diese sind beim praktisch wichtigen Phänomen der *Deepfakes*[342] besonders relevant. Eine in ihrer Persönlichkeit verletzte Person kann zu ihrem Schutz gegen jeden, der an der Verletzung mitwirkt, das Zivilgericht anrufen (Art. 28 Abs. 1 ZGB). Auf diesem Weg kann die betroffene Person namentlich verlangen, dass eine bestehende Persönlichkeitsverletzung beseitigt oder eine drohende Persönlichkeitsverletzung verboten wird (Art. 28*a* Abs. 1 Ziff. 1 und 2 ZGB). Diese Regelung ist technologieneutral.[343] Es handelt sich bei der Persönlichkeit (inkl. Recht am eigenen Bild und der eigenen Stimme) um ein absolut geschütztes Rechtsgut. Artikel 28 ZGB enthält ein «stillschweigendes, allgemeines Verbot unbefugter Beeinträchtigungen fremder Persönlichkeitssphären».[344] Persönlichkeitsverletzungen mittels KI-Anwendung werden von den geltenden Regeln deshalb ohne Weiteres erfasst. Von Bedeutung ist, dass Bildmanipulationen auch schon vor der Verbreitung von KI-Anwendungen existierten. So hatte sich das Bundesgericht auch in der Vergangenheit bereits mit Fotomontagen und Bildmanipulationen zu befassen, wobei dort nicht in Frage gestellt wurde, dass auch manipulierte Bilder die Persönlichkeitsrechte der Betroffenen tangieren können.[345] Dies wurde auch in einem aktuellen erstinstanzlichen Urteil zu einem mittels KI erstellen Video gestützt: Ein Politiker verwendete zu Wahlkampfzwecken ein mit einer KI-Anwendung erstelltes, gefälschtes Video einer politischen Konkurrentin. Das Basler Zivilgericht stufte dies als Persönlichkeitsverletzung ein.[346] Gestützt auf Artikel 28*a* Absatz 1 Ziffern 1 und 2 ZGB kann sowohl das Löschen als auch das Sperren von rechtswidrigen Inhalten im Internet angeordnet werden. Auch Klagen auf Schadenersatz, Genugtuung sowie Gewinnherausgabe sind denkbar (Art. 28*a* Abs. 3 ZGB).

6.3.4 Fazit

Allgemein ist das Schweizer Zivilrecht und namentlich auch das Haftpflichtrecht mit seinen offenen Generalklauseln in der Lage, technische Entwicklungen aufzufangen und stellt den

[341] 23.4316 Mo. RK-S «Modernisierung des Gewährleistungsrechts» und 23.4345 Mo. RK-N «Modernisierung des Gewährleistungsrechts».

[342] Vgl. 23.3563 Mo. Mahaim «Deepfakes regulieren».

[343] S. zur Technologieneutralität der Schweizer Gesetzgebung und der Erfassung von Deepfakes im Strafrecht auch unten: Ziff. 6.6.

[344] HEINZ HAUSHEER/REGINA E. AEBI-MÜLLER, Das Personenrecht des Schweizerischen Zivilgesetzbuches, 5. Aufl., Bern 2020, Rz. 490 ff., 544.

[345] Vgl. Urteil des BGer 5A_553/2012 vom 14. April 2014; Urteil des BGer 5A_376/2013 vom 29. Oktober 2013.

[346] Medialex newsletter 01/24, abrufbar unter: https://medialex.ch/2024/02/07/newsletter-01-24/ (abgerufen am 31. Juli 2024).

Gerichten ein Instrumentarium zur Verfügung, um im Einzelfall zu gerechten Lösungen zu gelangen. Durch die bestehenden Gefährdungshaftungen und Versicherungspflichten im Strassenverkehr und der Luftfahrt können zudem Haftungslücken in zentralen Anwendungsbereichen ausgeschlossen werden. Auch die gerichtliche Durchsetzung der Ansprüche ist auf Grundlage der bestehenden Regeln grundsätzlich möglich. Eine Übernahme der diskutierten EU-Richtlinie über KI-Haftung könnte aber die gerichtliche Durchsetzung von zivilrechtlichen Ansprüchen eventuell erleichtern und auch dazu beitragen, die Vorgaben der KI-Konvention des Europarats (namentlich Art. 8, 9 und 14, siehe Kapitel 4.3) im privaten Sektor besser umzusetzen. Eine isolierte Einführung der Richtlinie – ohne gleichzeitige Übernahme des KI-Gesetzes – ist jedoch nicht denkbar. Auch bleiben die zusätzlichen Analysen und Ergebnisse der Diskussionen in der EU zur Richtlinie über KI-Haftung – deren Zukunft ungewiss ist – abzuwarten.

Aufgrund der technischen Entwicklungen der Produkte – nicht nur, aber auch im Zusammenhang mit KI – zeichnet sich ein allgemeiner Modernisierungsbedarf in Bezug auf das Produktehaftpflichtgesetz ab.

6.4 Allgemeines Vertragsrecht

6.4.1 UNCITRAL Model Law on Automated Contracting

Die UN Commission on International Trade Law (UNCITRAL) hat im Juli 2024 das UNCITRAL Model Law on Automated Contracting verabschiedet.[347] Dieses wurde zuvor in der «Working Group IV: Electronic Commerce» der UNCITRAL – unter aktiver Mitwirkung der Schweiz – ausgearbeitet.[348] Die Bestimmungen des Modellgesetzes sehen vor, dass die Gültigkeit von Verträgen nicht allein aufgrund des Umstandes verneint werden darf, dass an dessen Abschluss ein automatisiertes System beteiligt war. Weitere Regeln betreffen die Zuordnung von durch automatisierte Systeme generierten Nachrichten. Eine Übernahme des Modellgesetzes für die Schweiz, bzw. ob dieses durch das geltende Recht bereits erfüllt wird, wird zu gegebener Zeit geprüft werden.

6.4.2 Schweizer Recht

6.4.2.1 Zurechnung von Willenserklärungen und vertragliche Haftung

Vertragliche Bindung setzt einen tatsächlichen oder *normativen* Konsens voraus, einen ausdrücklich oder *vertrauenstheoretisch* erklärten Rechtsfolgewillen. Entscheidend ist, ob ein

[347] S. die Medienmitteilung unter: https://unis.unvienna.org/unis/pressrels/2024/unisl362.html (abgerufen am 26. August 2024).

[348] Sämtliche Arbeitsdokumente sind abrufbar unter: https://uncitral.un.org/en/working_groups/4/electronic_commerce (abgerufen am 26. August 2024).

Verhalten an den Tag gelegt wurde, aus dem die Gegenpartei in guten Treuen auf das Vorhandensein eines solchen Willens schliessen durfte.[349] Schon heute werden im Online-Bereich Verträge durch Einsatz von Software mit einem hohen Automationsgrad abgeschlossen, ohne dass sich darauf Probleme bei der Zuordnung von Willenserklärungen ergäben. Vereinzelt wird in der Literatur diskutiert, ob sich in Zukunft bei zunehmender Autonomie von KI-Anwendungen Regelungslücken ergeben könnten, wenn sich beispielsweise Willenserklärungen nicht mehr klar dem Verantwortungsbereich einer Vertragspartei zuordnen lassen.[350]

Wird ein Anspruch aus Vertragsverletzung geltend gemacht, besteht insofern eine Erleichterung für geschädigte Personen, als sie das Verschulden der vertragsverletzenden Person nicht nachweisen müssen. Vielmehr muss die Person, welche den Vertrag nicht oder nicht gehörig erfüllt hat, nachweisen, dass ihr kein Verschulden zur Last fällt (Art. 97 Abs. 1 OR). Auch hier wird in der Lehre aber bereits diskutiert, ob man sich inskünftig bei der Fehlfunktion einer KI-Anwendung (zu) leicht durch den Nachweis, dass keine eigene Sorgfaltspflichtverletzung begangen wurde, exkulpieren könnte. Zur Diskussion gestellt wird insofern eine Angleichung an die Hilfspersonenhaftung nach Artikel 101 Absatz 1 OR, welche eine Zuordnung von Fehlern autonomer Anwendungen an die Partei, welche die Anwendung einsetzt, erlauben würde.[351] Es ist gegenwärtig jedoch nicht davon auszugehen, dass hier tatsächlich eine Regelungslücke besteht, lässt sich im Einzelfall doch beispielsweise auch der Einsatz einer KI-Anwendung als Sorgfaltspflichtverletzung qualifizieren.[352] Es darf auch nicht ausser Acht gelassen werden, dass KI-Anwendungen – im Gegensatz zu Personen, die zur Erfüllung einer vertraglichen Pflicht eingesetzt werden – keine eigene Rechtspersönlichkeit haben. Letztlich handelt es sich dabei um technische Hilfsmittel, die von einer Vertragspartei eingesetzt werden und deren Risikosphäre zuzuordnen sind. Eine vorschnelle Analogie zu Stellvertreter- oder Hilfspersonenregeln könnte Haftungslücken entstehen lassen und scheint Stand heute nicht sachgerecht.

6.4.2.2 Smart Contracts

Der Bundesrat hat sich in seinem Bericht «Rechtliche Grundlagen für Distributed Ledger-Technologie und Blockchain in der Schweiz – Eine Auslegeordnung mit Fokus auf dem Finanzsektor» vom 14. Dezember 2018 zu *Smart Contracts* geäussert.

Ein Smart Contract ist ein Computerprotokoll, meist basierend auf einem dezentralen Blockchainsystem, das die automatisierte Vertragserfüllung zwischen zwei oder mehr Parteien mit vorgängig codierten Angaben ermöglicht.[353] Laut dem Erfinder des Konzepts ist die einfachste Form eines Smart Contract der Verkaufsautomat, der die Ware freigibt, sobald der

---

[349] BGE **116** II 695 E. 2.

[350] Vgl. MICHAEL MARTIN KIANIČKA, Die Agentenerklärung. Elektronische Willenserklärungen und künstliche Intelligenz als Anwendungsfalls der Rechtsscheinhaftung, Zürich 2012; MALTE GRÜTZMACHER/JÖRN HECKMANN, Autonome Systeme und KI – vom vollautomatisierten zum autonomen Vertragsschluss? Die Grenzen der Willenserklärung, Computer und Recht 2019, 553 ff.

[351] CHRISTAPOR YACOUBIAN, Digitale Systeme als «Erfüllungsgehilfen» - Relevanz der fehlenden Rechtsfähigkeit, AJP 2023, 412 ff.; CHRISTAPOR YACOUBIAN, Der Analogieschluss zu Art. 101 Abs. 1 OR und die Grenzen zulässiger Rechtsfortbildung, AJP 2024, 195 ff.; MELINDA F. LOHMANN/THERESA PRESSLER, Algorithmische Vertragserfüllung (Teil 1) – Eine zukunftsgewandte Betrachtung der Rechtsfiguren der Erfüllungsgehilfin und der Substitutin unter Analyse des Urteils BGer 4A_305/2021 vom 2. November 2021, SJZ 2023, 879 ff., 884 ff.

[352] Vgl. CORINNE-WIDMER-LÜCHINGER, Apps, Algorithmen und Roboter in der Medizin: Haftungsrechtliche Herausforderungen, HAVE 2019, 3 ff.

[353] Definition in Anknüpfung an die relativ einheitliche Lehre, siehe u. a. LEE BACON/GEORGE BAZINAS, « Smart Contracts »: The next big Battleground?, Jusletter IT 18. Mai 2017, 2; MARKUS KAULARTZ/JÖRG HECKMANN, Smart contracts – Anwendungen der Blockchain-Technologie, Computer und Recht 2016, 618 ff., 618; STEPHAN D. MEYER/BENEDIKT SCHLUPPI, « Smart Contracts » und deren Einordnung in das schweizerische Vertragsrecht, recht 2017, 204 ff., 207; ROLF H. WEBER, Smart Contracts: Vertrags- und verfügungsrechtlicher Regelungsbedarf?, sic! 2018, 291 ff. 291 f.

Preis bezahlt worden ist.[354] Trotz seiner Bezeichnung ist ein Smart Contract, so ist sich die Lehre weitgehend einig, kein Vertrag im Sinne des Obligationenrechts, sondern eine Vertragserfüllungs-«Technologie».[355] Probleme können sich durch die Starrheit der vorprogrammierten Vertragsabwicklung geben, welche die im Schweizer Recht vorgesehenen Anpassungsmöglichkeiten von Verträgen im Einzelfall – beispielsweise durch die *clausula rebus sic stantibus* – verunmöglichen. Nach heutigem Stand empfiehlt die Lehre Parteien beim Abschluss eines Smart Contract geeignete Mechanismen für möglicherweise wechselnde Lebensumstände sowie zur Streitbeilegung vorzusehen.[356] Die im Blockchainkontext verbreitete Anonymität der Parteien stellt ein grosses faktisches Hindernis für die gerichtliche Durchsetzung von allfälligen Ansprüchen dar. Dieses Problem kann gesetzgeberisch jedoch nicht gelöst werden. Der Bundesrat ist in seinem Bericht «Rechtliche Grundlagen für Distributed Ledger-Technologie und Blockchain in der Schweiz» zum Schluss gelangt, dass es im Bereich Smart Contract sicher weitere Entwicklungen geben wird, die aber erst angelaufen sind und keine Regulierung aufdrängen.

6.4.3 Würdigung

Die Diskussionen zum Einsatz von KI-Anwendungen in Vertragsbeziehungen stehen noch am Anfang, auch haben sich in der Praxis bislang soweit ersichtlich keine Probleme ergeben. Die Entwicklungen sind auf jeden Fall weiterzuverfolgen, ein Eingreifen des Gesetzgebers schiene jedoch zum jetzigen Zeitpunkt nicht angezeigt.

6.5 Arbeitsrecht

6.5.1 Einleitung

In diesem Kapitel werden arbeitsrechtliche Fragen vertieft. Der Einsatz von KI in der Personalverwaltung hat stark zugenommen und ist heute in vielen Unternehmen üblich. Dies wirft neue und spezifische rechtliche Fragen auf, insbesondere im Hinblick auf die zahlreichen Vorschriften zum Schutz der Arbeitnehmerinnen und Arbeitnehmer, die für diesen Bereich kennzeichnend sind. Diese Entwicklungen haben bereits ihren Niederschlag in spezifischen Regelungen oder Gesetzesinitiativen auf EU-Ebene oder in der Schweiz gefunden.

---

[354] NICK SZABO, Formalizing and securing relationships on public networks, First Monday Internet Journal, 1997, 1.

[355] ANDREAS FURRER, Die Einbettung von Smart Contracts in das schweizerische Privatrecht, Anwalts Revue 2018, 103 ff., 109; GABRIEL OLIVIER BENJAMIN JACCARD, Smart contracts and the role of Law, Jusletter 23. November 2017, Rz. 8–9; STEPHAN D. MEYER/BENEDIKT SCHLUPPI, Smart Contracts (Fn. 353), 208; ROLF H. WEBER, Leistungsstörungen und Rechtsdurchsetzung bei Smart Contracts: Eine Auslegeordnung möglicher Problemstellungen, Jusletter 4. Dezember 2017, Rz. 2. Zu den Kontroversen: HANS RUDOLF TRÜEB, Smart Contracts, in: Pascal Grolimund u. a. (Hrsg.), Festschrift für Anton K. Schnyder, Zürich 2018, 723 ff., 725.

[356] Siehe nur ROLF H. WEBER, Leistungsstörungen (Fn. 355), Rz. 33 ff.

6.5.2 Auf europäischer Ebene

6.5.2.1 Richtlinie zur Verbesserung der Arbeitsbedingungen in der Plattformarbeit

Dem BJ sind keine Gesetzgebungsvorhaben auf EU-Ebene bekannt, die sich allgemein mit dem Einsatz von KI im Arbeitsverhältnis befassen. Hingegen ist die EU dabei, eine Richtlinie zur Verbesserung der Arbeitsbedingungen in der Plattformarbeit zu verabschieden, die auch Regeln für den Einsatz von Algorithmen in diesem Zusammenhang enthält.[357] Die Verhandlungsführer des Parlaments und des Rates haben am 8. Februar 2024 eine politische Einigung über den Text erzielt.[358] Der Rat der EU hat diese Einigung am 11. März 2024[359] und das Europäische Parlament am 24. April 2024[360] bestätigt. Für die endgültige Verabschiedung des Textes fehlt nur noch die förmliche Annahme durch den Ministerrat und die Unterzeichnung durch die Mitgesetzgeber.[361]

Ziel der Richtlinie ist die Verbesserung der Arbeitsbedingungen und des Schutzes von personenbezogenen Daten von Plattformbeschäftigten, insbesondere durch Förderung von Transparenz, Fairness und Rechenschaftspflicht beim algorithmischen Management im Kontext der Plattformarbeit (Art. 1 Abs. 1 Bst. b). Sie enthält ferner Vorschriften zur Verbesserung des Schutzes natürlicher Personen bei der Verarbeitung ihrer personenbezogenen Daten, indem sie Massnahmen für das algorithmische Management vorsieht (Art. 1 Abs. 2). Die Richtlinie gilt für Plattformtätigkeiten, die in der Europäischen Union erbracht werden, unabhängig vom Niederlassungsort der Plattform und unabhängig von dem ansonsten anwendbaren Recht (Art. 1 Abs. 3). Plattformen mit Sitz in der Schweiz könnten daher zur Anwendung dieser Richtlinie verpflichtet sein, wenn die Plattformbeschäftigten in der EU Leistungen für die Schweiz erbringen. Die Arbeit auf einer digitalen Plattform umfasst nach der Definition der Richtlinie Dienstleistungen, die auf elektronischem Weg, wie über eine Website oder eine mobile Anwendung angeboten werden (Art. 2 Abs. 1 Nr. 1 Bst. a). Unternehmen, die ihre Arbeit

---

[357] Vorschlag für eine Richtlinie des Europäischen Parlaments und des Rates zur Verbesserung der Arbeitsbedingungen in der Plattformarbeit, 9. Dezember 2021, KOM(2021) 762 endgültig/final, abrufbar unter: https://eur-lex.europa.eu/legal-content/DE/ALL/?uri=CELEX%3A52021PC0762 (abgerufen am 28. August 2024).

[358] Vgl. https://www.europarl.europa.eu/news/en/press-room/20240205IPR17417/provisional-deal-on-first-eu-wide-rules-for-platform-workers (abgerufen am 28. August 2024) und https://data.consilium.europa.eu/doc/document/ST-7212-2024-ADD-1/de/pdf (abgerufen am 28. August 2024).

[359] Vgl. https://www.consilium.europa.eu/de/press/press-releases/2024/03/11/platform-workers-council-confirms-agreement-on-new-rules-to-improve-their-working-conditions/ (abgerufen am 28. August 2024).

[360] Vgl. https://www.europarl.europa.eu/news/de/press-room/20240419IPR20584/le-parlement-adopte-la-directive-sur-le-travail-des-plateformes (abgerufen am 28. August 2024).

[361] Das Parlament hat am 4. Juli 2024 eine Berichtigung angenommen (deutsche Fassung vom 9. Juli 2024 abrufbar unter: https://www.europarl.europa.eu/doceo/document/TA-9-2024-0330-FNL-COR01_DE.pdf), mit redaktionellen oder formalen Änderungen an dem am 24. April 2024 angenommenen Text. Die Unterschiede zu dem am 24. April angenommenen Text sind in den Fussnoten angegeben.

auf herkömmliche Weise organisieren, sind daher von der Richtlinie nicht erfasst. Zu den kumulativen Definitionsmerkmalen einer digitalen Plattform gehört auch der Einsatz von Instrumenten zur Überwachung oder automatisierten Entscheidungsfindung (Art. 2 Abs. 1 Nr. 1 Bst. d). Auch Instrumente zur Entscheidungsunterstützung gehören dazu (Art. 2 Abs. 1 Nr. 9[362]).

Kapitel III (Art. 7 ff.) befasst sich mit dem algorithmischen Management und sieht unter anderem die folgenden Rechte und Pflichten vor:

- Artikel 7 Absatz 1 verbietet die Erhebung oder Verarbeitung von Daten bezüglich des Privatlebens der Arbeitnehmerin oder des Arbeitnehmers mittels Systemen zur Überwachung oder automatisierten Entscheidungsfindung, wie z. B. private Gespräche, Vorhersagen über die Art und Weise der Ausübung ihrer oder seiner Grundrechte, oder Daten über sensible persönliche Merkmale (z. B. Rasse oder ethnische Herkunft, Gesundheitszustand) oder die Verarbeitung biometrischer Daten oder von Daten über den emotionalen oder psychischen Zustand.

- Artikel 8 hält fest, dass die Verarbeitung personenbezogener Daten durch ein automatisiertes Überwachungs- oder Entscheidungssystem voraussichtlich ein hohes Risiko für die Rechte und Freiheiten natürlicher Personen im Sinne von Art. 35 Abs. 1 DSGVO zur Folge hat, und sieht die Pflicht zur Durchführung einer Folgenabschätzung vor. Die Folgenabschätzung ist den Arbeitnehmervertretern auszuhändigen (Art. 8 Abs. 2).

- Artikel 9 sieht eine Informationspflicht über die Nutzung automatisierter Überwachungs- oder Entscheidungssysteme vor, auch wenn die Entscheidungen keine wesentlichen Auswirkungen auf Personen haben, die auf einer Plattform arbeiten.

- Artikel 10 sieht eine Pflicht zur Überwachung und Evaluierung von Systemen für automatisierte Entscheidungen vor. Jede Entscheidung in Bezug auf Aussetzung oder Beendigung des Vertragsverhältnisses ist von Menschen zu treffen (Art. 10 Abs. 5).

- Artikel 11 sieht die Pflicht zur Erläuterung einer von einem automatisierten System getroffenen Entscheidung sowie das Recht auf Überprüfung und gegebenenfalls Berichtigung der Entscheidung vor (Art. 11 Abs. 2 und 3).

- Artikel 12 enthält Verpflichtungen zur Bewertung und Vermeidung von Gesundheits- und Sicherheitsrisiken.

- Artikel 13 sieht eine Pflicht zur Unterrichtung und Anhörung der Arbeitnehmervertretung bei Entscheidungen über die Einführung oder wesentliche Änderung automatisierter Systeme vor. Die Arbeitnehmervertretung ist darüber hinaus Adressat der Informationspflicht nach Art. 9, wird in die Überwachung und Bewertung von KI-Systemen einbezogen (Art. 10) und kann Entscheidungen solcher Systeme überprüfen lassen (Art. 11).

[362] Art. 2 Abs. 1 Bst. j gemäss Berichtigung.

- Schliesslich sieht Artikel 21 die Pflicht zur Vorlage aller relevanten Beweismittel vor, und die Artikel 22 und 23 verlangen Schutz vor Vergeltungsmassnahmen oder Entlassung wegen Geltendmachung von Rechten aus der Richtlinie.

- Die Bestimmungen der Richtlinie sind zwingend, d. h. ungünstigere Regelungen sind weder in den nationalen Gesetzen der Mitgliedstaaten noch in Gesamtarbeitsverträgen zulässig (Art. 26 Abs. 2). Günstigere Regelungen sind hingegen möglich. Die Richtlinie weicht in einigen Punkten bewusst von der DSGVO ab. Insbesondere schliesst sie die Einwilligung als Rechtfertigungsgrund für eine Datenverarbeitung aus (vgl. E. 40).

6.5.2.2 EU-Verordnung zu KI und Arbeitsrecht

Die KI-Verordnung (vgl. Ziff. 5) enthält trotz ihres allgemeinen Charakters Regelungen, die speziell für die Arbeitswelt relevant sind. Sie stuft die Anwendungen für die Arbeitswelt nach vorgegebenen Risikostufen ein.

So werden Systeme zur Erkennung von Emotionen als verbotene Anwendungen eingestuft, wenn sie am Arbeitsplatz eingesetzt werden, es sei denn, die Anwendung erfolgt aus medizinischen Gründen oder zu Sicherheitszwecken (Art. 5 Abs. 1 Bst. f). Eine Reihe von Anwendungen wird als hochriskant eingestuft (Art. 6 Abs. 2 und Anhang III Nr. 4): Systeme zur Anwerbung oder Auswahl natürlicher Personen, einschliesslich der gezielten Veröffentlichung von Stellenangeboten, der Analyse und Filterung von Bewerbungen und der Bewertung von Bewerbern; Systeme zum Treffen von Entscheidungen, welche die Arbeitsbedingungen, die Beförderung oder die Beendigung des Arbeitsverhältnisses beeinflussen, oder die dazu bestimmt sind, Aufgaben auf der Grundlage des individuellen Verhaltens, der Persönlichkeitsmerkmale oder der persönlichen Eigenschaften zuzuweisen, sowie Überwachungs- und Bewertungssysteme. Die mit der Einstufung in die Hochrisikokategorie verbundenen Pflichten sowie die Ausnahmen von dieser Einstufung wurden bereits im Abschnitt über die KI-Verordnung erläutert (vgl. Ziff. 5.2.7).

Es stellt sich die Frage, ob der Arbeitgeber Adressat der Pflichten aus der KI-Verordnung sein kann, insbesondere der Pflichten gemäss Kapitel 3 der Verordnung über die Hochrisikokategorie. Der Arbeitgeber wird grundsätzlich die Person sein, die das KI-System im Sinne von Art. 3 Nr. 4 KI-Verordnung einsetzt (Betreiber). So sieht Art. 26 Abs. 7 der Verordnung u. a. vor, dass der Arbeitgeber die Arbeitnehmervertreter und die betroffenen Arbeitnehmenden darüber informiert, dass sie der Verwendung eines Hochrisiko-KI-Systems unterliegen werden (vgl. zu weiteren Pflichten zu Lasten des Betreibers Ziff. 5.2.7.3.3). Gestützt auf Art. 25 Abs. 1 KI-Verordnung kann jedoch auch der Arbeitgeber als Anbieter betrachtet werden, da jede Person, die wesentliche Änderungen an einem Hochrisiko-KI-System vornimmt oder dessen Zweckbestimmung ändert, als Anbieter mit den entsprechenden Pflichten gilt.

Auch diskutiert wurde die Wechselwirkung zwischen diesen allgemeinen Regelungen und den auf EU-Ebene vorgesehenen arbeitsrechtlichen Bestimmungen.[363] Zunächst ist klar, dass

[363] Vgl. insbesondere AUDE CEFALIELLO/MIRIAM KULLMANN, Offering false security: How the draft artificial intelligence act undermines fundamental workers rights, European Labour Law Journal 2022, 542 ff.

die Verordnung einen horizontalen und damit subsidiären Charakter hat und alle arbeitsrechtlichen Verpflichtungen weiterhin gelten. Darüber hinaus kann es sein, dass bei der Konzeption des Systems, in das KI integriert wird, Massnahmen erforderlich sind, um es mit den bestehenden arbeitsrechtlichen Vorschriften in Einklang zu bringen. Diese Massnahmen obliegen jedoch dem Arbeitgeber, der nicht notwendigerweise Anbieter und Entwickler des Systems ist. Der Arbeitgeber, der sich für die Einführung eines KI-Systems entscheidet, muss sicherstellen, dass er seinen arbeitsrechtlichen Verpflichtungen nachkommt, was bedeutet, dass er über die Informationen und Mittel verfügen muss, um die Funktionsweise des Systems gegebenenfalls anzupassen.

6.5.3 Schweizerisches Arbeitsrecht: Aktueller Stand und Diskussion

6.5.3.1 Motion 23.4492 Gysi «Künstliche Intelligenz am Arbeitsplatz. Mitwirkungsrechte der Arbeitnehmenden stärken»

Die am 22. Dezember 2023 eingereichte Motion 23.4492 verlangt, «die Mitwirkungsrechte der Arbeitnehmenden beim Einsatz von Künstlicher Intelligenz (KI) und algorithmischen Systemen am Arbeitsplatz auf gesetzlicher Ebene zu stärken, wenn diese Systeme für Empfehlungen, Prognosen, Entscheidungen usw. benutzt werden, die Arbeitnehmende betreffen oder Arbeitnehmendendaten verwenden». Hierfür «sollen das Mitspracherecht ausgeweitet, Informationsrechte gestärkt, kollektive Klagerechte geschaffen sowie Sanktionsmöglichkeiten geprüft werden». Die Motion basiert auf einer Stellungnahme der Gewerkschaft syndicom und der NGO Algorithmwatch, die wiederum auf einer von diesen Organisationen in Auftrag gegebenen Studie beruht.[364]

6.5.3.2 Einsatz von KI in der schweizerischen Arbeitswelt

Eine 2018 durchgeführte und 2020 wiederholte Umfrage bei Schweizer Grossunternehmen mit fünf Fallstudien bietet einen Überblick über die Verbreitung des Einsatzes von KI in Schweizer Unternehmen.[365] Rund zwei Drittel der Befragten gaben an, solche Instrumente zu nutzen, wobei zwischen 2018 und 2020 ein Anstieg zu verzeichnen ist.[366]

Die am häufigsten genannten Einsatzbereiche waren: Rekrutierung, Mitarbeiterbindung und -entwicklung, Leistungsmanagement, Arbeitsplatzgestaltung, Compliance sowie Auswahl- und Entlassungsentscheidungen. Die Studie zeigt auch die am häufigsten genannten Verwendungszwecke wie Mitarbeiterbindung und -entwicklung (63 %), Leistungsmanagement (47 %)

---

[364] ISABELLE WILDHABER/ISABEL EBERT, Beteiligung der Arbeitnehmenden beim Einsatz von ADM-Systemen am Arbeitsplatz, November 2023, abrufbar unter: https://syndicom.ch/fileadmin/user_upload/Web/Website/Dossiers/KI/2023_Rechtsgutachten_final.pdf (abgerufen am 28. August 2024).

[365] Diese Studie wurde durchgeführt im Rahmen des Nationalen Forschungsprogramms «Big Data» (PNR 75), vgl. https://www.nfp75.ch/de (abgerufen am 28. August 2024).

[366] ISABELLE WILDHABER/ISABEL EBERT, Beteiligung der Arbeitnehmenden (Fn. 364), 5.

und Rekrutierung (39 %). Weit verbreitet sind beispielsweise die Vorabprüfung von Bewerbungen bei der Einstellung oder Instrumente zur Überwachung von Computereingaben oder Internetnutzung. In diesen Fällen stellt sich die Frage, welche Rolle KI spielt oder ob diese Werkzeuge überhaupt unter die Definition von KI fallen.[367] Die vorliegende Studie konzentriert sich auf Systeme zur Unterstützung der Entscheidungsfindung (Teilautomatisierung) oder auf vollautomatisierte Systeme.

6.5.3.3 Herausforderungen durch KI im schweizerischen Arbeitsrecht

Die folgenden Herausforderungen wurden im Zusammenhang mit KI am Arbeitsplatz in der oben genannten Studie oder in anderen Publikationen identifiziert:

- Datenschutz[368]: Diese Systeme können potenziell in grossem Umfang und kontinuierlich Personendaten von Arbeitnehmenden sammeln. Die Sammlung kann undifferenziert erfolgen und die Zwecke können vielfältig oder unklar sein.

- Gesundheitsschutz: Instrumente zur Leistungsmessung oder -überwachung können die psychische Gesundheit und das Wohlbefinden von Beschäftigten beeinträchtigen, wenn diese sich ständig bewertet fühlen.[369] Andererseits können Überlastungssituationen entstehen, wenn ein KI-Tool die Aufgabenerfüllung anhand von realitätsfernen Ausführungsparametern überwacht.

- Verzerrungen bei der Bewertung, starre Entscheidungsfindung: Die gemessenen oder in das System eingegebenen Parameter berücksichtigen nicht unbedingt alle Besonderheiten einer Situation (Hindernisse auf dem üblichen Weg, die es unmöglich machen, Lasten in der vorgesehenen Zeit zu transportieren; ein Stau, der es nicht erlaubt, eine Strecke in einer Standardzeit zurückzulegen). Der Einsatz von KI führt auch eine zahlenbasierte Leistungslogik ein, die wichtige, aber nicht quantifizierbare Elemente wie persönliche Qualitäten und soziale Kompetenzen oder nicht messbare Schwierigkeiten ausser Acht lässt.[370] Eine

---

[367] ISABELLE WILDHABER/ISABEL EBERT, Beteiligung der Arbeitnehmenden (Fn. 364), 10.

[368] ISABELLE WILDHABER, Die Roboter kommen - Konsequenzen für Arbeit und Arbeitsrecht, RDS 2016, 315 ff., 346 f.; ISABELLE WILDHABER, Répercussions de la robotique et de l'intelligence artificielle sur le lieu de travail, in: Jean-Philippe Dunand/Pascal Mahon/Aurélien Witzig (Hrsg.), La révolution 4.0 au travail - Une approche multidisciplinaire, Genf/Zürich/Basel 2019, 201 ff., 228 f.; WOLFGANG DÄUBLER, Digitalisierung und Arbeitsrecht, Frankfurt am Main 2022, §9, N 5 ff.

[369] ISABELLE WILDHABER/ISABEL EBERT, Beteiligung der Arbeitnehmenden (Fn. 364), 18: 22 % der Studienteilnehmer gaben an, dass das Verhalten der Arbeitnehmenden und wie sie zusammenarbeiten überwacht werde. Siehe auch AUDE CEFALIELLO/MIRIAM KULLMANN, Offering false security (Fn. 363), 559.

[370] Vgl. dazu JEAN-CHRISTOPHE SCHWAAB, Les nouvelles tendances en matière d'évaluation du personnel et le droit du travail, Revue de droit du travail et d'assurance-chômage, 2019, 103 ff., 114 f.

Software, die die Tastatureingaben registriert, könnte zu unbegründeten negativen Bewertungen führen, wenn die Person an einem schwierigen Thema arbeitet und häufig innehält, um nachzudenken. Auch Diskussionen unter Kollegen, ihre Qualität und ihr positiver Beitrag lassen sich mit KI-gestützten Instrumenten nur schwer erfassen. Diese Grenzen der automatisierten Entscheidungsfindung am Arbeitsplatz haben dazu geführt, dass über verschiedene Massnahmen nachgedacht wurde, wie z. B. die Notwendigkeit der Entscheidungsfindung oder -kontrolle durch einen Menschen, die Entscheidungsautonomie des Menschen gegenüber den Ergebnissen der Maschine oder das Recht, die von einem KI-System getroffene oder auf dessen Bewertungen beruhende Entscheidung erläutern oder von einem Menschen überprüfen zu lassen.[371] Wie in Ziff. 6.5.2 dargelegt, wurden diese Rechtsbehelfe bereits in die von der EU erlassenen Rechtsvorschriften aufgenommen. Dies ist teilweise auch in der Schweiz mit dem neuen DSG der Fall (vgl. Ziff. 6.5.3.4).

- Diskriminierende Verzerrungen: Die Frage des Diskriminierungspotenzials ist im Zusammenhang mit der Integration von KI in Entscheidungsprozesse am Arbeitsplatz, insbesondere bei der Personalrekrutierung, viel diskutiert worden.[372] Dieses Thema wird in der vorliegenden Analyse gesondert behandelt (vgl. Ziff. 4.3.2.5), aber trotzdem auch im vorliegenden Zusammenhang erwähnt. Verzerrungen ergeben sich insbesondere daraus, dass die vorhandenen Datensätze und Entscheidungsmengen, die von Instrumenten mit integrierter KI zum «Lernen» genutzt werden, die in der Gesellschaft bestehenden diskriminierenden Verzerrungen enthalten können.

- Transparenz und Begründung von Entscheidungen: In der Lehre wird darauf hingewiesen, dass das Phänomen der Erteilung von Anweisungen durch automatisierte Systeme in der Arbeitswelt Realität geworden ist.[373] Diese Anweisungen müssen jedoch den rechtlichen Rahmen des Arbeitsrechts einhalten, was bedeutet, dass man wissen muss, auf welcher Grundlage die Entscheidung getroffen wurde, und daher auch, wie die KI-Systeme programmiert sind. Arbeitnehmende, die eine Anweisung anfechten wollen, wissen dies jedoch nicht. Dies gilt auch für jede andere Entscheidung, die von einem KI-System getroffen wird oder auf einem KI-System beruht. Auch dieses Thema wurde weiter oben allgemein behandelt (vgl. Ziff. 4.3.2.3).

---

[371] AIDA PONCE DEL CASTILLO, Le travail à l'ère de l'IA: pourquoi la réglementation est nécessaire pour protéger les travailleurs, etui. 2020, 13; JEREMIAS ADAMS-PRASSL/ HALEFOM ABRAHA/AISLINN KELLY-LYTH et al., Regulating algorithmic management: A blueprint, European Labour Law Journal 2023, 124 ff., 139–140, 143.

[372] Unter anderen ISABELLE WILDHABER, Répercussions de la robotique (Fn. 368), 213 ff. Vgl. auch https://www.rts.ch/info/sciences-tech/13464935-la-place-de-lintelligence-artificielle-dans-le-recrutement.html , https://www.humanrights.ch/fr/nouvelles/discrimination-algorithmique-protection , https://algorithmwatch.ch/de/findhr/ (abgerufen am 28. August 2024).

[373] ISABELLE WILDHABER, Répercussions de la robotique (Fn. 368), 210 f.

6.5.3.4 Anwendbare Bestimmungen im Schweizer Recht

Die Regeln, die zur Lösung dieser Probleme herangezogen werden können, sind unterschiedlicher Natur:

- Zunächst sind die Datenschutzbestimmungen zu nennen. Das DSG gilt für Arbeitsverhältnisse in der Privatwirtschaft. Artikel 328*b* OR regelt diese Frage zusätzlich zu den allgemeinen Bestimmungen speziell für den Arbeitsvertrag. Artikel 328*b* erster Satz OR nennt zwei mögliche Zwecke für die Datenerhebung durch den Arbeitgeber: die Abklärung der Eignung des Arbeitnehmers für die Arbeitsleistung und die Durchführung des Arbeitsvertrages. Diese Bestimmung gehört zum relativ zwingenden Recht. Der zweite Satz verweist auf die Anwendbarkeit des DSG.

  Die Frage, ob der Arbeitgeber Daten zu anderen als den in Artikel 328*b* OR vorgesehenen Zwecken beschaffen darf, war lange Zeit umstritten, insbesondere im Zusammenhang mit der Einwilligung des Arbeitnehmenden, die im DSG einen Rechtfertigungsgrund darstellt (Art. 31 Abs. 1). Während sich ein Grossteil der Lehre einig war, dass die Einwilligung eine Datenbearbeitung zu anderen Zwecken nicht rechtfertigen kann,[374] hat das Bundesgericht diese Möglichkeit kürzlich bejaht:[375] Die beiden in Artikel 328*b* OR genannten Fälle begründen eine Vermutung der Rechtmässigkeit der Datenbearbeitung, wobei sich die Bearbeitung zu anderen Zwecken auf die allgemeinen Rechtfertigungsgründe des DSG stützen muss. Trotz dieser offenen Haltung des Bundesgerichts bestehen in der Lehre starke Vorbehalte gegen die Möglichkeit einer freiwilligen Einwilligung im Sinne von Artikel 6 Absatz 6 DSG im Arbeitsrecht oder in jeder Situation eines strukturellen Ungleichgewichts zwischen den Parteien.[376]

  Die durch KI aufgeworfenen Probleme lassen sich mit den allgemeinen Regeln (Art. 328*b* OR und das DSG) abdecken. Zunächst setzen die Vorschriften über den Zweck der Datenbearbeitung gewisse Grenzen. So ist der Arbeitgeber, der ein KI-System einsetzt, an die in Artikel 328*b* OR festgelegten Zwecke gebunden. Er muss die Beschaffung von Daten zu anderen Zwecken rechtfertigen. Der Zweck muss zudem gemäss Artikel 6 Absatz 3 DSG bestimmt und für die betroffene Person erkennbar sein. Auch die übrigen Grundsätze des DSG müssen eingehalten werden, insbesondere die Verhältnismässigkeit (Art. 6 Abs. 2 DSG). Die Einhaltung dieser Grundsätze führt bei

---

[374] Siehe insbesondere den Bericht des Bundesrates vom 16. November 2016 in Erfüllung des Postulats 12.3166 Meier-Schatz, Rechtliche Folgen der zunehmenden Flexibilisierung des Arbeitsplatzes, Ziff. 7.7.1 und Referenzen, abrufbar unter: https://www.newsd.admin.ch/newsd/message/attachments/80239.pdf (abgerufen am 28. August 2024).

[375] Urteil des BGer, 4A_518/2020, E. 4.2.4.

[376] CÉLIAN HIRSCH, Droit du travail et intelligence artificielle: défis des décisions automatisées pour les employeurs, in: Valérie Défago/Jean-Philippe Dunand/Pascal Mahon/Samantha Posse/David Raedler (Hrsg.), La protection des données dans les relations de travail à la lumière de la nouvelle loi fédérale sur la protection des données, Genf/Zürich 2024, 95 ff., 112 und die dortigen Referenzen; ISABELLE WILDHABER/ISABEL EBERT, Beteiligung der Arbeitnehmenden (Fn. 364), 16.

KI-Systemen zu einem Spannungsfeld, namentlich aufgrund der mit der Entwicklung der Systeme einhergehenden Beschaffung von grossen Datenmengen, der Vielfalt der verfolgten Zwecke oder der Verwendung zu anderen als den ursprünglich vorgesehenen Zwecken. So zeigt die oben genannte Studie, dass die Einführung von KI-Systemen von einigen Arbeitgebern gestoppt wurde, weil diese Systeme die oben genannten Grundsätze nicht einhielten.[377] Darüber hinaus sind verschiedene Bestimmungen des DSG zu erwähnen, die für die Arbeitswelt besonders relevant sind. Erstens kann es sich bei den gesammelten Daten um besonders schützenswerte Personendaten im Sinne von Artikel 5 Buchstabe c DSG handeln, wie z. B. die Angabe einer Gewerkschaftsmitgliedschaft oder Gesundheitsdaten, und/oder sie können als Grundlage für ein Profiling mit hohem Risiko dienen. Für die Beschaffung solcher Daten sowie für das Profiling mit hohem Risiko sind besondere Voraussetzungen zu erfüllen, unter anderem gegebenenfalls die ausdrückliche Einwilligung der betroffenen Person (Art. 6 Abs. 7 Bst. a DSG) und die Pflicht zur Durchführung einer Datenschutz-Folgenabschätzung, wenn besonders schützenswerte Personendaten in grossem Umfang bearbeitet werden (Art. 22 Abs. 1 und 2 Bst. a DSG).

Zu erwähnen ist auch Artikel 21 Absätze 1 und 2 DSG, der sich speziell auf KI-Systeme bezieht und spezifische Informationsrechte und -pflichten im Falle einer automatisierten Einzelentscheidung vorsieht (vgl. Ziff. 4.3.2.3). Es sei daran erinnert, dass der Arbeitgeber grundsätzlich von seinen Pflichten befreit ist, wenn er die ausdrückliche Einwilligung des Arbeitnehmenden einholt (Art. 21 Abs. 3 Bst. b DSG), unter Vorbehalt der oben erwähnten Problematik der Einwilligung im Arbeitsrecht. Der Begriff der automatisierten Einzelentscheidung ist relativ eng gefasst, da er keine Systeme umfasst, die eine Bewertung liefern, die als Grundlage für eine von einem Menschen getroffene Entscheidung dient.[378]

- Die Vorschriften über den Gesundheitsschutz können in mehrfacher Hinsicht zum Tragen kommen. Der Gesundheitsschutz ist vor allem in Artikel 328 OR und in Artikel 6 ArG geregelt. Konkretisiert wird die Pflicht zum Gesundheitsschutz in den Verordnungen zum ArG.

  Von besonderer Bedeutung ist in diesem Zusammenhang Artikel 26 ArGV 3, der in Absatz 1 Überwachungs- oder Kontrollsysteme, die dazu bestimmt sind, das Verhalten der Arbeitnehmenden am Arbeitsplatz zu überwachen, verbietet. Darunter sind alle technischen Systeme zu verstehen, mit denen Daten über das Verhalten der Arbeitnehmenden aufgezeichnet werden können, wie Kameras, Geräte zum Abhören von Telefongesprächen oder zur Überwachung der Tätigkeit am Computer, Ortungssysteme oder

---

[377] ISABELLE WILDHABER/ISABEL EBERT, Beteiligung der Arbeitnehmenden (Fn. 364), 8.

[378] CÉLIAN HIRSCH, Droit du travail et intelligence artificielle (Fn. 376), 104 f., obwohl die Rechtsprechung der EU zu einer breiten Auslegung zu tendieren scheint; ISABELLE WILDHABER/ISABEL EBERT, Beteiligung der Arbeitnehmenden (Fn. 364), 15.

Informatikwerkzeuge mit künstlicher Intelligenz.[379] Solche Überwachungssysteme sind jedoch zulässig, wenn sie zu rechtmässigen Zwecken eingesetzt werden, z. B. zur Gewährleistung der Sicherheit des Betriebsgeländes oder zur Qualitäts- oder Leistungskontrolle, sofern sie dem Grundsatz der Verhältnismässigkeit entsprechen und die betroffenen Arbeitnehmenden darüber informiert wurden.[380]

In diesem Zusammenhang sind noch weitere Pflichten relevant: Artikel 6 Absatz 2 ArG bestimmt, dass der Arbeitgeber die Arbeitnehmenden vor Überbeanspruchung zu schützen hat, und Artikel 2 Absatz 1 ArGV 3 legt fest, dass sich der Schutz auf die physische und psychische Gesundheit bezieht und insbesondere die Vermeidung von übermässig starker Beanspruchung umfasst. Dem Arbeitgeber obliegt eine Informations- und Anleitungspflicht (Art. 5 Abs. 1 ArGV 3) sowie die Pflicht zur Information und Beratung der Arbeitnehmenden oder ihrer Vertreter in Fragen des Gesundheitsschutzes (Art. 48 Abs. 1 Bst. a ArG).

- Das Mitwirkungsgesetz[381] regelt die Information und die Mitsprache der Arbeitnehmenden im Betrieb. Es handelt sich um ein Rahmengesetz, das durch eine Reihe von Einzelgesetzen konkretisiert wird, von denen einige in Artikel 10 aufgeführt sind. Das Gesetz sieht in Betrieben mit mindestens 50 Arbeitnehmenden einen Anspruch auf Vertretung vor (Art. 3). Die Arbeitnehmervertretung wiederum hat Anspruch auf Information über alle Angelegenheiten, deren Kenntnis Voraussetzung für eine ordnungsgemässe Erfüllung ihrer Aufgaben ist (Art. 9 Abs. 1). Über dieses allgemeine Informationsrecht hinaus sind die Mitwirkungsrechte in Spezialgesetzen geregelt. Es ist jedoch kein allgemeines Konsultationsrecht vorgesehen. Gemäss Artikel 10 sind in vier Bereichen besondere Mitwirkungsrechte vorgesehen: Arbeitssicherheit und Arbeitnehmerschutz, Übergang von Betrieben, Massenentlassungen, Anschluss an eine Einrichtung der beruflichen Vorsorge.

- Relevant sind auch die Grenzen des Weisungsrechts: Das Weisungsrecht nach Artikel 321*d* OR muss mit dem zwingenden Recht, insbesondere mit Artikel 328 OR (Persönlichkeitsschutz), dem Vertrag und dem Grundsatz von Treu und Glauben vereinbar sein.[382] Jegliche von einer KI erteilte Weisung muss daher in diesem Rahmen liegen, andernfalls die Arbeitnehmerin oder der Arbeitnehmer deren Befolgung verweigern darf. Insbesondere kann sich der Arbeitgeber nicht auf einen automatisierten Entscheidungsmechanismus berufen, um sich zu rechtfertigen, oder die Verantwortung auf den Systementwickler abwälzen.[383] Für den Arbeitgeber bedeutet dies, dass er in der Lage sein

---

379 SECO, Wegleitung zur ArGV 3 (Fn. 126), Art. 26, 326-2.

380 BGE **130** II 425, E. 4.4; SECO, Wegleitung zu ArgGV 3 (Fn. 126), Art. 26, 326-1.

381 SR **822.14**

382 BGE **132** III 115, E. 5.2.

383 ISABELLE WILDHABER, Die Roboter kommen (Fn. 368), 330 f.

muss, die Rechtmässigkeit der getroffenen Entscheidung nachzuweisen, was voraussetzt, dass er weiss, wie das KI-System funktioniert und programmiert ist, und dass er in der Lage ist, dies dem Arbeitnehmenden zu erläutern.

- Schliesslich ist auch der allgemeine Persönlichkeitsschutz relevant (Art. 328 OR): Jede Entscheidung über die zu verrichtende Arbeit, die auf einer von der KI generierten Bewertung beruht, kann die Persönlichkeit des Arbeitnehmers verletzen, wenn sie nicht auf der objektiven Realität der zu verrichtenden Arbeit beruht. Dies kann auch bei einer Zuweisung, die die Situation des Arbeitnehmers verschlechtert, oder bei einer Sanktion der Fall sein.

6.5.3.5 Würdigung

Die bevorstehende Verabschiedung der Richtlinie über Plattformarbeit wird zu einer unterschiedlichen Regelung im Arbeitsrecht der EU und im schweizerischen Recht führen. Das Schweizer Recht siehe keine Spezialregelungen für das algorithmische Management von Arbeit vor, weder speziell für digitale Plattformen, die Arbeitsleistungen anbieten, noch allgemein für alle Arbeitsverhältnisse.

Die Stellungnahme der Gewerkschaft syndicom und von Algorithmwatch,[384] auf die sich die oben zitierte Studie von WILDHABER/EBERT stützt, zeigt Lücken auf. Diese Positionen werden in der Motion 23.4492 Gysi aufgegriffen. Ihre Forderung zielt hauptsächlich auf eine Stärkung der Mitbestimmung der Arbeitnehmenden.

Einige Forderungen zielen auf eine Stärkung der Mitwirkungsrechte im Allgemeinen ab, ohne speziell auf die aufkommende KI einzugehen. Andere Forderungen beziehen sich speziell auf den Einsatz von KI. Dies ist der Fall bei möglichen Regelungen zur Information und Mitsprache der Arbeitnehmenden bei der Einführung und Nutzung von KI-Systemen durch den Arbeitgeber. Derzeit sind solche allgemeinen Informationsrechte in Artikel 9 Absatz 1 des Mitwirkungsgesetzes an das Vorliegen von Spezialgesetzen geknüpft. Bei der Einführung eines KI-Systems sind nur die Vorschriften über den Gesundheitsschutz relevant (Art. 48 Abs. 1 Bst. a ArbG; vgl. Ziff. 6.5.3.4). Mitwirkungsrechte ergeben sich aus dem geltenden Recht somit allenfalls indirekt bei einer möglichen Gesundheitsbeeinträchtigung. Hinzu kommt, dass Artikel 9 Absatz 1 des Mitwirkungsgesetzes zwar ein Informationsrecht vorsieht, dieses aber sehr allgemein formuliert ist. Die Datenschutzbestimmungen begründen zwar ein Informationsrecht der Arbeitnehmenden, aber nur ein individuelles, und kein kollektives, während KI-Systeme aber Fragen aufwerfen, die alle Arbeitnehmenden betreffen.[385] Diese kollektive Dimension betrifft zwar nicht nur Arbeitsverhältnisse, das Arbeitsrecht enthält aber Bestimmungen zur kollektiven Interessenvertretung der Arbeitnehmenden.

Schliesslich zeigt die Analyse, dass die allgemeinen arbeitsrechtlichen Bestimmungen die Einrichtung und den Einsatz von KI-Systemen durch den Arbeitgeber auf verschiedenen Ebe-

---

[384] Download verfügbar unter https://syndicom.ch/fr/themes/dossier/intelligenceartificielleia/algorithmes-au-travail (abgerufen am 28. August 2024).

[385] In diesem Sinn: ISABELLE WILDHABER/ISABEL EBERT, Beteiligung der Arbeitnehmenden (Fn. 364), 16 f.

nen regeln. So muss sich der Arbeitgeber nacheinander fragen, ob sein System allenfalls aufgrund von Artikel 26 ArGV 3 verboten ist, ob es den Anforderungen des Datenschutzes, insbesondere in Bezug auf die Zweckbindung oder die Verhältnismässigkeit, genügt, oder ob es spezifischeren Verpflichtungen im Zusammenhang mit der Bearbeitung von besonders schützenswerten Daten oder der Erstellung von Profilen mit hohem Risiko untersteht. Darüber hinaus muss er die in Artikel 21 DSG festgelegten Verpflichtungen in Bezug auf automatisierte Einzelentscheidungen einhalten. Schliesslich muss er sicherstellen, dass er die von oder mit Hilfe von KI-Systemen getroffenen Entscheide begründen kann, z. B. um die Rechtmässigkeit von Weisungen an Arbeitnehmende zu überprüfen oder um eine Kündigung zu begründen, wenn ein Arbeitnehmender dies verlangt (Art. 335 Abs. 2 OR). Die derzeitige Rechtslage hat den Vorteil, dass sie relativ viele Grenzen setzt, aber den Nachteil, dass jede einschlägige Norm nur einen Teilaspekt des Problems abdeckt und auf einem anderen Begriff beruht, wie z. B. «Überwachungs- oder Kontrollsystem zur Verhaltensüberwachung», «automatisierte Einzelentscheidung» oder «Profiling mit hohem Risiko». Diese Vielfalt macht die Rechtslage kompliziert und wird dazu führen, dass KI-Systeme unterschiedlichen Regeln unterliegen, je nachdem, ob sie unter eine, mehrere oder keine der in den verschiedenen Vorschriften definierten Kategorien fallen. Darüber hinaus besteht Unsicherheit in Bezug auf die Rechtfertigung der Datenbearbeitung durch Einwilligung gemäss DSG in arbeitsvertraglichen Beziehungen.

EU-Recht: Im Vergleich zur neuen EU-Richtlinie über Plattformarbeit besteht in der Schweiz eine Gesetzeslücke. Ebenso fehlt im Schweizer Recht ein Äquivalent zu den spezifischen Regeln der KI-Verordnung in Bezug auf das Arbeitsverhältnis.

Zwar gewährleisten die allgemeinen arbeits- und datenschutzrechtlichen Bestimmungen hier einen gewissen Arbeitnehmerschutz, jedoch findet das in der EU eingeführte Normensystem, das KI-Systeme von der Konzeption bis zum Einsatz über die gesamte Lebensdauer begleitend regelt, im geltenden Schweizer Recht keine Entsprechung. Auch die spezifischen Bestimmungen über die Mitwirkung der Arbeitnehmenden haben im schweizerischen Recht keine Entsprechung, obwohl die bestehenden Vorschriften über den Gesundheitsschutz und die Sicherheit am Arbeitsplatz einen Teil der Situationen abdecken.

Parlamentarische Vorstösse: Die Motion 23.4492 Gysi und die ihr zugrunde liegenden Dokumente weisen ebenfalls auf Lücken hin. Ein Teil dieser Lücken ist nicht KI-spezifisch. Ein gesondertes Recht auf Information und Mitwirkung bei der Einführung und Anwendung von KI-Systemen könnte aber angesichts des sehr allgemein gehaltenen Mitwirkungsgesetzes sinnvoll sein.

Die Analyse des regulatorischen Bedarfs im Arbeitsrecht muss mit der allgemeinen Analyse des Regelungsbedarfs im Bereich der KI koordiniert werden. Mit der Umsetzung des oben identifizierten übergreifenden Regelungsbedarfs, insbesondere im Fall einer möglichen Ratifizierung der KI-Konvention (vgl. Ziff. 4.6), könnte bereits eine Reihe von Herausforderungen in der Arbeitswelt angegangen werden. Zu denken ist insbesondere an die Grundsätze der Transparenz und Aufsicht (Art. 8), der Gleichstellung und Nichtdiskriminierung (Art. 10) und des Datenschutzes (Art. 11). Der Anwendungsbereich der Konvention auf privatrechtliche Beziehungen stellt jedoch eine Einschränkung dar. Die Konvention ist

nämlich auf Situationen beschränkt, in denen Grundrechte, Demokratie oder Rechtsstaatlichkeit betroffen sind (vgl. Ziff. 4.2.3.1). Eine zweite Einschränkung betrifft die Mitwirkung der Arbeitnehmenden und die damit zusammenhängenden arbeitsrechtlichen Regelungen.

Sollte sich die Schweiz für eine Annäherung an die KI-Verordnung entscheiden und beispielsweise bestimmte problematische Anwendungen im privaten Sektor verbieten, hätte dies auch Auswirkungen auf das Arbeitsrecht. Sollte die Schweiz diesen Weg hingegen nicht gehen, stellt sich die Frage, ob eine Übernahme der spezifischen arbeitsrechtlichen Regelungen sinnvoll wäre.

Ein spezifischer arbeitsrechtlicher Regelungsbedarf scheint daher derzeit nur punktuell zu bestehen. Die Entwicklung des allgemeinen rechtlichen Rahmens sollte jedoch beobachtet werden, um gegebenenfalls mit spezifischen Massnahmen eingreifen zu können.

6.6 Strafrecht

6.6.1 Grundsätzliche Anwendbarkeit

Grundsätzlich ist das schweizerische Strafgesetzbuch (StGB) technologieneutral ausgestaltet und bleibt unabhängig von der technologischen Vorgehensweise oder des spezifischen Instrumentariums der Täterschaft anwendbar.[386] Die materielle Reichweite des Schutzes ist daher grundsätzlich geeignet und darauf ausgerichtet, um den Einsatz unterschiedlicher Technologien durch einen Täter oder einen Täterkreis zu erfassen. Bedient sich die Täterschaft KI-Systemen wie zum Beispiel der Deepfake-Technologie, um ein Delikt gegen die Ehre oder den Privatbereich zu begehen, sind die entsprechenden Straftatbestände (Art. 173 ff. StGB, insbesondere auch der Identitätsmissbrauch gem. Art. 179^{decies} StGB) anwendbar. Verfolgt die Täterschaft durch den Einsatz von KI-Systemen weitergehende Absichten, zum Beispiel das Erlangen eines Vermögensvorteils mittels Täuschung, sind die entsprechenden Straftatbestände aus dem Bereich des Vermögensstrafrechts auch beim Einsatz einer solchen neuen, sich in Entwicklung befindlichen Technologie anwendbar (z. B. Betrug gem. Art. 146 StGB oder die Warenfälschung gem. Art. 155 StGB). Solange die Täterschaft KI-Systeme einsetzt und dabei tatbestandsmässig handelt, bleibt die Verantwortung grundsätzlich beim Menschen oder Täterkreis, der diese Technologie nutzt. Sind die Technologien in bestimmten Konstellationen besonders raffiniert oder nimmt die Qualität der Täuschung zu, wie im Beispiel der erwähnten Deepfakes, kann der Täterschaft im Rahmen der Strafzumessung im Einzelfall durchaus eine besonders arglistige Vorgehensweise oder Täuschungsabsicht vorgeworfen und eine Erhöhung der auszufällenden Strafe in Betracht gezogen werden.

Auch die strafbaren Handlungen gegen die sexuelle Integrität sind grundsätzlich technologieneutral ausgestaltet. Für Deepfakes ist namentlich Artikel 197 StGB (Pornografie) relevant.

[386] Siehe dazu die Stellungnahme des Bundesrats vom 16. August 2023 zur Mo. 23.3563, Deepfakes regulieren, abrufbar unter: https://www.parlament.ch/de/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20233563 (abgerufen am 29. August 2024).

Grundsätzlich sind im Kontext von Artikel 197 Absätze 4 und 5 StGB unter anderem die Herstellung, der Besitz und die Verbreitung von pornographischen Deepfakes, die als «harte Pornografie» gelten, strafbar.[387] Unabhängig von der angewandten Technik gelten sämtliche Handlungen als tatbestandsmässig, die Pornografie hervorbringen, d. h. nicht nur Neuanfertigungen, sondern auch Vervielfältigungen, Reproduktionen oder Bearbeitungen von anderweitig erstellten pornografischen Darstellungen.[388] Das Strafgesetzbuch misst der sexuellen Integrität von Kindern und Jugendlichen in diesem Kontext eine besondere Bedeutung zu und pönalisiert auch Gegenstände und Vorführungen, die *nicht tatsächliche* sexuelle Handlungen mit Minderjährigen zum Inhalt haben (also zum Beispiel mittels Datenverarbeitung generierte Bilder oder Darstellungen; Art. 197 Abs. 4 und 5 StGB). Das heisst, dass auch mittels Herstellung eines KI-Systems frei erfundene pornografische Darstellungen mit Minderjährigen i. d. R. als tatbestandsmässig zu qualifizieren sind. Zudem bilden die Artikel 187 und/oder 197 StGB einen Anknüpfungspunkt für die Strafbarkeit von Straftaten gegen Minderjährige im Ausland (Art. 5 StGB).

Des Weiteren sind bei tatbestandsmässigem Verhalten beispielsweise auch die öffentliche Aufforderung zu Verbrechen oder zur Gewalttätigkeit (Art. 259 StGB) oder eine Diskriminierung und Aufruf zu Hass (Art. 261bis StGB) anwendbar, unabhängig davon, unter Zuhilfenahme welcher Technologie eine Person oder eine Gruppe von Personen herabgesetzt oder diskriminiert wird oder mittels welcher spezifischen Vorgehensweise zu einem Vergehen mit Gewalttätigkeiten aufgefordert wird.

6.6.2 Strafrechtliche Verantwortlichkeit

6.6.2.1 Allgemeine Bemerkungen

Das Schweizer Strafrecht und damit auch die strafrechtliche Verantwortlichkeit fussen grundsätzlich auf dem sogenannten Schuldprinzip.[389] Die Schuld ist als eines der zentralsten Elemente des Schweizer Strafrechts Voraussetzung für eine Bestrafung (Art. 19 Abs. 1 StGB) sowie Grundlage für die Strafzumessung (Art. 47 Abs. 1 StGB). Demzufolge gibt es grundsätzlich keine Strafe ohne Schuld («*nulla poena sine culpa*»). Entsprechend knüpft die strafrechtliche Verantwortlichkeit grundsätzlich an der Schuldfähigkeit des Menschen an und erfasst in diesem Sinne nicht eine Technologie, eine Maschine oder ein System an sich (wie z. B. einen Roboter) als Rechtssubjekt. Die Frage, ob eine autonome Technologie als solche ein schuldfähiges und damit strafrechtliches Rechtssubjekt darstellen kann oder soll und ob

---

[387] Ausführlicher dazu siehe BRIGITTE TAG/MARTIN WYSS, Die strafrechtliche Einordnung von pornografischen Deepfakes, Jusletter 29. April 2024, 1 ff.

[388] ULRICH WEDER, in: Andreas Donatsch (Hrsg.), StGB/JStG, Mit weiteren Erlassen und Kommentar zu den Strafbestimmungen des SVG, BetmG, AIG und OBG, Orell Füssli Kommentar, Zürich 2022, Art. 197 N 18, m.H.; PK StGB-TRECHSEL/BERTOSSA, Art. 197 N 15; BRIGITTE TAG/MARTIN WYSS, Die strafrechtliche Einordnung (Fn. 387), 11.

[389] BGE **123** IV 1, E. 2.

die grundsätzlichen Prinzipien des Strafrechts in diese Richtung verändert und angepasst werden sollen, wurde zwar in der Lehre bereits aufgeworfen, wird jedoch überwiegend abgewiesen.[390] Ob ein solcher Regulierungsansatz im Strafrecht mit einem Mehrwert verbunden wäre, ist durchaus kritisch zu betrachten und muss grundsätzlich in Frage gestellt werden. KI-Systeme sind grundsätzlich nämlich nicht in der Lage, schuldfähig zu handeln. Während es einem menschlichen Täter zum Tatzeitpunkt möglich ist, sich aufgrund seines Willens für oder gegen ein Unrecht zu entscheiden oder ein solches in Kauf zu nehmen, «entscheidet» ein KI-System vielmehr aufgrund von Algorithmen und Datensätzen.[391]

Da ein KI-System folglich nicht selbst strafrechtlich belangt werden kann, stellt sich die Frage nach der Zurechnung und den Übergängen der Verantwortlichkeit im Zusammenhang mit durch KI-Systemen ausgelösten Schädigungen vielmehr zwischen der Ebene des Nutzers und des Herstellers (und ggf. des Halters) von automatisierten (bis hin zu autonomen) KI-Systemen. Dabei würde das Herstellerunternehmen als solches im Kontext des Unternehmensstrafrecht dann relevant, wenn die Schädigung in geschäftlicher Verrichtung im Rahmen des Unternehmenszwecks ausgelöst wurde und keinem Mitarbeiter individuell zugerechnet werden kann (Art. 102 StGB).

Grundsätzlich sind durch Hersteller in Bezug auf KI-Systeme dieselben Sorgfaltspflichten einzuhalten wie bei herkömmlichen technischen Produkten. Je nach Art des KI-Systems, dessen Einsatzbereichs und der daraus folgenden Gefährdung von strafrechtlich geschützten Rechtsgütern können solche Sorgfaltspflichten u.U. allerdings variieren.

Eine Herausforderung beim Einsatz von KI dürfte dabei auch die Beurteilung der strafrechtlichen Verantwortlichkeiten sein, wenn eine Maschine weitgehend automatisiert (bis hin zu autonom) die Kontrolle in einer Situation übernimmt und autonom algorithmenbasierte «Entscheidungen» oder Handlungen vornimmt, welche bisher ausschliesslich durch Menschen ausgeführt wurden (z. B. im Medizinalbereich, im Strassenverkehr oder auch im Rahmen eines modernen Waffeneinsatzes).

Je komplexer ein KI-basiertes System ist, umso schwerer dürfte im Schadensfall feststellbar und durch Gerichte eruierbar sein, wo der Fehler zu lokalisieren ist und wer diesen Fehler schuldhaft verursacht hat (sog. Blackbox-Problematik). Anders als bei herkömmlichen Pro-

---

[390] Zur Möglichkeit einer originären strafrechtlichen Verantwortlichkeit von Technik m. w. V. siehe MONIKA SIMMLER/NORA MARKWALDER, Roboter in der Verantwortung? – Zur Neuauflage der Debatte um den funktionalen Schuldbegriff, Zeitschrift für die gesamte Strafrechtswissenschaft 01/2017, 20 ff., 22; NORA MARKWALDER/MONIKA SIMMLER, Roboterstrafrecht, Zur strafrechtlichen Verantwortlichkeit von Robotern und künstlicher Intelligenz, Aktuelle Juristische Praxis 02/2017, 171 ff., 172.

[391] ANNA LOHMANN, Strafrecht im Zeitalter von Künstlicher Intelligenz, Der Einfluss von autonomen Systemen und KI auf die tradierten strafrechtlichen Verantwortungsstrukturen, Baden-Baden 2021.

dukten gibt es bei KI-Produkten nämlich verschiedene Entwicklungs- und Automatisierungsstufen, die teilweise eine weitgehende «Selbständigkeit» des Produkts ermöglichen. Darin liegt ein Stück weit der Clou solcher Produkte. Das heisst, dass gewisse Technologien ihr datengetriebenes maschinelles Lernen und die Datenauswertung auf Algorithmen basieren, die sich *autonom* fortentwickeln und anpassen können.[392] Die diesbezüglichen Vorgänge in einem KI-basierten (bzw. «smarten») Produkt sind daher nicht immer *ex ante* vorhersehbar und zuweilen nicht einmal *ex post* transparent und nachvollziehbar.[393] Die konkrete Zuordnung von vorhersehbaren Gefahren und Ursachen und damit auch der strafrechtlichen Verantwortung dürfte daher im Einzelnen oft sehr komplex sein und von einer Vielzahl von spezifischen Umständen eines Falles abhängen. So zum Beispiel vom konkreten Verhalten der Nutzerin oder des Nutzers und der Voraussehbarkeit und Vermeidbarkeit von Systemmängeln. Daher bleiben in Bezug auf diese Elemente schwierige Fragen bestehen, die nicht durchwegs pauschal zu beantworten sind. Angesichts der nach wie vor massgeblichen Entwicklungen im Zusammenhang mit KI-Systemen werden sich Abgrenzungsfragen wohl abhängig von der konkreten technologischen Entwicklung, ihres Einsatzes und konkreten Gefährdungsbereichs näher beantworten und herauskristallisieren.

6.6.2.2 Beispiel: Einsatz von KI-Systemen beim automatisierten Fahren

Beim automatisierten Fahren beispielsweise übernehmen Fahrzeuge weitgehend die Steuerung, was im Falle fehlerhafter Objekterkennung und Kollisionsvermeidung durch die Maschine zu Unfällen führen kann. Solche Unfälle, bei denen nicht nur die Fahrerin oder der Fahrer, sondern auch KI-Systeme im Einsatz sind, stellen eine Herausforderung für die strafrechtliche Verantwortungszurechnung dar. Grundsätzlich muss die oder der Fahrzeuglenkende das Fahrzeug gemäss Artikel 31 SVG jederzeit beherrschen. Dieser Grundsatz kann sich im Zusammenhang mit KI-Systemen allerdings als wesentlich komplexer erweisen und wird im Hinblick auf automatisierte Fahrzeuge zumindest teilweise relativiert.[394] Die Zurechenbarkeit eines Schadens hängt jeweils von den Einflussmöglichkeiten der fahrzeuglenkenden Person ab. Ein Verschulden ist folglich zu bejahen, wenn diese Person es trotz Pflicht, Möglichkeit und Zumutbarkeit unterlässt, das Fehlverhalten eines KI-basierten Systems zu unterbrechen oder zu verhindern bzw. dessen Schaden zu mindern oder rückgängig zu machen. Im Bereich des automatisierten Fahrens besteht insgesamt eine Tendenz, dass sich die Strafbarkeit mit fortschreitender Fahrzeugautomatisierung zunehmend (auch) zum Hersteller hin verlagert (sog. «accountability shift»).[395] Das heisst die strafrechtliche Verantwortung

---

[392] MARCO SCHREYER/ANITA GIERBL/T. FLEMMING RUUD/DAMIAN BORTH, Stichprobenauswahl durch Anwendung von Künstlicher Intelligenz, Expert Focus 2/22, 10 ff., 13 ff.; OMLOR SEBASTIAN, Methodik 4.0 für ein KI-Deliktsrecht, in: Zeitschrift zum Innovations- und Technikrecht 04/2020, 221 ff., 221.

[393] LEA BACHMANN, Prozedurale Entlastung von Herstellern «smarter» Produkte im Strafrecht?, Schweizerische Zeitschrift für Strafrecht 140/2022, 77 ff., 78.

[394] So werden bei automatisierten Fahrzeugen der Stufen 3 und 4 die Lenkerinnen und Lenker teilweise von den Aufmerksamkeits- und Beherrschungspflichten befreit. Die oder der Fahrzeuglenkende von Fahrzeugen der Stufe 3 darf sich zwar bei aktiviertem System anderen Dingen widmen, bleibt aber verpflichtet, die Kontrolle über das Fahrzeug wieder zu übernehmen, sobald das Automatisierungssystem dazu auffordert. Siehe dazu BBl **2021** 3026, S. 12.

[395] NADINE ZURKINDEN, Vertrauen in Fahrzeugautomatisierung als strafmindernder Umstand? Anmerkungen zur Urteilsbegründung des Regionalgerichts Emmental-Oberaargau vom 30. Mai 2018, PEN 17 16 DIP, Jusletter 3. Dezember 2018, 1 ff.; CHRISTOF RIEDO/STEFAN MAEDER, Die Benutzung automatisierter Motorfahrzeuge aus strafrechtlicher Sicht, in: Thomas Probst/Franz Werro (Hrsg.), Strassenverkehrsrechts-Tagung vom 21.–22. Juni 2016, 85 ff., 94.

kann, je nach Fall, sowohl beim Nutzer als auch beim Hersteller (und gemäss dem SVG je nachdem auch beim Halter[396]) liegen.

Der Nutzer eines automatisierten Fahrzeugs kann strafrechtlich zur Verantwortung gezogen werden, wenn er seine gebotenen Sorgfaltspflichten verletzt; das heisst, wenn er trotz Kenntnis oder fahrlässig ein automatisiertes Fahrzeug verwendet oder nicht angemessen überwacht und er in der Pflicht war, einen Fehler zu erkennen und zu verhindern.[397]

Der Hersteller von automatisierten Fahrzeugen trägt demgegenüber (grundsätzlich unabhängig vom Automatisierungsgrad) die Last der Produkthaftung[398] bzw. gewisse Sorgfaltspflichten im Hinblick auf das erlaubte Risiko, das mitunter durch technische Reglemente und Typengenehmigungen definiert wird. Grundsätzlich geht man in diesem Zusammenhang davon aus, dass der Hersteller das Risiko der Gefährdung durch sorgfältige Programmierung und Instruktion weitgehend beherrschen kann.[399] Demzufolge kann der Hersteller für Schäden haftbar gemacht werden, die durch Fehler oder Mängel des Produkts entstehen und ihm ein Verschulden an der Fehlfunktion nachgewiesen werden kann. Dabei geht es nicht nur um Sachschäden, sondern auch um mögliche Schädigungen bzw. Verletzungen von Leib und Leben wie z. B. Körperverletzung oder Tod. Wenn eine Fehlfunktion bekannt war und nicht behoben wurde oder ein Hersteller zum Beispiel um die Schwierigkeiten eines systemgesteuerten Fahrzeugs wusste, stehende Objekte zu erkennen, und er den potentiellen Kunden nicht umfassend instruiert hat, und dadurch ein Unfall generiert wird, kann der Hersteller strafrechtlich grundsätzlich im Zusammenhang mit einer Sorgfaltspflichtverletzung zur Verantwortung gezogen werden.[400] Bei KI-Systemen, die, wie es häufig der Fall sein wird, nach dem Inverkehrbringen aufgrund der breiten Nutzung trainiert werden und die Nutzer durch die Auswahl des Lernverfahrens, der Trainingsdaten und der Dauer des Lernprozesses auf die KI einwirken,[401] kann sich entsprechend die Kontroll- und Einflussmöglichkeit der Hersteller verringern.[402] In Bezug auf automatisierte Fahrzeuge ist allerdings darauf hinzuweisen, dass die Software des KI-Systems jeweils Bestandteil der Typengenehmigung ist und diese die durch das Training erlangten Erkenntnisse nicht selber verwerten darf. Vielmehr muss der Hersteller auf der Basis der neuen Erkenntnisse die Software beispielsweise durch Updates weiterentwickeln und unter Umständen eine neue Typengenehmigung erlangen. Ist die Einflussnahme auf oder das Zusammenwirken der Nutzerin oder des Nutzers mit dem KI-System

---

396 Der Halter hat auch bei automatisierten Fahrzeugen eine originäre Halterverantwortlichkeit, d. h. eine strafrechtliche Verantwortlichkeit für die Überwachung der Gefahrenquellen und das Motorfahrzeug (Art. 93, Abs. 2, Bst. b; 95, Abs. 1, Bst. e und 96, Abs. 3, SVG). Der Halter ist damit grundsätzlich u.a. (mit-)verantwortlich für die Betriebssicherheit des Fahrzeugs.

397 Ausführlicher dazu sowie auch zur Frage bis zu welchem Automatisierungsgrad ein Nutzer noch als Fahrzeugführer i.S.d. SVG verstanden werden kann: CHRISTOF RIEDO/STEFAN MAEDER, Die Benutzung automatisierter Motorfahrzeuge (Fn. 395), 91 f.

398 Relevant in diesem Kontext sind das Bundesgesetz über die Produktesicherheit (PrSG, SR 930.11) und das Produktehaftpflichtgesetz (PrHG, SR 221.112.944).

399 Siehe Überblick bei MELINDA F. LOHMANN, Ein zukunftsfähiger Haftungsrahmen für Künstliche Intelligenz, Warum die Schweiz ihr Produkthaftungsrecht aktualisieren muss, Haftung und Versicherung 04/2021, 120.

400 NADINE ZURKINDEN, Vertrauen in Fahrzeugautomatisierung (Fn. 395), 14 f.

401 HERBERT ZECH, Entscheidungen digitaler autonomer Systeme: Empfehlen sich Regelungen zu Verantwortung und Haftung?, Gutachten für den 73. Deutschen Juristentag, München 2020, A 35 ff.; MELINDA F. LOHMANN, Roboter als Wundertüten – eine zivilrechtliche Haftungsanalyse, Aktuelle Juristische Praxis 2017, Sonderheft Roboterrecht, 152 ff., 158.

402 HERBERT ZECH, Entscheidungen (Fn. 401), A 89; MELINDA F. LOHMANN, Roboter (Fn. 401), 158.

allerdings schadensursächlich, können sich diese dennoch auf die strafrechtliche Verantwortlichkeit der Nutzerin oder des Nutzers auswirken.[403]

6.6.3 Herausforderungen im Rahmen der Rechtsdurchsetzung

Die obigen Ausführungen zu den strafrechtlichen Bestimmungen zeigen, dass das geltende materielle Recht aufgrund seiner Technologieneutralität in seiner Grundausrichtung ein grundsätzlich adäquates Instrumentarium bietet, um Strafbarkeiten rund um den Einsatz von KI-Systemen zu erfassen. Das Strafrecht bleibt unabhängig von der technologischen Vorgehensweise oder des spezifischen Instrumentariums einer Täterschaft anwendbar und operabel. Des Weiteren bietet das geltende Strafrecht auch einen Rahmen, der grundsätzliche Sorgfaltspflichten und Verantwortlichkeiten (für Hersteller, Halter oder Nutzer) statuiert. Allerdings bleibt, wie in anderen Rechtsbereichen auch, darauf hinzuweisen, dass es sich bei Straftaten, die unter Zuhilfenahme von KI-Systemen begangen werden oder anderweitig unter Mitwirkung von KI zu strafrechtlich relevanten Schädigungen führen, zuweilen als schwierig erweisen dürfte, die Täterschaft dahinter zu eruieren oder – bei Vorliegen entsprechender Sorgfaltspflichten – die Voraussehbarkeit und Vermeidbarkeit von Systemmängeln nachzuweisen. Die Identifizierung der Täter, die Suche nach Beweisen sowie die klare Eruierung, die Lokalisierung und der Nachweis von Fehlerquellen und Sorgfaltspflichtverletzungen bei automatisierten Produkten wird zweifelsohne auch im strafrechtlichen Kontext eine besondere Herausforderung darstellen. Bei KI-Systemen stellen sich im Hinblick auf die Transparenz und die Nachvollziehbarkeit ihrer automatisierten Entscheidungs- und Lernschritte gerade in der Praxis erhebliche Herausforderungen. Die Schwierigkeiten werden also nicht unbedingt im materiellen Recht selbst begründet, sondern oftmals in der Rechtsdurchsetzung, wenn entweder die Täterschaft, die Technologien und die darauf auffindbaren Beweismittel oder auch die Hersteller sich im Ausland befinden oder *post facto* die genaue Schadensursache nicht konkret ermittelt werden kann (Blackbox-Problematik). Die Schweizer Rechtsdurchsetzung kann ferner auch faktisch an ihre Grenzen kommen, wenn Täter oder Entwickler im Ausland weniger strengen Rechtsvorschriften unterworfen sind oder die komplexen Funktionsweisen und Kausalzusammenhänge von KI-Systemen durch die Staatsanwaltschaft oder Richter schwer zu eruieren, nachzuvollziehen und zu begründen sind. Das heisst, dass einerseits die Blackbox-Problematik von automatisierten Systemen und andererseits die weitgehende Anonymität von Tätern im Ausland (z. B. bei Deepfakes) in der Praxis durchaus Schwierigkeiten mit sich bringen können, die nicht zwingend durch weitergehende materielle Normen angegangen werden können.

Ist die Täterschaft oder ein Hersteller wie beispielhaft angesprochen im Ausland, stellen sich auf der Ebene der Beweisermittlung häufig Schwierigkeiten bezüglich einer funktionierenden, raschen internationalen Zusammenarbeit, und weniger im nationalen Recht selbst. Die Schweiz hat sich folglich an der Ausarbeitung eines Übereinkommens der Vereinten Nationen

[403] NADJA BRAUN BINDER/THOMAS BURRI/MELINDA FLORINA LOHMANN/MONIKA SIMMLER/FLORENT THOUVENIN/KERSTIN NOËLLE VOKINGER, Künstliche Intelligenz: Handlungsbedarf im Schweizer Recht, Jusletter 28. Juni 2021, 21.

über die Cyberkriminalität beteiligt und ist Vertragspartei der entsprechenden Europaratskonvention vom 23. November 2001[404]. In diesen Abkommen geht es im Hinblick auf eine verbesserte Rechtsdurchsetzung grundsätzlich um die Sicherung und die internationale Zusammenarbeit im Rahmen grenzüberschreitender digitaler Kriminalität. Obschon das erwähnte Übereinkommen über Cyberkriminalität des Europarats, das für die Schweiz am 1. Januar 2012 in Kraft getreten ist, diesbezüglich bereits einige dahingehende internationale Instrumente bietet, sind weitere Entwicklungen noch in Bewegung. So stellen auch die e-Evidence-Regeln der EU einen wichtigen Schritt im Bereich des grenzüberschreitenden Zugriffs auf Daten als Beweismittel im Rahmen von Strafverfahren in Europa dar.[405] Ähnliches verfolgt das System des US CLOUD Act.[406] Die Schweiz setzt sich mit den Entwicklungen in der EU und den USA auseinander und wird zu gegebener Zeit unter Berücksichtigung der politischen und rechtlichen Möglichkeiten entscheiden, wie sie damit umgehen will.

6.6.4 Zusammenfassung und Ausblick

Die obigen Ausführungen zeigen, dass das materielle Strafrecht im Hinblick auf die strafrechtlichen Verantwortlichkeiten und Sorgfaltspflichten beim Einsatz von KI einen grundsätzlich adäquaten und im Einzelfall durchaus anwendbaren Rechtsrahmen bietet. Dabei knüpft das geltende Recht sowohl für Vorsatz- als auch für Fahrlässigkeitsdelikte an das weiterhin und auch im Hinblick auf den Umgang mit KI-Systemen relevante Schuldprinzip an.[407] Das heisst, es kommt im Einzelnen jeweils auf das strafrechtlich relevante, vorwerfbare Verschulden der oder des potenziell Verantwortlichen an. Insbesondere bei Vorsatzdelikten, die unter Zuhilfenahme von KI-Technologien wie z. B. von Deepfakes auf strafrechtlich relevante Schädigungen oder Vermögensvorteile abzielen, stellt i. d. R. weniger die materielle Rechtslage, sondern vielmehr die Rechtsdurchsetzung eine Herausforderung dar, da Täter ihre digitalen Spuren oftmals verschleiern und sich im Ausland oder hinter Deepfakes «verstecken» können.

Für die fahrlässige Strafbarkeit im Zusammenhang mit einem KI-System spielt es hingegen primär eine Rolle, ob deren Einsatz, die Herstellung oder die Inverkehrbringung per se als Sorgfaltspflichtwidrigkeit oder als sogenanntes erlaubtes Risiko qualifiziert wird sowie ob und welche Sorgfaltspflichten und Fehlerquellen *ex ante* erkennbar und vermeidbar waren.[408] Damit ist eine verschuldensunabhängige Kausalhaftung im Schadensfall im Strafrecht grundsätzlich ausgeschlossen.[409] Eine Person kann folglich für eine Sorgfaltspflichtverletzung nur

---

[404] SR **0.311.43**.

[405] Siehe den Bericht zur e-Evidence-Vorlage der EU, Gutachten des Bundesamts für Justiz vom 24. Oktober 2023 (abrufbar unter: www.bj.admin.ch > Publikationen & Service > Berichte, Gutachten und Verfügungen > Berichte und Gutachten > Bericht zur e-Evidence-Vorlage der EU, abgerufen am 29. August 2024).

[406] Siehe den Bericht zum US CLOUD Act, Gutachten des Bundesamts für Justiz vom 17. September 2021 (abrufbar unter: www.bj.admin.ch > Publikationen & Service > Berichte, Gutachten und Verfügungen > Berichte und Gutachten > Bericht zum US CLOUD Act, abgerufen am 29. August 2024) sowie das darauf basierende «Data Access Agreement» zwischen den USA und dem Vereinigten Königreich, das im Oktober 2022 in Kraft trat.

[407] MONIKA SIMMLER, Strafrechtliche Verantwortung im Zeitalter autonomer Technik: Vom Individual- zum Unternehmensstrafrecht?, in: Daniel Fink u. a. (Hrsg.), Strafjustiz zwischen künstlicher Intelligenz und prädiktiven Algorithmen, Basel 2021.

[408] BRAUN BINDER u. a., Künstliche Intelligenz: Handlungsbedarf im Schweizer Recht (Fn. 403), 23. Zum erlaubten Risiko beim Einsatz von KI-basierten Systemen bei selbstfahrenden Fahrzeugen: NADINE ZURKINDEN, Strafrecht und selbstfahrende Autos – ein Beitrag zum erlaubten Risiko, recht 2016, 144 ff.

[409] In bestimmten Bereichen, wie z. B. bei der zivilrechtlichen Haftung im Zusammenhang mit automatisierten Fahrzeugen, können beim Einsatz von künstlicher Intelligenz hingegen gewisse Konstellationen einer verschuldensunabhängigen Nutzer- oder Halterhaftung entstehen (siehe

strafrechtlich verantwortlich gemacht werden, wenn diese für sie erkennbar und vermeidbar gewesen wäre. Wo die jeweiligen Grenzen (insb. zwischen der Nutzer-, der Hersteller- und der Halterverantwortlichkeiten) konkret verlaufen, ist im Einzelnen allerdings nicht durchwegs klar[410] bzw. wird es schwierig und unter Umständen auch nicht sinnvoll sein, diese Abgrenzung im Hinblick auf zunehmend selbständige Technologien und Prozesse pauschal festzulegen. Demzufolge bleiben in Bezug auf die konkrete Auslegung und Anwendung dieser Sorgfaltsplichten in Zukunft durchaus Herausforderungen bestehen. In dieser Hinsicht wird insbesondere auch die Rechtsprechung wichtig bleiben.

Die jüngeren Entwicklungen im Bereich des automatisierten Fahrens sprechen grundsätzlich dafür oder deuten zumindest darauf hin, dass sich strafrechtliche Verantwortlichkeiten gerade mit Blick auf die zunehmende Fahrzeugautomatisierung stärker auf den Hersteller oder Personen, die im Hersteller-Unternehmen arbeiten, verschieben.[411] Hier werden naheliegenderweise gerade in Bezug auf weitgehend automatisierte bis hin zu autonomen Systemen wohl weiterhin praktische Abgrenzungsfragen im Hinblick auf die Risiko- und Einflusssphären von Herstellern und Nutzern[412] (und je nach Bereich auch Haltern) sowie in Bezug auf die konkreten Sorgfaltspflichten im Kontext der Unternehmensstrafbarkeit oder der fahrlässigen Mittäterschaft bestehen bleiben.[413] In dieser Hinsicht kann eine Klärung der Sorgfaltspflichten bei der Herstellungssicherheit, der Zulassung und beim Einsatz von KI beispielsweise via konkreter Zulassungs- oder Sicherheitsanforderungen eine bessere Voraussehbarkeit und damit auch Rechtssicherheit schaffen. Damit gemeint sind insbesondere Sorgfaltspflichten für den Einsatz und das Inverkehrbringen von KI-Systemen durch Private oder in Unternehmen sowie im Zusammenhang mit der Produktesicherheit.

Im Zusammenhang mit dem Einsatz von KI-Systemen im Strassenverkehr hat der Bundesrat an seiner Sitzung vom 18. Oktober 2023 die Vernehmlassung über zwei neue Verordnungen eröffnet, mit denen er das automatisierte Fahren regeln will. Die Vernehmlassung dauerte bis

---

Art. 58 SVG). Diese sind jedoch von strafrechtlichen Verantwortlichkeiten zu unterscheiden. Dazu MELINDA F. LOHMANN, Automatisierte Fahrzeuge im Lichte des Schweizer Zulassungs- und Haftungsrechts, Baden-Baden 2016, 211 ff.

410 BRAUN BINDER u. a., Künstliche Intelligenz: Handlungsbedarf im Schweizer Recht (Fn. 403), 23. Zu ähnlichen Fragen, die sich in Bezug auf (autonome) soziale Roboter stellen, siehe MONIKA SIMMLER/OLIVIA ZINGG, Rechtliche Aspekte sozialer Roboter, Gutachten im Auftrag der TA-SWISS, 2011, 19 ff. und 48 ff.

411 Siehe dazu NADINE ZURKINDEN, Vertrauen in Fahrzeugautomatisierung (Fn. 395), 1 ff. mit Verweis auf die Urteilsbegründung des Regionalgerichts Emmental-Oberaargau vom 30. Mai 2018, PEN 17 16 DIP. Zu dieser Verlagerungstendenz auch: MONIKA SIMMLER, Strafrechtliche Verantwortung (Fn. 407).

412 BRAUN BINDER u. a., Künstliche Intelligenz: Handlungsbedarf im Schweizer Recht (Fn. 403), 21.

413 Zu den entsprechenden Herausforderungen MONIKA SIMMLER/OLIVIA ZINGG, Rechtliche Aspekte sozialer Roboter (Fn. 410), 55 ff.; MONIKA SIMMLER, Strafrechtliche Verantwortung (Fn. 407); BRAUN BINDER u. a., Künstliche Intelligenz: Handlungsbedarf im Schweizer Recht (Fn. 403), 23.

zum 2. Februar 2024. Einerseits geht es um die im vorliegenden Kontext relevante Verordnung über das automatisierte Fahren (AFV) und andererseits um die Verordnung über die Finanzhilfen zur Förderung neuartiger Lösungen für den Verkehr auf öffentlichen Strassen (ÖStFV).[414] Dadurch sollen verschiedene Sorgfaltspflichten für den Hersteller, den Händler, den Fahrzeughalter und Nutzer statuiert und die Verletzung besonders wichtiger Pflichten mit Busse bestraft werden. Unter dem Hinweis auf die absehbare dynamische Entwicklung im Bereich des automatisierten Fahrens werden folglich die Auswirkungen dieser Regulierungen vorab abzuwarten und ihr Mehrwert und Zusammenspiel mit internationalen Entwicklungen in diesem Bereich zu evaluieren sein.

Zusammenfassend lässt sich sagen, dass das Schweizer Strafrecht grundsätzlich mit seinem technologieneutralen Ansatz ein generell geeignetes Instrument bietet, um den Einsatz von KI-Systemen insbesondere bei vorsätzlichen Straftaten zu erfassen.

Ein Grossteil der Herausforderungen liegt folglich primär bei (praktischen) Abgrenzungsfragen in Bezug auf die Verantwortlichkeiten, auf der Ebene der Rechtsanwendung sowie generell bei der Rechtsdurchsetzung. Dabei zieht die Konkretisierung von Sorgfaltspflichten im Umgang mit KI-Systemen, wie sie in der erwähnten Verordnung zum automatisierten Fahren (AFV) vorgesehen sind, in eine richtige Richtung, um zumindest in diesem Bereich mehr Klarheit zu schaffen. Ein darüberhinausgehendes Eingreifen des Gesetzgebers schiene daher zum jetzigen Zeitpunkt verfrüht und müsste jeweils auch im Kontext eines entsprechenden Anwendungsbereichs evaluiert werden. Die in der AFV anvisierten Sorgfaltspflichten würden zudem bereits in ihrer aktuellen Form dazu beitragen, die Grundsätze der Transparenz und Rechenschaftspflicht (Art. 8 und 9) umzusetzen, die sich aus der Konvention des Europarats über KI ergeben (siehe die Kapitel 4.3.2.3 und 4.3.2.4), falls die Schweiz dieses Abkommen ratifiziert. Eine erhöhte Transparenz könnte in der Folge wiederum – jeweils im Rahmen des Zumutbaren und Möglichen – zumindest teilweise auch der Blackbox-Problematik entgegenwirken.

Angesichts der sich in Bewegung befindlichen Entwicklungen auf schweizerischer, europäischer und internationaler Ebene werden die Analyse und daraus resultierende Folgerungen im Hinblick auf einen konkreten, über das Bestehende und Anvisierte hinausgehende Regulierungsbedarf weiterhin wichtig sein. So hat beispielsweise der Europarat im Anschluss an seine 78. Plenarsitzung des European Committee on Crime Problems (CDPC) im November 2020 auch eine Regulierung von KI-Systemen im Kontext des Strafrechts anvisiert.[415] Es ist folglich nicht auszuschliessen, dass neben der KI-Konvention auch themenspezifischere Diskussionen und Entwicklungen folgen könnten.

[414] Medienmitteilung vom 18. Oktober 2023, Der Bundesrat will automatisiertes Fahren ermöglichen, abrufbar unter: www.admin.ch > Dokumentation > Medienmitteilungen > Der Bundesrat will automatisiertes Fahren ermöglichen (abgerufen am 29. August 2024).

[415] Siehe Artificial Intelligence and Criminal Law - European Committee on Crime Problems (https://www.coe.int/en/web/cdpc/artificial-intelligence-and-criminal-law, abgerufen am 29. August 2024).

7 Schlussfolgerungen

Die Analyse hat bestätigt, dass das Schweizer Recht bereits Bestimmungen enthält, die auch auf KI-Systeme anwendbar sind. KI entwickelt sich also nicht in einem rechtlichen Vakuum.

An dieser Stelle ist darauf hinzuweisen, dass das BJ nicht das gesamte bestehende Bundesrecht umfassend analysiert hat. Die Analyse enthält daher meist nur Zwischenergebnisse, die noch vertieft werden müssen.

Wie in der Methodik erwähnt und zur Vermeidung der Wiederholung der Analyse im Bericht der interdepartementalen Arbeitsgruppe «Künstliche Intelligenz» von 2019[416] wurde der regulatorische Handlungsbedarf hauptsächlich unter der Annahme einer Ratifizierung der KI-Konvention des Europarats durch die Schweiz beurteilt. Die Analyse stellt zudem rechtliche Überlegungen zur KI-Verordnung im Falle einer Annäherung der Schweizer Gesetzgebung an das EU-Recht. Dies hängt jedoch von politischen Entscheidungen ab. Im Rahmen der Analyse wurden auch weitere ausgewählte Rechtsgebiete untersucht, in denen Entwicklungen und Herausforderungen bezüglich KI bestehen.

Hinsichtlich KI-Konvention scheint das Schweizer Recht im Falle einer Ratifizierung eine teilweise Umsetzung ermöglichen zu können. Es müssten jedoch regulatorische Eingriffe erwogen werden insbesondere zur Stärkung der Transparenz, des Rahmens für das Risiko- und Folgenmanagement sowie der Aufsichtsmechanismen (vgl. Ziff. 4.6). In Bezug auf den Privatsektor hat die Analyse gezeigt, dass sich der Anwendungsbereich der KI-Konvention auf die Fälle beschränkt, in denen eine direkte oder indirekte horizontale Wirkung der Grundrechte zwischen Privaten besteht oder in Zukunft erkannt wird (vgl. Ziff. 4.2.3.1).

Eine Annäherung des schweizerischen Rechts an die KI-Verordnung würde ein umfassendes Eingreifen des Gesetzgebers erfordern. Die KI-Verordnung enthält sehr spezifische Vorschriften für die verschiedenen KI-Systeme, je nach den von ihnen ausgehenden Risiken, sowie zahlreiche Verpflichtungen für die verschiedenen Akteure. Derzeit sind noch viele Fragen offen, insbesondere politische Fragen wie die, ob die Einführung gleichwertiger Rechtsvorschriften einen effizienteren Zugang zum EU-Markt ermöglichen würde, z. B. durch die gegenseitige Anerkennung von KI-Rechtsvorschriften im Rahmen einer Ausweitung des MRA auf den Bereich der KI (vgl. Ziff. 5.4). Weitergehende Analysen, die z. B. die Konturen einer möglichen schweizerischen Gesetzgebung in Anlehnung an die KI-Verordnung präzisieren könnten, wären ebenfalls notwendig, wenn der politische Wille dazu vorhanden ist, in diese Richtung zu gehen. Sie würden sich auch auf die EU-Richtlinien im Bereich des Privatrechts beziehen (vgl. Ziff. 6.3 ff.).

In den untersuchten ausgewählten Rechtsgebieten (vgl. Ziff. 6) hat die Analyse gezeigt, dass sich zwar gewisse Fragen stellen, die sich jedoch mit den geltenden Vorschriften grundsätzlich beantworten lassen. Ein besserer Schutz könnte auch durch die Verabschiedung allgemeiner Normen zur Umsetzung der KI-Konvention erreicht werden, z. B. durch die Stärkung der Transparenz von KI-Systemen.

---

[416] Bericht «Herausforderungen der künstlichen Intelligenz» (Fn. 1).

Um die rechtliche Analyse zu vertiefen und den Handlungsbedarf zu präzisieren, müssen zunächst die politischen Antworten auf die beiden oben genannten Arbeitshypothesen, d. h. die Ratifizierung der KI-Konvention durch die Schweiz und eine allfällige Annäherung an die europäische KI-Verordnung, vorliegen.

Abkürzungsverzeichnis

ABl.	Amtsblatt der Europäischen Union
Abs.	Absatz
AFV	Verordnung über das automatisierte Fahren
AHVG	Bundesgesetz über die Alters- und Hinterlassenenversicherung, SR **831.10**
ArG	Bundesgesetz über die Arbeit in Industrie, Gewerbe und Handel, SR **822.11**
ArGV 3	Verordnung 3 zum Arbeitsgesetz, SR **822.113**
Art.	Artikel
ASTRA	Bundesamt für Strassen
BAKOM	Bundesamt für Kommunikation
BBl	Bundesblatt
BehiG	Bundesgesetz über die Beseitigung von Benachteiligungen von Menschen mit Behinderungen, SR **151.3**
BFS	Bundesamt für Statistik
BGE	Amtliche Sammlung der Entscheidungen des Schweizerischen Bundesgerichts
BGer	Bundesgericht
BGÖ	Bundesgesetz über das Öffentlichkeitsprinzip der Verwaltung, SR **152.3**
BJ	Bundesamt für Justiz
BPR	Bundesgesetz über die politischen Rechte, SR **161.1**
BSK [Gesetz]- AUTOR	Basler Kommentar
Bst.	Buchstabe
BV	Bundesverfassung der Schweizerischen Eidgenossenschaft, SR **101**

CAI	Committee on Artificial Intelligence / Ausschuss für künstliche Intelligenz
CEN	Europäisches Komitee für Normung
CENELEC	Europäisches Komitee für elektrotechnische Normung
CNAI	Kompetenznetzwerk für künstliche Intelligenz
CO	Loi fédérale complétant le Code civil suisse (Livre cinquième : Code des obligations), RS **220**
CP	Code pénal suisse, RS **311.0**
CR [Gesetz]-AUTOR	Commentaire romand
Cst.	Constitution fédérale de la Confédération suisse, RS **101**
DSG	Bundesgesetz über den Datenschutz, SR **235.1**
DSV	Verordnung über den Datenschutz, SR **235.11**
DSGVO	Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung)
DV	Direktion für Völkerrecht
E.	Erwägung
EDA	Eidgenössisches Departement für auswärtige Angelegenheiten
EDI	Eidgenössisches Departement des Innern
EDÖB	Eidgenössischer Datenschutz- und Öffentlichkeitsbeauftragter
EGMR	Europäischer Gerichtshof für Menschenrechte

EJPD	Eidgenössisches Justiz- und Polizeidepartement
EMBAG	Bundesgesetz über den Einsatz elektronischer Mittel zur Erfüllung von Behördenaufgaben, SR **172.019**
EMRK	Europäische Konvention zum Schutze der Menschenrechte und Grundfreiheiten, SR **0.101**
et al.	*et alii*
ETSI	Europäisches Institut für Telekommunikationsnormen
EU	Europäische Union
EuGH	Gerichtshof der Europäischen Union
EUV	Vertrag über die Europäische Union
f.	folgende
ff.	fortfolgende
Fn.	Fussnote (innerhalb dieses Dokuments)
GIG	Bundesgesetz über die Gleichstellung von Frau und Mann, SR **151.1**
GPAI-Modell	general-purpose AI model / KI-Modell mit allgemeinem Verwendungszweck
Hrsg.	Herausgeber
IGE	Eidgenössisches Institut für Geistiges Eigentum
KI	künstliche Intelligenz
KI-Konvention	Rahmenübereinkommen des Europarats über künstliche Intelligenz und Menschenrechte, Demokratie und Rechtsstaatlichkeit
KI-Verordnung	Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU)

	Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz)
Konvention 108+	Aktualisiertes Übereinkommen des Europarats zum Schutz des Menschen bei der automatischen Verarbeitung personenbezogener Daten
LFG	Bundesgesetz über die Luftfahrt, SR **748.0**
LPD	Loi fédérale sur la protection des données, RS **235.1**
Mo.	Motion
MRA	Abkommen zwischen der Schweizerischen Eidgenossenschaft und der Europäischen Gemeinschaft über die gegenseitige Anerkennung von Konformitätsbewertungen, SR **0.946.526.81**
N / Nr	Nummer
NDB	Nachrichtendienst des Bundes
OECD	Organisation for Economic Cooperation and Development / Organisation für wirtschaftliche Zusammenarbeit und Entwicklung
OR	Bundesgesetz betreffend die Ergänzung des Schweizerischen Zivilgesetzbuches (Fünfter Teil: Obligationenrecht), SR **220**
ÖStFV	Verordnung über die Finanzhilfen zur Förderung neuartiger Lösungen für den Verkehr auf öffentlichen Strassen
PatG	Bundesgesetz über die Erfindungspatente, SR **232.14**
PrHG	Bundesgesetz über die Produktehaftpflicht, SR **221.112.944**
Ref.	Referenz

RTVG	Bundesgesetz über Radio und Fernsehen, SR **784.40**
RTVV	Radio und Fernsehverordnung, SR **784.401**
SECO	Staatssekretariat für Wirtschaft
SGK [Gesetz]-AUTOR	St. Galler Kommentar
SiK-N	Sicherheitspolitische Kommission des Nationalrates
SKMR	Schweizerisches Kompetenzzentrum für Menschenrechte
SR	Systematische Rechtssammlung
StGB	Schweizerisches Strafgesetzbuch, SR **311.0**
SVG	Strassenverkehrsgesetz, SR **741.01**
TA-SWISS	Stiftung für Technologiefolgen-Abschätzung
UNCITRAL	United Nations Commission on International Trade Law / Kommission der Vereinten Nationen für internationales Handelsrecht
UNO-Pakt I	Internationaler Pakt über wirtschaftliche, soziale und kulturelle Rechte, SR **0.103.1**
UNO-Pakt II	Internationaler Pakt über bürgerliche und politische Rechte, SR **0.103.2**
URG	Bundesgesetz über das Urheberrecht und verwandte Schutzrechte, SR **231.1**
usw.	und so weiter
UVEK	Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation
VDSZ	Verordnung über Datenschutzzertifizierungen, SR **235.13**
vgl.	vergleiche

VIG	Bundesgesetz über das Vernehmlassungsverfahren, SR **172.061**
VwVG	Bundesgesetz über das Verwaltungsverfahren, SR **172.021**
ZG	Zollgesetz, SR **631.0**
ZGB	Schweizerisches Zivilgesetzbuch, SR **210**
Ziff.	Ziffer
ZPO	Schweizerische Zivilprozessordnung, SR **272**
z. B.	zum Beispiel

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Justiz- und Polizeidepartement EJPD

Bundesamt für Justiz BJ
Direktionsbereich Öffentliches Recht
Fachbereich Rechtsetzungsprojekte I

Anhang 1

Die nachstehende Tabelle soll, soweit möglich, einen Vergleich zwischen der KI-Konvention und der KI-Verordnung ermöglichen. Die Tabelle erhebt keinen Anspruch auf Vollständigkeit. Die Artikel der KI-Konvention sind in der linken Spalte aufgeführt, während die rechte Spalte die Bestimmungen der KI-Verordnung, die einen Zusammenhang mit den betreffenden Artikeln der KI-Konvention haben, enthält. Von der KI-Verordnung werden nur die relevantesten Bestimmungen aufgeführt.

KI-Konvention des Europarats	**KI-Verordnung der Europäischen Union**
Art. 1 Abs. 1 Ziel und Zweck	Art. 1 Abs. 1 Gegenstand
Art. 1 Abs. 2 Abgestufte und differenzierte Massnahmen	Risikobasierter Ansatz (Art. 5, Art. 6 ff., Art. 50, Art. 51 ff.)
Art. 2 Begriffsbestimmung «System künstlicher Intelligenz»	Art. 3 Nr. 1 Die Definition ist im Wesentlichen die gleiche.
Art. 3 Anwendungsbereich	Art. 2 Anwendung im öffentlichen und privaten Sektor. Wie die KI-Konvention gilt auch die KI-Verordnung nicht für KI-Systeme, die ausschliesslich für militärische Zwecke oder Zwecke der nationalen Sicherheit oder Verteidigung eingesetzt werden. Ausschluss von Forschung und Entwicklung. Die KI-Verordnung sieht noch andere Ausnahmen vor.
Art. 4 Schutz der Menschenrechte	Art. 1 Abs. 1 Eines der Ziele der KI-Verordnung ist es, das Funktionieren des EU-Binnenmarktes in Bezug auf KI-Produkte zu verbessern und gleichzeitig den Schutz der Grundrechte zu gewährleisten. Sehr wenige individuelle Rechte, aber Ziel des Grundrechtsschutzes durch Festlegung von Produktsicherheitsregeln. Art. 77 Befugnisse der für den Schutz der Grundrechte zuständigen Behörden

Art. 5 Integrität demokratischer Prozesse und Achtung der Rechtsstaatlichkeit	Art. 1 Abs. 1 Eines der Ziele der KI-Verordnung ist es, das Funktionieren des EU-Binnenmarktes in Bezug auf KI-Produkte zu verbessern und gleichzeitig den Schutz der Grundrechte zu gewährleisten. Sehr wenige individuelle Rechte, aber Ziel des Grundrechtsschutzes durch Festlegung von Produktsicherheitsregeln.
Art. 6 Allgemeiner Ansatz	Nicht relevant
Art. 7 Menschenwürde und individuelle Autonomie	Art. 5 Verbotene Praktiken im KI-Bereich
Art. 8 Transparenz und Aufsicht	Art. 8 ff. Anforderungen an Hochrisiko-KI-Systeme, siehe insbesondere Art. 11 Technische Dokumentation, Art. 12 Aufzeichnungspflichten, Art. 13 Transparenz und Bereitstellung von Informationen für die Betreiber Art. 26 Pflichten der Betreiber (z. B. Informationspflicht). Art. 50 Transparenzpflichten Art. 49 und 71 Registrierung und Datenbank Art. 86 Recht auf Erläuterung der Entscheidungsfindung im Einzelfall
Art. 9 Rechenschaftspflicht und Verantwortung	Pflichtenkatalog für die verschiedenen Akteure in der Lieferkette von KI-Systemen (Anbieter, Betreiber usw.). Art. 17 Abs. 1, Bst. m Rechenschaftsrahmen für die interne Verantwortlichkeit Art. 99 ff. Sanktionen
Art. 10 Gleichstellung und Nichtdiskriminierung	Art. 10 Daten und Daten-Governance, insb. Art. 10 Abs. 2 Bst. f Art. 15 Abs. 4 Genauigkeit, Robustheit und Cybersicherheit
Art. 11 Privatsphäre und Schutz personenbezogener Daten	Art. 10 Daten und Daten-Governance Art. 15 Abs. 4 Genauigkeit, Robustheit und Cybersicherheit
Art. 12 Zuverlässigkeit	Art. 14 Menschliche Aufsicht Art. 15 Genauigkeit, Robustheit und Cybersicherheit

	Art. 40 ff. Harmonisierte Normen Art. 56 ff. Praxisleitfäden Art. 95 Verhaltenskodizes
Art. 13 Sichere Innovation	Art. 57 ff Massnahmen zur Innovationsförderung
Art. 14 Rechtsmittel	Art. 85 Beschwerde Art. 86 Recht auf Erläuterung der Entscheidungsfindung im Einzelfall
Art. 15 Verfahrensgarantien	Art. 26 Pflichten der Betreiber (z. B. Informationspflicht) Art. 50 Transparenzpflichten Art. 86 Recht auf Erläuterung der Entscheidungsfindung im Einzelfall
Art. 16 Rahmen für das Risiko- und Folgenmanagement	Art. 9 Risikomanagementsystem Art. 11 Technische Dokumentation Art. 12 Aufzeichnungspflichten Art. 16 Qualitätsmanagementsystem Art. 27 Grundrechte-Folgenabschätzung für Hochrisiko-KI-Systeme Art. 5 Verbotene Praktiken
Art. 17 Nichtdiskriminierung	Nicht relevant, betrifft die Umsetzung der KI-Konvention
Art. 18 Rechte von Menschen mit Behinderungen und von Kindern	Nicht relevant, betrifft die Umsetzung der KI-Konvention
Art. 19 Öffentliche Konsultation	Art. 67 Beratungsforum Beteiligung aller relevanten Interessenträger an der Ausarbeitung harmonisierter Normen (Art. 40 ff.), Praxisleitfäden (Art. 56), Verhaltenskodizes (Art. 95) usw.
Art. 20 Digitale Kompetenzen und Fähigkeiten	Art. 4 KI-Kompetenz Art. 13 Bereitstellung von Informationen für die Betreiber
Art. 21 Schutz der bestehenden Menschenrechte	Die KI-Verordnung stellt an mehreren Stellen klar, dass sie zusätzlich zum bestehenden Rechtsrahmen gilt und diesen nicht einschränken soll. Dies gilt insbesondere für den Datenschutz (vgl. E. 10 der Verordnung).

Art. 22 Umfassenderer Schutz	Nicht relevant
Art. 23 Konferenz der Vertragsparteien	Nicht relevant
Art. 24 Berichterstattungspflicht	Nicht relevant
Art. 25 Internationale Zusammenarbeit	Art. 57 Abs. 13 Grenzüberschreitende Zusammenarbeit Reallabore
Art. 26 Wirksame Aufsichtsmechanismen	Art. 28 ff. Notifizierende Behörden und notifizierte Stelle Art. 40 ff. Normen und Konformitätsbewertung Art. 64 ff. Governance Art. 74 ff. Marktüberwachung Art. 99 ff. Sanktionen
Art. 27 ff. Schlussbestimmungen	Nicht relevant