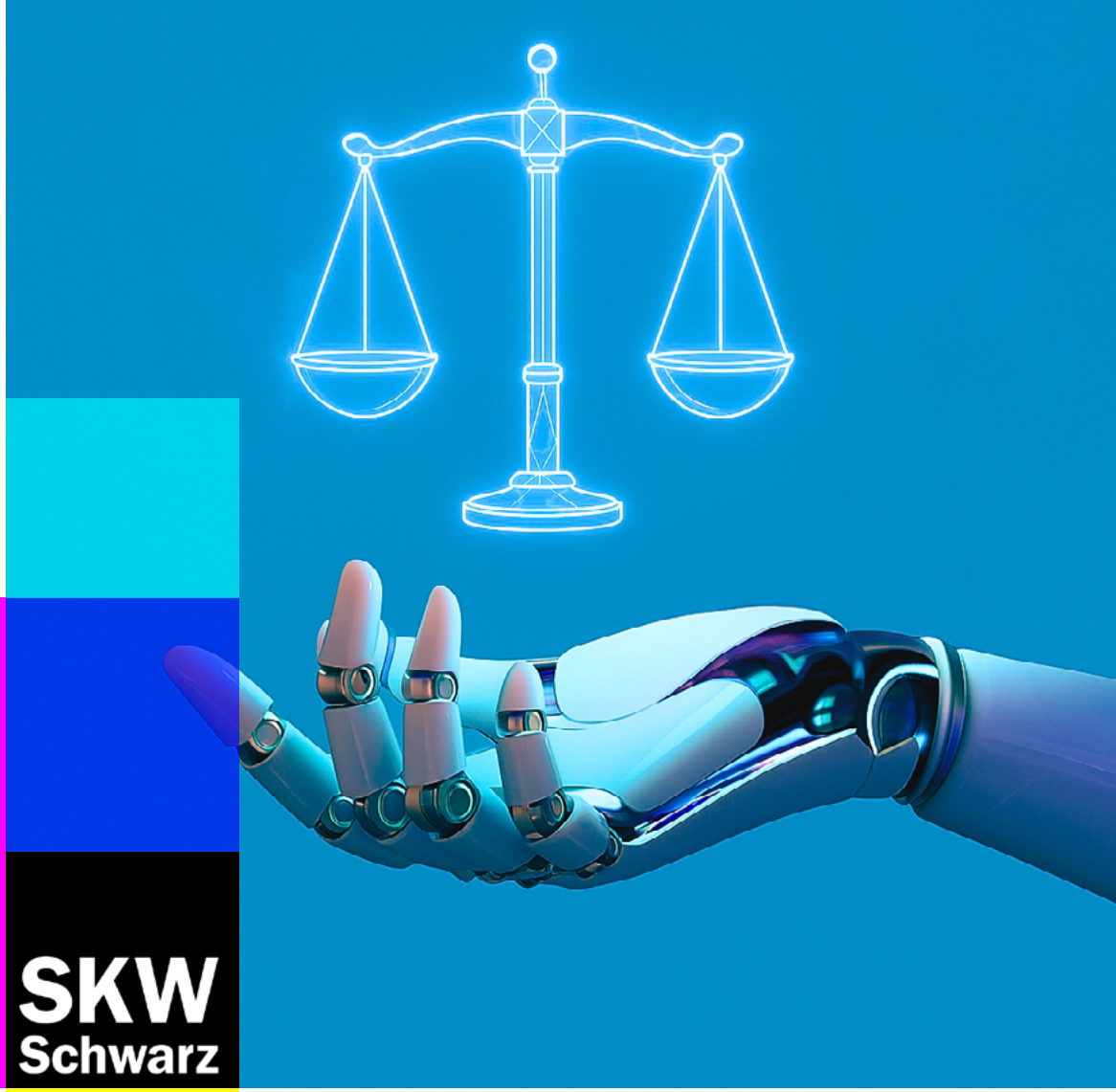
Whitepaper



SKW Schwarz

April 2025

Verbotene KI-Praktiken

Leitlinien der Europäischen Kommission

Leitlinien der Europäischen Kommission zu verbotenen KI-Praktiken

Der Europäische Gesetzgeber verfolgt mit Einführung der KI-VO hehre Ziele: (i) die Verbesserung des Funktionierens des Binnenmarkts, (ii) die Unterstützung von Innovation und (iii) die Förderung der Einführung menschenbezogener und vertrauenswürdiger künstlicher Intelligenz (KI). Gleichzeitig gilt es, (iv) Gesundheit, (v) Sicherheit, (vi) Grundrechte, (vii) Demokratie, (viii) Rechtsstaatlichkeit und (ix) Umweltschutz vor den schädlichen Auswirkungen von KI-Systemen zu bewahren.

Dies soll mit einem risikobasierten Ansatz gelingen, aus dem sich ein umfangreicher Katalog an verbotenen Praktiken im KI-Bereich ergibt (Art. 5 KI-VO). Obwohl die Rechtsfolge von Art. 5 KI-VO ein Verbot darstellt, enthält sie einige unbestimmte Rechtsbegriffe. Gemäß Art. 96 KI-VO ist die Kommission daher verpflichtet durch Leitlinien einzelne Vorschriften näher zu konkretisieren. Rechtzeitig mit der Anwendbarkeit der KI-VO hinsichtlich des Verbotes des Art. 5 KI-VO am 2. Februar 2025 veröffentlichte die Europäische Kommission daher die Leitlinien zu den verbotenen Praktiken aus der KI-VO (→). Diese sind zwar nicht verbindlich, wirken jedoch mittels Definitionen und Ausführungen über Anwendungsverhältnisse auf eine wirksame Praxis hin.

Dieses Whitepaper fasst die Kernaussagen der Leitlinien zusammen, erläutert diese und die jeweiligen Vorschriften der KI-VO mit dem Ziel, die Auslegung und Anwendung des Art. 5 KI-VO zu vereinfachen.

Leitlinien der Europäischen Kommission zu verbotenen KI-Praktiken

Inhalt

I) Anwendungsbereich

Zunächst konkretisieren die Leitlinien den Anwendungsbereich der KI-VO in Bezug auf Art. 5 KI-VO.

1. Sachlicher Anwendungsbereich

Sachlich bezieht sich Art. 5 KI-VO auf das Inverkehrbringen, die Inbetriebnahme und die Verwendung von bestimmten KI-Systemen. Das Inverkehrbringen soll dabei die erstmalige Bereitstellung eines KI-Systems umfassen, was sowohl das Bereitstellen zum Vertrieb als auch die Nutzung auf dem Unionsmarkt im Rahmen einer gewerblichen Tätigkeit beinhalten soll, unabhängig davon, ob dies gegen Entgelt oder kostenlos erfolgt. Die Inbetriebnahme meint die Lieferung eines KI-Systems an den Betreiber oder zur eigenen Verwendung in der Union für den beabsichtigten Zweck. Mit der Zweckbestimmung ist u.a. die Verwendung gemeint, für die ein KI-System vom Anbieter vorgesehen ist, einschließlich des spezifischen Kontexts und der Bedingungen für die Verwendung. Hinsichtlich der Verwendung soll es auf ein weites Verständnis zu jedem Zeitpunkt des Lebenszyklus nach Inverkehrbringen der KI ankommen. Umfasst werden soll also ebenfalls jeder Missbrauch, der auf eine verbotene Praxis hinauslaufen kann.

2. Persönlicher Anwendungsbereich

Im Hinblick auf den persönlichen Anwendungsbereich sind vor allem die Anbieter und Betreiber der KI-Systeme für die Verbotsvorschrift relevant. Nach Art. 3 Abs. 3 KI-VO sind Anbieter natürliche oder juristische Personen, Behörden, Einrichtungen oder sonstige Stellen, die eine KI entwickeln oder entwickeln lassen und es unter ihrem eigenen Namen oder ihrer Handelsmarke in Betrieb nehmen (gleich ob entgeltlich oder unentgeltlich). Betreiber sind gemäß Art. 3 Abs. 4 KI-VO natürliche oder juristische Personen, Behörden, Einrichtungen oder sonstige Stellen, die ein KI-System in eigener Verantwortung verwenden, es sei denn, die KI wird im Rahmen einer persönlichen und nicht beruflichen Tätigkeit verwendet. Die Verwendung in eigener Verantwortung meint, dass die Verantwortung für die Entscheidung über den Einsatz des Systems und die Art und Weise seiner tatsächlichen Nutzung bei der Person liegt. Betreiber einer KI können auch gleichzeitig als Anbieter auftreten.

Da die Compliance-Pflichten der KI-VO während des gesamten Lebenszyklus der KI einzuhalten sind, kommt es für die Zuständigkeit der etwaigen Maßnahmen zur Einhaltung der Konformität mit Art. 5 KI-VO bei Auseinanderfallen der beiden Personen darauf an, wer in der Wertschöpfungskette an geeigneterer Stelle steht, um diese Pflichten zu verantworten. So wird der Anbieter dafür Sorge zu tragen haben, dass keine KI in einer Art entwickelt wird, die nur in verbotener Art und Weise anzuwenden ist, der mithin bereits von ihren Funktionalitäten her der Verbotstatbestand immanent ist. Der Betreiber auf der anderen Seite trägt die Verantwortung, grundsätzlich rechtskonforme KI nicht zu verbotenen Zwecken einzusetzen.

3. Ausnahmen des Anwendungsbereichs

In Art. 2 KI-VO werden Ausnahmen des Anwendungsbereichs der Verordnung statuiert, die für die praktische Verwendung der Verbote des Art. 5 KI-VO relevant sind. Die Leitlinien bieten einen Überblick über dessen Auswirkungen.

3.1 Nationale Sicherheit, Verteidigung und militärische Zwecke

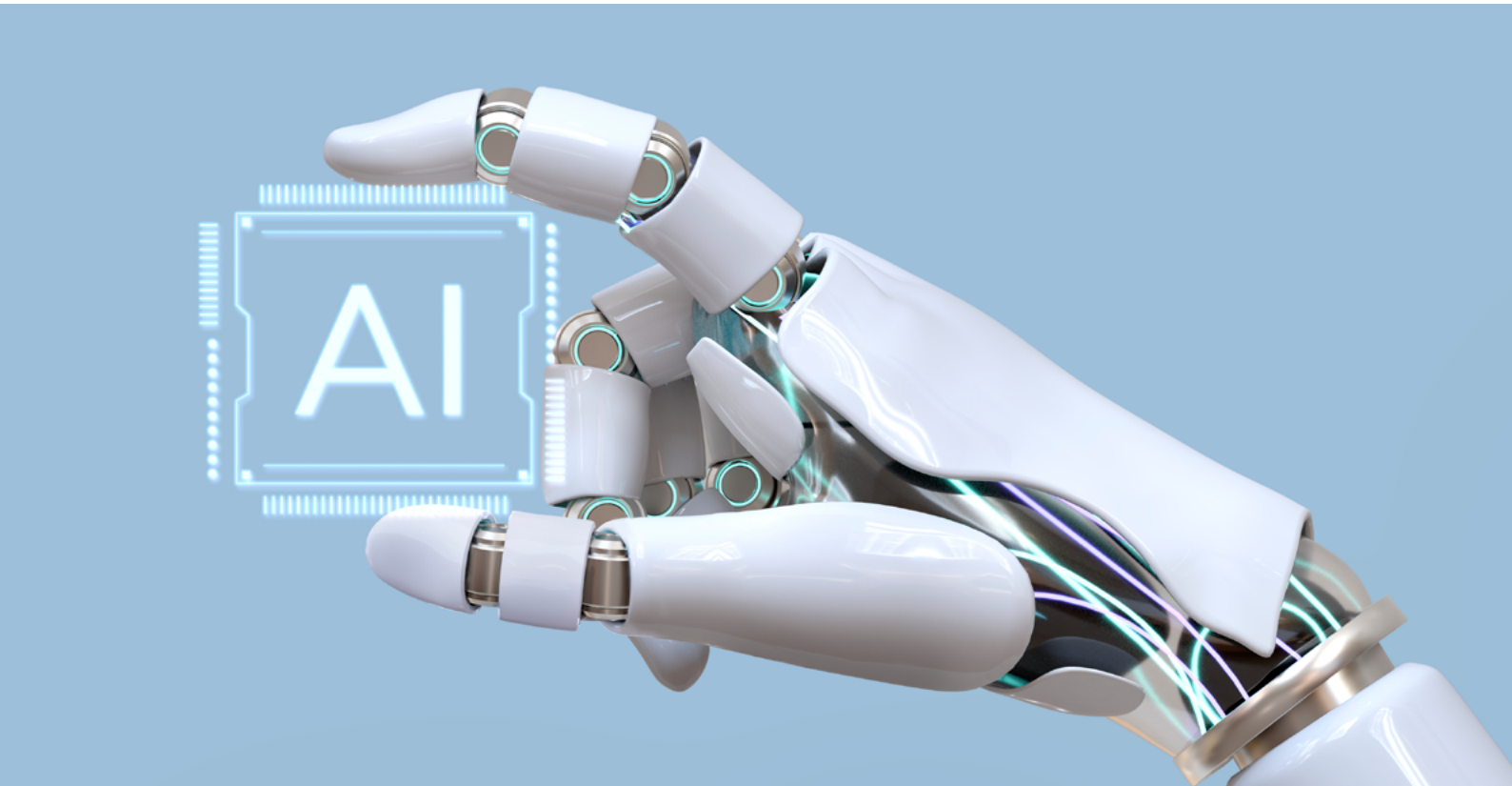
Gemäß Art. 2 Abs. 3 KI-VO gilt die Verordnung nicht für KI-Systeme, die ausschließlich im Bereich nationale Sicherheit, Verteidigung und militärische Zwecke eingesetzt werden. Hierbei soll es nicht auf die Einrichtungen ankommen, die diese Tätigkeit ausüben, sondern darauf, zu welchem Zweck das KI-System verwendet wird.

Unter **nationale Sicherheit** ist zu verstehen: Das vorrangige Interesse am Schutz der wesentlichen Funktionen des Staates und der grundlegenden Interessen der Gesellschaft. Umfasst ist ebenfalls die Verhütung und Ahndung von Handlungen, die geeignet sind, die verfassungsmäßigen, politischen, wirtschaftlichen oder sozialen Grundstrukturen eines Landes ernsthaft zu destabilisieren und insbesondere eine unmittelbare Bedrohung für die Gesellschaft, die Bevölkerung oder den Staat selbst darzustellen.

Für den Fall, dass die KI sowohl einen der oben genannten Zwecke, als auch einen Zweck verfolgt, der nicht von dem Anwendungsbereich ausgenommen ist – etwa zivile Rechts- oder Strafverfolgung - (sog. „dual use“), legt Erwägungsgrund 24 der KI-VO fest, dass die Verordnung sehr wohl Anwendung findet und somit die Compliance-Pflichten zu wahren sind. Eine klare Differenzierung ist allerdings im Hinblick auf Art. 5 Abs. 1 d), h) KI-VO sinnvoll, wenn die KI für Strafverfolgungszwecke eingesetzt wird. Dort ist Aufgabe der Polizei und anderer Strafverfolgungsbehörden, Straftaten zu verhüten, aufzudecken, zu ermitteln und zu verfolgen oder strafrechtliche Sanktionen zu vollstrecken. Hierbei handelt es sich nicht nur um Fragen der nationalen Sicherheit, sondern auch um die Sicherstellung und den Schutz der Rechtsstaatlichkeit, mithin eines der wesentlichen Gewährleistungsziele der KI-VO. Damit wird die Verordnung bei der Verwendung von KI für diese Zwecke uneingeschränkt Anwendung finden.

3.2 Strafverfolgung und justizielle Zusammenarbeit mit Drittländern

Gemäß Art. 2 Abs. 4 KI-VO ist die Verordnung nicht anwendbar für den Einsatz von KI im Rahmen der internationalen Zusammenarbeit im Bereich der Strafverfolgung oder justiziellen Zusammenarbeit, sofern ein Drittland beteiligt ist und angemessene Garantien und Schutzmaßnahmen entsprechend der Union etabliert hat. Damit dieser Ausschluss greift, müssen Rahmenbedingungen der Zusammenarbeit angemessene Schutzmaßnahmen hinsichtlich der Grundrechte und Grundfreiheiten beinhalten, welche durch Marktüberwachungsbehörden zu bewerten sind.



3.3 Forschung und Entwicklung

Laut Art. 2 Abs. 8 KI-VO gilt die Verordnung nicht für Forschungs-, Test- und Entwicklungstätigkeiten im Vorfeld der Inbetriebnahme der KI. Hintergrund ist der marktorientierte Ansatz der Verordnung. Während der Forschungs- und Entwicklungsphase, sollen die KI-Entwickler die Freiheit haben zu experimentieren und neue Funktionalitäten zu testen. Diese Herangehensweise kann dadurch gerechtfertigt werden, dass die Verfeinerung in der Entwicklungszeit unerlässlich ist, um die erforderlichen Sicherheits- und Ethikstandards zu erfüllen. Das wird deutlich, indem die Leitlinien festlegen, dass eine Erprobung unter realen Bedingungen nicht unter den Ausschluss fällt.

3.4 Persönliche, außerberufliche Tätigkeit

Sofern die KI im persönlichen und somit außerberuflichen Kontext eingesetzt wird, statuiert Art. 2 Abs. 10 KI-VO, dass die Verordnung keine Anwendung findet. Sobald die Person einen wirtschaftlichen Nutzen aus der Verwendung zieht oder die Tätigkeit in anderer Weise beruflich, geschäftlich, gewerblich oder freiberuflich ausübt, soll eine berufliche Tätigkeit angenommen werden. Die Ausnahme soll nicht für kriminelle Tätigkeiten gelten.

3.5 Freie und quelloffene Lizenzen

Gemäß Art. 2 Abs. 12 KI-VO gilt die Verordnung grundsätzlich nicht für KI-Systeme, die unter freien und quelloffenen Lizenzen bereitgestellt werden.

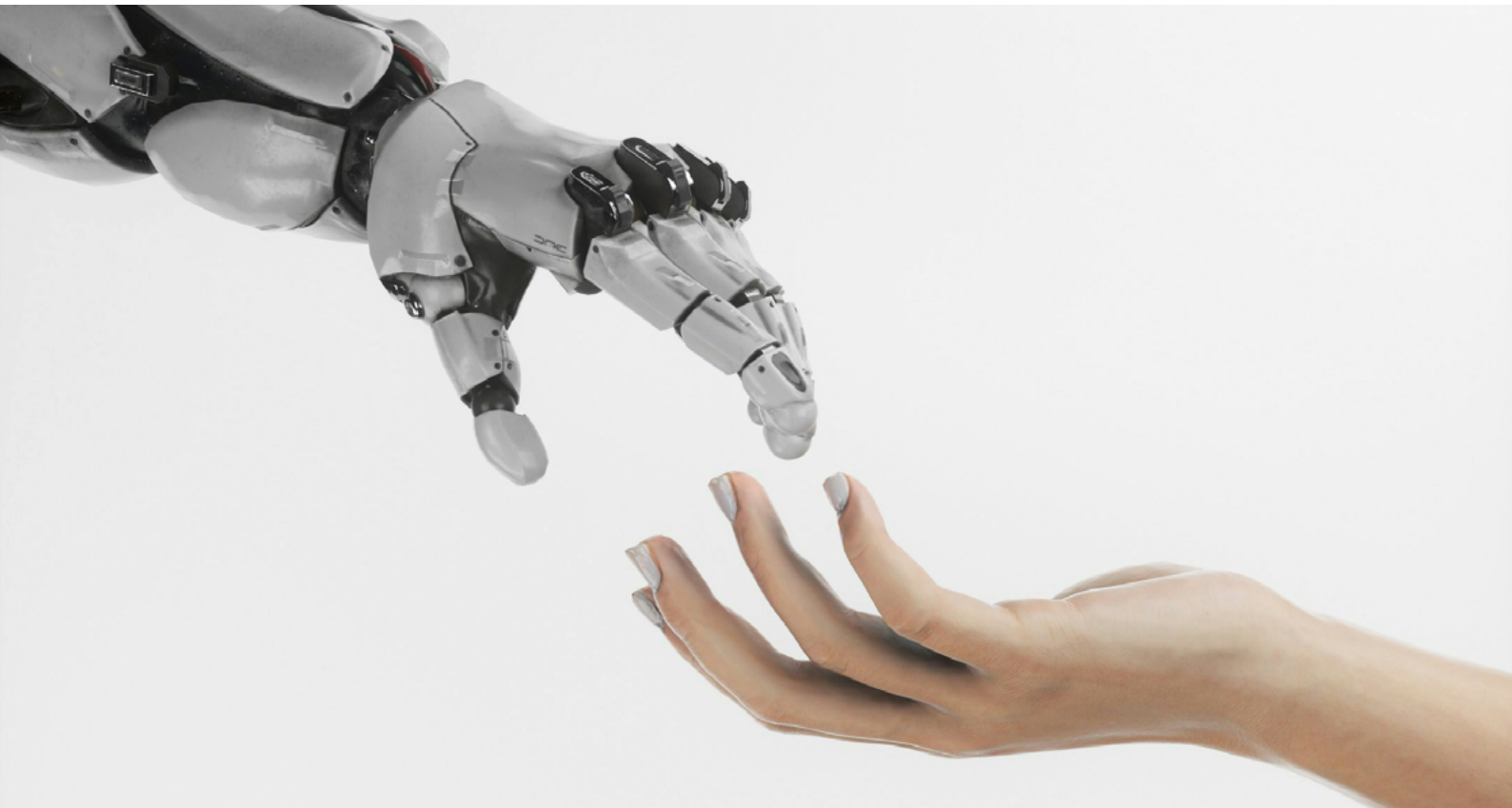
Hier besteht allerdings eine Ausnahme: Die Anforderungen an Hochrisiko-KI-Systeme sowie die Verbotstatbestände des Art. 5 KI-VO oder die Transparenzvorschriften des Art. 50 KI-VO sind auch auf KI-Systeme unter freien und quelloffenen Lizenzen anzuwenden, wenn diese in Verkehr gebracht oder in Betrieb genommen werden.

3.6 Anforderungen an hochriskante KI-Systeme

Die Verwendung von als hochriskant eingestuften KI-Systemen kann in einigen Fällen leicht die Voraussetzungen von Art. 5 KI-VO erfüllen und in einem Verbot resultieren. Umso mehr ist darauf zu achten, dass, sollte die Prüfung des KI-Systems ergeben, dass dieses System nicht dem Anwendungsbereich des Art. 5 KI-VO unterfällt und damit nicht grundsätzlich verboten ist, die Prüfung und Evaluierung insoweit fortgesetzt wird, ob die Anforderungen der Art. 6 ff. KI-VO an Hochrisiko-KI-Systeme erfüllt sind. In diesem Fall sind im Rahmen des Risikomanagements die Einhaltung von Schutzmaßnahmen besonders sorgfältig zu betrachten.

3.7 KI-Systeme mit allgemeinem oder bestimmtem Verwendungszweck

Die Verbote von Art. 5 KI-VO gelten sowohl für KI-Systeme mit allgemeinem, als auch mit bestimmtem Verwendungszweck. Um zu verhindern, dass das KI-System durch ihren Einsatz in der Praxis einen Schaden auslöst, wird von den Verantwortlichen erwartet, dass sie wirksame und überprüfbare Maßnahmen ergreifen, um Schutzvorkehrungen einzubauen und mögliche Schäden zu verhindern. In vertraglichen Beziehungen mit den Anwendern (bspw. AGB) soll die Nutzung des KI-Systems für verbotene Praktiken ausgeschossen und Anwendungshinweise bereitgestellt werden.



4. Wechselwirkung zwischen den Verboten und anderem Unionsrecht

Die Verbote aus Art. 5 KI-VO zielen auf frühe Phasen des Lebenszyklus des KI-Systems (schon Inverkehrbringen und Inbetriebnahme) ab, um so einen möglichst umfangreichen Schutz zu erzielen. Andere im Unionsrecht geltende Verbote werden nicht durch die KI-VO berührt. Da bei der Verwendung von KI-Systemen häufig personenbezogene Daten verarbeitet werden, sind stets die Anforderungen der europäischen Rechtsakte in Bezug auf Datenschutz zu beachten. Beispielhaft nennen die Leitlinien die Datenschutz-Grundverordnung (DSGVO) und die Strafverfolgungsrichtlinie (LED). Die weiteren Verbots- und Haftungsvorschriften der Union in diesem Kontext bleiben anwendbar und es wird ausdrücklich darauf hingewiesen, dass die Verbote aus Art. 5 KI-VO und deren Ausnahmen nicht zur Umgehung oder als Rechtfertigungen für Verstöße gegen Verpflichtungen aus anderen unionsrechtlichen Vorschriften verwendet werden dürfen.

Insbesondere dient Art. 5 KI-VO niemals als Erlaubnistatbestand im Sinne der DSGVO. Daraus, dass der Einsatz von KI nach Art. 5 KI-VO nicht verboten und damit grundsätzlich erlaubt ist, kann nicht geschlussfolgert werden, dass dies zugleich eine datenschutzrechtliche Rechtfertigung beinhaltet. Umgekehrt wird die Verarbeitung personenbezogener Daten durch eine verbotene KI niemals durch die DSGVO gerechtfertigt sein.

5. Durchsetzung von Art. 5 KI-VO

Mit der Durchsetzung der Vorschriften aus der KI-VO sind die von den Mitgliedsstaaten benannten Marktaufsichtsbehörden sowie der Europäische Datenschutzbeauftragte betraut. In Deutschland wird die Aufsicht voraussichtlich durch die Bundesnetzagentur (BNetzA) wahrgenommen werden. Sofern eine etwaige Beschwerde eingereicht wird, befähigt sie dies zur Durchsetzung eines Verbots. Um eine einheitliche Durchsetzung der Verbote durchzusetzen, ist vorgesehen, dass alle Marktüberwachungsbehörden ein Schutzklauselverfahren der Union einhalten, bei dem die Kommission entscheidet, ob das KI-System unter den Katalog verbotener Praktiken fällt.

In der KI-VO sind Geldbußen nach einem abgestuften Ansatz vorgesehen, dessen Höhe sich nach der Schwere des Verstoßes richtet. Die Nichteinhaltung der Verbote aus Art. 5 KI-VO zieht als schwerster Verstoß die höchste Geldbuße in Höhe von bis zu EUR 35.000.000,00 bzw. 7% des weltweiten Vorjahresumsatzes nach sich.

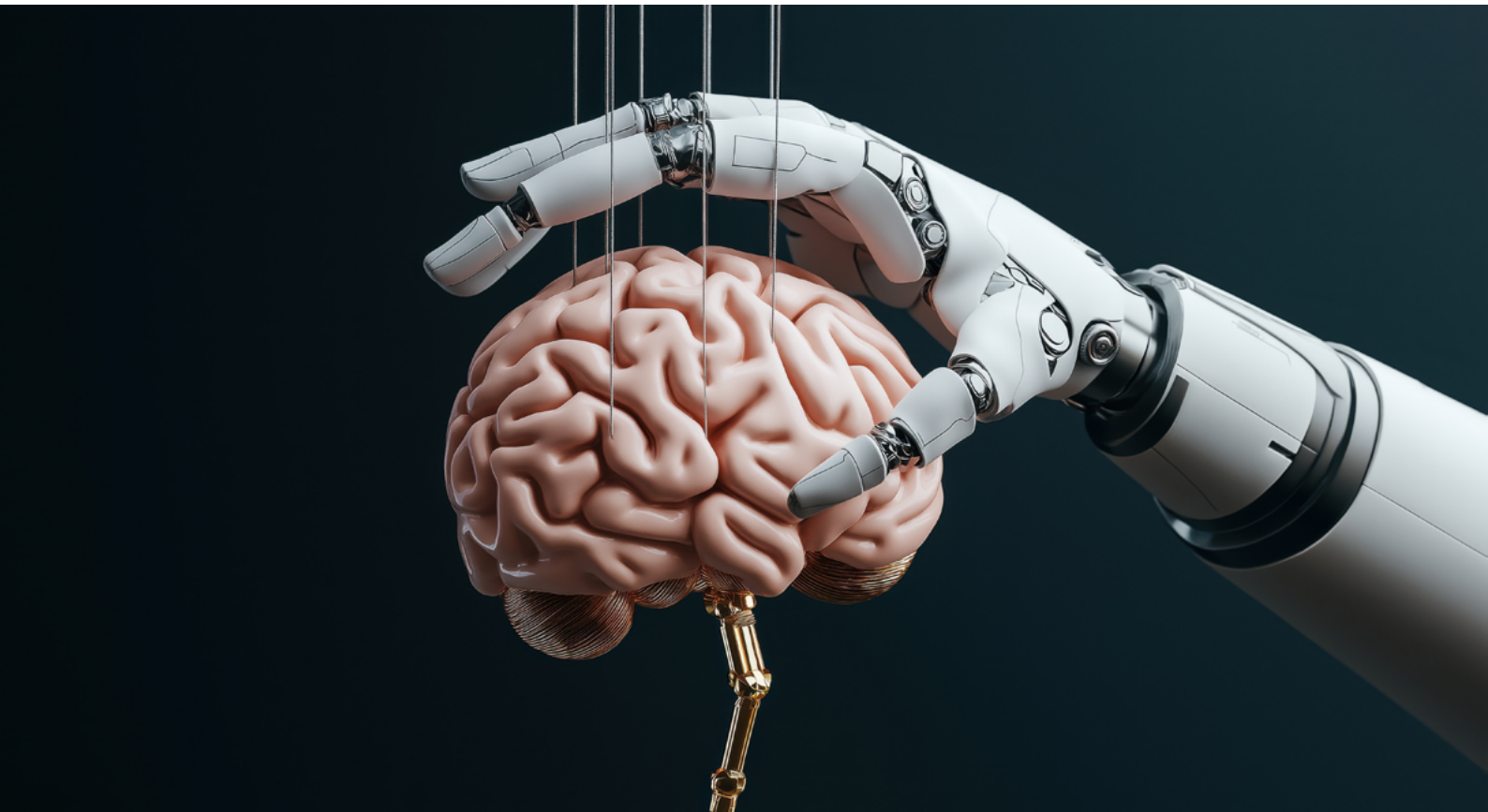
II) Die Verbote des Art. 5 KI-VO

Die Verbotsvorschrift enthält einige unbestimmte und auslegungsbedürftige Voraussetzungen. Im Folgenden werden die Begriffe und Anwendungsverhältnisse erläutert, um die Anwendung auf die Praxis zu erleichtern.

1. Verbote zum Einsatz manipulativer KI gemäß Art. 5 Abs. 1 a), b) KI-VO

Die ersten beiden Verbote in Art. 5 Abs. 1 KI-VO schützen Einzelpersonen und besonders schutzbedürftige Gruppen vor den erheblichen schädlichen Auswirkungen von KI-gestützter Manipulation und Ausbeutung. Sie betreffen KI-Systeme, die unterschwellige, gezielt manipulative oder täuschende Techniken einsetzen, um das Verhalten natürlicher Personen oder Personengruppen wesentlich zu beeinflussen (Art. 5 Abs. 1 a) KI-VO) oder die gezielt Schwachstellen aufgrund von Alter, Behinderung oder sozioökonomischer Situation ausnutzen (Art. 5 Abs. 1 b) KI-VO).

Zu prüfen ist stets der Kausalzusammenhang zwischen der gezielten Ausnutzung und dem möglichen Schaden. Dieser kann sowohl materieller Natur sein (z. B. finanzielle Verluste) als auch immaterieller Natur (z. B. psychische oder physische Beeinträchtigungen).



1.1 Verbot Art. 5 Abs. 1 a) KI-VO: Manipulative KI

Art. 5 KI-VO

1. Die folgenden AI-Praktiken sind verboten:

→ **(a)** das Inverkehrbringen, die Inbetriebnahme oder die Verwendung eines KI-Systems, das Techniken der unterschwelligen Beeinflussung außerhalb des Bewusstseins einer Person oder absichtlich manipulative oder täuschende Techniken mit dem Ziel oder der Wirkung einsetzt, das Verhalten einer Person oder einer Gruppe von Personen wesentlich zu verändern, indem ihre Fähigkeit, eine fundierte Entscheidung zu treffen, deutlich beeinträchtigt wird, wodurch sie veranlasst wird, eine Entscheidung zu treffen, die sie andernfalls nicht getroffen hätte, und zwar in einer Weise, die dieser Person, einer anderen Person oder einer Gruppe von Personen erheblichen Schaden zufügt oder mit hinreichender Wahrscheinlichkeit zufügen wird.

Das Verbot erfasst KI-Systeme, die manipulative Techniken einsetzen. Die Tatbestandsmerkmale der unterschwelligen, gezielt manipulativen oder täuschende Techniken überschneiden sich in Teilen.

→ **Unterschwellige Techniken** umgehen die bewusste Wahrnehmung und beeinflussen Entscheidungen, ohne dass Betroffene es bemerken. Sie nutzen visuelle, auditive oder taktile Reize, die zu subtil sind, um erfasst zu werden – insbesondere in Bereichen wie virtueller Realität oder Neurotechnologie.

→ **Manipulative Techniken** untergraben gezielt die Entscheidungsfreiheit einer Person.

→ **Täuschende Techniken** beruhen auf falschen oder irreführenden Informationen und stehen in engem Zusammenhang mit Art. 50 Abs. 4 KI-VO, der eine Kennzeichnung von „Deep Fakes“ vorschreibt.
Klarstellend sei angemerkt: Das Verbot in Art. 5 KI-VO kann jedoch auch bei Einhaltung dieser Kennzeichnungspflicht greifen.

Für die Anwendung des Verbots ist entscheidend, dass das Verhalten einer Person oder Gruppe erheblich beeinflusst wird – also eine erhebliche Verzerrung der Entscheidungsfreiheit vorliegt. Die Regelung weist Parallelen zu den EU-Vorschriften über unlautere Geschäftspraktiken auf und die Leitlinien verweisen ausdrücklich darauf.

Das Verbot unterscheidet zwei Fälle: (i) KI-Systeme, die das Verhalten gezielt beeinflussen sollen (z. B. ein Chatbot, der unterschwellige Werbesignale nutzt, um Kaufentscheidungen zu steuern) sowie (ii) KI-Systeme, die unbeabsichtigt eine erhebliche Beeinflussung bewirken (z. B. ein Gesundheits-Chatbot, der fehlerhafte Empfehlungen gibt und dadurch gesundheitliche Schäden verursacht).

Damit das Verbot greift, muss das KI-System erheblichen Schaden verursachen oder mit hoher Wahrscheinlichkeit verursachen können. Schäden können physischer, psychischer, finanzieller oder wirtschaftlicher Natur sein. Die Erheblichkeitsschwelle wird je nach Intensität, Reversibilität und Betroffenheit vulnerabler Gruppen – etwa Kinder – im Einzelfall geprüft.

Zur Vermeidung von Risiken empfehlen die Leitlinien Maßnahmen wie Transparenz über die Funktionsweise der KI, die Einhaltung relevanter Vorschriften und die Berücksichtigung des aktuellen Stands der Technik.

1.2 Verbot Art. 5 Abs. 1 b) KI-VO: Ausnutzung persönlicher Schwachstellen

Art. 5 KI-VO

1. Die folgenden AI-Praktiken sind verboten:

→ **(b)** das Inverkehrbringen, die Inbetriebnahme oder die Verwendung eines KI-Systems, das eine Vulnerabilität oder Schutzbedürftigkeit einer natürlichen Person oder einer bestimmten Gruppe von Personen aufgrund ihres Alters, einer Behinderung oder einer bestimmten sozialen oder wirtschaftlichen Situation mit dem Ziel oder der Wirkung ausnutzt, das Verhalten dieser Person oder einer dieser Gruppe angehörenden Person in einer Weise wesentlich zu verändern, die dieser Person oder einer anderen Person erheblichen Schaden zufügt oder mit hinreichender Wahrscheinlichkeit zufügen wird;

Dieses Verbot schützt besonders vulnerable Gruppen vor der gezielten Ausnutzung ihrer Schutzbedürftigkeit. Erfasst werden KI-Systeme, die gezielt die kognitive, emotionale oder physische Anfälligkeit von Kindern, älteren Menschen oder sozial benachteiligten Personen ausnutzen, um deren Verhalten zu beeinflussen und ihnen dadurch Schaden zuzufügen.

Im Hinblick auf Kinder und Jugendliche können KI-Systeme – wie interaktive digitale Spielzeuge oder Social-Media-Algorithmen – gezielt manipulative Mechanismen nutzen, um exzessiven Konsum oder risikobehaftetes Verhalten zu fördern. Ein Beispiel sind Empfehlungssysteme sozialer Netzwerke, die junge Nutzer gezielt zu übermäßigem Medienkonsum oder gesundheitsgefährdendem Verhalten (z.B. Essstörungen oder gefährliche Mutproben) animieren.

Ältere Menschen sind dagegen eher gefährdet durch KI-gestützte Betrugssysteme wie der digitale „Enkeltrick 2.0“, bei dem täuschendechte KI-generierte Stimmen oder Nachrichten ältere Menschen zur Geldüberweisung verleiten.

1.3 Außerhalb des Geltungsbereichs

Die Leitlinien grenzen zwischen unzulässiger Manipulation und legitimer Überzeugungsarbeit ab. Manipulation ist dadurch gekennzeichnet, dass sie verdeckt erfolgt und die Autonomie der betroffenen Person untergräbt. Überzeugungsarbeit bleibt hingegen zulässig, wenn sie transparent ist, sachliche Informationen bereitstellt und die Entscheidungsfreiheit respektiert.

Ein wesentliches Kriterium für die Abgrenzung ist, ob sich die Einflussnahme gezielt auf eine besonders schutzbedürftige Gruppe bezieht oder allgemeiner Natur ist. Ein KI-System, das gezielt die Schwachstellen von älteren Menschen oder Kindern ausnutzt, fällt unter das Verbot, während ein System, das allgemeine psychologische Prinzipien nutzt, nicht zwangsläufig untersagt ist.

Die Praxis zeigt, dass Anbieter und Betreiber von KI-Systemen sicherstellen müssen, dass ihre Systeme weder unterbewusste Beeinflussung noch gezielte Manipulation einsetzen. Besondere Vorsicht ist bei der Verwendung von KI-Systemen gegenüber vulnerablen Gruppen wie Kindern und älteren Menschen geboten. Hier sollten zusätzliche Schutzmaßnahmen implementiert werden, um Missbrauch und Schäden zu verhindern.

1.4 Zusammenspiel und Abgrenzung der Artikel 5 Abs. 1 a) und b) KI-VO

Wie grenzen sich die Anwendungsbereiche der Artikel 5 Abs. 1 a) und b) KI-VO nun voneinander ab?

→ **Art. 5 Abs. 1 a) KI-VO** verbietet KI-Systeme, die manipulative oder täuschende Techniken nutzen, insbesondere wenn diese unterhalb der Wahrnehmungsschwelle wirken und die kognitive Autonomie von Personen untergraben.

→ **Art. 5 Abs. 1 b) KI-VO** hingegen schützt spezifisch besonders schutzbedürftige Gruppen (Kinder, ältere Menschen, sozial benachteiligte Personen) und verbietet die gezielte Ausnutzung ihrer Schwachstellen.

→ **Wenn beide Vorschriften anwendbar erscheinen,** kommt es darauf an, ob die Beeinflussung unabhängig von der Schutzbedürftigkeit erfolgt (dann greift Buchstabe a) oder ob eine gezielte Schwachstellen-Ausnutzung vorliegt (dann greift Buchstabe b). Manipulative Praktiken gegenüber Gruppen, die nicht ausdrücklich durch Buchstabe b geschützt sind, können unter Buchstabe a fallen, sofern sie gezielt auf deren spezifische Schwachstellen abzielen.

1.4 Zusammenspiel und Abgrenzung der Artikel 5 Abs. 1 a) und b) KI-VO

Es ist zulässig, dass dieselbe Praxis, die ein Verbot nach Art. 5 Abs. 1 a) oder b) KI-VO darstellt, auch ein Verstoß nach anderen Rechtsvorschriften des Unionsrechts darstellen und nach beiden Regelwerken sanktioniert werden kann. Hierdurch kann ein umfassendes Sicherheitsniveau gewährleistet werden, welcher im Einklang mit den verschiedenen Zielen, Geltungsbereichen und Adressaten der Rechtsakte steht. Die Leitlinien enthalten daher Ausführungen zu den Wechselwirkungen mit anderem Unionsrecht.

Hinsichtlich des Unlauterkeitsrechts beziehen sich die Leitlinien auf die Richtlinie 2005/29/EG. Sowohl das Verbot nach Art. 5 Abs. 1 a), b) KI-VO als auch die Richtlinie über unlautere Geschäftspraktiken will den Verbraucher vor irreführenden oder aggressiven Geschäftspraktiken schützen. Das Verbot nach der KI-VO geht letztlich sowohl im Anwendungsbereich als auch in den Sanktionen weiter; eine parallele Anwendung ist vorgesehen.

2. Das Verbot sozialer Bewertung (Social Scoring) nach Art. 5 Abs. 1 c) KI-VO

Das Verbot in Art. 5 Absatz 1 c) KI-VO zielt auf solche inakzeptablen KI-gestützten Social Scoring-Praktiken ab, die Einzelpersonen oder Gruppen auf der Grundlage ihres Sozialverhaltens oder ihrer persönlichen Merkmale bewerten oder klassifizieren. Diese Praktiken führen zu Benachteiligungen oder Schlechterstellungen in sozialen Kontexten, die in keinem Zusammenhang mit den Umständen stehen, unter denen die Daten ursprünglich erzeugt oder erhoben wurden oder im Hinblick auf die Tragweite des sozialen Verhaltens unverhältnismäßig oder ungerechtfertigt ist. Neben positiven Ergebnissen, sind die negativen Folgen von Social Scoring meist diskriminierend und ungerecht, möglicherweise führen sie sogar zu einem Ausschluss aus der Gesellschaft oder zu inakzeptablen sozialen Kontroll- und Überwachungspraktiken. Die Menschenwürde, Grundrechte wie das Recht auf Nichtdiskriminierung und Gleichbehandlung, Datenschutz und der Schutz des Privat- und Familienlebens sowie einschlägige soziale und wirtschaftliche Rechte sind laut Erwägungsgrund 31 besonders von dieser Praktik berührt und werden entsprechend von dem Verbot geschützt.

Aber was ist unter den einzelnen Tatbestandsmerkmalen des Art. 5 Abs. 1 c) KI-VO zu verstehen?

2.1 Hauptbegriffe und Voraussetzungen des Verbots

Art. 5 KI-VO

1. Die folgenden AI-Praktiken sind verboten:

- **(c)** das Inverkehrbringen, die Inbetriebnahme oder die Verwendung von KI-Systemen zur Bewertung oder Einstufung von natürlichen Personen oder Gruppen von Personen über einen bestimmten Zeitraum auf der Grundlage ihres sozialen Verhaltens oder bekannter, abgeleiteter oder vorhergesagter persönlicher Eigenschaften oder Persönlichkeitsmerkmale, wobei die soziale Bewertung zu einem oder beiden der folgenden Ergebnisse führt:
  - **i)** Schlechterstellung oder Benachteiligung bestimmter natürlicher Personen oder Gruppen von Personen in sozialen Zusammenhängen, die in keinem Zusammenhang zu den Umständen stehen, unter denen die Daten ursprünglich erzeugt oder erhoben wurden;
  - **ii)** Schlechterstellung oder Benachteiligung bestimmter natürlicher Personen oder Gruppen von Personen in einer Weise, die im Hinblick auf ihr soziales Verhalten oder dessen Tragweite ungerechtfertigt oder unverhältnismäßig ist.

Mithin geben die Leitlinien Auslegungshinweise hinsichtlich der jeweiligen Begrifflichkeiten vor, auf die im Folgenden näher eingegangen werden soll.

2.1.1 „Bewertung“ oder „Klassifizierung“

Die Begriffe „Bewertung“ und „Klassifizierung“ sind zu differenzieren:

→ Unter den Begriff der **Bewertung** wollen die Leitlinien eine gewisse Form der Beurteilung fassen – hier kommt es daher darauf an, dass durch eine Evaluation bestimmte Rückschlüsse gezogen werden.

→ Die **Klassifizierung** ist begrifflich weiter gefasst – es handelt sich hierbei eher um eine auf persönlichen Merkmalen beruhende Kategorisierung. Eine Klassifizierung muss nicht zwingend zu einem Urteil führen.

Der hier verwendete Bewertungsbegriff soll den Leitlinien nach außerdem ein Oberbegriff für den im Datenschutz verwendeten „Profiling“ Begriff darstellen – also der Verwendung von Informationen und die Bewertung der Merkmale oder Verhaltensmuster einer Person, um sie einer bestimmten Kategorie oder Gruppe zuzuordnen und die Fähigkeit zur Erfüllung bestimmter Aufgaben oder wahrscheinliches Verhalten zu analysieren oder vorherzusagen.

2.1.2 „Über einen bestimmten Zeitraum“

Weiter ist es erforderlich, dass die Bewertung oder Klassifizierung auf Daten beruht, die über einen längeren Zeitraum und nicht nur einmalig oder punktuell gesammelt wurden.

2.1.3 „Auf Grundlage von Sozialverhalten oder persönlichen Eigenschaften“, bzw. „Persönlich-keitsmerkmalen“

Die für die Bewertung relevanten Daten müssen das Sozialverhalten oder bekannte, vermutete oder vorausgesagte persönliche Eigenschaften bzw. Persönlichkeitsmerkmale betreffen.

Das Sozialverhalten soll weit verstanden werden und im Allgemeinen Handlungen, Verhalten, Gewohnheiten und Interaktionen innerhalb der Gesellschaft o.ä. umfassen. Die Begründung für das Verbot liegt in diesem Fall unter anderem darin, dass die Daten hierüber meist dezentral gesammelt und aus verschiedenen Quellen kombiniert werden, was zu einer verstärkten Überwachung und Verfolgung von Personen führen kann.

Sofern die Bewertung auf persönlichen Eigenschaften oder Persönlichkeitsmerkmalen beruht, sind hiervon eine Vielzahl von Informationen über eine Person, wie etwa das Geschlecht, sexuelle Orientierung, Ethnie, familiäre Situation, Adresse, Einkommen, etc. umfasst. Hinsichtlich der Terminologie können die Begriffe „persönlichen Eigenschaften“ und „Persönlichkeitsmerkmale“ weitgehend synonym verwendet werden, wobei „Persönlichkeitsmerkmale“ zusätzlich die Erstellung spezifischer Profile von Einzelpersonen als Persönlichkeiten implizieren kann.

Weiterhin soll eine Differenzierung erfolgen, je nachdem ob die Merkmale bekannt, abgeleitet oder vorausgesagt sind. Bekannte Merkmale beruhen daher auf Informationen, die der KI als Input zur Verfügung gestellt wurden – sie sind in der Regel überprüfbar. Wie es der Name schon sagt, werden die abgeleiteten Merkmale von Informationen abgeleitet, wobei dies meist durch die KI vorgenommen wird. Die vorausgesagten Merkmale werden auf der Grundlage von Mustern mit weniger als 100%iger Genauigkeit geschätzt. Die Differenzierung hier ist ausschlaggebend für die Genauigkeit und Fairness der Bewertung.

2.1.4 Folge: „Benachteiligung“ oder „Schlechterstellung“

Die soziale Bewertung muss zu einer Benachteiligung oder Schlechterstellung in einem sozialen Kontext führen, der keinen Bezug zu den ursprünglichen Umständen der Datenerhebung hat. Zudem ist die Schlechterstellung im Hinblick auf die Tragweite des sozialen Verhaltens unverhältnismäßig oder ungerechtfertigt.

In Bezug auf die Kausalität ist erforderlich, dass der von der KI erstellte Score nachteilige Auswirkungen mit sich bringt und somit ursächlich für die Benachteiligung bzw. Schlechterstellung ist. Dieses Erfordernis soll den Leitlinien nach schon erfüllt sein, wenn die nachteilige Folge nicht schon eingetreten ist, das KI-System aber dazu bestimmt oder in der Lage ist, ein solches Ergebnis hervorzubringen. Nicht erforderlich sei, dass das durch das KI-System durchgeführte Scoring die einzige Ursache für die Benachteiligung oder Schlechterstellung ist. Mithin unschädlich ist, wenn eine andere Organisation den Score verwendet, als die, die ihn erstellt hat.

Zunächst gilt es, sich den Begriff der Benachteiligung näher anzuschauen. Gemeint ist, dass die Person oder Personengruppe aufgrund dieser Bewertung im Vergleich zu anderen weniger günstig behandelt wird, ohne dass ein besonderer Schaden oder eine besondere Beeinträchtigung vorliegen muss. Eine Schlechterstellung setzt im Gegensatz dazu voraus, dass ein bestimmter Schaden oder ein Nachteil erlitten wird.

Nachfolgend unterscheiden die Leitlinien hier zwischen zwei Szenarien, die gleichzeitig erfüllt sein können.

Szenario 1

Szenario 1 umfasst die Benachteiligung oder Schlechterstellung in einem sozialen Kontext, unabhängig von dem Kontext, in dem die Daten gesammelt wurden. Die Benachteiligung oder Schlechterstellung wirkt sich also in einem anderen Lebensbereich aus, als in dem, aus dem die gesammelten Daten über Sozialverhalten oder Persönlichkeitsmerkmale stammen. Hinzu kommt, dass die Daten für die Bewertung verwendet werden, ohne dass es einen offensichtlichen Zusammenhang zum Zweck der Bewertung oder Klassifizierung gibt. Somit erfolgt eine generalisierte Überwachung von Personen. Diese Praktik ist besonders kritisch, da sie in den meisten Fällen entgegen den berechtigten Erwartungen der Betroffenen und unter Verstoß gegen Unionsvorschriften, wie etwa dem Datenschutzrecht erfolgt. Das Vorliegen dieser Voraussetzung ist ebenfalls am Einzelfall unter Berücksichtigung aller Umstände zu prüfen.

Szenario 2

Szenario 2 bezieht sich darauf, dass die Benachteiligung oder Schlechterstellung nicht gerechtfertigt oder verhältnismäßig hinsichtlich der Schwere des sozialen Verhaltens ist. Zur Beurteilung dessen ist auf den allgemeinen Verhältnismäßigkeitsgrundsatz abzustellen. Weiterer Indikator für die Verhältnismäßigkeit soll der Vergleich zwischen Auswirkungen der Benachteiligung und Schwere des Sozialverhaltens der Person im Hinblick auf ein legitimes Ziel sein.

2.1.5 Nutzung des KI-Systems oder Bereitstellung von öffentlichen oder privaten Personen

Aber was genau ist nun verboten? Art. 5 Abs. 1 c) KI-VO verbietet Social Scoring unabhängig davon, ob das KI-System oder der Score von öffentlichen oder privaten Personen bereitgestellt oder verwendet wird, schließlich ist das Gefahrenpotential weitgehend gleich. Während im öffentlichen Sektor besonders das Machtungleichgewicht und die Abhängigkeit von öffentlichen Diensten beachtlich ist, so sind ähnliche Strukturen beispielsweise bei der Bepreisung von Versicherungen oder der Beurteilung der Kreditwürdigkeit von Personen zu sehen.

Um hier ein zulässiges Scoring erreichen zu können, liegt es an den Anbietern oder Betreibern, eine transparente Funktionsweise und Bereitstellung von Informationen über die Art der Daten und Datenquellen zu gewährleisten und nachzuweisen. Offensichtlich trägt auch die Einhaltung der geltenden Rechtsvorschriften sowie das Etablieren geeigneter Schutzmaßnahmen zur Rechtmäßigkeit bei.

2.2 Ausnahmen

Die Bewertung juristischer Personen ist nicht von dem Verbot umfasst, wenn die Bewertung nicht auf persönlichen oder Persönlichkeitsmerkmalen von Einzelpersonen beruht. Sehr wohl in den Anwendungsbereich fällt der Fall, dass juristische Personen auf Grundlage eines Gesamtergebnisses bewertet werden, welches sich auf die Evaluation einer Gruppe natürlicher Personen stützt und das Ergebnis dann eben diese Personen direkt betrifft (bspw. alle Arbeitnehmenden eines Unternehmens).

Weiterhin nicht von dem Verbot erfasst werden Bewertungspraktiken, die für einen bestimmten Zweck im Einklang mit etwaigen Rechtsvorschriften durchgeführt werden und verhältnismäßig sind. Beispielhaft führen die Leitlinien hier die Kreditwürdigkeitsprüfung oder Risikobewertung als wesentliche Aspekte der Dienstleistung im Finanz- und Versicherungssektor an.

2.3 Zusammenspiel mit anderen Unionsrechtsakten

Auch dieser Teil des Verbotes von Art. 5 KI-VO hängt eng mit anderen Rechtsvorschriften der EU zusammen. Die Leitlinien richten besonderen Augenmerk auf die Einhaltung des Verbraucherschutzes nach der UGP-Richtlinie, die Einhaltung der DSGVO, des EU Antidiskriminierungsrechts und die Verbraucherkreditrichtlinie (EU) 2023/2225. Diese Vorschriften können durch das Social Scoring berührt werden und sind daher ergänzend zu berücksichtigen.

Social Scoring soll nicht verboten sein, wenn die durch die Bewertung verfolgten Zwecke rechtmäßig sind und im Einklang mit dem nationalen Recht und Unionsrecht stehen. Mithin kann die soziale Bewertung auch durch Gesetze legitimiert werden, die die Rahmenbedingungen festlegen und sicherstellen, dass die aus dem Scoring resultierende Schlechterstellung oder Benachteiligung gerechtfertigt und verhältnismäßig ist.

3. Das Verbot der Risikobewertung gemäß Art. 5 Abs. 1 d) KI-VO

Art. 5 Abs. 1 d) KI-VO regelt das Verbot, KI-Systeme zur Bewertung des Risikos der Begehung einer Straftat einzusetzen. Hierbei geht es um eine individuelle Risikobewertung und Vorhersage von Strafdelikten, ausschließlich auf Grundlage von Profiling oder der Bewertung von Persönlichkeitsmerkmalen und Eigenschaften. Die Chancen der Verwendung solcher Systeme liegen darin, dass so die Effizienz der Strafverfolgung gesteigert und ein proaktiver Ansatz zur Aufdeckung, Abschreckung und Vorhersage von Straftaten ermöglicht werden kann. Das Risiko liegt aber darin, dass so Voreingenommenheit verstärkt werden und bestimmte relevante Umstände des Einzelfalls übersehen werden können. Weiter führen die Leitlinien an, dass diese Praktik auch das Vertrauen in die Strafverfolgung und das Justizsystem untergraben könnte. Schließlich erwähnt Erwägungsgrund 42 der KI-VO, dass im Hinblick auf die Unschuldsvermutung, Personen stets nach ihrem tatsächlichen Verhalten zu beurteilen sind. Eine Risikobewertung zur Begehung von Ordnungswidrigkeiten soll laut den Leitlinien nicht verboten sein, da dessen Verfolgung weniger grundrechtsinvasiv sei.

3.1 Begriffsbestimmungen für die Voraussetzungen des Verbots

Art. 5 KI-VO

1. Die folgenden AI-Praktiken sind verboten:

→ **(d)** das Inverkehrbringen, die Inbetriebnahme für diesen spezifischen Zweck oder die Verwendung eines KI-Systems zur Durchführung von Risikobewertungen in Bezug auf natürliche Personen, um das Risiko, dass eine natürliche Person eine Straftat begeht, ausschließlich auf der Grundlage des Profiling einer natürlichen Person oder der Bewertung ihrer persönlichen Merkmale und Eigenschaften zu bewerten oder vorherzusagen; dieses Verbot gilt nicht für KI-Systeme, die dazu verwendet werden, die durch Menschen durchge-führte Bewertung der Beteiligung einer Person an einer kriminellen Aktivität, die sich bereits auf objektive und überprüfbare Tatsachen stützt, die in unmittelbarem Zusammenhang mit einer kriminellen Aktivität stehen, zu unterstützen;

Das Gesetz verbietet KI-Systeme zu Risikobewertungen, ausschließlich auf Grundlage der Erstellung eines Profils einer Person oder der Bewertung ihrer Persönlichkeitsmerkmale. Weiterhin nennt es eine ausdrückliche Ausnahme des Verbots. Der Wortlaut enthält also unbestimmte Rechtsbegriffe und Voraussetzungen, die der Erläuterung bedürfen.

3.1.1 Risikobewertung oder Vorhersage der Wahrscheinlichkeit, dass eine Person eine Straftat begeht

Die hier gemeinten Bewertungen und Prognosen werden auch oft als individuelle „Kriminalitätsvorhersagen“ bezeichnet. Diese Begrifflichkeit zeichnet aus, dass im Allgemeinen eine KI-Technologie und Analysemethode auf eine große Menge von Daten angewandt werden und in Kombination mit kriminologischen Theorien zur Vorhersage von Kriminalität getroffen wird. Diese Vorhersage ist dann die Grundlage für Maßnahmen der Polizei oder Strafverfolgungsbehörde zur Bekämpfung, Kontrolle und Verhütung von Kriminalität.

3.1.2 Bewertung oder Vorhersage auf Grundlage des Profils oder der Persönlichkeitsmerkmale einer Person

Eine weitere Bedingung für das Verbot ist, dass die Risikobewertung oder Vorhersage ausschließlich auf der Erstellung eines Profils oder der Beurteilung ihrer Persönlichkeitsmerkmale und Eigenschaften beruht. Dies gilt unabhängig davon, ob sich das Verbot auf ein Individuum oder eine Gruppe bezieht. Hinsichtlich der Profilerstellung bezieht sich Art. 5 Abs. 1 d) KI-VO ausdrücklich auf den „Profiling“ Begriff der DSGVO. Wie schon oben erläutert, sind Persönlich-keitsmerkmale derart subjektiv und ohne allgemein anerkannte Taxonomie, wodurch das Verbot seine Rechtfertigung finden soll.

3.1.3 Ausschließlichkeitsmerkmal

Das Verbot sieht es vor, dass es hinsichtlich der Bewertung nur greifen soll, wenn sie ausschließlich auf den oben genannten Merkmalen beruht. Das Verbot ist also relativ eng zu sehen, greift aber laut den Leitlinien jedenfalls dann, wenn das KI-System zur Unterstützung der menschlichen Einschätzung einer Person eingesetzt wird, die bereits auf objektiv überprüfbaren Tatsachen beruht, die in direktem Zusammenhang mit einer kriminellen Tätigkeit stehen – also bspw. wenn bereits ein begründeter Verdacht gegen die betreffende Person besteht. Entscheidend ist, ob dem Einsatz der KI eine menschliche Einschätzung vorgeht.

3.1.4 Die ausdrückliche Ausnahme der menschlichen Vorbeurteilung

In seinem letzten Satz sieht Art. 5 Abs. 1 d) KI-VO ausdrücklich vor, dass das Verbot nicht gilt, wenn die KI zur Unterstützung der menschlichen Beurteilung dient. Das KI-System würde aber trotzdem als KI-System mit hohem Risiko eingestuft werden, sodass es den besonderen Anforderungen und Sicherheitsvorkehrungen von Art. 14 und Art. 26 KI-VO unterliegt. Die menschliche Bewertung soll sicherstellen, dass die Risikovorhersage durch die KI auf objektiv überprüfbaren Fakten beruht und ggf. eingegriffen werden kann, um negative Risiken oder Folgen zu vermeiden.

3.2 Private Akteure

Auch Private Akteure können neben den Strafverfolgungsbehörden von dem Verbot erfasst werden, wenn diese gesetzlich mit der Ausübung öffentlicher Gewalt und öffentlicher Befugnisse zur Verhütung, Ermittlung, Feststellung oder Verfolgung von Straftaten, oder auch der Vollstreckung strafrechtlicher Sanktionen betraut sind.

3.3 Außerhalb des Anwendungsbereichs

Die ortsbezogene oder georäumliche Kriminalitätsvorhersage beruht auf dem jeweiligen Ort der Straftat oder auf der Wahrscheinlichkeit, dass in diesen Gebieten Straftaten begangen werden. Da diese Beurteilung sich nicht auf eine bestimmte Person bezieht, soll sie nicht von dem Verbot erfasst sein. Das Verbot gilt auch nicht, wenn sich die Prognose auf juristische Personen wie Unternehmen oder NGO's bezieht. Auch die Vorhersage der Begehung von Ordnungswidrigkeiten fällt nicht in den Anwendungsbereich, da deren Verfolgung in der Regel weniger grundrechtsinvasiv ist.

4. Das Verbot der Erstellung einer Datenbank zur ungezielten Gesichtserkennung, Art. 5 Abs. 1 e) KI-VO

Diese Vorschrift verbietet das Inverkehrbringen, die Inbetriebnahme zu diesem Zweck oder die Verwendung von KI-Systemen, die Datenbanken zur Gesichtserkennung durch das ungezielte Auslesen von Gesichtsbildern aus dem Internet oder aus Videoaufnahmen erstellen oder erweitern. Das Verbot findet seinen Ursprung laut Erwägungsgrund 43 der KI-VO in der erheblichen Schwere des Grundrechtseingriffs, der hiermit einhergeht. Betroffen sind unter anderem das Recht des Einzelnen auf Privatsphäre, Datenschutz und das Recht, anonym zu bleiben.

4.1 Hauptbegriffe und Voraussetzungen des Verbots

Auch hier gilt es zunächst die Tatbestandsmerkmale des Art. 5 Abs. 1 e) KI-VO genauer zu betrachten: Erklärungsbedürftige Voraussetzungen des Verbots sind insoweit die Datenbank, ungezieltes Scraping und das Internet oder Videoüberwachungsaufnahmen als Quelle.

4.1.1 Datenbank zur Gesichtserkennung

Eine Datenbank ist den Leitlinien nach, jede Sammlung von Daten oder Informationen, die speziell für eine schnelle Suche und Abfrage durch einen Computer organisiert ist. Eine Gesichtserkennungsdatenbank ist in der Lage, ein menschliches Gesicht mit Bildern oder Videos einer Datenbank abzugleichen und festzustellen, ob eine wahrscheinliche Übereinstimmung zwischen den beiden besteht. Nicht erforderlich ist, dass die Datenbank allein dem Zweck der Gesichtserkennung dient.

4.1.2 Ungezieltes Scraping von Gesichtsbildern

Art. 5 KI-VO

1. Die folgenden AI-Praktiken sind verboten:

→ **(e)** das Inverkehrbringen, die Inbetriebnahme für diesen spezifischen Zweck oder die Verwendung von KI-Systemen, die Datenbanken zur Gesichtserkennung durch das ungezielte Auslesen von Gesichtsbildern aus dem Internet oder von Überwachungsaufnahmen erstellen oder erweitern;

Scraping meint die Verwendung von Webcrawlern, Bots oder anderen Mitteln, um Daten oder Inhalte aus verschiedenen Quellen (z.B. Videoüberwachung, Websites, soziale Medien) automatisch zu extrahieren. Dieses Vorgehen ist als ungezielt anzusehen, wenn wahllos möglichst viele Daten und Informationen gesammelt werden, ohne gezielt auf ein Individuum und einen konkreten Personenkreis abzustellen. Anders herum ist gezieltes Scraping, also die Sammlung von Bildern und Videos einer bestimmten Person – bspw. zur Opferidentifizierung, offensichtlich nicht von dem Verbot erfasst.

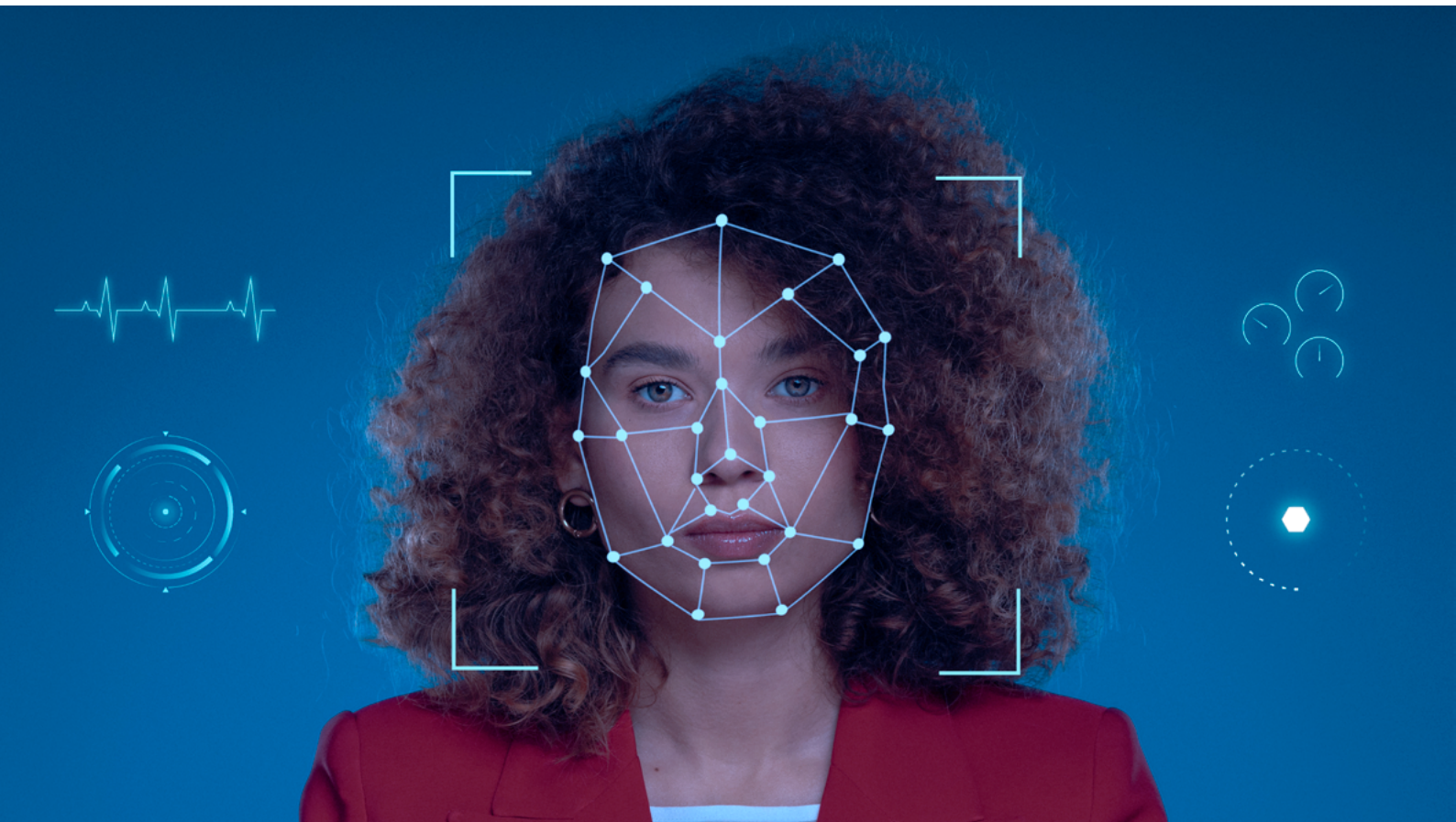
4.1.3 Aus dem Internet oder Videoüberwachungsaufnahmen

Sofern es sich bei der Quelle um das Internet handelt, gilt der Grundsatz, dass allein die Tatsache, dass eine Person Bilder mit Gesicht in den sozialen Medien veröffentlicht, nicht bedeutet, dass sie automatisch ihr Einverständnis hinsichtlich der Aufnahme dieser Bilder in eine Gesichtserkennungsdatenbank gegeben hat.

Die Überwachungskamera als Quelle zeichnet häufig Material an öffentlichen Orten wie Flughäfen, Straßen, Parks, etc. auf. Gerade dort kann von einem (vermuteten) Einverständnis nicht ausgegangen werden, wird der Betroffene doch häufig überhaupt nicht wissen, dass er aufgezeichnet worden ist.

4.2 Außerhalb des Geltungsbereichs

Ausgenommen aus dem Geltungsbereich ist das ungezielte Auslesen von anderen biometrischen Daten als Gesichtsbildern, oder wenn nicht ein KI-System an dem Scraping beteiligt ist. Gesichtsbilddatenbanken, die nicht zur Erkennung von Personen verwendet werden, sondern bspw. lediglich vom Trainieren von KI-Modellen, sind nicht von dem Anwendungsbereich erfasst. Das Verbot gilt auch nicht für KI-Systeme, die mittels der gesammelten Bilder neue Bilder von fiktiven Personen erzeugen.



5. Das Verbot der KI-Systeme zur Emotionserkennung nach Art. 5 Abs. 1 f) KI-VO

Art. 5 KI-VO

1. Die folgenden AI-Praktiken sind verboten:

→ **(f)** das Inverkehrbringen, die Inbetriebnahme für diesen spezifischen Zweck oder die Verwendung von KI-Systemen zur Ableitung von Emotionen einer natürlichen Person am Arbeitsplatz und in Bildungseinrichtungen, es sei denn, die Verwendung des KI-Systems soll aus medizinischen Gründen oder Sicherheitsgründen eingeführt oder auf den Markt gebracht werden;

In Erwägungsgrund 44 der KI-VO wird deutlich, dass sich dieses Verbot vor allem darin begründet, dass laut der Leitlinien die Wirksamkeit und Genauigkeit von Emotionserkennung angezweifelt wird. Solchen Systemen mangele es vor allem an Zuverlässigkeit, Spezifität und Verallgemeinerbarkeit.

Ein wichtiger Aspekt, der in den Leitlinien hervorgehoben wird, ist die Transparenzpflicht für den Einsatz von Emotionserkennungssystemen. Gemäß Art. 50 Abs. 3 KI-VO müssen Anbieter und Betreiber solcher Systeme bestimmte Transparenzanforderungen erfüllen.

5.1 Begriffsbestimmung und Voraussetzungen

Die verschiedenen Voraussetzungen wie die Emotionserkennung, die Verwendung am Arbeitsplatz oder Bildungseinrichtung sowie die explizite Ausnahme bedürfen der besseren Handhabbarkeit wegen einer Klarstellung.

5.1.1 Erkennen von Emotionen

Das Verbot erfordert eine präzise Unterscheidung zwischen KI-Systemen zur Emotionserkennung und solchen, die dazu bestimmt sind, Emotionen oder Absichten natürlicher Personen auf Grundlage ihrer biometrischen Daten zu erkennen und abzuleiten. Es bezieht sich also nur auf solche Systeme, die explizit darauf ausgelegt sind, Rückschlüsse auf die Emotionen der natürlichen Person zu ziehen. Grundlage dieser Ableitung bilden biometrischen Daten. Biometrische Daten sind personenbezogene Daten, die sich aus spezifischen technischen Verarbeitungen in Bezug auf die physischen, physiologischen oder Verhaltensmerkmale einer natürlichen Person ergeben, wie z.B. Gesichtsbilder oder daktyloskopische Daten. Beispiele aus den Leitlinien, die Emotionserkennung betreffen, sind Gesichtsausdrücke, Gesten oder Merkmale der Stimme.

Die Begriffe „Emotionen“ oder „Absichten“ sind weit zu verstehen und umfassen eine Vielzahl von Gefühlen wie Glück, Traurigkeit, Wut, Überraschung, Ekel, Verlegenheit, Aufregung, Scham, Verachtung, Zufriedenheit und Vergnügen – aber eben nicht Schmerz oder Müdigkeit.

Die bloße Erfassung von offensichtlichen Ausdrücken, Gesten oder Bewegungen stellt keine Emotionserkennung dar, es sei denn, sie werden zur Identifizierung oder Ableitung von Emotionen verwendet. Zum Beispiel ist die Beobachtung, dass eine Person lächelt, keine Emotionserkennung. Die Schlussfolgerung, dass diese Person glücklich ist, stellt hingegen eine Emotionserkennung dar.

5.1.2 Bereichsbeschränkung

Das Verbot ist auf die Bereiche Arbeitsplatz und Bildungseinrichtungen beschränkt. Diese Abgrenzung basiert vor allem auf dem Machtungleichgewicht und der vulnerablen Position der betroffenen Person in diesem Kontext. Der Arbeitsplatz umfasst das gesamte Arbeitsverhältnis, einschließlich des Einstellungsverfahrens. Bildungseinrichtungen umfassen alle Ebenen der Bildung.

5.1.3 Ausnahmen

Es gibt jedoch Ausnahmen von diesem Verbot, insbesondere wenn der Einsatz eines KI-Systems aus medizinischen oder Sicherheitsgründen beabsichtigt ist, wie z.B. Systeme für therapeutische Zwecke. Unter Therapeutischen Verwendungen sind in der Regel Anwendungen von CE-gekennzeichneten Medizinprodukten zu verstehen. Die allgemeine Überwachung des Stresslevels am Arbeitsplatz fällt nicht unter die Ausnahme für Gesundheits- oder Sicherheitsaspekte, so sollen Systeme zur Erkennung von Burnout oder Depressionen weiterhin verboten sein.

Arbeitgeber und Bildungseinrichtungen dürfen Emotionserkennungssysteme nur dann aus medizinischen oder Sicherheitsgründen einsetzen, wenn ein expliziter Bedarf besteht – z.B. wenn dieser von entsprechenden Experten angezeigt wird. Es ist dabei zu beachten, dass die gesammelten Daten ausschließlich für den festgelegten Zweck genutzt werden dürfen. Zudem sind Systeme zur Erkennung körperlicher Zustände wie Schmerz oder Müdigkeit von der Definition der Emotionserkennung ausgeschlossen. So fallen Systeme zur Überwachung der Ermüdung von Berufspiloten oder Fahrern zur Unfallverhütung nicht unter das Verbot.

Ein weiteres Beispiel für die Ausnahme aus medizinischen Gründen ist die Verwendung von Emotionserkennung zur Unterstützung von Menschen mit Autismus oder zur Verbesserung der Zugänglichkeit für blinde oder gehörlose Personen. Dagegen fällt der Einsatz von Emotionserkennung zur Bewertung des Wohlbefindens, der Motivation oder der Zufriedenheit von Schülern oder Mitarbeitern nicht unter diese Ausnahme und bleibt somit verboten.

5.2 Außerhalb des Anwendungsbereichs

Emotionserkennungssysteme, die nicht auf biometrischen Daten beruhen, wie etwa die Stimmungsanalysen von Texten, sollen nicht unter dieses Verbot fallen. Ebenso unterliegen Emotionserkennungssysteme, die außerhalb des Arbeitskontext oder von Bildungseinrichtungen eingesetzt werden – z.B. im kommerziellen Kontext zur Kundenansprache – nicht dem Verbot.

Systeme zur „Crowd Control“, die beispielsweise allein das allgemeine Geräusch- und Stimmungsniveau an einem bestimmten Ort analysieren, fallen nicht unter das Verbot. Wenn solche Systeme jedoch auch die Emotionen von Einzelpersonen ableiten (z.B. wütende Gesichter erkennen), werden sie unter das Verbot fallen, soweit sie am Arbeitsplatz oder in Bildungseinrichtungen eingesetzt werden. Auch Systeme, die im medizinischen Bereich eingesetzt werden, wie z. B. Pflegeroboter oder die Verwendung von Emotionserkennungssystemen durch Ärzte bei einer Untersuchung an ihrem Arbeitsplatz, sowie Spracherkennungssysteme zur Analyse von Notrufen, fallen in der Regel ebenfalls nicht unter das Verbot.

Emotionserkennungssysteme, die nicht unter das Verbot fallen (weil sie beispielsweise außerhalb des Arbeitsplatzes oder von Bildungseinrichtungen eingesetzt werden), werden gemäß Art. 6 Abs. 2 und Anhang III Nummer 1 c) der KI-VO in der Regel als Hochrisiko-KI-Systeme einzustufen sein und müssen dann die entsprechenden Anforderungen erfüllen.

5.3 Öffnungsklausel für Mitgliedsstaaten

Mitgliedstaaten können günstigere nationale Rechtsvorschriften beibehalten oder einführen. So könnten Mitgliedstaaten beispielsweise Gesetze erlassen, die die Verwendung von Emotionserkennungssystemen im Arbeitsbereich für medizinische Zwecke gänzlich untersagen.

6. Das Verbot gemäß Art. 5 Abs. 1 g) KI-VO zur biometrischen Kategorisierung

Art. 5 KI-VO

1. Die folgenden AI-Praktiken sind verboten:

→ **(g)** das Inverkehrbringen, die Inbetriebnahme für diesen spezifischen Zweck oder die Verwendung von Systemen zur biometrischen Kategorisierung, mit denen natürliche Personen individuell auf der Grundlage ihrer biometrischen Daten kategorisiert werden, um ihre Rasse, ihre politischen Einstellungen, ihre Gewerkschaftszugehörigkeit, ihre religiösen oder weltanschaulichen Überzeugungen, ihr Sexualleben oder ihre sexuelle Ausrichtung zu erschließen oder abzuleiten; dieses Verbot gilt nicht für die Kennzeichnung oder Filterung rechtmäßig erworbener biometrischer Datensätze, wie z. B. Bilder auf der Grundlage biometrischer Daten oder die Kategorisierung biometrischer Daten im Bereich der Strafverfolgung;

Art. 5 Abs. 1 g) KI-VO verbietet KI-Systeme zur biometrischen Kategorisierung, um so Rückschlüsse auf sensible Merkmale wie Ethnie, politische Meinung, religiöse oder philosophische Überzeugung, sexuelle Orientierung, etc. zu ziehen. Schon Erwägungsgrund 30 der KI-VO statuiert, dass das Verbot nicht gelten soll, wenn die biometrischen Daten rechtmäßig erworben wurden.

6.1 Begriffsbestimmungen

Die folgenden Erläuterungen der relevanten Begrifflichkeiten innerhalb der Vorschrift sollen dem leichteren Verständnis dienen.

6.1.1 Biometrische Kategorisierungssystem

Ein biometrisches Kategorisierungssystem ist ein KI-System, das natürliche Personen anhand ihrer biometrischen Daten in vordefinierte Gruppen einordnet. Zu den relevanten biometrischen Daten gehören bspw. Gesichtsmerkmale, Hautfarbe, DNA oder Verhaltensaspekte wie die Analyse des Tippverhaltens oder des Gangs. Ziel einer solchen Kategorisierung ist es, spezifische Merkmale einer Person zu identifizieren und auf dieser Basis Rückschlüsse auf sensible Informationen zu ziehen, wie etwa politische Meinungen, Gewerkschaftszugehörigkeit, religiöse oder philosophische Überzeugungen, Sexualleben oder sexuelle Orientierung.

Es gibt jedoch auch Ausnahmen von diesem Verbot. Eine Kategorisierung ist nicht verboten, wenn sie als Zusatz zu einem anderen kommerziellen Dienst verwendet wird und aus objektiv technischen Gründen unbedingt erforderlich ist. Das bedeutet, dass eine Kategorisierung nur dann zulässig ist, wenn sie für die Erbringung des Hauptdienstes unabdingbar ist. Ein praktisches Beispiel für eine verbotene Nutzung wäre ein KI-System auf einer Social-Media-Plattform, das Benutzer auf der Grundlage biometrischer Daten in Fotos nach ihrer mutmaßlichen politischen Orientierung kategorisiert, um ihnen gezielt politische Nachrichten zu senden. In diesem Fall wäre die Kategorisierung nicht notwendig und somit unzulässig.

6.1.2 Individuelle Kategorisierung auf Basis biometrischer Daten

Wichtig ist dabei, dass das Verbot der biometrischen Kategorisierung ausschließlich auf die individuelle Zuordnung von Personen zu bestimmten Gruppen abzielt. Eine Kategorisierung, die auf einer ganzen Gruppe basiert, fällt nicht unter das Verbot. Es geht also nur um die Kategorisierung einzelner Personen anhand ihrer biometrischen Daten, nicht um die Analyse von Gruppen als Ganzes.

6.2 Außerhalb des Anwendungsbereichs

Bestimmte Situationen fallen nicht unter das Verbot. So erfasst das Verbot nicht die Kennzeichnung, Filterung oder Kategorisierung von rechtmäßig erfassten biometrischen Datensätzen, einschließlich im Bereich der Strafverfolgung. Ein Beispiel hierfür wäre die Sortierung von Bildern nach Haar- oder Augenfarbe. Ziel ist es, sicherzustellen, dass die Daten alle demografischen Gruppen gleichmäßig repräsentieren und um Diskriminierung zu verhindern.

6.3 Verknüpfung mit dem Datenschutzrecht

Art. 5 Abs. 1 g) KI-VO schränkt die Möglichkeiten einer rechtmäßigen Verarbeitung personenbezogener Daten im Einklang mit Unionsdatenschutzrecht (DSGVO, LED, EUDPR) weiter ein. In diesem Kontext wird auf Art. 11 Abs. 3 LED verwiesen, der Profiling verbietet, wenn es zur Diskriminierung aufgrund besonderer Kategorien personenbezogener Daten führt.

7. Das Verbot der biometrischen Echtzeit-Fernidentifizierungssysteme nach Art. 5 Abs. 1 h) KI-VO

Art. 5 KI-VO

1. Die folgenden AI-Praktiken sind verboten:

→ **(h)** die Verwendung biometrischer Echtzeit-Fernidentifizierungssysteme in öffentlich zugänglichen Räumen zu Strafverfolgungszwecken, außer wenn und insoweit dies im Hinblick auf eines der folgenden Ziele unbedingt erforderlich ist:

- → **i)** gezielte Suche nach bestimmten Opfern von Entführung, Menschenhandel oder sexueller Ausbeutung sowie die Suche nach vermissten Personen;
- → **ii)** Abwenden einer konkreten, erheblichen und unmittelbaren Gefahr für das Leben oder die körperliche Unversehrtheit natürlicher Personen oder einer tatsächlichen und bestehenden oder tatsächlichen und vorhersehbaren Gefahr eines Terroranschlags;
- → **iii)** Aufspüren oder Identifizieren einer Person, die der Begehung einer Straftat verdächtigt wird, zum Zwecke der Durchführung von strafrechtlichen Ermittlungen oder von Strafverfahren oder der Vollstreckung einer Strafe für die in Anhang II aufgeführten Straftaten, die in dem betreffenden Mitgliedstaat nach dessen Recht mit einer Freiheitsstrafe oder einer freiheitsentziehen-den Maßregel der Sicherung im Höchstmaß von mindestens vier Jahren bedroht ist.

Unterabsatz 1 Buchstabe h gilt unbeschadet des Artikels 9 der Verordnung (EU) 2016/679 für die Verarbeitung biometrischer Daten zu anderen Zwecken als der Strafverfolgung.

Gemäß Art. 5 Abs. 1 h) KI-VO sind KI-Systeme verboten, die dazu dienen, Personen aus der Ferne – ohne nennenswerte zeitliche Verzögerung – durch Abgleich der biometrischen Daten mit den in einer Referenzdatenbank gespeicherten biometrischen Daten zu identifizieren. Begründet wird dies in Erwägungsgrund 32 der KI-VO mit dem dadurch erzeugten Gefühl der ständigen Überwachung und dem massiven Eingriff in die Grundrechte. Außerdem bringen die begrenzte Überprüfbarkeit und mangelnde Möglichkeit zur Korrektur dieser Systeme erhebliche Risiken mit sich.

7.1 Begriffsbestimmungen

Die Begriffe „Biometrische Echtzeit-Fernidentifizierungssysteme“, „öffentlich zugängliche Räume“ und „Strafverfolgungszwecke“ bedürfen einer Erläuterung.

7.1.1 Biometrische Echtzeit-Fernidentifizierungssysteme

Biometrische Echtzeit-Fernidentifizierungssysteme sind KI-Systeme, die dem Zweckdienen, Personen aus der Ferne durch Abgleich der biometrischen Daten einer Person mit den in einer Referenzdatenbank gespeicherten biometrischen Daten zu identifizieren. Beispiel: Videobasierte Gesichtserkennung.

7.1.2 Öffentlich zugängliche Räume

Als öffentlich zugängliche Räume im Sinne dieser Vorschrift zählen u.a. Geschäfte, Restaurants, Cafés, Banken, Schwimmbäder, Fitnessstudios, Stadien, Bus- und U-Bahn-Haltestellen, Bahnhö-fe, Flughäfen, Transportmittel, Kinos, Konzert- und Konferenzsäle (vgl. Erwägungsgrund 19 der KI-VO).

7.1.3 Strafverfolgungszwecke

Dies umfasst Tätigkeiten der Strafverfolgungsbehörden oder in deren Auftrag zur Verhütung, Ermittlung, Aufdeckung oder Verfolgung von Straftaten oder zur Strafvollstreckung, einschließlich des Schutzes vor und der Abwehr von Gefahren für die öffentliche Sicherheit.

Das Verbot adressiert nicht nur Strafverfolgungsbehörden selbst, sondern auch private Stellen, die im Auftrag der Strafverfolgungsbehörden handeln. Die Leitlinien der EU-Kommission benennen beispielhaft öffentliche Verkehrsunternehmen, die von den Strafverfolgungsbehörden aufgefordert werden, die Sicherheit in den öffentlichen Verkehrsnetzen unter deren Anweisungen und Aufsicht zu gewährleisten. Zudem Sportverbände, die von den Strafverfolgungsbehörden ersucht werden, unter deren Anweisungen und Aufsicht für die Sicherheit bei Sportveranstaltungen zu sorgen.

Private Stellen, die mit Strafverfolgungsbehörden zu Strafverfolgungszwecken zusammenarbeiten, sollten sich daher mit Art. 5 Abs. 1 h) KI-VO auseinandersetzen und prüfen, (i) ob die von ihnen ggf. bereits einleiten oder geplanten KI-Systeme in den Anwendungsbereich

des Verbotes in Art. 5 Abs. 1 h) KI-VO fallen und (ii) wenn ja, ob eine Ausnahme des Verbotes zu ihren Gunsten einschlägig ist. Auch Software-Hersteller im Bereich der biometrischen Echtzeit-Fernidentifizierungssystemen sollten sich mit den Möglichkeiten und Grenzen der Norm befassen, um das eigene Geschäftsmodell weiterhin auf rechtlich belastbare Füße zu stellen.

7.2 Ausnahmen des Anwendungsbereichs

Biometrische Echtzeit-Fernidentifizierungssysteme zu Strafverfolgungszwecken sind dann erlaubt, wenn und insoweit ihr Einsatz zu folgenden Zielen „unbedingt erforderlich“ ist:

(1) gezielte Suche nach Opfern von Entführung, Menschenhandel oder sexueller Ausbeutung sowie die Suche nach vermissten Personen
(2) Abwenden einer konkreten und erheblichen Gefahr für das Leben oder die körperliche Unversehrtheit von Personen oder einer Gefahr eines Terroranschlags;
(3) Auffinden einer Person, die im Verdacht steht, eine Straftat begangen zu haben, die im Anhang II der KI-VO aufgeführt wird und die im betreffenden Mitgliedstaat mit Freiheitsstrafe von im Höchstmaß mindestens vier Jahren geahndet werden könnte.

Unter Bezugnahme auf Ausnahme (3) wäre folgender Einsatz eines Echtzeit-Fernidentifizierungssysteme denkbar: Ein Flughafenbetreiber setzt im Auftrag und in Zusammenarbeit mit den Polizeibehörden Live-Gesichtserkennungstechnologien ein, um das Flughafengebäude zu überwachen und gesuchte Personen mit ausstehenden Haftbefehlen wegen schwerer Straftaten zu identifizieren. An verschiedenen Punkten im Flughafengebäude verwendet der Flughafenbetreiber Live-Videomaterial von Personen, die vor einer mobilen Kamera vorbeigehen, um ihre Gesichter mit einer Fahndungsliste gesuchter Personen zu vergleichen.

7.3 Die Öffnungsklausel aus Art. 5 Abs. 5 KI-VO

Art. 5 Abs. 5 der KI-VO enthält eine Öffnungsklausel. Danach können die Mitgliedstaaten die Möglichkeiten des Einsatzes von biometrischen Echtzeit-Fernidentifizierungssystemen zu Strafverfolgungszwecken „ganz oder teilweise“ durch nationales Recht regeln. Die KI-VO bildet insoweit den europarechtlichen Rahmen, an dem die nationalen Rechtsgrundlagen für den Einsatz von biometrischen Echtzeit-Fernidentifizierungssystemen zu messen sind. Gemäß Art. 5 Abs. 3 der KI-VO muss im nationalen Recht geregelt werden, bei welchen Delikten und zum Schutz welcher Rechtsgüter Echtzeit-Fernidentifizierungssysteme zum Einsatz kommen können. In Deutschland ist das nichts Neues, da verfassungsrechtlich ohnehin eine nationale Rechtsgrundlage geboten ist, die das Bestimmtheits- und Verhältnismäßigkeitsgebot beachtet.

III) Ausblick und Praxishinweis

Die Leitlinien der Europäischen Kommission zu verbotenen Praktiken im KI-Bereich verdeutlichen die strengen Anforderungen, die an Anbieter und Betreiber von KI-Systemen gestellt werden und stellen eine wichtige Orientierungshilfe für Unternehmen und Behörden dar. Denn die Auslegung der Verbote ist für Unternehmen und Behörden nicht einfach und der Interpretationsspielraum des Art. 5 KI-VO ist groß. Die Anforderungen des Art. 5 KI-VO zielen darauf ab, die Autonomie, Entscheidungsfreiheit und die Schutzbedürftigkeit von Personen zu wahren und Missbrauch zu verhindern. Die verbotenen Praktiken, wie manipulative Techniken, Social Scoring und Risikobewertung von Straftaten, stellen erhebliche Eingriffe in die Grundrechte dar und müssen daher besonders sorgfältig überwacht und reguliert werden.

Es ist zu erwarten, dass die Europäische Kommission und die Mitgliedstaaten kontinuierlich an der Weiterentwicklung und Präzisierung der Vorschriften im KI-Bereich arbeiten werden. Dies könnte zusätzliche Klarstellungen und Anpassungen der bestehenden Regelungen umfassen, um neuen technologischen Entwicklungen und Herausforderungen gerecht zu werden. Empfehlenswert ist weiterhin das Unternehmen und Behörden die Rechtsprechung des EuGH und nationaler Gerichte zu verbotenen KI-Praktiken aufmerksam verfolgen.

Anbieter und Betreiber von KI-Systemen sollten zunächst überprüfen, ob ihr KI-System von dem Anwendungsbereich der in Art. 5 KI-VO aufgelisteten Verbote erfasst werden. Sofern dies der Fall ist, sollte in Erwägung gezogen werden, ob die etwaigen Ausnahmen greifen.

Bei dem Einsatz von KI-Systemen gegenüber besonders schutzbedürftigen Gruppen wie Kindern und älteren Menschen ist besondere Vorsicht geboten. So sollten zusätzliche Schutzmaßnahmen implementiert werden, um Missbrauch und Schäden zu verhindern.

Angesichts der Komplexität der Regelungen ist es ratsam, rechtliche Beratung in Anspruch zu nehmen, um sicherzustellen, dass alle gesetzlichen Anforderungen erfüllt sind und das eigene Geschäftsmodell auf rechtlich belastbaren Füßen steht. Durch die Beachtung dieser Hinweise können Anbieter und Betreiber von KI-Systemen nicht nur rechtliche Risiken minimieren, sondern auch das Vertrauen der Nutzer in ihre Technologien stärken. Ebenso sollte das Zusammenspiel der Regelungen des Art. 5 KI-VO mit anderem Unionsrecht, insbesondere der DS-GVO berücksichtigt werden. Denn die KI-Verordnung findet neben anderen Unionsrechtsvorschriften Anwendung.

Ihre Ansprechpartner bei uns

Jan-Dierk Schaal
Partner

+49 (0)40 3 34 01 - 340
j.schaal@skwschwarz.de

Dr. Oliver Hornung
Partner

+49 (0)69 63 00 01 - 65
o.hornung@skwschwarz.de

Moritz Mehner
Partner

+49 (0)89 2 86 40 - 206
m.mehner@skwschwarz.de

Dr. Christoph Krück
Counsel

+49 (0)89 2 86 40 - 268
c.krueck@skwschwarz.de

Fabian Bauer
Counsel

+49 (0)69 63 00 01 - 82
f.bauer@skwschwarz.de

Weitere Autorin

Franziska Sofie Wulf
Wissenschaftl. Mitarbeiterin

10719 Berlin
Kranzler Eck
Kurfürstendamm 21
T +49 30 8892650-0
F +49 30 8892650-10

60598 Frankfurt/Main
Mörfelder Landstraße 117
T +49 69 630001-0
F +49 69 6355-22

20457 Hamburg
Willy-Brandt-Straße 59
T +49 40 33401-0
F +49 40 33401-530

80333 München
Wittelsbacherplatz 1
T +49 89 28640-0
F +49 89 28094-32

skwschwarz.de