www.jusletter.ch

David Rosenthal

Der EU AI Act – Verordnung über künstliche Intelligenz

Der EU AI Act bewegt die Gemüter: Kritiker sehen ihn als Regulierungsmonster, die Macher loben ihn als Leuchtturm der globalen KI-Regulierung, der die Innovation fördern wird. Manche Unternehmen haben schon mit seiner Umsetzung begonnen, obwohl die meisten seiner Regelungen noch einige Zeit nicht gelten werden. Immerhin kann er auch Unternehmen in der Schweiz betreffen. Eine erste Kommentierung.

Beitragsart: Beiträge
Rechtsgebiete: Datenschutz, Urheberrecht, Informatik und Recht, Europarecht und Internationales Recht, Verwaltungsrecht

Zitiervorschlag: David Rosenthal, Der EU AI Act – Verordnung über künstliche Intelligenz, in: Jusletter 5. August 2024

Inhaltsübersicht

[1] Der «EU AI Act» (**AIA**) oder zu Deutsch die «Verordnung über künstliche Intelligenz» der Europäischen Union (Verordnung (EU) 2024/1689) ist am 12. Juli 2024 im Europäischen Amtsblatt publiziert worden und am 1. August 2024 in Kraft getreten.[1] Zuweilen ist auch vom «KI-Gesetz» der EU die Rede. Es wird bis zu 36 Monate dauern, bis all seine Bestimmungen gelten. Es zeichnet sich allerdings schon jetzt ab, dass er weit über die Europäische Union hinaus Wirkung entfalten wird. Wer einen Service oder ein Produkt mit künstlicher Intelligenz (KI) einsetzt, wird sich daran orientieren, jedenfalls wenn er auf Abnehmer in der EU oder auf dem Weltmarkt hofft. Dem ging ein mehrjähriges Tauziehen voraus, das zuletzt im Februar 2024 in einem vorläufig finalen Text endete.

1. Regelungskonzept

[2] Der AIA ist – wie die EU-Datenschutz-Grundverordnung (**DSGVO**) – als Verordnung direkt anwendbares EU-Recht.[2] Die Durchsetzung erfolgt teilweise in den Mitgliedstaaten (die jeweils eine Marktaufsichtsbehörde sowie eine Behörde zur Regelung der Verfahren der Stellen zur Beurteilung der Produktekonformität bezeichnen müssen), teilweise zentral (so ist die EU-Kommission für General Purpose AI Models zuständig).[3]

[3] Wie bei der DSGVO gibt es zentrale Institutionen, z.B. das «European Artificial Intelligence Board» (**EAIB**), für welches jeder Mitgliedstaat einen Delegierten bestellt.[4] Es soll ferner ein «Scientific Panel of Independent Experts» geschaffen werden.[5] Das zentrale «AI Office» soll als Teil der Europäischen Kommission die Marktaufsichtsbehörden unterstützen, aber auch Templates zur Verfügung stellen und beispielsweise bei der Einhaltung der Vorgaben im Zusammenhang

1 https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=OJ:L_202401689.

2 Vgl. zu den Schnittstellen zwischen dem AI Act und der DSGVO Kathrin Schürmann/Björn Möller, «Die Zukunft gestalten: KI-VO im Fokus des Datenschutz- und Risikomanagements», in: Datenschutz-Berater, 12/2023, S. 327 f.

3 Siehe Rn. 75 ff.

4 Art. 65 f. AIA; eine Übersicht der Aufgaben, welche die Mitgliedstaaten haben, findet sich in Kai Zenner, «The EU AI Act: responsibilities of the Member States», 23. Juli 2024, https://www.kaizenner.eu/_files/ugd/88fe02_90c3a8e3c8494a5aa00a8e13c8d3e06b.pdf, archiviert unter https://perma.cc/YN7G-FKAU.

5 Art. 68 AIA.

mit Allzweck-KI-Modellen mitwirken.[6] Der Europäischen Kommission kommen auch unmittelbar Aufsichtsaufgaben zu.[7]

[4] Ob und wann der AIA nicht nur für EU-Mitgliedstaaten, sondern auch für die EWR-Staaten gelten wird, ist noch nicht klar. Der AIA gilt für private Organisationen wie auch für Behörden, welche die jeweiligen Voraussetzungen erfüllen.[8] Ausnahmen gelten für die Bereiche Militär, Verteidigung und nationale Sicherheit, da diese den Mitgliedstaaten noch weitgehend exklusiv vorbehalten sind.[9]

[5] Der AIA ist zwar lang (ca. 140 Seiten mit den Erwägungen und Anhängen) und – wie fast jeder EU-Erlass – mühsam in der Lektüre, aber nicht besonders kompliziert. Kurz:

1. Es wird zwischen «AI Systems»[10] (**AIS**, KI-Systeme) und «General Purpose AI Models»[11] (**GPAIM**, was mit «Allzweck-KI-Modellen» übersetzt werden könnte – gemeint sind primär Modelle für generative KI wie etwa «GPT» oder «Llama») unterschieden.
2. Es wird der persönliche und örtliche Geltungsbereich definiert, wobei vor allem unterschieden wird zwischen Anbietern von AIS und GPAIM («Provider») und Betreibern der AIS («Deployer»).[12]
3. Einige wenige KI-Anwendungen werden verboten.[13]
4. Einige KI-Anwendungen werden als «hohes Risiko» definiert; sie werden über die betreffenden AIS reguliert («High-Risk AI System», **HRAIS**): Für ein HRAIS müssen gewisse Anforderungen erfüllt sein (z.B. Qualitäts- und Risikomanagement, Dokumentation, Konformitätserklärung), um deren Einhaltung sich vor allem die Provider kümmern müssen; einige wenige Pflichten werden auch Deployern auferlegt.[14] Es wird geschätzt, dass nur etwa fünf bis zehn Prozent aller AIS in diese Kategorie fallen werden; wir gehen davon aus, dass es in den meisten Unternehmen noch weniger sein werden, da die meisten Unternehmen einen Bogen um HRAIS machen werden – jedenfalls, in der Rolle eines Anbieters.[15]
5. Für alle anderen AIS werden einige allgemeinere Pflichten sowohl für Provider wie auch für

6 Art. 53 f. AIA, Art. 64 AIA, Art. 62 Abs. 3 lit. a AIA; siehe dazu Kai Zenner, «The EU AI Act: responsibilities of the Commission», 10. Juli 2024, https://www.kaizenner.eu/_files/ugd/88fe02_0e02577d0c314b61b17a589476d52066.pdf, archiviert unter https://perma.cc/CH3U-GVL3.

7 Bspw. Art. 37 Abs. 1 AIA, Art. 53 Abs. 4 AIA.

8 Art. 2 Abs. 1 lit. a–f AIA i.V.m. Art. 3 AIA; er erfasst diese Stellen aufgrund ihrer Funktion.

9 Art. 2 Abs. 3 AIA.

10 Art. 3 Ziff. 1 AIA.

11 Art. 3 Ziff. 63 AIA.

12 Art. 2 AIA; siehe Rn. 20 ff. und Rn. 29 ff.

13 Art. 5 AIA; siehe Rn. 42 ff.

14 Kapitel III AIA; siehe Rn. 51 ff.: vgl. hierzu kritisch Angela Müller, «Der Artificial Intelligence Act der EU: Ein risikobasierter Ansatz zur Regulierung von Künstlicher Intelligenz», Zeitschrift für Europarecht (EuZ) 1/2022, S. 18 (der Beitrag bezieht sich noch auf eine frühere Version des AIA).

15 Vgl. zur Kategorisierung anhand des Risikolevels Angela Müller (Fn. 14), S. 15 f.

Deployer definiert, primär zur Transparenz.[16]
6. Bei den GPAIM wird unterschieden zwischen:[17]

- gar nicht erfasst (sehr wenige Fälle),
- «normal» (mit einigen Pflichten für die Provider) und
- «systemische Risiken» (mit zusätzlichen Pflichten für die Provider).

7. Die Europäische Kommission führt ein zentrales Register der HRAIS, und es gibt diverse Regelungen (inklusive Meldepflichten), um Zwischenfälle im Zusammenhang mit HRAIS im Blick zu behalten und darauf bei Bedarf zu reagieren.[18]
8. Mit der Durchsetzung sind diverse nationale und EU-weite Behörden beschäftigt. Es entsteht ein kompliziertes Geflecht an Zuständigkeiten und Abstimmungsverfahren.[19]
9. Es sind begleitende Instrumente vorgesehen, wie etwa

- «regulatory sandboxes»[20] (Recht, bei der Entwicklung neuer AIS für frühe Rechtssicherheit Aufsichtsbehörden einbeziehen zu dürfen),
- Regelungen für das Testen von AIS in der «echten Welt»[21],
- Verhaltenskodizes und
- diverse Bestimmungen zur Schaffung von Standards, Benchmarks und Templates.

Wie relevant sie in der Praxis wirklich sein werden, wird sich zeigen. Den grössten praktischen Nutzen dürften die Templates haben, welche insbesondere das AI Office erstellen soll. Je nachdem, wie praxisnah und detailliert sie ausfallen, werden sie die Compliance der Unternehmen entweder erschweren oder erleichtern. Die «regulatory sandboxes»[22] wiederum dürften eher ein Schlag ins Wasser sein: Es ist nicht wirklich zu erkennen, wie sie die Innovation fördern sollten, auch wenn das ihr erklärtes Ziel ist. Es entspricht allgemeiner Erfahrung, dass ein frühzeitiger Einbezug von Aufsichtsbehörden bei neuen Techniken die Innovation leider oft behindert denn fördert, weil diese naturgemäss selbst über weniger technische und praktische Erfahrungen in den relevanten Fragen verfügen und die Tendenz haben, bei der Auslegung von gesetzlichen Vorgaben strengere Massstäbe als vom Gesetzgeber angedacht anzuwenden. Dies hat die Umsetzung der DSGVO bewiesen; bis heute versuchen zahlreiche EU-Aufsichtsbehörden den Datenschutz

16 Art. 50 AIA; Ausführungen zur Gewährleistung von Transparenz beim Einsatz von KI bietet der Blog-Beitrag «Wie Unternehmen beim Einsatz von KI Transparenz gewährleisten können» von David Rosenthal, abrufbar unter https://www.vischer.com/know-how/blog/teil-16-wie-unternehmen-beim-einsatz-von-ki-transparenz-gewaehrleisten-koennen/ (Stand: 22. Juli 2024); vgl. ausführlich zum Transparenzbegriff Bettina Blawert, «Transparenz nach der DSGVO und der KI-VO-E – Ein Rechtsvergleich mit Empfehlungen zur Umsetzung», in: Datenschutz-Berater, 04/2023, S. 115–118; Ebenfalls ausführlich zur Rolle der Transparenz im AI Act Daria Bohatchuk/Alfred Früh, «Transparenz im Fokus der Europäischen KI-Verordnung», in: Jusletter 12. Februar 2024, S. 3 ff.

17 Art. 51 AIA.

18 Art. 49 Abs. 1 AIA i.V.m. Art. 71 Abs. 1 AIA.

19 Siehe Rn. 75 ff.

20 Art. 57 ff. AIA.

21 Art. 60 AIA.

22 Vgl. dazu auch The Organization for Economic Co-operation and Development (OECD), «Regulatory Sandboxes in Artificial Intelligence», in: OECD Digital Economy Papers, Juli 2023, No. 356, S. 14 ff. https://doi.org/10.1787/8f80a0e6-en.

deutlich über das gesetzliche Niveau hinaus auszubauen. Es zeichnet sich ab, dass manche von ihnen auch für die Umsetzung des AIA zuständig sein werden. Insofern wird der AIA nicht wie von der Kommission erhofft oder propagiert die Innovation im Bereich KI fördern, sondern sie in erster Linie verteuern und erschweren und damit die Markteintrittshürden für kleine Unternehmen erhöhen.[23] Positiv ist an der Regulierung des AIA aus dem Blickwinkel der Innovation primär, dass sie für gewisse Rechtssicherheit sorgt.

10. Es gibt einige Bestimmungen zur Durchsetzung, die Untersuchungs- und Eingriffsbefugnisse durch die Aufsichtsbehörden, Administrativ-Bussen (meist etwas geringer als bei der DSGVO) und ein Auskunftsrecht für Personen in der EU, über die mit Hilfe einer KI Entscheidungen getroffen wurden.[24]

11. Bestimmungen zum Inkrafttreten und ganz wenige Übergangsbestimmungen.[25]

[6] GPAIM sind auffällig milder und weniger eingehend reguliert als AIS und zudem separat abgehandelt, was unter anderem dem Umstand geschuldet ist, dass sie erst im Laufe der Beratungen als Regulierungsobjekt aufgenommen wurden; der ursprüngliche Entwurf des AIA geht auf Mitte 2021 zurück, also noch vor dem Hype um «ChatGPT» & Co. und den diesen Anwendungen zugrunde liegenden grossen Sprachmodellen oder «Large Language Models» (**LLM**). Das GPT LLM ist ein klassisches Beispiel für ein GPAIM.

[7] Der AIA ist primär eine Marktzugangs- und Produkteregulierung, wie wir sie bei diversen Produkten mit erhöhten Risiken (z.B. Medtech) kennen. Sie grenzt den AIA zur DSGVO ab, welche primär Verhalten (Bearbeiten von Personendaten) und Betroffenenrechte reguliert. Es gibt zwar einige wenige allgemeine Verhaltensgrundsätze für den Einsatz von KI und ein Betroffenenrecht, aber sie sind punktuell auf bestimmte Anwendungsfälle ausgerichtet; auch die «verbotenen» KI-Anwendungen sind im Grunde sehr eng definiert. Es wird vor allem reguliert, welche Begleitmassnahmen für als besonders risikoreich beurteilte AIS getroffen werden müssen, und es wird geregelt, wie sie in Verkehr gebracht, eingesetzt und überwacht werden müssen.

[8] Manche HRAIS werden dabei AIS sein, die Teil bereits regulierter Produkte sind, in welchen Fällen viele der vom AIA vorgesehenen Pflichten in ähnlicher Weise schon gelten; der AIA verweist auch regelmässig auf diese und sieht eine kombinierte Umsetzung vor (etwa was das Risiko- und Qualitätsmanagement, die Dokumentation und die Konformitätserklärungen betrifft, aber auch in Bezug auf die behördliche Aufsicht). Offensichtlich war der Gesetzgeber bemüht, die Bürokratie nicht allzu sehr ausufern zu lassen; ob ihm dies gelungen ist, erscheint allerdings fraglich. Immerhin ist durchgängig erkennbar, dass ein risikobasierter Ansatz gelten soll.[26] Viele der Vorgaben sind eher generisch gehalten, wodurch in der Praxis Spielraum bei der Umsetzung besteht – soweit übereifrige Aufsichtsbehörden hier nicht wie bei der DSGVO ihr Unwesen treiben werden.

[9] Der AIA hält immer wieder fest, dass er zusätzlich zum bestehenden Recht gilt, d.h. dieses in keiner Weise einschränken will.[27] Das betrifft insbesondere die DSGVO.[28] Der AIA stellt für

---

23 Ähnlich Ruth Fulterer, «Fachleute aus der ganzen Welt testen in St. Gallen KI-Regeln», in: NZZ Nr. 179, August 2023, S. 23, m.w.H.

24 Siehe Rn. 60 und Rn. 75 ff.

25 Siehe Rn. 79 ff.

26 Vgl. Angela Müller (Fn. 14), S. 7 ff.

27 Bspw. Art. 26 Abs. 10 UAbs. 4 AIA.

28 Art. 2 Ziff. 7 AIA, Erw. 10 AIA.

sich auch keine Rechtsgrundlage für die Bearbeitung von Personendaten dar, mit einigen wenigen Ausnahmen – namentlich den Fall, in welchem es für Tests von HRAIS (aber nicht anderen AIS) hinsichtlich eines möglichen «Bias» unbedingt erforderlich (sinnvoll genügt nicht) ist, besondere Kategorien von Personendaten zu bearbeiten.[29] In der Sache selbst enthält er kaum Regelungen im Bereich des Datenschutzes; es fällt aber auf, dass mit dem EU Data Protection Supervisor einer Datenschutzbehörde die Aufgabe der KI-Marktaufsicht über die Institutionen der EU zukommen soll. Inwiefern die EU-Mitgliedstaaten die KI-Marktaufsicht ebenfalls ihren bereits bestehenden Datenschutzbehörden anvertrauen werden, wird sich zeigen. Wir schätzen, dass dies in ungefähr der Hälfte der Fälle so sein wird.

[10] Was ebenfalls auffällt, ist, dass der Gesetzgeber offenbar einige Mühe hatte, den staatlichen Einsatz von AIS im Bereich der Strafverfolgung und ganz speziell bei Systemen zur biometrischen Fern-Identifizierung von Menschen in der Öffentlichkeit zu regeln (also z.B. anhand ihres Gesichts oder Gangs über Kameras in der Öffentlichkeit).[30] Der Einsatz dieser Systeme ist in gewissen Fällen erlaubt und im AIA viel ausführlicher geregelt als alle anderen Anwendungen (wobei zu beachten ist, dass der Bereich der Wahrung der nationalen Sicherheit von vornherein vom Geltungsbereich des AIA ausgenommen ist).[31] Für Unternehmen und betroffene Personen in der Schweiz wird dies weniger relevant sein. Darum gehen wir hier nicht näher auf diese und andere staatliche Anwendungen ein.

2. Geltungsbereich: Was ist ein KI-System

[11] Der Geltungsbereich des AIA ist leider in verschiedener Hinsicht alles andere als klar – und er ist überaus breit definiert.[32]

[12] Das beginnt bereits mit der Definition von AIS.[33] Sie enthält fünf Elemente, wovon drei auf fast jede IT-Anwendung zutreffen, nämlich vereinfacht gesagt: dass ein System (i) maschinenbasiert ist, (ii) aus einem Input abgeleitet wird, wie ein Output zu erzeugen ist, und (iii) dass dieser Output etwas ausserhalb der Anwendung bewirken kann. Das trifft wohl auf jede klassische Tabellenkalkulation oder Bildbearbeitungssoftware zu, da auch sie aus Input (Zahlen und Formeln, Bilder) einen Output generieren (Ergebnis der Berechnungen, mit Filtern bearbeitete Bilder) und diese etwas bewirken können. Auch jeder Chip hat Kanäle für Input und solche für Output. Das Kriterium des möglichen Einflusses auf die Umwelt erscheint ebenfalls uferlos, weil ein IT-System, das keinen Einfluss auf seine Umwelt hat, sinnfrei sein wird.

[13] Hinzu kommen als weitere Kriterien: Dass (iv) das System sich auch nach seiner Implementierung anpassen kann (es ist ausdrücklich die Rede von «sein kann» bzw. «may» – die Lernfähigkeit eines Systems nach seiner Inbetriebnahme ist nach unserem Verständnis also kein zwingendes Kriterium; wäre es zwingend, wären sehr viele KI-Anwendungen, die wie «ChatGPT» klar erfasst sein sollen, keine AIS mehr, weil ihre Modelle nicht laufend angepasst werden, sondern

29 Art. 10 Abs. 5 AIA; für einen Vergleich der Schweizer Regelungen mit dem AI Act in Bezug auf Personendaten siehe auch Rayan Houdrouge/Kathryn Kruglak, «Are Swiss data protection rules ready for AI?», in: Jusletter 27. November 2023, S. 17 ff.

30 Ähnlich Angela Müller (Fn. 14), S. 17 f.

31 Art. 26 Abs. 10 AIA.

32 Vgl. hierzu auch Angela Müller (Fn. 14), S. 14 f.

33 Art. 3 Ziff. 1 AIA.

den Benutzern quasi im «eingefrorenen» Zustand übergeben werden). Und dass das System (v) so ausgestaltet ist, dass es sich mehr oder weniger autonom verhält.

[14] Dieses letzte Element der Definition, also die mindestens teilweise Autonomie, erscheint somit als das einzige wirklich relevante Unterscheidungsmerkmal, das KI-Systeme von allen anderen Systemen unterscheidet. Allerdings ist auch nicht vollständig klar, was «Autonomie» bedeutet. Im Kern soll es wohl um eine Abgrenzung zu Systemen gehen, deren Output aus dem Input vollständig nach von Menschen formulierten Regeln erzeugt wird, also einem vollständig (d.h. statisch oder deterministisch) ausprogrammierten System («Wenn-Dann-Systeme») – im Gegensatz zu Systemen, die beispielsweise eine Mustererkennung auf Basis eines Trainings betreiben. Es ist nicht mehr eine statische bzw. deterministische Programmierung, die allein bestimmt, welcher Output aus welchem Input resultiert, sondern ein Vorgang des maschinellen Lernens; die Entscheidlogik stammt nur indirekt vom Menschen, welcher der Maschine erklärt, wie sie lernen und basierend auf dem Gelernten im Einzelfall selbst entscheiden muss; es wird also nicht jeder Entscheid vom Menschen vorgegeben. In diese Richtung äussern sich auch die Erwägungen. Während die meisten bei deterministischem Code an solchen von menschlichen Programmierern denken, kann dieser allerdings auch von einer KI programmiert sein. Ein von einer KI deterministisch programmiertes System ist somit kein AIS, weil ihm die Autonomie fehlt. Das führt zur spannenden Frage, wie auf diese Weise die Regelungen des AIA umgangen werden könnten, indem z.B. ein AIS beauftragt wird, basierend auf seinem Wissen ein System zu entwickeln, das zwar über eine deterministische Programmierung verfügt, aber dennoch so komplex ist, dass wir Menschen seine Funktion bzw. Entscheidungslogik nicht mehr verstehen (wir gehen davon aus, dass sich die Verfasser des AIA über solche Dinge keine Gedanken gemacht haben; hätten sie es getan, hätten sie auch dies ausgeschlossen). Vorderhand werden wir aber davon ausgehen müssen, dass die Anwendbarkeit des AIA verhindert werden kann, indem nur Systeme eingesetzt werden, die nicht autonom im vorstehenden Sinn agieren, auch wenn sie zum selben Ergebnis führen oder ebenso problematische Dinge tun wie ein AIS.

[15] Vereinfacht gesagt ist also jedes IT-System, welches seine Funktion (genauer: wie es aus einem bestimmten Input einen bestimmten Output erzeugt) nicht nur auf Basis einer Vorprogrammierung erfüllt, sondern mindestens teilweise auf Basis eines Trainings, ein AIS. Hiermit schliesst sich auch der Kreis jener, die das «Schliessen» von einem Input auf einen Output («ableiten», «*inference*») als das wesentliche Kriterium erachten;[34] es ist dasselbe Verständnis, nur aus einem anderen Blickwinkel. Wir erachten nicht die Tatsache der Ableitung als entscheidend, sondern die Art und Weise, wie sie geschieht – eben mindestens teilweise basierend auf einem Training, nicht einer klassischen Programmierung, und damit eben mindestens teilweise autonom.

[16] Dies macht den Fächer enorm breit. Erfasst ist beispielsweise jedes System, das ein neuronales Netzwerk[35] oder sonst ein Verfahren des maschinellen Lernens beinhaltet, um gestützt darauf Output zu generieren bzw. Entscheide zu treffen. AIS sind somit selbst so banale Anwendungen wie eine Zeichenerkennung (OCR) in einem Kopierapparat oder PDF-Leseprogramm, der Einsatz eines Fingerabdrucksensors an einem Computer oder Mobiltelefon zu dessen Entsperrung oder

34 Vgl. dazu Martina Arioli, «Risikomanagement nach der EU-Verordnung über Künstliche Intelligenz», in: Jusletter IT, 4. Juli 2024, S. 3, Fn. 4.

35 Eine leicht verständliche Erläuterung, was ein neuronales Netzwerk ist und wie es funktioniert und trainiert wird, bietet der Blog-Beitrag «Was in einem KI-Modell steckt und wie es funktioniert» von David Rosenthal, abrufbar unter https://www.vischer.com/know-how/blog/teil-17-was-in-einem-ki-modell-steckt-und-wie-es-funktioniert/.

der Algorithmus, der in einer Internet-Suchmaschine unsere Eingaben interpretiert, um diese in einen passenden Suchbefehl umzuwandeln. An vielen Orten werden Unternehmen und erst recht die «normalen» Anwenderinnen und Anwender gar nicht realisieren, dass sie schon seit vielen Jahren mit AIS arbeiten. Selbst der Einsatz vergleichsweise einfacher und übersichtlicher Methoden wie etwa «Random Forest»[36] machen aus einem IT-System ein AIS, jedenfalls wenn «Autonomie» wie in den Erwägungen dargelegt so verstanden wird, dass ein System seine Entscheide nicht ausschliesslich nach einer von einem Programmierer getroffenen Entscheidung fällt. Wo genau die Abgrenzung KI/Nicht-KI zu ziehen ist, ist allerdings nicht trivial. Verwendet eine Software das Verfahren der «linearen Regression»,[37] um z.B. anhand von bisherigen Temperatur- und Besucherzahlen eines Freibads vorherzusagen, bei welcher Temperatur wieviele Gäste zu erwarten sind, so wird das kaum jemand als KI bezeichnen, sondern als reine Statistik: Die bisherigen Erfahrungswerte und Prognosetemperatur sind der Input für die Formel des Programms und die berechnete Gästezahl deren Output; es könnte auch gesagt werden, dass «Training» und Auswertung erst durch den Anwender erfolgt und nicht denjenigen, der das System bereitstellt. In Tat und Wahrheit können allerdings auch klassische Methoden des *Machine Learnings* als mathematische Funktionen beschrieben werden, nur sind sie oft komplexer. Eine scharfe Abgrenzung ermöglicht also das Kriterium der «Autonomie» bzw. der Differenzierung Training vs. Programmierung nicht.

[17] Während der AIA die allermeisten AIS nicht reguliert, dürfte sich seine Definition dessen, was «KI» ist, trotz der Unschärfen vielerorts durchsetzen; er ist zudem ähnlich zur Definition von KI der OECD.[38] Umso wichtiger wird es sein, dass Unternehmen und Aufsichtsbehörden bei der Regulierung von KI sich der Breite des Begriffs bewusst sind, einschliesslich der Tatsache, dass er sehr viele Anwendungen umfasst, an die sie beim Verfassen ihrer Regelwerke gar nicht denken und die dafür auch nicht passen.

[18] Die Definition von GPAIM ist noch diffuser als jene von AIS. Was ein «AI Model» ist, wird gar nicht gesagt. Ein «general purpose» (oder auf Deutsch «Allzweck-») KI-Modell wird zu einem solchen, wenn es für viele unterschiedliche Aufgaben eingesetzt werden kann und in dieser allgemeinen Form auf den Markt gebracht wird, so dass es in eine Vielzahl von Systemen und Anwendungen integriert werden kann.[39]

[19] Eine Praxisbemerkung noch zum Schluss: Die meisten Bestimmungen im AIA beziehen sich jeweils auf ein AIS, also ein «System». Die meisten dieser Bestimmungen im AIA verlangen jedoch zusätzlich einen bestimmungsgemässen Einsatzzweck dieses Systems, also beispielsweise für eine bestimmte Hoch-Risiko-Anwendung (z.B. Berechnung eines Bonitätsscore). In der Praxis wird die Kombination allerdings nicht mehr als System, sondern als Anwendung oder «Use Case» bezeichnet. Die Unterscheidung wird insofern von praktischer Relevanz sein, als im Unternehmen bei guter Governance jeweils entsprechende Verzeichnisse ihrer KI-Aktivitäten führen

---

36 Vgl. dazu etwa den Blog-Beitrag von IBM, abrufbar unter https://www.ibm.com/topics/random-forest; bei diesem Verfahren wird ein System dadurch trainiert, dass es basierend auf einer Anzahl von Datensätzen (z.B. über Preise von Liegenschaften) eine Mehrheit (sog. Ensembles) von Entscheidbäumen entwickelt, die dann im Einsatz für Prognosen benutzt werden können (z.B. was eine Liegenschaft einbringen wird).

37 Es wird versucht, zu berechnen, wie in einem Koordinatensystem eine Linie durch eine Wolke von Datenpunkten gelegt werden muss, damit vorhergesagt werden kann, wo weitere Punkte statistisch gesehen liegen müssen, wenn der Parameter der einen oder anderen Achse geändert wird.

38 Vgl. dazu https://oecd.ai/en/wonk/definition.

39 Art. 3 Ziff. 63 AIA.

werden, und sich hierbei die Frage stellt, ob diese nach Systemen oder eben nach Anwendungen geführt werden. Auch wenn der AIA nach Systemen unterscheidet, erweist sich für die Zwecke der KI-Governance häufig eine Inventarisierung nach Anwendung, d.h. nach Use Case, als nützlicher, weil ein System ohne Anwendung i.d.R. nicht angeschafft wird (selbst generische Tools wie «ChatGPT» haben eine solche), die Anwendung die Rechtsfolgen definiert und jede Anwendung typischerweise intern auch einen eigenen «Eigner» hat.

3. Geltungsbereich: Wer gilt als Provider (Anbieter)?

[20] Die beiden wichtigsten Rollen, die eine Organisation unter dem AIA separat oder zugleich haben kann, sind wie bereits erwähnt jene des Providers (zu Deutsch der «Anbieter») und jene des Deployers (zu Deutsch der «Betreiber»). Es gibt zwar weitere Rollen wie die der Importeure («Einführer»), der Distributoren («Händler») und Produktehersteller sowie des EU-Vertreters («Bevollmächtigter»), aber sie können letztlich dem Provider zugerechnet werden. Die Aufsichtsbehörden können gegen alle vorgehen, und alle müssen mit ihnen kooperieren.

[21] Es ist von zentraler Bedeutung, dass ein Unternehmen für jedes AIS oder GPAIM, mit dem es zu tun hat, ermittelt, welche Rolle es diesbezüglich hat, da sich aus dieser Rolle die Pflichten unter dem AIA ergeben. Die meisten Pflichten werden dabei dem Provider auferlegt (siehe unten).

[22] Provider ist in erster Linie derjenige, der (i) ein AIS oder GPAIM entwickelt (selbst oder im Auftrag) und (ii) der es unter seinem Namen oder seiner Marke in den Verkehr bringt oder in Betrieb nimmt:[40]

- Inverkehrbringen («placing on the market») bezieht sich gemäss Definition nur auf den EU-Markt und erfasst nur die erste Bereitstellung im Markt.[41]
- Inbetriebnahme («putting into service») bezieht sich ebenfalls nur auf das Gebiet der EU. Gemeint ist die Bereitstellung des AIS zum Erstgebrauch (d.h. erstmaligen Gebrauch) direkt an einen Deployer, aber auch zum (erstmaligen) Eigengebrauch (für GPAIM gilt diese Variante nicht, da sie nach der Konzeption des AIA nur eine Vorstufe eines AIS sind).[42]
- In beiden Fällen ist vermutlich eine Ausrichtung auf die EU gemeint, d.h. sie muss das Ziel sein; ein «non-intentional spill-over» genügt nicht.
- Keine Rolle spielt dabei, ob dies entgeltlich oder unentgeltlich geschieht.

[23] Bevor das geschieht, ist die Forschung, die Entwicklung und das Testen an und von AIS vom Geltungsbereich des AIA ausgenommen (Gegenausnahme sind einzig «Freilandversuche», also das Testen in der «realen Welt»; für sie gibt es ein separates Kapitel im AIA).[43]

[24] Diese Definition hat Folgen. Wer ein von ihm entwickeltes AIS nicht in der EU in den Verkehr bringt und auch nicht für den dortigen Eigen- oder Fremdeinsatz liefert, kann jedenfalls gemäss Wortlaut kein Provider sein und unterliegt somit nicht dem AIA, selbst wenn sein AIS den Weg

40 Art. 3 Ziff. 3 AIA.

41 Art. 3 Ziff. 9 AIA.

42 Art. 3 Ziff. 11 AIA.

43 Art. 2 Abs. 8 AIA; Art. 60 AIA.

dorthin findet oder dort Wirkung erzielt. Das gilt selbst für Unternehmen, die sich in der EU befinden. Das erscheint aufgrund des Wortlauts der Definitionen klar, sorgt aber für eine Lücke, weil der räumliche Geltungsbereich an sich weiter gefasst worden ist: Gemäss diesem sollen auch jene Provider erfasst sein, die sich im Ausland befinden, sofern der Output des AIS bestimmungsgemäss in der EU genutzt wird.[44] Diese Regelung soll die Umgehung des AIA durch im Ausland ver- und betriebene AIS mit EU-Wirkung verhindern,[45] greift aber gemäss Wortlaut bei Providern ins Leere, weil die Legaldefinition des Providers bereits einen EU-Marktbezug voraussetzt. Es wird interessant sein zu sehen, wie die Aufsichtsbehörden mit diesem gesetzgeberischen Versehen umgehen werden, da die Absicht des Gesetzgebers zwar klar ist, der Wortlaut aber ebenso.

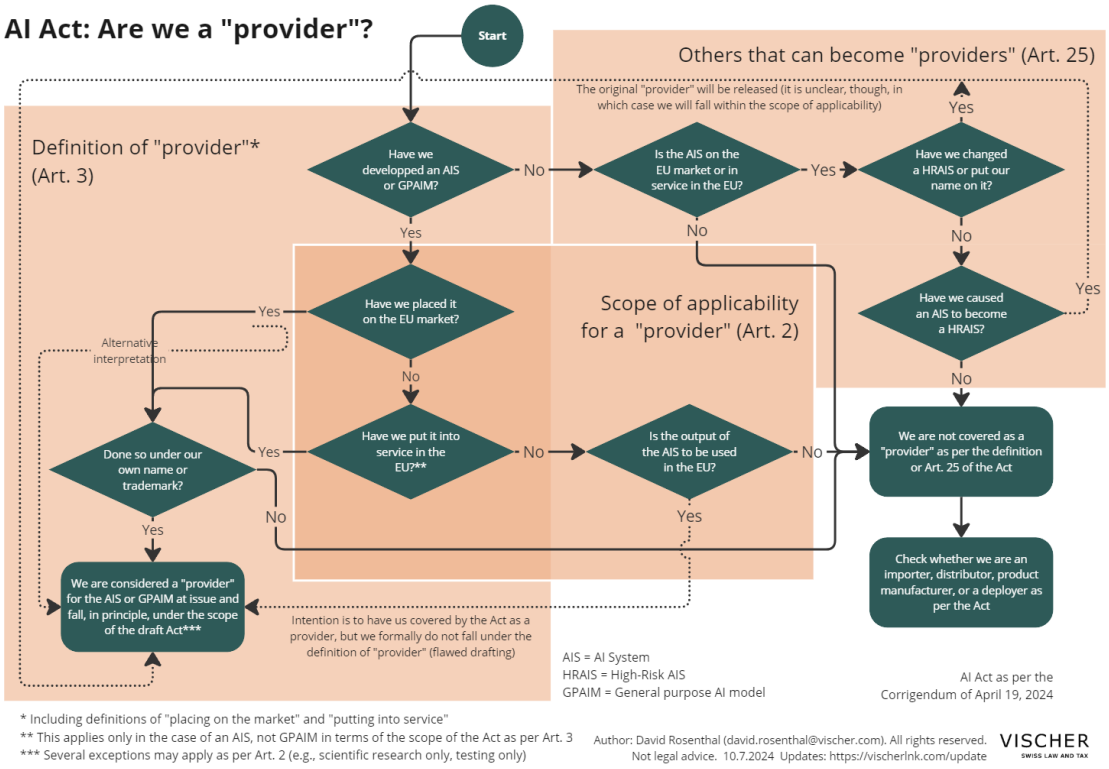
[25] Einen Anwendungsbereich für die vorstehende Regelung gibt es aber so oder so: Als Provider gilt ausnahmsweise auch jeder, der sich die Rolle des Providers quasi anmasst. Gemeint sind jene, die (i) ihren Namen oder ihr Markenzeichen an einem HRAIS anbringen (was auch immer dies bedeutet), das schon auf dem EU-Markt ist, (ii) ein solches HRAIS wesentlich ändern (es aber weiterhin ein HRAIS ist) oder (iii) ein AIS auf dem EU-Markt entgegen seiner ursprünglichen Bestimmung so verändern bzw. einsetzen, dass es zum HRAIS wird.[46] Ein Beispiel ist die Verwendung eines Allzweck-Chatbots für eine Hoch-Risiko-Anwendung (z.B. wenn «ChatGPT» in der EU zur Analyse von CVs von Stellenbewerbern genutzt wird). Der ursprüngliche Provider gilt dann nicht mehr als solcher. In diesem Fall wird die Spezialregelung solche Provider erfassen, falls diese ihren Sitz ausserhalb der EU haben, der Output der KI aber bestimmungsgemäss in der EU genutzt wird.[47] Ist dies nicht der Fall, stellt sich wiederum die Frage, ob diese «abgeleiteten» Provider noch in den Geltungsbereich des AIA fallen, wenn sie selbst das AIS nicht auf den EU-Markt gebracht und es dort auch nicht zum Einsatz gebracht haben. Das wäre eine weitere Lücke im AIA.

44 Art. 2 Abs. 1 lit. c AIA.

45 Erw. 22 AIA.

46 Art. 25 Abs. 1 AIA.

47 Vgl. dazu Martina Arioli (Fn. 34), S. 15.



[26] In der Literatur wurde teilweise vertreten, dass der AIA jedes AIS reguliere, welches Personen in der EU betrifft, direkt oder indirekt. Dies trifft nach der hier vertretenen Ansicht nicht zu. Es ist zwar richtig, dass der AIA auf alle betroffenen Personen («affected persons») in der EU ausdrücklich Anwendung findet.[48] Das bedeutet jedoch nicht automatisch, dass er auch auf alle Provider, Deployer und weitere Personen Anwendung findet; ihr persönlicher und örtlicher Geltungsbereich ist vom AIA separat definiert.[49] Wäre automatisch jedes AIS (mitsamt seinen Providern, Deployern und weiteren beteiligten Stellen) im Geltungsbereich des AIA, wären die vorstehend diskutierten differenzierten Regeln gar nicht nötig gewesen. Die Aufnahme der betroffenen Personen im Geltungsbereich des AIA ist vielmehr nötig, damit sie ihre (wenigen) Rechte unter dem AIA geltend machen können.

[27] Nur für HRAIS und nur rudimentär geregelt sind die Fälle, in denen ein AIS von **mehreren Stellen** stammt bzw. wenn das AIS eines Providers ein AIS eines anderen Providers enthält. In diesen Fällen soll mit dem Sublieferanten ein Vertrag abgeschlossen werden, der seinem Kunden als Provider die Einhaltung des AIA ermöglicht; der Sublieferant wird vermutlich nach Massgabe seiner eigenen Markthandlungen selbstständig erfasst sein.[50] Wer wiederum ein AIS in ein anderes Produkt verbaut, welches er dann unter eigenem Namen im EU-Markt anbietet oder in der EU in Verkehr bringt, gilt unter dem AIA als Produkthersteller («Product Manufacturer») und – falls es sich um ein HRAIS handelt (z.B. weil es eine Sicherheitsfunktion eines regulierten Produkts

48 Art. 2 Abs. 1 lit. g AIA.

49 Art. 2 Abs. 1 lit. a–c AIA.

50 Art. 25 Abs. 4 UAbs. 1 AIA.

übernimmt) – als dessen Provider im Sinne des AIA.[51] Mit anderen Worten: Der Provider kann nicht nur derjenige sein, der ein AIS (i) als Erstes entwickelt, sondern auch (ii) derjenige, der es weiterentwickelt und in ein übergeordnetes AIS oder sonstiges Produkt verbaut sowie (iii) wer sich gemäss (i) oder (ii) darstellt.

[28] Was passiert, wenn jemand ein AIS (mit)entwickelt, es aber nicht unter seinem Namen oder seiner Marke, sondern derjenigen eines anderen in der EU in Verkehr bringt oder in Betrieb nimmt, ist ebenfalls nicht ganz klar. Ein Provider kann diese Stelle formal gesehen nicht sein, weil ein zwingendes Kriterium der Legaldefinition nicht erfüllt ist. Der AIA kennt zwar gewisse Formen der «gemeinsamen» Verantwortlichkeit, sie geht aber nicht so weit wie im Datenschutz, wo es auch möglich ist, dass zwei Stellen für eine Verarbeitungsaktivität gemeinsam verantwortlich sein können, obwohl sie arbeitsteilig zusammenwirken. Etwas Ähnliches müsste an sich auch hier gelten, nur fehlt soweit ersichtlich die Rechtsgrundlage dafür; es ist nur das Zusammenwirken mehrerer Stellen geregelt, bei denen jeder auch für sich allein als Provider gilt. Über Hilfskonstruktionen kann der Fall jedoch gelöst werden:

- Wer unbefugt den Namen oder die Marke eines anderen auf dem von ihm entwickelten AIS angibt und sich damit im Grunde anmasst, wird in Bezug auf die Verantwortlichkeit als Provider dementsprechend wohl so behandelt werden, als wäre es auch sein eigener Name oder seine eigene Marke. Er müsste als Provider betrachtet werden – schon nur aufgrund des Prinzips von Treu und Glauben.
- Wer hingegen an einem AIS (mit)entwickelt, jedoch befugt den Namen oder die Marke eines anderen anbringt, kann unseres Erachtens mit guten Gründen vertreten, dass er lediglich für den anderen gehandelt hat, der als Provider gilt; die Nennung seines Namens oder das Anbringen seiner Marke – oder die Erlaubnis an einen Dritten, dies zu tun – ist gewissermassen die Erklärung, die Verantwortung des Providers für das AIS zu übernehmen.

4. Geltungsbereich: Wer gilt als Deployer (Betreiber)?

[29] Anwender von AIS werden unter dem AIA als Deployer erfasst und mit einigen wenigen Pflichten belegt (dazu unten). Deployer ist, wer ein AIS in eigener Verantwortung («under its authority») einsetzt.[52] Ausgenommen sind jene Personen, die ein AIS (nur) für persönliche, nicht berufliche Aktivitäten einsetzen («personal non-professional activity», wobei auch hier die Gesetzesredaktion unsorgfältig erfolgte – die Ausnahme existiert gleich auf zwei Ebenen und ist leicht verschieden formuliert[53]).

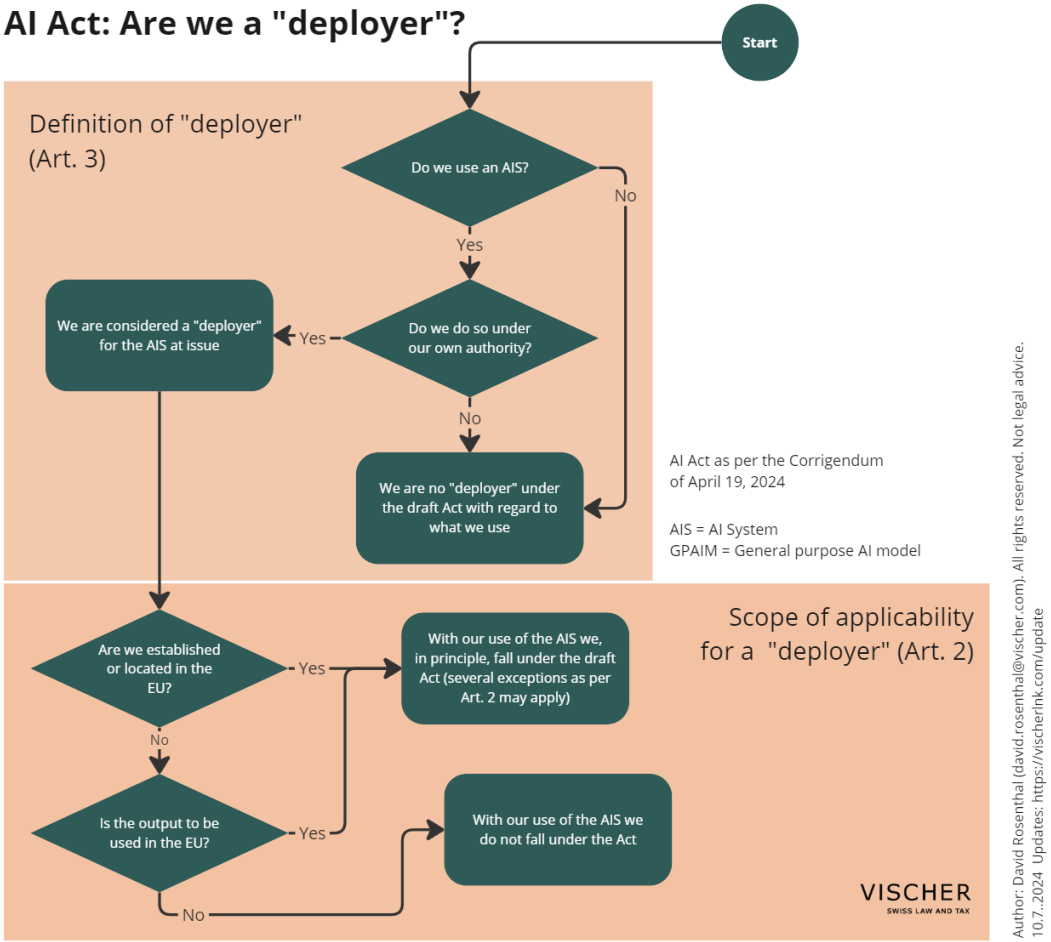
[30] Das Kriterium «under its authority» grenzt den blossen Genuss eines AIS bzw. dessen Output vom Einsatz eines AIS als eigenes Werkzeug ab. Hier wird eine gewisse Kontrolle erforderlich sein, damit das Kriterium erfüllt ist, denn Kontrolle führt zur Verantwortung. Wenn ein Benutzer auf einer Website einen Kundendienst-Chatbot benutzt, dann hat er diese Kontrolle nicht; jedenfalls ist dies nicht so vorgesehen. Er kann Fragen stellen und das Unternehmen steuert, wie

51 Art. 2 Abs. 1 lit. e AIA; Art. 25 Abs. 3 AIA.

52 Art. 3 Ziff. 4 AIA.

53 Art. 2 Ziff. 10 AIA vs. Art. 3 Ziff. 4 AIA.

der Chatbot sie beantwortet. Anders wäre es erst, wenn es dem Benutzer gelingt, den Chatbot zu «hacken» und zu nicht geplanten Äusserungen zu bringen. Stellt ein Unternehmen hingegen seinen Kunden im Rahmen einer Dienstleistung eine KI-Funktionalität zur Verfügung, die diese bestimmungsgemäss hinsichtlich ihres Einsatzes in gewissen Grenzen kontrollieren sollen und können, dann dürfte das Kriterium «under its authority» wohl erfüllt sein. Bietet eine Firma wie OpenAI ihren Kunden den Chatbot «ChatGPT» an, so unterscheidet sich dieser vom erwähnten Kundendienst-Chatbot dadurch, dass es der Kunde von OpenAI ist, der dem Chatbot die Anweisungen erteilen soll, worüber er sprechen soll – und nicht mehr das Unternehmen, das ihn betreibt. Wo genau die Grenze verläuft (z.B. bei themenspezifischen Chatbots), ist freilich auch hier noch nicht klar. Es bietet sich an, hier auf die Praxis zum vergleichbaren Konzept des «Verantwortlichen» bzw. «Controllers» im Datenschutz zurückzugreifen, das aus der Kontrolle über den Zweck und die (wesentlichen) sonstigen Parameter einer Datenbearbeitungsaktivität auch die diesbezügliche (alleinige oder gemeinsame) Verantwortung ableitet.



[31] In den räumlichen Geltungsbereich fallen all jene Deployer, die entweder in der EU sind (ob mit einer Niederlassung oder weil sie sich als natürliche Personen dort aufhalten) oder, falls nicht,

diejenigen, deren Output der von ihnen betriebenen AIS in der EU «verwendet» wird.[54] Letztere Regel soll gemäss den Erwägungen die Umgehung des AIA durch Anbieter in Drittländern (wie der Schweiz oder den USA) verhindern, die Daten in der EU sammeln oder von dort erhalten, sie mit AIS im EU-Ausland bearbeiten und das Ergebnis zur Nutzung in die EU zurücksenden, ohne dass das AIS in der EU auf den Markt gebracht oder dort zum Einsatz gebracht wird.[55] Das Beispiel wirft selbst Fragen auf, da ein Unternehmen mit Sitz in der EU, welches ein Unternehmen im EU-Ausland bittet, in seinem Auftrag Daten mit einem AIS zu bearbeiten, wohl trotzdem als dessen Deployer gelten wird, so wie im Datenschutz die Handlungen des Auftragbearbeiters dem Verantwortlichen zugerechnet werden. Viel relevanter dürfte die Ausnahmeregelung für jene Unternehmen sein, die sich im EU-Ausland befinden (z.B. in der Schweiz oder in den USA) und AIS für sich selbst einsetzen. So oder so soll die Regelung vom Einsatz des AIS betroffene natürliche Personen auf dem Gebiet der EU schützen. Gemäss den Erwägungen ist es erforderlich, dass die Verwendung des KI-Outputs in der EU beabsichtigt («intended to be used in the Union») und nicht bloss zufällig war.[56]

[32] Was genau als «**Verwendung» von KI-Output** in der EU gilt, wird nicht näher ausgeführt. Ob es hierzu bereits genügt, dass im EU-Ausland erzeugter KI-Output eine Wirkung auf Personen in der EU hat, ist denkbar, aber unseres Erachtens eher fraglich. Einige Beispiele:

- Wer im EU-Ausland auf seiner Website KI-generierte Inhalte publiziert und diese (auch) auf ein Publikum in der EU ausrichtet, dürfte erfasst sein. Dasselbe gilt für KI-generierte Inhalte auf Werbeunterlagen, die auch Kunden in der EU ansprechen sollen.
- Erfasst sein dürfte ebenso das Unternehmen, das seinen Kunden im EU-Raum durch ein AIS generierte Texte sendet.
- Findet ein von einem AIS im EU-Ausland erstelltes Dokument zufällig seinen Weg auch in die EU, dann ist der Deployer des AIS noch nicht erfasst.
- Nicht erfasst sein dürfte hingegen ein zwar von einem Computer, aber weder direkt noch indirekt von einer KI generierter Inhalt.
- Die Möglichkeit, dass KI-Output auf dem Weg des datenschutzrechtlichen Auskunftsrechts in die EU gelangen kann, genügt unseres Erachtens ebenfalls nicht zur Anknüpfung; dazu kann es zwar kommen, es wäre aber in der Regel keine bestimmungsgemässe Verwendung.
- Der Einsatz eines HRAIS oder die Durchführung einer nach AIA verbotenen Praktik durch einen Schweizer Arbeitgeber in Bezug auf seine Schweizer Mitarbeitenden dürfte selbst dann nicht dem AIA unterliegen, wenn er Mitarbeitende aus der EU beschäftigt, den Output der KI aber nur am Sitz des Unternehmens in der Schweiz verwendet.
- Setzt der Arbeitgeber dafür jedoch eine Cloud bzw. einen Cloud-Provider in der EU ein, sieht die Rechtslage allenfalls anders aus: Ob die Auslagerung des Betriebs eines AIS an einen Provider mit Sitz oder Rechenzentrum in der EU bedeutet, dass auch der Output der AIS als dort «verwendet» gilt, ist noch unklar. Wir glauben, dass dies nicht der Fall ist bzw. sein darf: Mit Blick auf den primären Schutzzweck, nämlich den Schutz der Gesundheit,

54 Art. 2 Abs. 1 lit. b und lit. c AIA.

55 Erw. 22 AIA.

56 Erw. 22 AIA.

der Sicherheit und der Grundrechte[57] von natürlichen Personen in der EU, kann die Auslagerung des IT-Betriebs in die EU für sich noch nicht genügen, weil allein dadurch diese in keiner Weise tangiert sind. Es gibt keinen Bezug zwischen dem Ort der Bearbeitung bzw. des Providers und der betroffenen Personen. Das Kriterium der Verwendung des Outputs in der EU soll seinem Zweck nach jene Fälle abdecken, in denen der Output zwar ausserhalb der EU erzeugt wird, in der EU aber auf betroffene Personen «trifft». Das ist hier nicht der Fall. Eine Niederlassung wird mit der blossen Beauftragung eines Dienstleisters in der EU erst recht nicht begründet (so auch nicht unter der DSGVO). In den sekundären Schutzzielen des AIA werden auch der Schutz der Demokratie, des Rechtssystems und der Umwelt genannt, aber auch diese rechtfertigen nicht wirklich die Erweiterung des Geltungsbereichs; sie käme noch am ehesten bezüglich dem Schutzziel des Umweltschutzes im Hinblick auf den Energieverbrauch der Rechenzentren für KI in Frage, aber diesbezüglich enthält der AIA selbst keine wirklich relevanten Regeln. Zu beachten ist immerhin, dass der Provider, der im Rahmen seiner Cloud-Dienste auch KI-Funktionalitäten anbietet, selbst dem AIA unterliegen kann, auch wenn sein Kunde als Deployer dies nicht tut. Diesen Fall haben wir im Bereich der DSGVO ebenfalls bei Providern mit Sitz im EWR; sie unterliegen der DSGVO, ihre Schweizer Kunden aber möglicherweise nicht.

[33] Ungeachtet der noch bestehenden Unklarheiten führt die Regelung zur Verwendung von KI-Output in der EU zu einer erheblichen extraterritorialen Wirkung des AIA. Dies bedeutet, dass sich auch viele Schweizer Unternehmen mit den vom AIA den Deployern auferlegten Kennzeichnungspflichten befassen sollten. In Bezug auf eine etwaige Rolle als Provider können sie sich hingegen wie erwähnt jedenfalls einstweilen auf den Standpunkt stellen, dass sie nicht erfasst sind, wenn sie das AIS nicht in der EU in Verkehr bringen oder dort in Betrieb nehmen.

5. Weitere Abgrenzungsfragen für Unternehmen

[34] Weitere Abgrenzungsfragen werfen die Rollendefinitionen des AIA insbesondere dort auf, wo Unternehmen AIS selbst weiterentwickeln, im Verkehr mit Dritten einsetzen oder innerhalb einer Unternehmensgruppe weitergeben.

[35] Unklar ist, was als «**Entwickeln» eines AIS** («develop») gilt. Liegt ein solches beispielsweise bereits vor, wenn ein Unternehmen ein kommerzielles KI-Produkt für seine eigene Anwendung parametrisiert (z.B. mit entsprechenden Systemprompts ausstattet), ein *Fine Tuning* des Models vornimmt oder in eine weitere Anwendung integriert (z.B. Einbau einer im Markt angebotenen Chatbot-Software in die eigene Website oder App)? Wäre dem so, würden viele Anwender selbst zum Provider werden, da die Legaldefinition des Providers auch denjenigen erfasst, der AIS zum Eigengebrauch («for own use») zum Einsatz bringt, falls dieser Einsatz bestimmungsgemäss in der EU und unter eigenem Namen erfolgt.[58] Ein solch weites Verständnis ist nach der hier vertretenen Auffassung zwar abzulehnen, aber es ist zu befürchten, dass die Aufsichtsbehörden hier ein breites Verständnis an den Tag legen werden und mindestens jene Handlungen als Entwickeln betrachten, welche über Prompting, Parametrisierung und das Liefern von sonstigem Input hin-

57 Vgl. zur grundrechtlichen Perspektive ausführlich Angela Müller (Fn. 16), S. 19 f.

58 Art. 3 Ziff. 3 AIA i.V.m. Art. 3 Ziff. 11 AIA.

ausgehen. Demnach wäre ein *Fine Tuning* des Modells eines AIS erfasst (weil gewissermassen das Trainingswissen und damit die Programmierung verändert wird), die Verwendung von «Retrieval Augmented Generation» (RAG) hingegen nicht (weil hier lediglich der Input verändert wird). Denkbar wäre auch ein Ansatz, der sich an der Gefahr orientiert, der vom Beitrag im Einzelfall ausgeht, was allerdings zu einer erheblichen Rechtsunsicherheit führen würde, weil dieselbe Handlung je nach Systemzweck zu einer unterschiedlichen Qualifikation desjenigen führen würde, der sie vornimmt. Dabei ist zu berücksichtigen, dass auch der Deployer bis zu einem gewissen Mass in der Pflicht bleibt.

[36] Bis zur Klärung der Rechtslage ist es Unternehmen als Vorsichtsmassnahme zu empfehlen, beim Einsatz solcher AIS auch gegen aussen den Namen des Technologielieferanten anzugeben (z.B. «Powered by …» bei einem Chatbot, der auf der Website den eigenen Kunden als Service angeboten wird). So kann vertreten werden, dass das AIS – selbst wenn die Implementierung oder Optimierung als Entwicklung gelten sollte – nicht als unter eigenem Namen oder den eigenen Handelsmarken bzw. Markenzeichen zum Einsatz gebracht gilt («under its own name or trademark»). Dies sollte jedoch mit Erlaubnis des Technologielieferanten erfolgen, da er bereit sein muss, die Rolle des Providers und die damit verbundene Verantwortlichkeit zu übernehmen (d.h. die Anpassungen ihm alleine zugerechnet werden).[59] Eine weitere mögliche Vorsichtsmassnahme ist die Einschränkung der erlaubten Verwendung eines AIS auf Personen ausserhalb der EU, weil die Legaldefinition des Providers dann mutmasslich nicht mehr greift.

[37] Unklar ist auch, wie die **Weitergabe von KI-Services oder KI-Technologie innerhalb einer Gruppe** von Unternehmen zu werten ist. Es sind hier verschiedene heikle Szenarien denkbar. Ein Szenario ist der Fall, in welchem die Gruppengesellschaft X ausserhalb der EU ein AIS vom Anbieter Y ausserhalb der EU einkauft und es dann den anderen Gruppengesellschaften auch in der EU zur Verfügung stellt. Ist das AIS bis dahin noch nicht auf dem EU-Markt eingeführt oder in der EU in Einsatz gebracht worden, kann X selbst ohne es weiterentwickelt zu haben zu dessen Händler («Distributor») oder allenfalls sogar Einführer («Importer») werden. Dies würde eine Reihe von Sonderpflichten mit sich bringen. Während X die Qualifikation als Einführer (und Provider) wohl verhindern kann, indem nur AIS in die EU weitergegeben werden, die dort schon auf dem Markt sind, schützt dies X nicht vor der Qualifikation als Händler. Hier kann immerhin vertreten werden, dass der Begriff voraussetzt, dass das Unternehmen Teil der in der Definition erwähnten Lieferkette («supply chain») von Y ist und dass diese einen konzerninternen Vertrieb eines AIS an Gruppengesellschaften von X vernünftigerweise nicht mehr umfasst. Sie machen das AIS auch nicht öffentlich zugänglich, wie dies ein Händler tut, sondern nur konzernintern.

[38] Zur Frage, ob ein Schweizer Unternehmen schon deshalb unter den AIA fallen kann, weil es sein KIS in einem **Rechenzentrum in der EU** betreiben lässt oder den Betrieb zwar in der Schweiz vornimmt, das Rechenzentrum aber von einem Cloud-Anbieter in der EU bereitgestellt wird, haben wir schon vorne im Zusammenhang mit der Verwendung des Outputs in der EU erörtert (Rn. 32). Es kann sich in diesen Fällen aber auch die Frage stellen, ob eine Auslagerung in eine EU-Cloud nicht als «Inbetriebnahme» des betreffenden KIS in der EU zu gelten hat, weil es hierzu genügt, dass das KIS zum erstmaligen Eigengebrauch in der EU bereitgestellt wird. Hat das Unternehmen das KIS selbst entwickeln (oder entwickeln lassen) und bringt es das KIS unter eigenem Namen zum Einsatz, wären in diesem Fall die beiden weiteren Voraussetzungen erfüllt

59 Siehe dazu Rn. 28.

und es läuft Gefahr, unter dem AIA als Provider des KIS zu gelten, selbst wenn es dieses nur für sich nutzt und seinen Sitz in der Schweiz hat. Um dieses Risiko zu reduzieren, sollten Schweizer Unternehmen KIS somit vorzugsweise auf Rechenzentren ausserhalb der EU betreiben, selbst wenn die Betreibergesellschaft (z.B. der Cloud-Provider) seinen Sitz selbst in der EU hat. Und noch etwas sollten Schweizer Unternehmen bedenken: Mitarbeitende, die sich beim Einsatz von AIS in der EU aufhalten und so zur Anwendbarkeit des AIA führen können (z.B. im Home-Office oder auf Reisen).

[39] Im Unternehmenskontext kann sich auch die Frage stellen, ab welchem Moment ein Unternehmen, das ein **Produkt mit einer KI-gestützten Funktionalität seinen Kunden zur Verfügung stellt**, selbst als Provider und die Kunden als dessen Deployer gelten. Die Hürden sind dabei nicht sehr hoch: Nehmen wir als Beispiel eine Bank ausserhalb der EU, die ihren EU-Geschäftskunden in ihrem Online-Banking und -Trading eine KI-gestützte Analyse ihrer Portfolios anbietet; dies ist Teil der kommerziellen Dienstleistung der Bank. Die Bank wird, wenn sie diese Funktionalität selbst entwickelt hat oder entwickeln lassen hat, aufgrund ihres eigenen Einsatzes bereits als Provider gelten, zumal sie es (jedenfalls nach einem weiten Begriffsverständnis) in der EU zum Einsatz bringt (siehe dazu aber Rz. 41) und dies unter ihrem eigenen Namen geschieht.

[40] Die Kunden, die die KI-Analyse-Funktion einsetzen, können ihrerseits zu Deployern werden und dabei dem AIA unterstellt sein, selbst wenn sich die Bank nicht darum kümmern sollte. Hierbei sind zwei Voraussetzungen zu beachten: Die Kunden müssen sich erstens in der EU befinden oder der Output der KI-Analyse in der EU verwendet werden. Sie müssen die KI-Funktion zweitens in eigener Verantwortung («under its authority») einsetzen (siehe dazu oben). Erhält der Bankkunde somit die nötige Kontrolle über die KI-Funktion, wird er (der als Geschäftskunde nicht unter die Ausnahme für persönliche, nicht berufliche Nutzungen fällt) aus eigener Verantwortung sicherstellen müssen und wollen, dass es sich um keine verbotene KI-Praktik handelt und allenfalls anwendbare Deployer-Pflichten eingehalten werden. Die Schwierigkeit in der Praxis kann dabei sein, dass es gerade bei Anbietern aus dem EU-Ausland für EU-Deployer möglicherweise gar nicht klar ist, wo und wann in den genutzten Produkten und Dienstleistungen AIS zum Einsatz kommen und was genau sie tun.

[41] In diesem Zusammenhang kann sich auch die Frage stellen, ob ein **Provider sachlogisch nur dann vorkommen kann, wenn es auch einen Deployer gibt**, dem der Provider das AIS zur Verfügung stellt. Die Antwort hierauf dürfte ein «Nein» sein. Ansonsten wären alle Anbieter, die KI-Dienste nur an Verbraucherinnen und Verbraucher in der EU anbieten, nicht vom AIA erfasst, weil solche Nutzerinnen und Nutzer kraft einer Ausnahmeregelung nicht als Deployer gelten.[60] Die Definition des Providers setzt keinen Deployer voraus: Es genügt zum einen die Bereitstellung in der EU für den *Eigengebrauch* in der EU («for own use in the Union»)[61]; um das zu erfüllen, reicht es womöglich bereits aus, dass der Anbieter die Nutzerinnen und Nutzer in der EU anspricht und sie den Dienst dort nutzen lassen, da es weiterhin der Dienst des Anbieters und somit «seine» Verwendung des AIS ist (die Schlussfolgerung, dass dies bereits einen Eigengebrauch *in der EU* darstellt, ist unseres Erachtens zwar nicht zwingend, dürfte aber im Hinblick auf den Schutzzweck des AIA mitunter vertreten werden; dies würde dazu führen,

60 Art. 3 Ziff. 3 Ziff. 4.

61 Art. 3 Ziff. 3 AIA i.V.m. Art. 3 Ziff. 11 AIA.

dass der bloss via Website einem EU-Publikum angebotene Chatbot bereits zum Eigengebrauch *in der EU* führen würde, d.h. ein Fernzugriff aus der EU als Gebrauch vor Ort qualifiziert; dem könnte entgegengehalten werden, dass der Gesetzgeber gerade nicht davon ausging, dass dies zur Erfassung genügt, weshalb er die Verwendung von KIS-Output in der EU als separates Anknüpfungskriterium definierte, welches hier jedenfalls beim Deployer greifen würde). Zum anderen lässt es das alternative Kriterium zur Provider-Qualifikation des Inverkehrbringens («placing on the market») genügen, dass ein AIS «zur Verwendung auf dem Unionsmarkt im Rahmen einer Geschäftstätigkeit» («for … use on the Union market in the course of a commercial activity») geliefert worden ist. Dieser Fall liegt hier wohl vor. Abschliessend geklärt ist aber auch diese Frage nicht.

6. Verbotene KI-Anwendungen

[42] Acht spezifische «KI-Praktiken» sind unter dem AIA ganz verboten.[62] Vom Verbot erfasst sind in den meisten Fällen, sowohl jene, die betreffende AIS in der EU auf den Markt oder erstmals in der EU zum Einsatz bringen als auch jene, die ein AIS so verwenden wollen (was, wie gezeigt, auch jene umfasst, die sie im Ausland nutzen, soweit der Output der AIS bestimmungsgemäss auch in der EU verwendet wird).

[43] Die Liste der verbotenen KI-Praktiken ist zwar sehr spezifisch, aber trotzdem nicht überall trennscharf abgegrenzt formuliert:

1. Verwendung von **unterschwelligen, absichtlich manipulativen oder täuschenden Techniken**, die darauf abzielen oder bewirken, dass das Verhalten wesentlich verzerrt wird oder die Fähigkeit der Person, eine informierte Entscheidung zu treffen, beeinträchtigt wird, falls dies zu einer Entscheidung führen kann, die mit hinreichender Wahrscheinlichkeit einen erheblichen Schaden verursacht oder verursachen kann.[63] Keine Absicht muss hinsichtlich der Schadensverursachung vorhanden sein.[64] Nicht erfasst sein sollen ansonsten zulässige Anwendungen im medizinischen Bereich.[65]

2. **Ausnutzen von Vulnerabilitäten oder der Schutzbedürftigkeit** von Personen aufgrund ihres Alters, einer Behinderung oder einer besonderen sozialen oder wirtschaftlichen Situation, um ihr Verhalten in einer Weise wesentlich in einer Weise zu verändern, die mit hinreichender Wahrscheinlichkeit einen erheblichen Schaden verursacht oder verursachen kann.[66]

3. **Biometrische Kategorisierung**, um Rückschlüsse auf die Rasse, die politische Meinung, die Gewerkschaftszugehörigkeit, die religiöse oder weltanschauliche Überzeugung, das Sexualleben oder die sexuelle Ausrichtung einer Person zu ziehen (d.h. auf der Grundlage biometrischer Daten).[67] Es geht also darum, dass gewissermassen von «Äusserlichkeiten» auf (in diesem Fall sen-

62 Art. 5 Abs. 1 AIA.

63 Art. 5 Abs. 1 lit. a AIA.

64 Erw. 29 AIA.

65 Erw. 29 AIA.

66 Art. 5 Abs. 1 lit. b AIA.

67 Art. 5 Abs. 1 lit. g AIA.

sible) «innere» Werte geschlossen wird, was insbesondere deshalb als verboten wird, weil es auch mit AIS unzuverlässig ist.[68] Keine biometrische Kategorisierung ist beispielsweise die Analyse von Text einer Person, wohl aber deren Stimme. Nicht verboten ist auch der Einsatz von AIS, um biometrische Merkmale als solche zu erkennen, z.B. alle Personen mit grünen Augen.

4. **Bewertung oder Kategorisierung** von Personen (einzeln oder in Gruppen) über einen bestimmten Zeitraum hinweg auf der Grundlage ihres **Sozialverhaltens** oder bekannter, abgeleiteter oder vorhergesagter **persönlicher Eigenschaften**, falls diese soziale Bewertung zu einer nachteiligen oder ungünstigen Behandlung führt, die in keinem Zusammenhang mit dem ursprünglichen Kontext der Daten steht oder aber im Hinblick auf ihr soziales Verhalten oder dessen Tragweite ungerechtfertigt oder unverhältnismässig ist.[69]

5. **Biometrische Identifizierung** aus der Ferne und in Echtzeit in öffentlich zugänglichen Räumen zum Zweck der Strafverfolgung, mit Ausnahme der gezielten Suche nach Opfern gewisser Straftaten, der Vorbeugung gewisser spezifischer Bedrohungen im Falle erheblicher und unmittelbar bevorstehender Bedrohungen oder der Lokalisierung oder Identifizierung von Verdächtigen bestimmter definierter Kategorien von Straftaten (Anhang II), vorbehaltlich zusätzlicher Bedingungen (z.B. gerichtliche Genehmigung, Erlaubnis nur für die Suche nach bestimmten Zielpersonen).[70]

6. Erstellung von Profilen oder Bewertung von Persönlichkeitsmerkmalen oder Eigenschaften von Personen, um das **Risiko der Begehung von Straftaten einzuschätzen** oder vorherzusagen, ausser zur Unterstützung der Risikobewertung von Personen, die an einer Straftat beteiligt sind.[71]

7. Aufbau oder Erweiterung einer **Gesichtserkennungsdatenbank** auf der Grundlage von ungezieltem Scraping im Internet oder CCTV-Aufnahmen (d.h. Überwachungskameras).[72] Hier scheint der Gesetzgeber Anbieter wie «Clearview» und «PimEyes» im Blick gehabt zu haben.

8. Das **Ziehen von Rückschlüssen auf Emotionen** (einschliesslich Absichten) von Personen am Arbeitsplatz oder in Bildungseinrichtungen, es sei denn, dies soll medizinischen oder sicherheitstechnischen Gründen dienen.[73] Hier kommt es jedenfalls gemäss dem Text nicht darauf an, ob die Emotionserkennung auf biometrischen Daten (z.B. Stimme, Gesichtsausdruck) oder anderen Daten (z.B. Inhalt von E-Mails oder anderen Texten) basiert. Dies ist jedoch nicht vollständig klar. Denn auch bei den HRAIS, wo Systeme zur Emotionserkennung ebenfalls erfasst sind, wird im Text nicht differenziert, jedoch geht aus den Erwägungen hervor, dass jedenfalls dort nur Emotionserkennung auf Basis biometrischer Daten gemeint ist.[74]

---

68 Vgl. auch Erw. 54 AIA.

69 Art. 5 Abs. 1 lit. c AIA.

70 Art. 5 Abs. 1 lit. h AIA.

71 Art. 5 Abs. 1 lit. d AIA.

72 Art. 5 Abs. 1 lit. e AIA.

73 Art. 5 Abs. 1 lit. f AIA.

74 Rn. 52.; vgl. ausführlich zur Anwendung Künstlicher Intelligenz im Rahmen der Emotionsverarbeitung Robert van den Hoven van Genderen/Rosa Ballardini, «AI and Emotional data between the Scylla and Charybdis of European Regulation», in: Jusletter IT 15. Februar 2024, S. 58 f.

[44] Zu beachten ist, dass diese Praktiken nur dann verboten sind, wenn sie mittels eines AIS ausgeführt werden. Wer also etwa eine Emotionserkennung am Arbeitsplatz einzig anhand selbst, also von Menschenhand programmierter Wenn-Dann-Regeln durchführt (keine Mustererkennung), ist nicht erfasst. Die Tatbestände sind zudem oft sehr spezifisch formuliert, so dass sich rasch Ausnahmen ergeben. Der Einsatz von AIS zur Überführung von Schülerinnen und Schülern, die bei einer Prüfung tatsächlich schummeln, ist vom Verbot der Erkennung von Emotionen und Absichten beispielsweise nicht erfasst (allerdings eine Hoch-Risiko-Anwendung); erfasst wäre aber ein AIS, das Schüler aufspüren soll, noch bevor sie schummeln. Detektiert eine AIS die Nervosität eines Schülers, damit die Lehrperson ihn besonders im Auge behalten kann, dann wäre dies eine verbotene Praktik, weil Nervosität eine Emotion darstellt. Erfasst wären auch Lügendetektoren. Grund für das Verbot ist übrigens auch hier die Unzuverlässigkeit der Erkennung und die damit verbundene Gefahr von Ungenauigkeiten und Diskriminierung[75].[76]

[45] Auch am Arbeitsplatz sind Beispiele im Bereich der Emotionserkennung denkbar, die in der Praxis vorkommen können und keineswegs ungewöhnlich sind. Ein Beispiel ist die Analyse von Gesprächen eines Call Centers. Zur Qualitätskontrolle und für Schulungszwecke werden z.B. gerne sog. Sentiment-Analysen[77] durchgeführt, bei denen die Stimmung der Anrufer ermittelt wird (positiv, neutral, negativ). Geschieht dies nur anhand der Stimme des Anrufers (z.B. live), so sind die Call-Center-Mitarbeitenden nicht betroffen (die Anwendung würde jedoch anderweitig vom AIA erfasst). Erfolgt die Auswertung jedoch nachträglich anhand eines Transkripts des Dialogs, so kann es je nach Programmierung vorkommen, dass in der Zusammenfassung für den Supervisor auch die Stimmung bzw. die Emotion des Call-Center-Agent angegeben wird. Dies wäre unzulässig, jedenfalls wenn nicht danach differenziert wird, ob die Emotionserkennung auf Basis biometrischer Daten erfolgt.[78]

[46] Wer wiederum ein AIS benutzt, um «Data Loss Prevention» (**DLP**) zu betreiben, führt keine verbotene Praktik aus, auch wenn der Diebstahl von Daten eine Straftat sein kann und DLP daher im weitesten Sinne auch die Beurteilung des Risikos einer Straftat umfasst, aber dies eben nicht das Ziel ist, sondern die Verhinderung eines ungewollten Datenabflusses, ganz gleich, ob es sich um eine Straftat handelt (die verbotene Praktik Nr. 6 oben zielt auf das «predictive policing» ab[79] und nicht auf die Verhinderung von Delikten, die im Gange sind). Das AIS muss aber so funktionieren, dass es Indizien auf Datendiebstahl aufspürt (z.B. Übermittlung bestimmter geschäftlicher Dateien an eine private E-Mail-Adresse) und nicht Hinweise auf Emotionen oder Absichten eines Mitarbeitenden, einen solchen zu begehen (z.B. Suche nach Mustern, wonach der Mitarbeitende eine Wut auf seinen Arbeitgeber hat, frustriert im Job ist oder innerlich gekündigt hat, also keine Loyalität mehr aufweist). Keine Emotion ist hingegen Müdigkeit. Wer also eine KI einsetzt, um übermüdete Mitarbeitende zu erkennen, darf das. Der Einsatz von KI zur Erkennung von Stress am Arbeitsplatz ist hingegen nicht erlaubt.

---

75 Vgl. ausführlich zu Diskriminierung beim Einsatz von künstlicher Intelligenz Florent Thouvenin/Stephanie Volz/Soraya Weiner/Christoph Heitz, «Diskriminierung beim Einsatz von Künstlicher Intelligenz (KI), Technische Grundlagen für Rechtsanwendung und Rechtsentwicklung», in Jusletter IT 4. Juli 2024, S. 17 ff.

76 Erw. 54 AIA.

77 Vgl. dazu etwa Mayur Wankhade/Annavarapu Chandra Sekhara Rao/Chaitanya Kulkarni, «A survey on sentiment analysis methods, applications, and challenges», in: Artificial Intelligence Review 55 (2022) S. 5731–5780, https://doi.org/10.1007/s10462-022-10144-1.

78 Dazu mehr in Rn. 52.

79 Menschen sollen aufgrund ihres tatsächlichen Verhaltens beurteilt werden, nicht nach ihren Absichten.

[47] Ein weiteres Beispiel: Bei der verbotenen Praktik Nr. 1 genügt es nicht, dass Personen mittels KI manipuliert werden. Es muss dies auch zu dem Zweck erfolgen, sie wesentlich in ihrem Verhalten zu beeinflussen, sie daran zu hindern, informierte Entscheidungen zu treffen, und das muss zudem zu einem Entscheid führen, der wesentlichen Schaden für die Person verursachen kann. So halten sogar die Erwägungen fest, dass «übliche und rechtmäßige Geschäftspraktiken, beispielsweise im Bereich der Werbung» («common and legitimate commercial practices») hier nicht erfasst sind.[80]

[48] Etwas weniger klar ist der Fall bei Praktik Nr. 4 (**Scoring**), welche recht breit verstanden werden kann, weil sie lediglich eine nachteilige Behandlung voraussetzt. Das Verbot greift aber nur und erst, wenn die Daten mit dieser Behandlung, um die es geht, nichts zu tun haben oder die Schlechterbehandlung unverhältnismässig oder unberechtigt ist. Spannend wird hier zu sehen sein, ob beispielsweise KI-gestützte Funktionen, die für jeden Besucher eines Online-Shops den aufgrund seines Verhaltens aus Sicht des Anbieters optimalen Preis berechnen, erfasst sein werden. Um dem Verbot entgegenzuwirken, müsste auch hier ein System eingesetzt werden, das nicht auf Basis eines Trainings urteilt, sondern regelbasiert agiert.

[49] Bei Praktik Nr. 4 wird sich weiter die Frage stellen, ob und wann sie KI-basierte Bonitätsbeurteilungen meint und damit verbieten würde. Die Zahlungsfähigkeit einer Person ist zwar kein soziales Verhalten und auch keine persönliche Eigenschaft, die Zahlungswilligkeit kann es aber sein. Erfasst wäre also, wer ein AIS auf den Markt bringt, welches zwecks Kreditgewährung Korrelationen zwischen dem Verhalten einer Person und ihrer mutmasslichen Zahlungsunwilligkeit aufzeigt und hierbei Daten aus einem Kontext verwendet, der mit der Zahlungswilligkeit der Person nichts zu tun hat oder deren Beizug unverhältnismässig oder ungerechtfertigt wäre. Dem würde der Anbieter einer Lösung für ein Bonitäts-Scoring entgegenhalten, dass diese auch die Zahlungsfähigkeit abdeckt. Hierauf könnte geantwortet werden, dass die Beurteilung «Zahlt seine Rechnungen mit erhöhter Wahrscheinlichkeit nicht» durchaus als «Social Score» verstanden werden kann. Diesfalls müsste zur Verwendung des AIS aufgezeigt werden, dass nur Daten verwendet werden, die im Kontext des Bezahlens von Rechnungen erhoben wurden und die Bonitätsbeurteilung tatsächlich einigermassen die richtigen trifft und die Konsequenzen vertretbar sind. Als HRAIS wäre eine solche Anwendung freilich so oder so erfasst (siehe nachfolgend).

[50] Ob der Gesetzgeber bei Praktik Nr. 4 tatsächlich so weit gehen wollte, wie dies aus der weit gefassten Regelung wie gezeigt geschlossen werden könnte, ist nicht klar. In den Erwägungsgründen hält er jedenfalls fest, dass das Verbot «nicht die rechtmäßigen Praktiken zur Bewertung natürlicher Personen berühren [sollte], die im Einklang mit dem Unionsrecht und dem nationalen Recht zu einem bestimmten Zweck durchgeführt werden.»[81] Die kryptische Formulierung, die im Englischen nicht klarer ist, ist wohl dahingehend zu verstehen, dass er nur extreme Fälle untersagen wollte – wie sie etwa in China praktiziert werden. Die Aussage könnte so verstanden werden, dass immer dann, wenn eine Beurteilung bzw. ein Profiling einer natürlichen Person erstens datenschutz- und auch sonst rechtskonform ist (nach dem Recht der EU) und zweitens die Daten erklärtermassen (auch) zu dem spezifisch Zweck erhoben worden sind, für den sie verarbeitet werden, der AIA die Bewertung nicht verbieten will. Gibt also der erwähnte Online-Shop in seiner Datenschutzerklärung an, was er mit den Daten der betroffenen Personen tun will, auch

80 Erw. 29 AIA.

81 Erw. 31 AIA.

wenn diese inhaltlich damit auf den ersten Blick nichts zu tun haben, und erfüllt er auch sonst die Vorgaben der DSGVO (soweit sie anwendbar ist), dann wäre der Einsatz eines AIS hierfür demnach nicht verboten.[82]

7. Hoch-Risiko-KI-Systeme

[51] Als «Hoch-Risiko KI-System» (HRAIS) gelten zwei Arten von AIS. Zunächst sind es grundsätzlich jene AIS, welche nach den im Anhang I zum AIA aufgeführten EU-Regularien sind, für welche vor dem Vertrieb oder Einsatz eine Konformitätsbewertung durch einen Dritten erforderlich sind, sowie jene AIS, die als Sicherheitskomponente in solchen regulierten Produkten eingesetzt werden (als Sicherheitskomponente gelten auch solche AIS, deren Ausfall zu einer Gefährdung der Gesundheit oder Sicherheit von Mensch oder Besitz führen).[83] Die Anforderungen des AIA ergänzen in diesen Fällen die Regeln, die für diese Produkte ohnehin gelten. Beispiele für solche Produkte sind Medizinprodukte, Spielzeug, Funkgeräte, Sportboote, Flugzeuge, Fahrzeuge, Eisenbahnen oder Aufzüge.[84]

[52] Weiter gelten als HRAIS all jene AIS, die in einem weiteren Anhang des AIA aufgeführt sind.[85] Wie schon bei den verbotenen Praktiken sind auch diese AIS sehr spezifisch definiert, weshalb die Liste regelmässig überprüft und nötigenfalls angepasst werden sollte. Derzeit umfasst die Liste folgende Anwendungen. Entscheidend ist jeweils, ob ein AIS bestimmungsgemäss für den betreffenden Zweck eingesetzt werden soll. Es werden im betreffenden Annex meist die Anwendungsgebiete definiert, und dann die einzelnen, erfassten **Fälle hohen Risikos** näher beschrieben:

1. **Biometrische Fernidentifizierung** mittels KI, die über die blosse Authentifizierung hinausgeht, oder **biometrische Kategorisierung** mittels KI, welche sensible oder geschützte Merkmale oder Eigenschaften verwenden, die aus solchen abgeleitet wurden.[86] Gemeint sind insbesondere Anwendungen zur Gesichtserkennung[87] über Kamerabilder, die in einer Zentrale verarbeitet werden, aber nicht Systeme zur biometrischen Zugangskontrolle oder Systeme, die sicherstellen sollen, dass nur autorisierte Personen Zugriff auf ein Computersystem erhalten.[88]

2. **Erkennung von Emotionen oder Absichten**[89] auf der Grundlage biometrischer Daten mittels KI.[90] Anhang III erwähnt nur «Emotionen», nicht Absichten («intentions»), aber letztere werden

82 Vgl. zum Profiling in der KI-Verordnung Jürgen Taeger, «EU-Regulierungen von Profiling und Microtargeting gegen Desinformation und Manipulation», in: Datenschutz-Berater, 2024, S. 138.

83 Art. 6 Abs. 1 AIA.

84 Anhang I des AIA.

85 Art. 6 Abs. 2 AIA.

86 Anhang III AIA, Ziff. 1 a)–c), Art. 3 Ziff. 39 AIA, Erw. 18 AIA; vgl. auch Jonas Pfister/Jessica Fleisch/Jakob Zanol, «New legal framework for AI-based facial recognition», in: Jusletter IT 27. April 2023, S. 404.

87 Vgl. ausführlich zu den Auswirkungen des AI Acts in Bezug auf Gesichtserkennungssoftwares Alexandra Wudel/Michael Schulz, «Der Artificial Intelligence Act – eine Praxisanalyse am Beispiel von Gesichtserkennungssoftware», in: HMD 59, S. 588–604 (2022), https://doi.org/10.1365/s40702-022-00854-z.

88 Erw. 54 AIA.

89 Art. 3 Ziff. 39 AIA.

90 Anhang III AIA, Ziff. 1 d).

in der Begriffsdefinition und in den Erwägungen immer mitgenannt.[91] Es spielt keine Rolle, ob Personendaten bearbeitet werden oder nicht. Beispiele könnten sein, die Emotionserkennung bei einem Verkaufsautomaten, einem Videospiel oder einem System zur medizinischen Diagnose. Als Emotionen gelten beispielsweise Glück, Trauer, Wut, Überraschung, Ekel, Verlegenheit, Aufregung, Scham, Verachtung, Zufriedenheit und Vergnügen. Die Erkennung von Schmerzen und Müdigkeit ist hingegen nicht erfasst, ebenso nicht das Erkennen von Gesten und Bewegungen, beispielsweise zur Steuerung eines Systems. Die Erwägungen halten ausserdem fest, dass biometrische Systeme, die ausschliesslich dazu bestimmt sind, um Massnahmen zur Cybersicherheit und zum Schutz personenbezogener Daten durchführen zu können, nie als HRAIS gelten sollen.[92]

Die Regelung zur Emotionserkennung ist insofern missverständlich, als dass in Anhang III von «KI-Systemen, die bestimmungsgemäß zur Emotionserkennung verwendet werden sollen» die Rede ist und nicht vom definierten Begriff «Emotionserkennungssystem». Das ist insofern von Belang, als letzterer Begriff zwingend voraussetzt, dass die Erkennung der Emotion auf der Grundlage von biometrischen Daten einer Person erfolgt, also beispielsweise aufgrund ihres Gesichtsausdrucks, ihrer Stimme oder ihrer Bewegungen. Gemeint ist aber in beiden Fällen dasselbe, d.h. auch die Regelung zu den HRAIS meint Emotionserkennung nur auf Basis biometrischer Daten, wie aus den Erwägungsgründen hervorgeht.[93] Das kann von erheblicher praktischer Relevanz sein, wenn es um eine Emotionserkennung auf der Basis des allgemeinen, nicht personenbezogenen Bedeutungsgehalts von Texten einer Person geht. Wenn also im oben erwähnten Call-Center-Beispiel für eine Sentiment-Analyse eine Unterhaltung zunächst transkribiert und erst in dieser Form von einem AIS analysiert wird, wäre dies kein HRAIS, wenn der Inhalt dessen, was jemand sagt, nicht mehr als ein biometrisches Datum gilt. Als ein solches gelten personenbezogene Daten zu den physischen, physiologischen oder verhaltenstypischen Merkmalen einer natürlichen Person.[94] Um eine Text-basierte Sentiment-Analyse als Emotionserkennung zu erfassen, müsste gezeigt werden, dass die von einer Person verwendeten Aussagen und Formulierungen «verhaltenstypische» Merkmale sind und für die Sentiment-Analyse ausgewertet werden. Dem wird normalerweise nicht so sein. Gemeint sind hier unseres Erachtens konstante, individuelle, messbare und bei jeder Person vorkommende Merkmale dieser Person wie ihre Stimme (d.h. bestimmte Aspekte davon, wie Akzent oder Betonungen), Unterschrift und Lippenbewegungen, ihr Tippverhalten auf Tastaturen oder ihre Gangdynamik. Bei Text wäre ein solches Merkmal eine für eine bestimmte Person charakteristische Formulierung wie z.B. eine bestimmte Art und Weise, wie sie ihre E-Mails «unterzeichnet» (in Analogie zur Unterschrift als verhaltenstypisches Merkmal). Der auf Basis allgemeiner, nicht personenbezogener Erfahrungswerte ermittelte Bedeutungsgehalt von geschriebenem Text ist unserer Ansicht nach jedoch kein solches Merkmal einer bestimmten Person und darf daher für eine Emotionserkennung wie im Beispiel der Sentiment-Analyse verwendet werden (z.B. Verärgerung, wenn die Person sagt, sie sei verärgert oder Formulierungen verwendet, die dies nach allgemeinem Sprachgebrauch ausdrücken, wie z.B. «Das kann doch nicht Ihr Ernst sein!»). Würde hingegen ein KI-Modell auf Texten einer bestimmten Person trainiert, die nach ihrem jeweiligen Gemütszustand bezeichnet sind, und würde dieses Modell später

91 Erw. 18 AIA, Art. 3 Ziff. 39 AIA.

92 Erw. 54 AIA.

93 Erw. 54 AIA.

94 Art. 3 Ziff. 34 AIA.

eingesetzt, um die Emotionen eben dieser Person anhand von weiteren Texten dieser Person zu ermitteln, quasi indem es aufgrund seines personenbezogenen Trainings «zwischen den Zeilen» lesen kann, läge eine Emotionserkennung im Sinne des AIA vor. Wo genau die Grenze verläuft, bleibt freilich auch damit nicht ganz klar.

3. KI soll als Sicherheitsbauteil bei der Verwaltung und dem **Betrieb kritischer digitaler Infrastrukturen** (wie z.B. gewisse Internet-, Telekom-, Kommunikations-, Cloud-Computing- und Rechenzentrums-Dienste), im Strassenverkehr oder bei der Versorgung mit Wasser, Gas usw. eingesetzt werden (nicht aber KI einzig zur Cybersicherheit).[95]

4. Einsatz in der allgemeinen und beruflichen **Bildung**, sofern (i) der Zugang, die Zulassung oder die Zuweisung zu Bildungsangeboten durch KI bestimmt werden soll, (ii) KI zur Bewertung der Lernergebnisse oder des (für den Zugang zur Bildung nötigen) Bildungsniveaus von Personen eingesetzt werden soll oder (iii) KI zur Überwachung oder Erkennung von verbotenem Verhalten bei Prüfungen eingesetzt werden soll.[96]

5. Einsatz im Bereich **Beschäftigung und Arbeitnehmer**, soweit (i) KI für die Einstellung oder Auswahl von sich auf eine Stelle bewerbenden Personen oder für die Platzierung von Stellenangeboten verwendet wird oder (ii) KI verwendet wird, um Entscheidungen zu treffen, die sich auf die Beschäftigungsbedingungen auswirken oder Beförderungen oder die Beendigung des Arbeitsverhältnisses betreffen, um Arbeit zuzuweisen auf der Grundlage von Verhalten oder anderen persönlichen Merkmalen und um Leistung und Verhalten zu beobachten und zu bewerten.[97]

6. KI wird verwendet, um (für oder als Behörde) zu beurteilen, ob einer bestimmten Person wesentliche **öffentliche Unterstützungsleistungen** und Dienste, einschliesslich der Gesundheitsversorgung, zur Verfügung stehen oder weiterhin zur Verfügung stehen werden.[98]

7. KI wird zur Bewertung der **Kreditwürdigkeit** einer Person oder Bestimmung ihres Bonitäts-Scores verwendet werden, jeweils ausser zur Aufdeckung von Finanzbetrug.[99]

8. KI wird verwendet, um **Notrufe** von Personen zu bewerten und zu klassifizieren, oder um Notrufe oder **Notfalldienste** oder die medizinische Versorgung zuzuweisen oder zu triagieren.[100] Die Verwendung von KI für medizinische Zwecke (wie z.B. die Beurteilung von Patienten) ist als solche nicht ausdrücklich als Anwendungsfall aufgeführt. Wer jedoch eine Software oder ein Gerät anbietet, welches KI beispielsweise für Diagnosezwecke verwendet oder zur Behandlung, wird typischerweise über die Regulierung von Medizinprodukten erfasst sein.[101] Nicht ganz klar ist, ob die Regelung, wonach derjenige zum Provider wird, der ein «normales» AIS zu einem

95 Anhang III AIA, Ziff. 2.

96 Anhang III AIA, Ziff. 3.

97 Anhang III AIA, Ziff. 4.

98 Anhang III AIA, Ziff. 5 a).

99 Anhang III AIA, Ziff. 5 b).

100 Anhang III AIA, Ziff. 5 d).

101 Vgl. Rn. 51.

HRAIS umfunktioniert,[102] im Fall greift, in welchem beispielsweise ein Arzt «ChatGPT» benutzt, um z.B. Aufzeichnungen seiner Patienten zu analysieren.

9. KI wird für **Risikobewertungen** und die **Preisgestaltung** von **Lebens- oder Krankenversicherungen** verwendet.[103]

10. Verwendung zu **Strafverfolgungszwecken**, wenn (i) KI zur Bewertung des Risikos einer Person, Opfer einer Straftat zu werden, verwendet werden soll, (ii) KI als Lügendetektor oder ähnliches Instrument verwendet werden soll oder (iii) KI zur Feststellung der Zuverlässigkeit von Beweismitteln verwendet werden soll (in jedem Fall mit Ausnahme der oben genannten verbotenen Praktiken).[104]

11. Verwendung zu **Strafverfolgungszwecken**, wenn (i) KI dazu dient, das Risiko einer Straftat oder einer erneuten Straftat einer Person nicht nur auf der Grundlage eines (automatisierten) Profils zu bewerten, (ii) KI dazu dient, persönliche Merkmale oder Eigenschaften oder frühere kriminelle Verhaltensweisen von Personen oder Gruppen zu bewerten, oder (iii) KI zur Erstellung von Profilen von Personen im Rahmen der Aufdeckung, Untersuchung oder Verfolgung von Straftaten verwendet werden soll.[105]

12. Im **Migrations-, Asyl- und Grenzkontrollwesen**, soweit (i) KI als Lügendetektor oder ähnliches Werkzeug verwendet werden soll, (ii) KI zur Bewertung des Risikos von Personen, die in die EU einreisen oder dies tun wollen, verwendet werden soll, (iii) KI zur Unterstützung der Prüfung von Anträgen auf Asyl, Visa, Aufenthaltsgenehmigungen und damit zusammenhängenden Beschwerden sowie zur Bewertung der damit zusammenhängenden Beweise verwendet werden soll, (iv) KI zum Aufdeckung, Anerkennung oder Identifizierung von Personen verwendet werden soll, ausser zur Überprüfung von Reisedokumenten.[106]

13. KI wird von einer **Justizbehörde** oder in deren Auftrag oder im Rahmen einer alternativen Streitbeilegung verwendet, um die Justizbehörde bei der Ermittlung und Auslegung von Sachverhalten und Rechtsvorschriften und deren Anwendung auf einen bestimmten Fall zu unterstützen.[107]

14. KI wird zur **Beeinflussung** des Ergebnisses **einer Wahl oder einer Abstimmung** eingesetzt, jedoch nicht, wenn Personen dem Ergebnis der KI nicht unmittelbar ausgesetzt sind (z.B. KI-Systeme, die zur Organisation, Optimierung und Strukturierung der Verwaltung oder Logistik politischer Kampagnen eingesetzt werden).[108]

[53] In der Praxis privater Unternehmen von Relevanz sein werden vor allem die Fälle 2, 3, 4, 5, 7, 8 und 9 sein. Die meisten Anwendungsfälle haben wir bisher im Arbeitsbereich und Bil-

102 Art. 25 Abs. 1 lit. c AIA.
103 Anhang III, Ziff. 5 c).
104 Anhang III, Ziff. 6 a)–c).
105 Anhang III, Ziff. 6 d)–e).
106 Anhang III, Ziff. 7.
107 Anhang III, Ziff. 8 a).
108 Anhang III, Ziff. 8 b).

dungssektor gesehen. Die Vorselektion oder Bewertung von Stellenbewerbern (z.B. wie gut sie in das gesuchte Profil passen oder welche Erkenntnisse sich aus dem Lebenslauf ziehen lassen) ist eine solche Anwendung, selbst wenn letztlich Menschen die Entscheidung treffen. Extremere Anwendungsfälle haben wir im Sportbereich gesehen, wo bereits heute AIS zur Beurteilung und Ausbildung von Profi- und teilweise auch Amateur-Athleten eingesetzt werden, zu denen die betreffenden Vereine und Organisationen nicht selten ein Arbeits- oder Ausbildungsverhältnis stehen.[109] Fragen stellen sich in der Praxis immer wieder auch in Bezug auf Systeme, welche Unternehmen zur Cybersicherheit (z.B. «Endpoint Detection and Response», EDR) und «Data Loss Prevention» (DLP) einsetzen: Sie überwachen den Netzwerkverkehr und das Verhalten der Mitarbeitenden und schlagen Alarm, wenn ungewöhnliche Vorkommnisse beobachtet werden, die auf Eindringlinge, Datendiebstahl oder sonstige Regelverstösse beobachtet werden. Als Erkennung von Emotionen und Absicht gelten Massnahmen zur Cybersicherheit zwar wie erwähnt nicht,[110] jedoch ist fraglich, ob solche Systeme nicht der «Beobachtung … des Verhaltens von Personen in … Beschäftigungsverhältnissen» dienen.[111] Dies dürfte tendenziell der Fall sein, jedenfalls wenn die Ausrichtung dieser Systeme primär nach innen gerichtet ist, wie dies beispielsweise bei einem DLP ist, welches den unerwünschten Datenabfluss durch Mitarbeitende erkennen und unterbinden soll (während z.B. EDR-Systeme primär auf Anzeichen von externen Bedrohungen wie Ransomware-Angriffe und Malware achten). Den Erwägungen zufolge erfolgt die Klassifizierung als HRAIS, weil die Überwachung der Leistung und des Verhaltens von Beschäftigten «deren Grundrechte auf Datenschutz und Privatsphäre» untergraben kann,[112] was bei DLP-Systemen sein kann. Gegen die Klassifizierung als HRAIS könnte eingewendet werden, dass mit Verhalten die Beobachtung des regulären Verhaltens eines Beschäftigten gemeint ist, nicht das Aufspüren eines ganz ausnahmsweise vorkommenden Fehlverhaltens, jedenfalls wenn der Alarm nur dann anschlägt, wenn bestimmte (nicht mittels KI-ermittelte) Schwellenwerte oder Trigger überschritten sind. Zu erwähnen ist schliesslich noch, dass der AIA auch eine «de minimis»-Regelung für HRAIS kennt. Sie kommt dann zum Tragen, wenn gezeigt werden kann, dass ein AIS kein erhebliches Risiko der Beeinträchtigung in Bezug auf die Gesundheit, Sicherheit oder Grundrechte natürlicher Personen birgt, indem es unter anderem das Ergebnis einer Entscheidungsfindung nicht wesentlich beeinflusst. Dann liegt kein HRAIS vor. Art. 6(3) AIA zählt hierzu einige Fälle auf, in welchen das so sein soll, etwa wenn ein KI-System nur eine eng gefasste Aufgabe durchführt oder bestimmte abgeschlossene Tätigkeiten betrifft. Will ein Anbieter diese Regelung beanspruchen, muss er dies vor Inverkehrbringen und Inbetriebnahme dokumentieren. Der Registrierungspflicht für HRAIS unterliegt er trotzdem.

[54] Der folgende One-Pager gibt einen Überblick über diese und die weiteren Anwendungsfälle und bisher diskutierten Definitionen des AIA (nur auf Englisch):[113]

109 Vgl. dazu Sven Hintermann/David Rosenthal im Blog-Beitrag «Fair Play mit KI: Wie Sportorganisationen Athletendaten nutzen dürfen», abrufbar unter https://www.vischer.com/know-how/blog/teil-11-fair-play-mit-ki-wie-sportorganisationen-athletendaten-nutzen-duerfen/.

110 Erw. 54 AIA; siehe auch Ziff. 2 in Rn. 52.

111 Anhang III, Ziff. 4 b).

112 Erw. 57 AIA.

113 Abrufbar unter https://vischerlnk.com/ai-act-uc.

[55] Liegt ein solches HRAIS vor, hat vor allem der Provider eine ganze Reihe von Aufgaben zu erfüllen, da primär er dafür verantwortlich ist, dass die Vorgaben aus Kapitel 2 der Regelungen zu den HRAIS in Bezug auf diese erfüllt werden. Diese **Vorgaben** für HRAIS sind:

- Es muss ein eingehendes Risikomanagement betrieben werden, d.h. insbesondere Risikobeurteilungen, die über den gesamten Lebenszyklus des HRAIS wiederholt werden, mit entsprechenden Tests des HRAIS und Massnahmen, um die identifizierten Risiken zu kontrollieren (vgl. hierzu das Open-Source-Werkzeug GAIRA[114]);[115]
- Die Daten, die für Training, Validierung und Tests verwendet werden, müssen bestimmten Qualitätskriterien genügen und müssen Gegenstand einer weitreichenden Data Governance sein, um beispielweise Bias, Fehler und Lücken zu erkennen;[116]
- Es muss eine detaillierte technische Dokumentation zum HRAIS erstellt und nachgeführt werden, die auch die Erfüllung der Vorgaben belegt;[117]
- Das HRAIS muss das, was es tut, automatisch in Logs angemessen protokollieren, damit sein korrektes Funktionieren über Zeit überwacht werden kann;[118]

114 https://vischerlnk.com/gaira.

115 Art. 9 AIA; vgl. dazu Martina Arioli (Fn. 34), S. 6 ff.; vgl. ausführlich zu einem Risikomanagement über den gesamten Lebenszyklus OECD, «Advancing Accountability in AI, Governing and managing risks throughout the lifecycle for trustworthy AI», in: OECD Digital Economy Papers, February 2023, No. 349, https://doi.org/10.1787/2448f04b-en.

116 Art. 10 AIA.

117 Art. 11 AIA.

118 Art. 12 AIA.

- Das HRAIS muss so ausgestaltet sein und mit Nutzungsanweisungen und weiteren Hinweisen geliefert werden, dass seine Anwender es richtig handhaben, seinen Output richtig verstehen und einschätzen können sowie vom Menschen überwacht und kontrolliert werden kann (was u.a. umfasst, dass das HRAIS jederzeit sicher gestoppt werden kann und in einer Weise zur Verfügung gestellt wird, dass Menschen dem HRAIS nicht einfach vertrauen);[119]
- Das HRAIS muss angemessen genau und zuverlässig arbeiten (wobei vorgesehen ist, dass die Europäische Kommission die Entwicklung entsprechender Standards fördern soll), und es muss angemessen fehlertolerant sein;[120]
- Das HRAIS muss über eine angemessene Sicherheit verfügen, insbesondere, um es gegen Angriffe von Dritten zu schützen, sowohl im Bereich der klassischen Informationssicherheit als auch bei KI-spezifischen Angriffen.[121,122]

[56] Um die Einhaltung dieser Anforderungen nachzuweisen und sicherzustellen und die Aufsicht zu ermöglichen, müssen Provider:

- Eine entsprechende Dokumentation führen (bis zehn Jahre nach Inverkehrbringen oder Inbetriebnahme);[123] den zuständigen Behörden sind auf Anfrage die nötigen Nachweise zu erbringen;[124]
- Ein Qualitätsmanagementsystem vorweisen;[125] hat der Provider Grund zur Annahme, dass das HRAIS nicht den Anforderungen entspricht, sind unverzüglich die nötigen Korrekturmassnahmen zu treffen (einschliesslich Rückrufe) und darüber zu informieren;[126]
- Die von ihren HRAIS im Einsatz generierten Logs aufbewahren (soweit sie sie haben, und zwar mindestens sechs Monate);[127]
- Das HRAIS in einer EU-Datenbank registrieren (ausser jene, die im Rahmen von kritischen Infrastrukturen eingesetzt werden);[128]
- Eine Konformitätsbeurteilung durch einen entsprechenden Dritten vornehmen lassen;[129]
- Das HRAIS mit einem Konformitätszeichen («CE») und ihren Kontaktangaben versehen;[130]

119 Art. 13 f. AIA.

120 Art. 15 AIA.

121 Art. 15 Abs. 1 und 5 AIA.

122 Vgl. etwa der Blog-Beitrag von David Rosenthal «Die andere Seite der Medaille: Wo wir KI vor Angreifern schützen müssen», abrufbar unter https://www.vischer.com/know-how/blog/teil-6-die-andere-seite-der-medaille-wo-wir-ki-vor-angreifern-schuetzen-muessen/.

123 Art. 16 lit. d AIA, Art. 11 AIA, Art. 17 f. AIA.

124 Art. 21 AIA.

125 Art. 16 lit. c AIA, Art. 17 AIA.

126 Art. 20 Abs. 1 AIA.

127 Art. 16 lit. e AIA, Art. 19 AIA.

128 Art. 16 lit. i AIA, Art. 49 AIA.

129 Art. 16 lit. g AIA, Art. 47 AIA.

130 Art. 16 lit. b und lit. h AIA, Art. 48 AIA.

- Ein System betreiben, mit welchem sie das HRAIS nach dem Inverkehrbringen in Bezug auf technische Entwicklungen und Risiken beobachten können (sog. Post-Market Monitoring-System);[131] und
- Eine Untersuchung sowie eine Meldung an die Aufsichtsbehörden vornehmen, wenn von einem HRAIS ein Risiko für die Gesundheit, die Sicherheit oder die grundlegenden Rechte von betroffenen Personen ausgeht oder es zu einem ernsthaften Zwischenfall kommt.[132]

[57] Ein Provider, der selbst nicht in der EU ansässig ist, dort aber ein HRAIS anbietet, muss einen **Vertreter in der EU** bezeichnen («Bevollmächtigter»), der über die nötigen Unterlagen verfügt, damit die EU-Behörden darauf zugreifen können.[133] Speziell daran: Er ist selbstständig **verpflichtet, das Mandat niederzulegen** und die Aufsichtsbehörde zu informieren, falls er Grund zur Annahme hat, dass der Provider seinen Pflichten nach dem AIA nicht nachkommt. Auch den Importeuren und Distributoren von HRAIS kommen bestimmte Pflichten zu.[134]

[58] Auch jene, die ein HRAIS «nur» in eigener Verantwortung einsetzen, also die Deployer, haben Pflichten unter dem AIA.[135] Sie müssen insbesondere sicherstellen, dass:

- Das HRAIS gemäss der Nutzungsanweisungen verwendet wird;[136]
- Das HRAIS von qualifizierten Menschen überwacht wird;[137]
- Die Logs, die das HRAIS automatisch erstellt, mindestens sechs Monate aufbewahrt werden;[138]
- Etwaiger Input für das HRAIS im Hinblick auf den Zweck geeignet und repräsentativ ist;[139]
- Der Provider im Rahmen von dessen Post-Market-Monitoring-Systems über den Betrieb des HRAIS informiert;[140]
- Die Aufsichtsbehörde und der Provider (und ggf. seinen Händler) informiert wird, wenn Grund zur Annahme besteht, dass von einem HRAIS ein Risiko für die Gesundheit, die Sicherheit oder die grundlegenden Rechte von betroffenen Personen ausgeht – hierzu ist das HRAIS im Betrieb zu überwachen;[141] es muss mit der Aufsichtsbehörde zusammengearbeitet werden;[142]
- Beim Einsatz von HRAIS am Arbeitsplatz die betroffenen Mitarbeitenden darüber informiert sind;[143]

131 Art. 72 AIA.
132 Art. 16 lit. j AIA, Art. 20 Abs. 2 AIA, Art. 73 AIA.
133 Art. 22 AIA.
134 Art. 23 f. AIA.
135 Art. 26 AIA.
136 Art. 26 Abs. 1 AIA.
137 Art. 26 Abs. 2 AIA.
138 Art. 26 Abs. 6 UAbs. 1 AIA.
139 Art. 26 Abs. 4 AIA.
140 Art. 26 Abs. 5 UAbs. 1 AIA.
141 Art. 26 Abs. 5 UAbs. 1 AIA.
142 Art. 26 Abs. 12 AIA.
143 Art. 26 Abs. 7 AIA.

- Betroffene Personen informiert sind, wenn ein HRAIS für sie betreffende Entscheidungen verwendet wird, selbst wenn das HRAIS nur unterstützend zum Einsatz kommt (dies geht somit weiter als unter der DSGVO);[144] und
- Eine **Grundrechte-Folgenabschätzung** durchgeführt (allenfalls in Kombination mit einer Datenschutz-Folgenabschätzung) und der Aufsichtsbehörde mitgeteilt worden ist, sofern es sich beim Deployer um eine Stelle des öffentlichen Rechts, um eine private Stelle, die «öffentliche Dienste» erbringt, oder um den Betrieb eines HRAIS zur Kreditwürdigkeitsprüfung, der Berechnung eines Bonitätsscores oder der Risikobewertung oder Preisbildung im Bereich von Lebens- und Krankenversicherungen handelt;[145] Ob «öffentliche Dienste» auch rein privatrechtliche Tätigkeiten meint, ist nicht völlig klar, aber es ist aufgrund der Erwägungen davon auszugehen; Erwähnt werden Dienste, die im öffentlichen Interesse, etwa in den Bereichen Bildung, Gesundheitsversorgung, Sozialdienste, Wohnungswesen und Justizverwaltung erbracht werden.[146]

[59] Diese Pflichten sind nicht abschliessend. Der AIA gibt in einer Art **Generalklausel** selbst bei rechtskonform eingesetzten HRAIS der Marktaufsicht die Kompetenz, weitere Massnahmen anzuordnen, falls dies zum Schutz der Gesundheit oder Sicherheit der Personen, deren grundlegender Rechte oder anderer öffentlicher Interessen erforderlich sein sollte.[147]

[60] Ferner sieht der AIA analog zur DSGVO eine Art **Auskunftsrecht** für betroffene Personen in der EU (nicht aber der Schweiz) vor, wenn ein Deployer auf Basis des Outputs eines HRAIS eine Entscheidung trifft, die auf die Person rechtliche oder vergleichbar wesentliche, für ihre Gesundheit, Sicherheit oder grundlegende Rechte nachteilige Auswirkungen hat.[148] Sie kann dann verlangen, dass der Deployer ihr erklärt, welche Rolle dem HRAIS im Entscheid zugekommen ist und was die wesentlichen Elemente des Entscheids waren. Das gilt für alle HRAIS der oben aufgezählten 15 Kategorien, ausgenommen Nr. 2. Die Ansprüche gemäss DSGVO bleiben vorbehalten; sie sind insofern enger, als sie sich auf vollständig automatisierte Entscheide beziehen.

[61] Diese Pflichten können aufgrund der extraterritorialen Geltung des AIA für Deployer von HRAIS auch für Unternehmen ausserhalb der EU von Bedeutung sein, selbst wenn der Provider sich nicht in der EU befindet, der Output des HRAIS aber in der EU verwendet wird (siehe oben). Ob und wie gut diese Pflichten beispielsweise gegen **Unternehmen in der Schweiz durchgesetzt** werden können und werden, ist eine andere Frage. Es ist – wie schon im Falle der DSGVO – eher **nicht damit zu rechnen**, da dem auch rechtliche Schranken im Schweizer Recht entgegenstehen.[149] Stellvertretende Sanktionen gegen den korrekt handelnden EU-Vertreter dürften dagegen nicht möglich sein.

144 Art. 26 Abs. 11 AIA.

145 Art. 27 AIA; vgl. dazu Martina Arioli (Fn. 34), S. 14 f.

146 Erw. 96 AIA.

147 Art. 82 Abs. 1 AIA.

148 Art. 86 AIA.

149 Art. 271 StGB, welcher grenzüberschreitende Zwangsmassnahmen auf Schweizer Boden ohne Rechts- oder Amtshilfe untersagt – und ein Folgeleisten.

8. Regelungen für KI-Modelle

[62] Erst spät im Gesetzgebungsprozess wurde der AIA auch um Regelungen für Provider bestimmter KI-Modelle erweitert (die Rolle des Deployers gibt es hier nicht). Allerdings sind nur für unterschiedlichste Zwecke einsetzbare Modelle («general purpose», «KI-Modelle mit allgemeinem Verwendungszweck») erfasst (**GPAIM**).[150] Wie genau sich diese gegenüber anderen Modellen abgrenzen, ist allerdings nicht klar. So könnte vertreten werden, dass ein Modell nur zur Transkription wie «Whisper» oder eines nur für Übersetzungen kein GPAIM ist. Ist ein GPAIM durch *Fine-Tuning* auf bestimmte Themen wie z.B. rechtliche Fragen «spezialisiert» worden, wird es aber vermutlich weiterhin als GPAIM gelten, weil es sich nach wie vor für viele unterschiedliche Anwendungen einsetzen lässt.

[63] Aus den Erwägungsgründen geht hervor, dass GPAIM nicht nur in ihrer eigentlichen Form, d.h. als Dateien, erfasst sind, sondern auch dort, wo sie über eine API (Anwendungsprogrammierschnittstelle) quasi als «Model-as-a-Service» angeboten werden.[151] Sie gelten gemäss den Erwägungen zudem nicht als AIS, d.h. auch die entsprechenden Pflichten greifen ins Leere, weil ihnen eine Benutzerschnittstelle fehlt (warum ein API nicht als solche gelten soll, wird nicht erklärt). Wenn also OpenAI nebst «ChatGPT» den Zugang zu GPT4 auch über ein API anbietet, muss sie die AIA-Vorgaben für AIS im Rahmen dieses API nicht beachten, während «ChatGPT» als AIS gilt und die Vorgaben zu beachten sind. Den Vorgaben für AIS ist dann derjenige unterworfen, der das Modell in seiner Anwendung nutzt (soweit sie als AIS gilt); er wird dann zu deren Provider, weil er die Anwendung entwickelt hat und er sie für sich einsetzt. Bringt er diese Anwendung im EU-Markt in Verkehr oder nimmt er sie bei sich in der EU in Betrieb, dürfte allerdings wohl auch das Modell als in Verkehr gebracht gelten.

[64] Wer ein GPAIM erstellt, aber nur intern nutzt, ist von den Vorgaben für GPAIM nicht erfasst, weil er erst dann als Provider gilt, wenn er das GPAIM in Verkehr bringt; das Kriterium der Inbetriebnahme für eigene Zwecke («putting into service») greift wohl wegen eines gesetzgeberischen Versehens nur bei AIS.[152] Die Ausnahmen hätten gemäss den Erwägungsgründen enger gefasst sein sollen.

[65] Provider von GPAIM haben insbesondere folgende Pflichten:

- Sie müssen zu Händen der Aufsichtsbehörden eine detaillierte **technische Dokumentation** des Modells führen und nachführen;[153]
- Sie müssen zu Händen der **Nutzer** des GPAIM (d.h. den Providern von AIS) eine weitaus weniger detaillierte **Dokumentation** des Modells führen und nachführen;[154]
- Sie müssen, falls sie sich ausserhalb der EU befinden, einen **Vertreter** in der EU benennen («Bevollmächtigter»);[155]
- Sie müssen interne Regeln (d.h. eine «Strategie») zur **Einhaltung des EU-Urheberrechts** einführen, einschliesslich der dessen sog. Text- and Data-Mining-Regelung (**TDM**) und de-

150 Kapitel 5 AIA.

151 Erw. 97 AIA.

152 Art. 3 Ziff. 3 AIA.

153 Art. 53 Abs. 1 lit. a AIA.

154 Art. 53 Abs. 1 lit. b AIA.

155 Art. 54 AIA.

ren Opt-out-Recht.[156] Die TDM-Regelung erlaubt Dritten im Sinne einer Schrankenbestimmung die Entnahme und Nutzung von Inhalten aus Datenbanken, zu denen sie rechtmässig Zugang haben, gibt den Inhabern der Rechte der darin enthaltenen Werke jedoch ein «opt-out»-Recht in Bezug auf Anwendungen ausserhalb der wissenschaftlichen Forschung.[157] Es ist schon unabhängig vom AIA nicht völlig klar, wie wichtig die TDM-Regelung für das Trainieren von KI-Modellen tatsächlich ist und wie genau sie operativ umzusetzen ist, insbesondere vor dem Hintergrund der Tatsache, dass sie national unterschiedlich umgesetzt ist.[158] Die Tragweite der Vorgabe des AIA ist ebenfalls nicht völlig klar.

Nach einer strengen Lesart könnte es mit der neuen Bestimmung noch schwieriger werden, KI-Modelle zu trainieren, denn wer später ein GPAIM in der EU anbieten will, müsste das EU-Urheberrecht beim Training seines Modells auch dann einhalten bzw. entsprechende Weisungen vorsehen, wenn er diesem Recht gar nicht unterliegt (so kann beispielsweise das Schweizer Urheberrecht oder das US-Recht mehr oder andere Schutzaufnahmen bieten, die ein Training von KI-Modellen gegenüber dem EU-Recht leichter möglich macht). Es würde verhindert werden, dass z.B. ein US-Anbieter sein Modell unter einem weniger strengen ausländischen Recht aufbaut und dann in der EU auf den Markt bringt. Dem EU-Urheberrecht würde damit für die Zwecke von GPAIM mit anderen Worten faktisch weltweite Geltung verschafft, da die meisten Anbieter von GPAIM diese auch in der EU anbieten wollen. Die Regelung würde somit EU-Anbieter von Allzweck-KI-Modellen vor einer Benachteiligung durch ausländische Anbieter schützen.

Nach einer weniger strengen Lesart wäre genau das Gegenteil der Fall. Das EU-Urheberrecht müsste demnach nur insoweit eingehalten werden, als es auch tatsächlich angewandt werden will. Nach dem sog. Schutzlandprinzip wäre dies nur dann der Fall, wenn das Training des KI-Modells in der EU erfolgt, nicht aber beispielsweise in der Schweiz oder in den USA, wo für die lokal erfolgten Verwendungen von Werken das lokale Recht zur Anwendung käme, das weniger streng ist als in der EU. Das EU-Urheberrecht käme nur aber immerhin dann zur Anwendung, wenn das KI-Modell später in der EU eingesetzt wird und dort einen Output generiert (der dann das EU-Urheberrecht nicht verletzen darf, wobei sich die Frage stellt, ob dies nicht einzig Sache des Verwenders des Modells ist). Diese Lesart überzeugt: Sie kann aus dem Text des AIA abgeleitet werden, der nur eine Strategie zur Einhaltung des EU-Urheberrechts verlangt, und nicht dessen *sinngemässe* Einhaltung ausserhalb der EU. Diese Lesart respektiert auch das international anerkannte Schutzlandprinzip.[159] Die geforderte Strategie bestünde somit darin sicherzustellen, dass kein Training des KI-Modells in der EU erfolgt und so das dortige Urheberrecht auch nicht verletzt wird. Die Erwägungen machen jedenfalls den Eindruck, dass es vor allem darum ging, die Stellen, die in der EU trainieren und sich dabei auf die TDM-Schrankenbestimmung stützen, angemessene Massnahmen zur Berücksichtigung von Opt-outs vornehmen.[160]

156 Art. 53 Abs. 1 lit. c AIA.

157 Die Schweiz kennt mit Art. 24d Urheberrechtsgesetz eine ähnliche, aber nicht identische Regelung. Sie erlaubt die Nutzung von Werken für die wissenschaftliche Forschung, ohne ein Opt-out-Recht zu statuieren.

158 Vgl. Ansgar Kaiser, Der fehlende Werkgenuss beim Text and Data Mining, in: Florent Thouvenin/Alexander Peukert/Thomas Jaeger/Christophe Geiger (Hrsg.), Kreation Innovation Märkte – Creation Innovation Markets, Festschrift Reto M. Hilty, Berlin/Heidelberg, 2024, S. 256, 259.

159 In der Schweiz in Art. 110 Abs. 1 IPRG festgehalten.

160 Erw. 105 AIA.

So oder so ist allerdings nicht ganz klar, was mit der etwas ungewöhnlichen Regelung auf sich hat, eine «Strategie … für die Einhaltung … auf den Weg» zu bringen. Dies kann so verstanden werden, dass zwar Bemühungen unternommen werden müssen, das (anwendbare oder nicht anwendbare) EU-Urheberrecht einzuhalten, diese Strategie aber nicht zwingend erfolgreich sein muss. Naheliegend wäre gewesen, eine Vorgabe zu formulieren, wonach ein Nachweis zu erbringen ist, dass das EU-Urheberrecht eingehalten worden ist. Das geschah nicht. Erste Rechteinhaber haben übrigens bereits Opt-out-Erklärungen öffentlich abgegeben;[161]

- Sie müssen öffentlich in summarischer Form («hinreichend detaillierte Zusammenfassung») darlegen, **welche Inhalte** sie für das Training ihrer Modelle benutzt haben.[162] Es muss somit nicht im Einzelnen belegt werden, welche Inhalte für das Training verwendet worden sind; und
- Sie müssen mit der Europäischen Kommission und den nationalen **Aufsichtsbehörden zusammenarbeiten**.[163]

[66] Der AIA entbindet zwar GPAIM unter einer **freien, offenen Lizenz** von den ersten drei Pflichten, nicht aber von den beiden letzten. Zudem wird der Begriff «Open Source» eng angelegt.[164]

[67] Für Provider von GPAIM, die ein «systemisches Risiko» mit sich bringen, gelten noch zusätzliche Vorgaben.[165] Aus Sicht des Gesetzgebers bringt ein GPAIM dann ein **systemisches Risiko** mit sich, wenn es auf die öffentliche Gesundheit, die Sicherheit, die Grundrechte oder sonst die Gesellschaft in der EU erhebliche Auswirkungen haben kann.[166] Das soll dann der Fall sein, wenn es als solches etwa von der Europäischen Kommission eingestuft wird oder sich der hohe Wirkungsgrad z.B. über entsprechende Benchmarks oder dergleichen zeigt. Letzteres wird dann angenommen, wenn für dessen Erstellung eine Rechenleistung von 10 hoch 25 FLOPS (= «Rechenoperationen mit Bruchzahlen») oder mehr zum Einsatz gekommen ist.[167] Die grossen LLM wie etwa GPT4 von OpenAI dürften zwar nach verbreiteter Annahme als solche GPAIM mit systemischen Risiken gelten, doch die genauen Zahlen sind nicht bekannt. Wir schätzen, dass sie noch knapp unter der genannten rechnerischen Grenze liegen. Der AIA ist in dieser Hinsicht allerdings so oder so offen formuliert; auch «kleinere» GPAIM können wie erwähnt zu GPAIM mit systemischem Risiko erklärt werden. Deren Provider müssen dann zusätzlich ihre Modelle im Hinblick auf die damit verbundenen Risiken evaluieren, entsprechende Massnahmen zur Behandlung der Risiken treffen und ernsthafte Zwischenfälle («serious incidents») erheben, do-

161 So etwa am 16. Mai 2024 die Sony Music Group, abrufbar unter https://www.sonymusic.com/sonymusic/declaration-of-ai-training-opt-out/; vgl. ausführlich zur urheberrechtlichen Einordnung des Trainings generativer KI-Modelle Sandra Marmy-Brändli/Isabelle Oehrli, «Das Training künstlicher Intelligenz», in: sic! Das Training künstlicher Intelligenz, 2023. S. 655–666.

162 Art. 53 Abs. 1 lit. d AIA.

163 Art. 53 Abs. 3 AIA.

164 Art. 53 Abs. 2 AIA.

165 Art. 55 AIA.

166 Art. 3 Ziff. 65 AIA.

167 Art. 51 Abs. 1 und 2 AIA.

kumentieren und den Aufsichtsbehörden melden, einschliesslich der diesbezüglich möglichen Massnahmen.[168] Ferner haben sie für eine angemessene Cybersicherheit der GPAIM zu sorgen.[169]

9. Weitere Pflichten für Provider und Deployer

[68] Der AIA definiert auch einige fallspezifische **Transparenzpflichten**, die für alle AIS gelten sollen, auch solche die keine HRAIS sind.[170] Insbesondere werden Pflichten zur Markierung und Kennzeichnung bestimmter KI-generierter Inhalte eingeführt.

[69] Provider müssen sicherstellen, dass:

- Personen, die mit einem AIS interagieren und dieses dafür vorgesehen ist, über eben diesen Umstand informiert werden, ausser es ist dies aus Sicht eines verständigen Anwenders offenkundig.[171] Ein typischer Verstoss gegen diese Regel können Chatbots sein, welche z.B. durch entsprechende Namen und Bilder den Eindruck erwecken, die Personen würden sich mit einem realen Menschen unterhalten. Relevant sein kann die Vorgabe aber auch bei Offline-Kommunikation, wie etwa bei Bots, welche automatische E-Mail-Antworten generieren, die nicht als solche gekennzeichnet sind.
- die von einem AIS synthetisch generierten Text-, Ton-, Video- und Bildinhalte als KI-generiert oder KI-manipuliert markiert sind.[172] Diese Markierung muss maschinenlesbar sein (aber nicht zwingend auch vom Menschen ohne Hilfsmittel erkennbar[173]). Es ist noch nicht klar, wie solche Markierungen bei Textinhalten erfolgen sollen (während es sehr viel einfacher ist, Bilder mit entsprechenden «Wasserzeichen» zu versehen). Das Anbieten von passenden «KI-Content-Detektoren» ist zwar keine Pflicht, aber solche werden sicher bald angeboten werden zur Identifizierung solcher Wasserzeichen bzw. «Watermarkings», die hoffentlich zuverlässiger sind als das, was heute auf dem Markt ist. Da sich der Gesetzgeber offenbar der Tatsache bewusst war, dass seine Definition von AIS sehr breit ist, hat er eine Ausnahme für AIS formuliert, wonach AIS, welche lediglich eine «unterstützende Funktion» bei normalen Bearbeitungsvorgängen haben, den Input oder seine Bedeutung nicht wesentlich verändern oder sie zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten gesetzlich zugelassen sind. Anbieter von automatischen Übersetzungsdiensten wie «DeepL» werden von dieser Ausnahme zu profitieren versuchen; sie basieren vollständig auf generativer KI. Ein weiteres Beispiel wäre ein Werkzeug, dass die Auflösung von Bildern mittels KI erhöht. Genau genommen ist die Pflicht zur Markierung von KI-Output aber so formuliert, dass sie auf fast jedes AIS zur Anwendung gebracht werden müsste, da fast jedes AIS in irgendeiner Form Text, Ton oder Bilder ausgibt (z.B. ein Text- oder Spracherkennungsprogramm, eine KI, welche Defekte an Fabrikaten aufspürt oder ein Recommender-System, das einem Käufer in einem Online-Shop weitere Produkte

168 Art. 55 AIA.

169 Art. 55 AIA.

170 Art. 50 AIA.

171 Art. 50 Abs. 1 AIA.

172 Art. 50 Abs. 2 AIA.

173 Erw. 133 AIA, welche Metadatenidentifizierungen, kryptographische Methoden und andere Techniken als zulässige Methoden vorsieht, die zweifellos nicht ohne Hilfsmittel interpretiert werden können.

empfiehlt). Das war natürlich nicht die Absicht des Gesetzgebers, der klar Output generativer KI im Blick hatte, und zwar Inhalte, bei welchen der Mensch sonst möglicherweise nicht erkennen würde, dass sie von einer Maschine stammen.[174] In diesem Sinne sollte die Bedingung, dass die Inhalte «synthetisch» sein müssen, nach dem Sinn und Zweck der Regelung dahingehend zu verstehen ist, dass sie den Inhalten eines Menschen nachempfunden sind (also nicht «menschlich»); gemeint ist also der Output generativer KI. Der von einem AIS berechnete Bonitätsscore eines Kunden wäre zwar auch ein Textinhalt, aber kein «synthetischer» im Sinne dieser Bestimmung. Schliesslich sei noch der Hinweis angebracht, dass die Pflicht zum *Watermarking* nur den Provider betrifft: Anwender (und Deployer) sind frei, solche Wasserzeichen zu entfernen.

[70] Deployer müssen sicherstellen, dass:

- betroffene Personen informiert werden, wenn sie bei diesen AIS zur Erkennung von Emotionen oder Absichten oder zur Kategorisierung anhand biometrischer Merkmale eingesetzt werden und dabei Personendaten bearbeitet werden.[175] Weil die Emotionserkennung auch hier auf Basis biometrischer Daten erfolgen muss (dies ergibt sich aus der Defition des Begriffs des Emotionserkennungssystems[176]), ist das in der Praxis häufige Anwendungsbeispiel der Sentiments-Analyse, die auf Text statt Stimme basiert, nicht erfasst.[177]
- von ihnen generierte «**Deep Fakes**» als solche erkennbar sind; für Kreativschaffende gibt es allerdings eine Ausnahme: Wenn *Deep Fakes* offenkundig Teil von Kunst, Literatur, Satire und dergleichen sind, darf der Hinweis so eingeschränkt werden, dass die Darstellung und der Genuss des Werks nicht beinträchtigt wird.[178] Es wird spannend zu sehen sein, inwiefern diese Ausnahme auch auf Werbung angewendet wird. Die Bestimmung soll übrigens nicht jene schützen, über die ein *Deep Fake* hergestellt wird (diese können sich gestützt auf das Datenschutzrecht oder Spezialbestimmungen[179] wehren), sondern jene, die damit getäuscht werden könnten: *Deep Fakes* werden nämlich nicht verboten, sondern müssen nur als solche ausgewiesen werden.[180]
- in publizierten Texten, die Themen von öffentlichem Interesse betreffen, darauf hingewiesen wird, dass diese KI-generiert oder manipuliert worden sind – es sei denn, der Text wurde von einem Menschen geprüft und ein Mensch oder eine juristische Person hat für dessen Publikation die redaktionelle Verantwortung übernommen.[181]

[71] Die vorstehenden Pflichten gelten nur für AIS, nicht für GPAIM. Wenn also ein Unternehmen wie OpenAI ein AIS wie «ChatGPT» anbietet, dann muss sie dafür sorgen, dass KI-Inhalte als solche maschinenlesbar markiert sind. Bietet sie den Zugang zu ihren Modellen via API (d.h. über

174 Erw. 133 AIA.

175 Art. 50 Abs. 3 AIA.

176 Art. 3 Ziff. 39 AIA.

177 Vgl. Rn. 52, Ziff. 2.

178 Art. 50 Abs. 4 UAbs. 1 AIA.

179 In der Schweiz z.B. Art. 129^{decies} AIA.

180 Erw. 134 AIA.

181 Art. 50 Abs. 4 UAbs. 2 AIA.

eine Schnittstelle für Computerprogramme) an, so muss sie dies nicht tun. Die Pflicht obliegt dann dem Unternehmen, das dieses API in einer eigenen Anwendung nutzt und dadurch zu einem AIS wird – falls dieses Unternehmen als Provider im Sinne des AIA gilt und in dessen Geltungsbereich fällt, weil es sein AIS in der EU in Verkehr bringt oder in Betrieb nimmt.

[72] Fällt ein AIS nicht unter diese Bestimmungen, ist es auch kein HRAIS und geht es nicht um eine verbotene Praktik, dann bestehen unter dem AIA grundsätzlich keine Pflichten für Provider, Deployer und die weiteren beteiligten Stellen bezüglich des Betriebs dieses AIS. Sie werden im AIA jedoch grundsätzlich ermuntert, sich freiwillig an **Verhaltenskodizes** zu halten, die dereinst im Bereich von AIS entstehen sollen, um ethische Grundsätze, einen sorgsamen Umgang mit der Umwelt[182], die Medienkompetenz im Umgang mit KI, Diversität und Inklusion sowie die Vermeidung negativer Auswirkungen auf vulnerable Personen sicherzustellen.[183] Ferner sieht ein erst im Laufe der Beratungen eingefügter Artikel eine Art **KI-Ausbildungspflicht** für Provider und Deployer vor, d.h. sie müssen nach besten Kräften dafür sorgen, dass die Personen, die mit AIS umgehen, angemessen darin ausgebildet, über die anwendbaren Pflichten gemäss AIA informiert sind und den Personen die Chancen und Risiken von AIS bewusst sind («AI literacy»).[184] Dies gilt für alle AIS, nicht nur HRAIS. Ob und inwieweit diese Vorgabe durchgesetzt wird, wird sich zeigen.

[73] Damit geht der AIA für den einen oder anderen wohl erstaunlich wenig weit in der Regulierung des Einsatzes von AIS, die weder ein HRAIS sind, noch eine verbotene Praktik betreffen – also in der Praxis in der Mehrheit der Fälle. Dies, obwohl in zahlreichen internationalen Initiativen, Erklärungen und sogar in der KI-Konvention des Europarats[185] immer wieder Vorgaben formuliert worden sind, die als wichtig für den verantwortungsvollen Einsatz von KI bezeichnet wurden, wie etwa den Grundsatz der Transparenz, der Nichtdiskriminierung, der Selbstbestimmung, der Fairness, der Verhinderung von Schaden, der Robustheit und Zuverlässigkeit von AIS, der Erklärbarkeit von KI und der menschlichen Aufsicht[186]. Bis auf ausgewählte Aspekte der Transparenz verlangt der AIA ihre Einhaltung zwingend nur im Bereich der HRAIS, und auch dort nicht wirklich konsequent und primär von Providern; er erlaubt wie gezeigt nicht einmal die Verarbeitung von besonderen Kategorien von personenbezogenen Daten, um ein «normales» AIS auf Bias zu testen und solchen zu eliminieren (zumal der AIA dies auch nicht verlangt). Es wird sich zeigen, ob die Durchsetzung dieser Grundsätze auf anderem Wege erfolgen (beispielsweise über das Datenschutz- oder Lauterkeitsrecht), über Leitlinien, die sich Organisationen freiwillig auferlegen, oder gar nicht.

182 Vgl. ausführlich OECD, «Measuring the environmental impacts of artificial intelligence compute and applications. The AI footprint», in: OECD Digital Economy Papers, November 2022, No. 341, S. 22 ff, https://doi.org/10.1787/7babf571-en.

183 Art. 95 Abs. 2 AIA.

184 Art. 4 AIA.

185 Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, verabschiedet am 17. Mai 2024, abrufbar unter https://rm.coe.int/1680afae3c.

186 Vgl. dazu auch Jan Pohle, «Innovativ – und jetzt? Der Verordnungsentwurf der Europäischen Kommission zur Regulierung künstlicher Intelligenz – eine Herausforderung für die Unternehmenscompliance», in Compliance-Berater, 2021, S. 374.

10. Anwendungsbeispiele

[74] In der Folge haben wir einige praktische Anwendungsbeispiele aufgeführt, um zu zeigen, auf wen welche Bestimmungen des AIA im Alltag zur Anwendung gelangen können:[187]

Fall	**Provider im Geltungsbereich des AIA***	**Deployer im Geltungsbereich des AIA***
Ein Unternehmen in der EU stellt Mitarbeitenden «ChatGPT» oder «Copilot» zur Verfügung. Sie verwenden es, um E-Mails, Vorträge, Blog-Beiträge, Zusammenfassungen, Übersetzungen und andere Texte zu erstellen und Bilder zu generieren.	Nein	Ja
Das Unternehmen befindet sich in der Schweiz. Es ist geplant, dass auch Personen in der EU die KI-generierten Inhalte erhalten (z.B. als E-Mails oder Texte auf der Website).	Nein	Ja
Ein Unternehmen in der EU hat ein unternehmenseigenes Chat-Tool basierend auf einem LLM entwickeln lassen und setzt es intern ein, um E-Mails, Vorträge, Blog-Beiträge, Zusammenfassungen, Übersetzungen und andere Texte zu erstellen und Bilder zu generieren.	Ja	Ja
Das Unternehmen befindet sich in der Schweiz. Es ist geplant, dass auch Personen in der EU die KI-generierten Inhalte erhalten (z.B. als E-Mails oder Texte auf der Website).	(Ja)[a)]	Ja
Ein Unternehmen in der EU setzt eine im Markt angebotene Fachanwendung bei sich zur automatischen Vorselektion von Online-Stellenbewerbungen ein.	Nein	Ja, HRAIS
Ein Unternehmen in der EU nutzt «ChatGPT» oder «Copilot», um Unterlagen von Stellenbewerbenden auf etwaige Probleme hin zu analysieren. Die Ergebnisse bleiben intern.	Ja[b)], HRAIS	Ja, HRAIS

[187] Die Beispiele sind als generalisierte, simplifizierte Illustration des Konzepts zu verstehen, d.h. im konkreten Einzelfall kann die Beurteilung eine andere sein.

Das Unternehmen befindet sich in der Schweiz und es geht nur um Arbeitsstellen in der Schweiz, wobei auch Bewerbende aus der EU berücksichtigt werden.	Nein	Nein[c)]
Ein Unternehmen in der EU stellt auf seiner Website einen selbsterstellten Chatbot zur Beantwortung von allgemeinen Anfragen zum Unternehmen bereit.	Ja	Ja
Das Unternehmen befindet sich in der Schweiz. Die Website richtet sich auch an Personen in der EU.	(Ja)[a)d)]	Ja
Ein Unternehmen in der EU nutzt das Produkt bzw. den Service eines Dritten, um den Chatbot auf seiner Website zu realisieren. Diesem werden Inhalte des Unternehmens in der Form einer Datenbank zur Verfügung gestellt (RAG), ein *Fine-Tuning* gibt es nicht. Das Unternehmen gibt nicht an, von wem der Chatbot stammt.	(Nein)[e)]	Ja
Das Unternehmen gibt auf der Website an, von wem der Chatbot stammt («powered by . . . »).[188]	Nein	Ja
Ein Unternehmen in der EU setzt lokal ein LLM ein, um Texte zu transkribieren. Das Python-Skript, um es zu nutzen, überträgt es von einer kostenlosen Vorlage aus dem Internet unverändert in den eigenen Computer.	(Nein)[f)]	Ja
Ein Unternehmen in der EU setzt einen auch für Kunden in der EU angebotenen Service eines US-Dienstleisters ein, um damit Avatare für Schulungsvideos zu generieren.	Nein	Ja
Das Unternehmen ist in der Schweiz. Die Schulungsvideos sind nur für den Einsatz in der Schweiz gedacht.	Nein	Nein

Bemerkungen:

a) Der Wortlaut der Definitionen spricht gegen eine Qualifikation als Provider, aber der Gesetzgeber wollte das Unternehmen in diesem Fall als Provider, für den der AIA gilt, erfassen, weil der Output in der EU genutzt wird.[189]

b) Es kommt hier die Spezialregelung zum Tragen, wonach ein AIS so genutzt wird, dass es zu einem HRAIS wird; in diesem Fall wird der Anwender zum Provider.[190]

c) Es liegt keine Verwendung der Outputs der KI in der EU vor, da es um Stellen in der Schweiz geht; die Nationalität oder Herkunft der Stellenbewerbenden darf richtigerweise keine Rolle spielen. Die Herkunft der Stellenbewerber (nicht die Nationalität) wäre nur dann relevant, wenn ihnen die Outputs der KI zugesandt werden (was hier nicht der Fall ist).

d) Der Chatbot ist ein AIS, das auch von Personen genutzt werden soll, die sich in der EU befinden. Damit könnte vertreten werden, dass das AIS auch in der EU benutzt bzw. wird hierfür

188 Vgl. Rn. 22 e contrario.

189 Vgl. Rn. 24.

190 Vgl. Rn. 25.

bereitgestellt wird (zum Eigengebrauch, weil die Personen in der Regel keine Deployer sein werden) und somit eine Inbetriebnahme vorliegt.[191]

e) Die Parametrisierung, Anbindung einer Datenbank und das Einbinden des AIS in die Website sollte richtigerweise noch nicht als Entwicklung gelten; klar ist die Rechtslage allerdings noch nicht.[192] Das AIS wird allerdings unter dem eigenen Namen des Unternehmens in Betrieb genommen.

f) Es lässt sich vertreten, dass die Übernahme des Skripts dem Installieren einer schon fertig entwickelten Software gleichkommt und daher kein Entwickeln darstellt; zudem wird das Unternehmen es nicht unter dem eigenen Namen implementieren. Wird das Skript geändert, könnte die Qualifikation anders sein.

11. Durchsetzung

[75] Der AIA setzt zur Durchsetzung seiner Vorgaben diverse Behörden auf Ebene der einzelnen Mitgliedstaaten und der EU ein. So sollen GPAIM von der Europäischen Kommission beaufsichtigt werden, während die Aufsicht in Bezug auf AIS bei den Mitgliedstaaten liegt.[193] Dabei wird wiederum zwischen dem System der Konformitätsbeurteilungen und -erklärungen und der eigentlichen Marktaufsicht unterschieden. Das System wird noch komplizierter durch den Umstand, dass dort, wo bereits eine Marktaufsicht besteht (bei regulierten Produkten und der Finanzindustrie), die Aufsicht grundsätzlich weiterhin durch die bisherigen Behörden erfolgen soll. Basiert ein AIS wiederum auf einem GPAIM desselben Providers (z.B. «ChatGPT»), so ist das neu zu schaffende, zentrale «AI Office» für die Marktaufsicht zuständig (es ist Teil der Europäischen Kommission). Wird ein solches AIS wiederum in einer Weise eingesetzt, dass es als HRAIS gelten muss, sind die nationalen Marktaufsichtsbehörden zuständig.[194]

[76] Den Marktaufsichtsbehörden werden durch den AIA weitreichende Untersuchungs- und Eingriffsbefugnisse eingeräumt. Das geht soweit, dass sie auch die Herausgabe des *Source-Codes* von AIS verlangen können.[195] Sie müssen ein AIS jedenfalls dann untersuchen, wenn sie Grund zur Annahme haben, dass ein Risiko für die Gesundheit, die Sicherheit oder die grundlegenden Rechte der betroffenen Personen besteht oder ein AIS fälschlicherweise nicht als HRAIS klassifiziert ist.[196] Auch formelle Verstösse (fehlende Deklaration etc.) müssen verfolgt werden, und selbstverständlich können beliebige betroffene Personen in der EU einen Verstoss gegen den AIA ihnen zur Anzeige bringen.[197]

---

191 Vgl. Rn. 41.

192 Vgl. Rn. 35.

193 Art. 88 AIA; Art. 74 AIA, Art. 79 Abs. 2 AIA, Art. 99 Abs. 1 AIA; zur Bedeutung der Zusammenarbeit einzelner nationaler Behörden bezüglich der Umsetzung des AI Acts siehe auch Fabian Teichmann/Sonia Boticiu, «Insights into the regulatory challenges of generative artificial intelligence», in: Jusletter 7. August 2023, S. 6.

194 Art. 75 AIA; das EDPB findet, die EU-Mitgliedstaaten sollten die Marktaufsicht über am besten alle Hochrisiko-Systeme den Datenschutzbehörden übertragen: https://www.edpb.europa.eu/news/news/2024/edpb-adopts-statement-dpas-role-ai-act-framework-eu-us-data-privacy-framework-faq_en (Stand 17. Juli 2024).

195 Art. 74 Abs. 13 AIA.

196 Art. 79 Abs. 2 AIA, Art. 80 Abs. 1 AIA.

197 Art. 83 Abs. 1 AIA, Art. 85 AIA.

[77] Die nationalen Marktaufsichtsbehörden sind naturgemäss auf ihr Territorium beschränkt. Unklar ist, welche Zuständigkeiten sie in Bezug auf Provider und Deployer in EU-Drittländern wie der Schweiz und den USA haben. Geraten die Marktaufsichtsbehörden der einzelnen Mitgliedstaaten in Bezug auf die zu treffenden Massnahmen miteinander in den Clinch, entscheidet die Europäische Kommission.

[78] Der AIA sieht selbstverständlich auch administrative Bussen vor. Wie unter der DSGVO sollen sie «wirksam, verhältnismäßig und abschreckend» sein.[198] Insgesamt wirkt der Strafrahmen etwas milder als jener der DSGVO. Lediglich der Strafrahmen für verbotene KI-Praktiken geht mit 7% des weltweiten jährlichen Umsatzes oder, falls höher, EUR 35 Millionen deutlich darüber hinaus.[199] Ansonsten liegt er bei maximal 3% oder EUR 15 Millionen oder in gewissen Fällen noch tiefer.[200] Anders als unter der DSGVO wird betont, dass sie die Interessen von KMU einschliesslich Start-ups berücksichtigen sollen, was sich unter anderem daran zeigt, dass bei diesen der Strafrahmen am jeweils niedrigeren Wert ermittelt werden soll, wenn es um AIS geht.[201] Wie unter der DSGVO kann und soll auch die fahrlässige Verletzung des AIA gebüsst werden.[202]

12. Übergangsbestimmungen

[79] Der AIA trat am 20. Tag nach seiner Publikation im Amtsblatt in Kraft, also am 1. August 2024.[203] Seine Regelungen gelten grundsätzlich nach einer Übergangsfrist von 24 Monaten, d.h. ab dem 2. August 2026. Es sind jedoch folgende Ausnahmen vorgesehen:

- Die verbotenen KI-Praktiken sind bereits nach sechs Monaten verboten und ab diesem Zeitpunkt gilt auch die Pflicht zur Ausbildung von Mitarbeitenden («AI Literacy»), d.h. ab dem 2. Februar 2025;[204]
- Die Regelungen zu GPAIM und den Stellen, die Konformitätserklärungen ausstellen können, gelten bereits nach zwölf Monaten, d.h. ab dem 2. August 2025, ebenso die Regelungen zur Schaffung der im AIA Act vorgesehenen Behörden und Gremien (wie z.B. das AI Office) sowie die Sanktionsbestimmungen;[205]
- Die Regelungen zu HRAIS aufgrund bestehender EU-Produkteregulierungen gelten für diese erst nach 36 Monaten, d.h. ab dem 2. August 2027.[206]

198 Art. 99 Abs. 1 AIA.
199 Art. 99 Abs. 3 AIA.
200 Art. 99 Abs. 4 AIA.
201 Art. 99 Abs. 1 und Abs. 6 AIA.
202 Art. 99 Abs. 7 lit. i AIA.
203 Art. 113 Abs. 1 AIA.
204 Art. 113 Abs. 3 lit. a AIA.
205 Art. 113 Abs. 3 lit. b AIA.
206 Art. 113 Abs. 3 lit. c AIA.

[80] Die Europäische Kommission hat im Rahmen des «AI Pact» die Wirtschaft aufgerufen, den AIA bereits vorgängig umzusetzen.[207] Wir haben den Eindruck, dass etliche Unternehmen diesem Aufruf durchaus auch nachkommen möchten.

[81] In übergangsrechtlicher Hinsicht gilt für verbotene KI-Praktiken, die zum Zeitpunkt des Inkrafttretens des AIA bestanden haben, keine Ausnahme.[208]

[82] Für HRAIS, die schon vor dem AIA auf den Markt gebracht wurden, gelten die neuen Regelungen nur und erst, wenn diese in wesentlicher Weise angepasst werden.[209] Für GPAIM, die schon vor dem AIA auf den Markt gebracht wurden, müssen die Vorgaben des AIA innert zwei Jahren nach Inkrafttreten der jeweiligen Regeln erfüllt werden.[210]

13. Schlussbemerkung und Handlungsempfehlung

[83] Ob der AI Act die hohen Erwartungen, die an diese KI-Regulierung gestellt werden, gerecht werden wird, muss sich noch zeigen. Der Erlass macht den Eindruck, dass der Gesetzgeber versucht hat, ein Arsenal an Abwehrmitteln gegen einen Feind in Position zu bringen, den er noch gar nicht wirklich kennt und von dem er auch nicht weiss, ob und wie er mit welcher Wirkung zuschlägt.

[84] Es wird betont, dass der AI Act risikobasiert ausgestaltet ist. Dies scheint in der Tat der Fall zu sein, zumindest wenn man bedenkt, dass die Regulierung des Einsatzes «normaler» KI-Anwendungen (vielleicht überraschenderweise) sehr zahm bleibt und nur sehr wenige und spezifische «harte» Bestimmungen vorsieht. Wenn es aber um die als besonders riskant erachteten Anwendungen geht, wird aus dem Vollen geschöpft und den Anbietern viel zugemutet – und manches, für das noch gar keine anerkannten «best practices» bestehen, etwa im Bereich des Umgangs mit Daten für KI-Modelle. Diese Hoch-Risiko-Anwendungen sind zum Glück vergleichsweise eng und vor allem abschliessend definiert, auch wenn die Liste im Laufe der Zeit angepasst werden kann. Immerhin: Anbieter von bereits regulierten Produkten gemäss Anhang I und im Bereich von kritischen Infrastrukturen werden sich besonders intensiv mit ihnen auseinandersetzen müssen.

[85] Für die meisten Unternehmen in der EU und zu einem geringeren Teil in der Schweiz wird der AI Act zwar Arbeit mit sich bringen, aber diese dürfte nicht ausufernd werden, wenn sie einigermassen im Griff haben, wo bei ihnen im Hause in eigenen und von Dritten genutzten Produkten und Dienstleistungen AIS zum Einsatz kommen. Sie werden dabei vor allem sicherstellen wollen, dass sie nicht in die «Minenfelder» der verbotenen oder als besonders riskant erachteten KI-Anwendungen geraten. Halten sie sich davon fern, dürften sie mit den wenigen verbleibenden Pflichten (vorwiegend zur Transparenz) aber vergleichsweise einfach zurechtkommen, und zwar selbst dann, wenn sie als «Provider» gelten, weil sie eine Anwendung selbst entwickelt oder weiterentwickelt haben. Eine Ausnahme mag die Pflicht zur Markierung von KI-generierten Inhalten sein; hier fehlen derzeit offenbar noch passende Standards und einsatzbereite Lösungen.

207 Vgl. https://digital-strategy.ec.europa.eu/en/policies/ai-pact.

208 Art. 111 Abs. 1 und Abs. 2 AIA.

209 Art. 111 Abs. 2 AIA.

210 Art. 111 Abs. 3 AIA.

Es darf gespannt erwartet werden, was die Marktführer hier anbieten werden, insbesondere für KI-generierte Texte.

[86] Schweizer Unternehmen kommen um den AI Act ebenfalls nicht herum, auch wenn er für sie nur in reduziertem Masse gilt und wie im Falle der DSGVO nicht davon auszugehen ist, dass sie im Fokus der EU-Aufsichtsbehörden sein werden (allerdings hat beispielsweise die Schweizer Finanzmarktaufsicht FINMA bereits informell erläutert, dass sie von Schweizer Finanzinstituten die Einhaltung des AI Act erwartet, soweit er selbst auf sie Anwendung finden will).[211] Zwar ist noch nicht in allen Fällen klar ist, ob der AI Act auf sie zur Anwendung kommt, weil bei seiner Formulierung Fehler gemacht worden sind. Es muss jedoch damit gerechnet werden, dass die Nutzung von AIS dann erfasst ist, wenn der Output bestimmungsgemäss an Empfänger in die EU geht oder sonst in der EU zum Einsatz kommt. Bietet ein Schweizer Unternehmen seinen Kunden im EU-Markt im Rahmen seiner Produkte, Dienstleistungen oder Website unter dem eigenen Namen KI-Funktionen an, die es mindestens teilweise selbst entwickelt hat, ist es ebenfalls erfasst. Solange es sich nicht um eine Hoch-Risiko-Anwendung handelt, sind die Pflichten allerdings wiederum moderat.

[87] Daraus ergibt sich, dass für die meisten Unternehmen der Aufwand vor allem darin bestehen wird, zu verstehen, was in welcher Form an KI im eigenen Unternehmen zum Einsatz kommt, um rechtzeitig reagieren zu können, falls sich ein Projekt in Richtung eines der genannten Minenfelder bewegt. Geht es um eine Hoch-Risiko-Anwendung, werden Unternehmen vor allem gut beraten sein, die Rolle des Providers und daher Eigenentwicklungen eher zu vermeiden, weil dem Provider ungleich mehr Pflichten auferlegt sind als dem reinen Deployer. Ganz so einfach ist das allerdings nicht, weil davon auszugehen ist, dass bereits Massnahmen wie etwa ein *Fine-Tuning* als «Entwickeln» gilt. Hier wird sich zeigen, inwiefern die Herausforderung so gelöst werden kann, dass Technologie-Lieferanten und -Integratoren diese Pflichten als «Provider» übernehmen und dies von ihren Kunden so auch kommuniziert werden darf.

[88] Das führt zu unserer Handlungsempfehlung: Unternehmen sollten alle eigenen und fremdgenutzten KI-Anwendungen erfassen und danach beurteilen, ob und in welcher Rolle (Provider, Deployer, Händler etc.) sie unter den AI Act fallen werden. Hierzu können im Markt verfügbare Hilfsmittel helfen.[212] Dabei ist auch zu berücksichtigen, dass die Anwendung auf Schweizer Unternehmen beschränkt ist; es genügt z.B. nicht, dass ein Produkt oder Service aus der EU zum Einsatz kommt oder mit Daten von Personen aus der EU gefüttert wird. Wir empfehlen Unternehmen nicht nur im Hinblick auf den AI Act, ein Verzeichnis ihrer KI-Anwendungen zu führen. Hier können auch die Ergebnisse in Bezug auf den AI Act eingetragen werden.

[89] Sind diese Abklärungen erfolgt, sind zum einen entsprechende Vorgaben zu erlassen, um sicherzustellen, dass diese KI-Anwendungen nicht ohne vorherige Prüfung in einer Weise genutzt werden, die zu einer anderen Qualifikation führt. Zum anderen sind in den Fällen, in denen der AI Act zur Anwendung kommt, die daraus resultierenden Pflichten zu ermitteln und ein Plan zu erstellen, wie diese in den vorgesehenen Fristen umgesetzt werden können. Für neue Vorhaben gilt diese Handlungsempfehlung analog. Bei vielen Schweizer Unternehmen hat sich zudem die Regelung eingebürgert, dass verbotene KI-Praktiken auch dann mittels Weisung im eigenen Unternehmen untersagt werden, wenn der AI Act formal nicht zur Anwendung gelangt.

---

211 Vgl. ausführlich zu den Auswirkungen des AI Acts auf die Schweiz Angela Müller (Fn 16), S. 21 ff.

212 Vgl. etwa der «AI Act Checker» im Open-Source-Tool GAIRA, abrufbar unter https://vischerlnk.com/gaira.

[90] Zum Schluss: Wird die Schweiz einen eigenen «AI Act» erlassen? Der Bundesrat will bis Ende 2024 verlautbaren, wo er gesetzgeberischen Handlungsbedarf im Bereich der künstlichen Intelligenz sieht. Mit einem Schweizer «AI Act» ist vorderhand aber nicht zu rechnen; eine direkte Übernahme erscheint nicht opportun, und für eine Schweizer Abwandlung dürfte der Aufwand für Verwaltung und Wirtschaft den Nutzen bei weitem übersteigen, ist doch damit zu rechnen, dass auch viele Schweizer Anbieter von relevanten Produkten den AI Act einhalten werden – die grossen internationalen Anbieter sowieso. Sinnvoll erscheinen jedoch die Anpassungen, die nötig sind, um den freien Warenverkehr mit der EU weiterhin sicherzustellen und Handelshemmnisse zu vermeiden, einschliesslich der gegenseitigen Anerkennung von Produktezulassungen.

David Rosenthal ist Partner der Kanzlei VISCHER und Lehrbeauftragter der Universität Basel und ETH Zürich.

Mit herzlichem Dank an Rahel Hirschi für die Unterstützung bei der Erarbeitung der Verweise auf den AI Act und Rona Lengen für die Aufarbeitung der Literatur zum AI Act.

Dieser Beitrag basiert auf einem Blog-Beitrag des Autors vom Februar 2024 auf https://vischer.com/ki, wurde jedoch in der Folge wesentlich angepasst und ergänzt.