Der EU-Kommissionsvorschlag einer Verordnung für künstliche Intelligenz aus rechts- und kriminalpolitischer Perspektive – Zur unumgänglichen aber nachbesserungsbedürftigen Harmonisierung

von Jana Engelhard und
Prof. Dr. Anja Schiemann *

Abstract
Am 21.4.2021 wurde von der Europäischen Kommission der Entwurf für eine Verordnung zur „Festlegung harmonisierter Vorschriften für künstliche Intelligenz" (KI-VO-E) vorgelegt. Ziel ist es, einen einheitlichen Rechtsrahmen für vertrauenswürdige künstliche Intelligenz zu schaffen. Durch den Verordnungsvorschlag zur künstlichen Intelligenz soll gewährleistet werden, dass KI-Systeme die Grundrechte der Union beachten und so Vertrauen und Rechtssicherheit geschaffen wird. Die KI-VO-E verfolgt einen sog. „risikobasierten Regulierungsansatz", so dass je nach Risiko der KI-Anwendung für die Europäischen Grundrechte höhere oder niedrigere Anforderungen etwa in Bezug auf die Sicherheit oder Transparenz aufgestellt werden. So sehr insgesamt die Harmonisierung zu begrüßen ist, so muss doch am Feinschliff der Verordnung noch gearbeitet werden, um zu überzeugen. Dies gilt auch für diverse Privilegierungen der Strafverfolgungsbehörden. Diesen wird zwar – zu Recht – mehr gestattet als Privaten, allerdings ist ein sachlicher Grund nicht immer zu erkennen.

On 21.4.2021, the European Commission presented the draft for a regulation to "establish harmonized rules for artificial intelligence" (AI Act). The aim is to create a uniform legal framework for trustworthy artificial intelligence. The proposed regulation on artificial intelligence aims to ensure that AI systems respect the fundamental rights of the Union, thus creating trust and legal certainty. The AI Act follows a so-called "risk-based regulatory approach", so that depending on the risk of the AI application, higher or lower requirements are set for European fundamental rights, for example in terms of security or transparency. As much as the harmonization is to be welcomed overall, the fine-tuning of the regulation still needs to be worked on to be convincing. This also applies to various privileges for law enforcement agencies. They are - rightly - allowed more than private individuals, but the underlying objective reason is not always apparent.

I. Rückblick

Die Bestrebungen der Europäischen Kommission, deutliche Akzente auf dem Feld der künstlichen Intelligenz (KI) zu setzen, reicht knapp 20 Jahre zurück. Bereits 2004 wurde verstärkt Forschung im Bereich der KI durch die Kommission gefördert und 2018 eine hochrangige Expertengruppe eingesetzt,[1] die 2019 Ethik-Leitlinien für vertrauenswürdige KI veröffentlichte.[2] Im Februar 2020 stellte die Kommission in ihrem Weißbuch zur KI ihre klare Zielvorstellung vor, nämlich ein Europäisches Konzept der Exzellenz und des Vertrauens im Bereich der KI zu schaffen.[3] Im Rahmen der sich anschließenden Diskussion gingen über 1.500 Beiträge ein, die, ebenso wie viele Studien und Dokumente sowie die Arbeiten der Expertengruppe, Eingang in den Verordnungsentwurf fanden.[4] Dieser Verordnungsvorschlag zur Festlegung harmonisierender Vorschriften für KI wurde dann am 21.4.2021 veröffentlicht, um einheitliche Standards innerhalb der Europäischen Union für künstliche Intelligenz festzulegen.[5] Mittlerweile hat sich auch der Rechtsausschuss des Europäischen Parlaments mit dem Entwurf beschäftigt und eine Stellungnahme verfasst.[6] Am 29.9.2022 hat der Bundestag erstmals einen Antrag der CDU/CSU-Fraktion mit dem Titel „Europäische KI-Verordnung – Raum lassen für Innovation und Wettbewerbsfähigkeit" beraten.[7] In dem Antrag wurde die Bundesregierung insbesondere dazu aufgefordert, ausreichend und regelmäßig zum Stand der Verhandlungen zur KI-Verordnung zu informieren und inhaltliche Anregungen des Deutschen Bundestags in die Verhandlungen einfließen zu lassen. Den Verhandlungen zur KI-Verordnung im Rat der EU solle endlich die Bedeutung eingeräumt werden, die dem Regelwerk mit

* *Jana Engelhard* ist Wissenschaftliche Mitarbeiterin an der Universität zu Köln und sowohl mit der Betreuung des Examenskurses Strafrecht betraut als auch im Projekt COMBI tätig. *Anja Schiemann* ist Universitätsprofessorin und Lehrstuhlinhaberin des Lehrstuhls für Strafrecht und Strafprozessrecht an der Universität zu Köln.

1 Meldung abrufbar unter: https://ec.europa.eu/commission/presscorner/detail/de/IP_18_1381 (zuletzt abgerufen am 22.10.2022).

2 Diese Guidelines können abgerufen werden unter: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (zuletzt abgerufen am 14.11.2022).

3 Das Weißbuch ist abrufbar unter: https://ec.europa.eu/info/files/white-paper-artificial-intelligence-european-approach-excellence-and-trust_de. Wichtige Meilensteine der Entwicklung sind abrufbar unter: https://digital-strategy.ec.europa.eu/de/policies/european-approach-artificial-intelligence (beides zuletzt abgerufen am 14.11.2022).

4 S. *Orssich*, EuZW 2022, 245 (255), so bspw. auch der Bericht der deutschen Datenethikkommission oder der Bericht der Enquete-Kommission KI des Deutschen Bundestags.

5 Der Vorschlag für eine Verordnung ist abrufbar unter: https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206 (zuletzt abgerufen am 14.11.2022).

6 Die Stellungnahme ist abrufbar unter: https://www.europarl.europa.eu/doceo/document/JURI-PA-719827_DE.pdf (zuletzt abgerufen am 22.10.2022).

7 Hier ging es um die BT-Drs. 20/3689 v. 27.9.2022.

weitreichenden Auswirkungen auf die durch KI entstehenden Chancen für Gesellschaft, Wirtschaft und Forschung zukomme.[8]

II. Überblick über den Verordnungsvorschlag zur KI

1. Einführung

Die angemahnten weitreichenden Auswirkungen liegen zum einen darin begründet, dass der „Artificial Intelligence Act" in Form einer Verordnung erlassen werden soll, die nach Art. 288 Abs. 2 AEUV ohne weiteren Umsetzungsakt zu unmittelbar anwendbarem Recht in Deutschland wird. Durch die unmittelbare Anwendbarkeit in jedem EU-Mitgliedstaat soll eine Fragmentierung des Binnenmarktes verhindert werden. Denn einige Mitgliedstaaten haben bereits eine Verabschiedung nationaler Vorschriften zur Regulierung von KI-Systemen in Erwägung gezogen. Die einheitliche Anwendbarkeit der Verordnung soll durch unterschiedliche länderspezifische Gesetze bedingten Binnenmarkthindernissen entgegenwirken und einen einheitlichen Rechtsrahmen schaffen.[9] Gleichzeitig soll durch die harmonisierten gesetzlichen Vorgaben für Bürger Vertrauen in die Nutzung von KI-Systemen geschaffen und Unternehmen motiviert werden, KI-Lösungen zu entwickeln.[10]

Anders als noch im Weißbuch[11] wird eine differenziertere Risikoeinteilung vorgenommen. KI-Systeme werden in vier Kategorien eingeteilt:

- Verbotene KI-Anwendungen (Art. 5 KI-VO-E)
- Hochrisiko-KI-Systeme (Art. 6-51 KI-VO-E)
- KI-Systeme mit geringem Risiko und mit besonderen Transparenzanforderungen (Art. 52-55 KI-VO-E)
- KI-Systeme mit minimalem Risiko (Art. 69 KI-VO-E)

2. Anwendungsbereich des Verordnungsentwurfs

a) Personeller und räumlicher Anwendungsbereich

Die Verordnung soll nach Art. 2 Abs. 1 KI-VO-E für Anbieter gelten, die KI-Systeme in der Union in Verkehr bringen oder in Betrieb nehmen. Dies soll unabhängig davon gelten, ob diese Anbieter in der EU oder einem Drittland niedergelassen sind. Weiterhin erstreckt sich der territoriale Anwendungsbereich auf Nutzer von KI-Systemen, die sich in der Union befinden. Darüber hinaus gilt das Verwendungsprinzip, nach dem auch solche Anbieter und Nutzer von KI-Systemen vom Anwendungsbereich der Vorschrift erfasst sind, die in einem Drittland niedergelassen oder ansässig sind, soweit das vom System hervorgebrachte Ergebnis in der Union verwendet wird. Die zentralen Begriffe Anbieter, Nutzer und KI-Systeme werden in Art. 3 KI-VO-E definiert. Anbieter ist eine natürliche oder juristische Behörde, Einrichtung oder sonstige Stelle, die ein KI-System entwickelt oder entwickeln lässt, um es unter ihrem eigenen Namen oder ihrer eigenen Marke entgeltlich oder unentgeltlich in Verkehr zu bringen. Dadurch, dass der KI-VO-E mit den Anbietern Entwickler von KI-Systemen in die Verantwortung nimmt, setzt er im Gegensatz zur DS-GVO bereits beim Entstehungsprozess an.[12]

Nutzer sind natürliche oder juristische Personen, Behörden, Einrichtungen oder sonstige Stellen, die ein KI-System in eigener Verantwortung verwenden, es sei denn, das KI-System wird im Rahmen einer persönlichen und nicht beruflichen Tätigkeit genutzt. Insofern soll die KI-VO zwar für Anbieter und professionelle Nutzer, nicht aber für private Endnutzer gelten.[13]

b) Sachlicher Anwendungsbereich – die Definition von KI-Systemen

Das im Zentrum des KI-VO-E stehende *System der künstlichen Intelligenz* wird zur begrifflichen Handhabbarmachung in Art. 3 Nr. 1 KI-VO-E als *„eine Software, die mit einer oder mehreren der in Anhang I aufgeführten Techniken und Konzepte entwickelt worden ist und im Hinblick auf eine Reihe von Zielen, die vom Menschen festgelegt werden, Ergebnisse wie Inhalte, Vorhersagen, Empfehlungen oder Entscheidungen hervorbringen kann, die das Umfeld beeinflussen, mit dem sie interagieren"* definiert. Die im Verweis auf Anhang I erfassten Techniken und Konzepte spezifiziert der KI-VO-E wie folgt:

- *Konzepte des maschinellen Lernens, mit beaufsichtigtem, unbeaufsichtigtem und bestärkendem Lernen unter Verwendung einer breiten Palette von Methoden, einschließlich des tiefen Lernens (Deep Learning)*
- *Logik- und wissensgestützte Konzepte, einschließlich Wissensrepräsentation, induktiver (logischer) Programmierung, Wissensgrundlagen, Inferenz- und Deduktionsmaschinen, (symbolischer) Schlussfolgerungs- und Expertensysteme*
- *Statistische Ansätze, Bayessche Schätz-, Such- und Optimierungsmethoden*

Mit der in Art. 3 Nr. 1 KI-VO-E erfassten Begriffsdefinition erfolgt die Abkehr der EU-Kommission von ihren erstmals im Jahr 2018, in der Mitteilung über KI in Europa, geprägten Ansätzen der Umgrenzung des KI-Begriffs.[14] Letztere beschreiben KI-Systeme abstrahiert als *„Systeme mit einem „intelligenten" Verhalten"*,[15] bevor exemplarisch konkretisierende technische Verfahren

---

8 BT-Drs. 20/3689, S. 3.

9 Vgl. *Ebers*, RDi 2021, 588 (589); *Engelmann/Brunotte/Lütkens*, RDi 2021, 317.

10 S. KI-VO-E, S. 1; abrufbar unter https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC_1&format=PDF (zuletzt abgerufen am 17.11.2022).

11 Hier gab es lediglich eine Grobeinteilung in hohes Risiko und kein hohes Risiko.

12 S. *Ebert/Spiecker*, NVwZ 2021, 1188.

13 Vgl. *Engelmann/Brunotte/Lütkens*, RDi 2021, 317 (319).

14 Vgl. Mitteilung KI für Europa, 2018, S. 1; abrufbar unter https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52018DC0237&from=de (zuletzt abgerufen am 14.11.2022).

15 A.a.O.

ebenso wie Anwendungsbeispiele aufgeführt werden. Als Ausgangspunkt der Ermittlung einer einheitlichen Terminologie diente auf EU-Ebene somit zunächst der vorwiegend psychologisch geprägte Intelligenzbegriff. Dieser ist nicht nur seinerseits keiner konsentierten Definition zugänglich,[16] sondern aufgrund des ihm ureigenen Menschenbezugs erweisen sich Einzelheiten betreffend dessen Übertragbarkeit auf technische Anwendungen als weitestgehend ungeklärt.[17] Eine bedeutsame Änderung des definitorischen Bezugspunkts, die EU-Kommission selbst spricht von einer *Präzisierung*,[18] hat der KI-Begriff durch die seitens der EU-Kommission eingesetzte hochrangige Expertengruppe für KI (HEG-KI) erfahren. Indem diese KI-Systeme Intelligenz sinngemäß als Software, die durch Dateninterpretation komplexe Ziele zu lösen imstande ist, beschreibt,[19] bewegt sie den Fokus weg vom Intelligenzbegriff und hin zu technisch-datenbasierten Verfahren, die die Funktionsweise von KI prägen. Begrüßenswert ist es im Angesicht dieser kursorischen Historie der Bemühungen zur Umgrenzung des KI-Begriffs auf EU-Ebene, dass sich die Kommission in Art. 3 Nr. 1 KI-VO-E für eine an technischen Eigenschaften orientierte Grundausrichtung der Begriffsbestimmung entschieden hat.[20]

Ein vergleichbarer *Paradigmenwechsel* bleibt indes im Hinblick auf die Extensität der gewählten KI-Definition zu vermissen, obschon der EU-Kommission die Nachteile einer zu weit gefassten Begriffsbestimmung hinlänglich bekannt sein dürften. Nicht nur betrifft diese Kritik bereits die KI-Regulierungsmodelle der EU-Kommission prä-KI-VO-E,[21] sondern im Weißbuch der EU-Kommission zur Regulierung künstlicher Intelligenz mahnt ebendiese selbst an, dass eine KI-Definition zwar einerseits flexibel und entwicklungsoffen, andererseits aber auch *„präzise [sein muss], um die erforderliche Rechtssicherheit zu gewährleisten"*.[22] An jener Präzision mangelt es in der Legaldefinition des Art. 3 Nr. 1 KI-VO-E, da diese aufgrund der nicht näher spezifizierten Inklusion logik- und wissensgestützter Systeme sowie statistischer Ansätze ebenso einfache deterministische Software wie Expertensysteme erfasst.[23] Begründen lässt sich der dahinter stehende technikneutrale Ansatz auch nicht allein mit dem Ziel der Schaffung einer möglichst zukunftstauglichen Definition,[24] da deren Reichweite vor allem Software mit umfasst, die keine spezifischen Risiken künstlicher Intelligenz wie z.B. Autonomie,[25] Lernfähigkeit[26] oder Intransparenz[27] aufweist. Insofern erlegt Art. 3 Nr. 1 KI-VO-E konventioneller Software strenge Regulierungsmaßnahmen auf, die technischen Fortschritt in diesem Bereich eher zu hemmen drohen. Ausräumen lässt sich diese Kritik auch nicht unter Bezugnahme auf die in Art. 5 KI-VO-E geregelten verbotenen Praktiken. Dieser stützt sich zwar nicht auf funktionale Eigenschaften der eingesetzten Technik, sondern auf Risiken, die aus ihrer Verwendung resultieren, sodass unerheblich sein könnte, ob die verbotene Praktik per KI-System oder konventioneller Software Anwendung findet.[28] Indes bezieht sich die Legaldefinition des Art. 3 Nr. 1 KI-VO-E ebenso auf den Einsatz von Hochrisiko-KI (Art. 6 ff. KI-VO-E), deren Regulierung sich in der KI-VO-E selbst unmittelbar an Grundrechtsgefährdungen knüpft, die aus den besonderen Eigenschaften künstlicher Intelligenz herrühren.[29] Gleiches ergibt sich allgemein aus den Gründen und Zielen des Regulierungsvorschlags, die KI-Spezifika *„…wie Undurchsichtigkeit, Komplexität, der sogenannte „Bias", ein gewisses Maß an Unberechenbarkeit und teilweise autonomes Verhalten"* aufgreifen, welche konventioneller Software nicht vergleichbar immanent sind.[30] Eine rein ergebnisbezogene Sichtweise ist somit nicht angezeigt. Ein Zugewinn an terminologischer Flexibilität wird hingegen durch die in Art. 4, 73 KI-VO-E angelegte Option der Anpassung des Anhang I zum KI-VO-E durch delegierte Rechtsakte erreicht, weil dies eine stete Adaption an technischen Fortschritt gewährleistet.[31] Entbehrlich wird eine eindeutig auf künstliche Intelligenz zugeschnittene Definition dadurch aber nicht.

3. Die risikobasierte Einstufung von KI-Systemen

Der Vorschlag der KI-VO folgt einem sog. risikobasierten Ansatz und stellt für unterschiedliche Risikokategorien von KI-Systemen unterschiedliche Anforderungen auf.[32] Wie bereits oben erwähnt, werden vier Risikostufen benannt.

---

16 Zu verschiedenen Definitionsansätzen des Intelligenzbegriffs vgl. nur *Gruber/Stamouli*, in: Pädagogische Psychologie, Intelligenz und Vorwissen, S. 28 f.; zu Unklarheiten des Intelligenzbegriffs im Kontext von KI vgl. nur *Kaplan,* Künstliche Intelligenz 2017, S. 15 f.; *Herberger,* NJW 2018, 2825 (2826 f.); *Nink,* Justiz und Algorithmen 2020, S. 146 m.w.N.

17 Ausführlich *Kaplan,* Künstliche Intelligenz, 2017, S. 15 ff.

18 Vgl. Europäische Kommission, COM(2020) 65 final – Weißbuch KI, 2020, S. 19.

19 Für den vollständigen Wortlaut der Definition vgl. HEG-KI, Guidelines, 2019, S. 47.

20 Ebenso *Roos/Weitz,* MMR 2021, 844 (845).

21 Vgl. nur *Dettling/Krüger,* MMR 2019, 211 (212); *Herberger,* NJW 2018, 2825 (2827) mit direktem Verweis auf das zu weite KI-Begriffsverständnis der EU-Kommission in Fn. 23.

22 Europäische Kommission, COM(2020) 65 final – Weißbuch KI, 2020, S. 19; *Lachenmann,* ZD-Aktuell 2020, 07021 spricht diesbezüglich von gängigen Definitionsmaßstäben.

23 So *Spindler,* CR 2021, 361 (362 f.), wonach die Legaldefinition in Art. 3 Nr. 1 KI-VO-E deswegen sogar extensiver als jene der HEG-KI ist; ebenfalls mangelnde Präzision attestieren KI-Bundesverband, Feedback AI-Act 2021, S. 1; *Ebers/Hoch et al.,* RDi 2021, 529; *Roos/Weitz,* MMR 2021, 844 (845); Bitkom e.V., Positionspapier zum EU-Regulierungsrahmen für KI 2021, S. 2.

24 So aber Europäische Kommission, COM(2021) 206 final – KI-VO-E 2021, S. 14.

25 Dazu *Werner,* NJOZ 2019, 1041.

26 Vgl. *Wischmeyer,* AöR 2018, 1 (4); *Kirn/Müller-Hengstenberg,* MMR 2014, 225 (226 f.).

27 Vgl. *Hagendorff,* Österreich Z Soziol 2019, 53 (54) bezogen auf maschinelles Lernen.

28 In diese Richtung *Ebert/Spiecker gen. Döhmann,* NVwZ 2021, 1188 (1189).

29 Vgl. Europäische Kommission, COM(2021) 206 final – KI-VO-E 2021, S. 12; so auch *Ebers/Hoch et al.,* RDi 2021, 529.

30 Europäische Kommission, COM(2021) 206 final – KI-VO-E 2021, S. 2.

31 Dazu auch *Rostalski/Weiss,* ZfDR 2021, 329 (331), die treffend von einer "dynamischen Konzeption" sprechen.

32 Vgl. auch *Kevekorde*, in: Hdb. Multimedia-Recht, Stand: 58. ErgLg. (März 2022), Teil 29.1 Rn. 48; *Engelmann/Brunotte/Lütkens*, RDi 2021, 317 (319); *Rostalski/Weiss*, ZfDR 2021, 329 (337).

a) Verbotene KI-Systeme mit unannehmbarem Risiko

Von vornherein verboten werden gem. Art. 5 KI-VO-E KI-Systeme mit unannehmbarem Risiko. Diese verletzen nach Ansicht der Europäischen Kommission Europäische Grundrechte und Werte, so dass diese schädlichen Anwendungsmöglichkeiten von KI-Systemen von vorherein verboten sind.[33] Insgesamt lassen sich bei der Aufstellung der verbotenen Praktiken in Art. 5 KI-VO-E drei Kategorien unterscheiden.

aa) Verhaltensmanipulation

Verboten werden zum einen bestimmte KI-Systeme, die zu Zwecken der Verhaltensmanipulation eingesetzt werden. Art. 5 Abs. 1 lit. a KI-VO-E benennt hier den Einsatz von Techniken der unterschwelligen Beeinflussung außerhalb des Bewusstseins einer Person, um das Verhalten einer Person in einer Weise wesentlich zu beeinflussen, die dieser Person oder einer anderen Person einen physischen oder psychischen Schaden zufügt oder zufügen kann.[34]

bb) Social Scoring

Als verbotene Praktik wird des Weiteren das sog. Social Scoring[35] genannt. Gem. Art. 5 Abs. 1 lit. c KI-VO-E wird die Inbetriebnahme oder die Verwendung von KI- Systemen durch Behörden oder in deren Auftrag verboten, die zur Bewertung oder Klassifizierung der Vertrauenswürdigkeit natürlicher Personen über einen bestimmten Zeitraum auf der Grundlage ihres sozialen Verhaltens oder bekannter oder vorhergesagter persönlicher Eigenschaften oder Persönlichkeitsmerkmale dient. Dabei muss die soziale Bewertung zu einer Schlechterstellung oder Benachteiligung in sozialen Zusammenhängen führen, die in keiner Beziehung zu der ursprünglichen Datenerhebung stehen oder ungerechtfertigt oder unverhältnismäßig sind.[36] Durch die Einschränkung „durch Behörden oder in deren Auftrag" ist das Social Scoring mittels KI-Systemen im rein privaten Sektor nicht verboten.

cc) Biometrische Echtzeit-Fernidentifizierungssysteme im öffentlichen Raum

Verboten ist schließlich der Einsatz von biometrischen Echtzeit-Fernidentifizierungssystemen in öffentlich zugänglichen Räumen für Zwecke der Strafverfolgung gem. Art. 5 Abs. 1 lit. d KI-VO-E. Definiert wird ein „biometrisches Fernidentifizierungssystem" in Art. 3 Nr. 36 KI-VO-E. Danach dienen diese Systeme dem Zweck, natürliche Personen aus der Ferne durch Abgleich biometrischer Daten einer Person mit den gespeicherten biometrischen Daten einer Referenzdatenbank zu identifizieren, ohne dass der Nutzer des KI-Systems vorher weiß, ob die Person anwesend sein wird und identifiziert werden kann. Nach Art. 3 Nr. 33 KI-VO-E sind biometrische Daten solche zu physischen, physiologischen oder verhaltenstypischen Merkmalen einer natürlichen Person, die eine eindeutige Identifizierung dieser Person ermöglichen oder bestätigen. Angesichts dieser Weite der Legaldefinition sind KI-Systeme zur Gesichtserkennung, aber auch zur Analyse von Bewegungsmustern oder Retina-Scans erfasst.[37] Verboten sind nur „Echtzeit"-Fernidentifizierungssysteme, worunter gem. Art. 3 Nr. 37 KI-VO-E für die Klassifizierung erforderlich ist, dass der Abgleich und die Identifizierung ohne erhebliche Verzögerung erfolgen, wobei auch die Identifizierung mit begrenzten kurzen Verzögerungen erfasst ist.[38] Dagegen fallen System zur nachträglichen biometrischen Fernidentifizierung gem. Anhang III Nr. 1 a KI-VO-E unter die hochriskanten KI-Systeme (hierzu unten unter b). Nicht erfasst werden sonstige Identifikationen, die der Nachidentifizierung dienen. Dies ist auch nach dem KI-VO-E möglich, da diese biometrischen Identifizierungen nur den KI-Systemen mit einem geringen Risiko zugeordnet werden (hierzu unter c). Warum die Nachidentifizierung ausgenommen ist, wird in der Begründung nicht weiter spezifiziert. Dieser hätte es aber bedurft, denn schließlich dient das Verbot dem Schutz vor einem „gläsernen Menschen" im Sinne eines Orwell´schen Überwachungsstaats.[39] Insofern kann die Fernidentifikation aufgrund biometrischer Daten als solche zu Verunsicherungen in der Bevölkerung und abschreckenden Effekten führen.[40]

Auch für die Verwendung biometrischer Echtzeit-Fernidentifizierungssysteme in öffentlich zugänglichen Räumen zu Strafverfolgungszwecken sind Ausnahmen vorgesehen, von denen ein Mitgliedsstaat gem. Art. 5 Abs. 4 KI-VO-E Gebrauch machen darf.[41] Hierdurch wird das Verbot nach Art. 5 Abs. 1 lit. d KI-VO-E erheblich aufgeweicht. Der Einsatz von Echtzeit-Fernidentifizierungssystemen kann danach erlaubt sein

- zur gezielten Suche nach potentiellen Opfern oder nach vermissten Kindern,
- zum Abwenden einer konkreten, erheblichen und unmittelbaren Gefahr für das Leben oder die körperliche Unversehrtheit natürlicher Personen oder eines Terroranschlags oder
- zum Erkennen, Aufspüren, Identifizieren oder Verfolgen eines Täters oder Verdächtigen einer Straftat im Sinne des Übergabebeschlusses über den Europäischen Haftbefehl und die Übergabeverfahren zwischen den Mitgliedsstaaten, der in dem betreffenden Mitgliedsstaat nach dessen

---

33 S. Europäische Kommission, COM(2021) 206 final – KI-VO-E 2021, S. 15.

34 Zu der Auslegung des Begriffs der „unterschwelligen Beeinflussung außerhalb des menschlichen Bewusstseins" vgl. ausführlich *Rostalski/Weiss*, ZfDR 2021, 329 (338).

35 Vgl. zum Begriff *Conrad/Hauser*, in: Auer-Reinsdorff/Conrad (Hrsg.), Hdb. IT- und Datenschutzrecht, 3. Aufl. (2019), § 36 Rn. 244 ff.

36 Ausf. zu Social Scoring und Art. 5 Abs. 1 Nr. 1 lit. c KI-VO-E vgl. *Linardatos*, GPR 2022, 58 (61); *Rostalksi/Weiss*, ZfDR 2021, 329 (340 f.).

37 S. *Valta/Vasel*, ZRP 2021, 142 (143); *Rostalski/Weiss*, ZfDR 2021, 329 (342).

38 S. *Rostalski/Weiss*, ZfDR 2021, 329 (342).

39 So *Valta/Vasel*, ZRP 2021, 142 (143); diese zitierend *Rostalski/Weiss*, ZfDR 2021, 329 (342).

40 So auch EDPB-EDPS, Joint Opinion on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence, 18.6.2021, S. 12; *Linardatos*, GPR 2022, 58 (62).

41 Insofern ist Art. 5 Abs. 1 lit. d KI-VO-E auch nicht unmittelbar anwendbar, sondern bedarf einer Umsetzung in nationales Recht.

Recht mit einer Freiheitsstrafe oder einer freiheitsentziehenden Maßregel der Sicherung im Höchstmaß von mindestens drei Jahren bedroht ist.

Die Verwendung der Identifikationssysteme steht unter einem Genehmigungsvorbehalt durch eine zuständige Justiz- oder Verwaltungsbehörde, es sei denn, es besteht Gefahr in Verzug (Art. 5 Abs. 3 KI-VO-E). Darüber hinaus normiert Art. 5 Abs. 2 S. 1 KI-VO-E abstrakte Kriterium der stets erforderlichen Interessenabwägung. Zu berücksichtigen sind danach das Ausmaß und die Wahrscheinlichkeit von Schäden, die Schwere des Eingriffs in Grundrechtspositionen sowie die Wahrscheinlichkeit und das Ausmaß der Folgen. Nach Art. 5 Abs. 2 S. 2 KI-VO-E sind zudem notwendige sowie verhältnismäßige Schutzvorkehrungen und Bedingungen für die Verwendung einzuhalten, insbesondere hinsichtlich zeitlicher, geografischer und personenbezogener Beschränkungen. Gemeint ist damit insbesondere die Pflicht zur Zweckbindung, Datenlöschung aber auch Datenvermeidung und Datensparsamkeit, die bereits aus der DS-GVO folgt.[42]

Durch die zahlreichen Ausnahmen vom Verbot hat die Vorschrift faktisch eher den Charakter eines Erlaubnisdenn eines Verbotstatbestands mit Vorbehalt einer Erlaubnis. Durch diese Öffnungsklausel ist es quasi vorprogrammiert, doch wieder länderspezifische unterschiedliche Ausgestaltungen der biometrischen Echtzeitidentifizierung im Rahmen der Strafverfolgung festzuschreiben und so statt der gewünschten Einheitlichkeit eine Fragmentierung im Bereich der Strafverfolgung voranzutreiben. Insofern sollten die Ausnahmetatbestände klarer normiert und der Kreis eng gezogen werden. Auch fehlt ein prinzipielles Verbot automatischer Erkennung sensibler Personenmerkmale wie Sexualität, Herkunft, Religion etc., die gar keine Relevanz für Maßnahmen der Strafverfolgung haben und insofern hätten ausgespart werden können.[43]

b) Hochrisiko-KI-Systeme

Im Zentrum des Verordnungsentwurfs steht die in Art. 6 ff. KI-VO-E erfasste Regulierung hochriskanter KI-Systeme. Entsprechend der Risiko-Methodik des KI-VO-E, ist eine Einstufung als hochriskant vorzunehmen, sofern *erhebliche* Sicherheits-, Gesundheitsrisiken, oder Risiken für die Grundrechte einer Person im Laufe des gesamten Lebenszyklus eines KI-Systems, d.h. von seiner Entwicklung bis hin zu seiner Verwendung, aufzutreten vermögen.[44] Eine Spezifizierung erfährt der normativ ausfüllungsbedürftige Erheblichkeitsbegriff durch die über eine Folgenbetrachtung hinausgehenden sektorbezogenen Ansätze des Art. 6 KI-VO-E.[45]

aa) Klassifizierung hochriskanter KI-Systeme

Art. 6 Abs. 1 lit. a KI-VO-E qualifiziert KI-Systeme als solche mit einem hohen Risiko, die entweder als Produkt selbst oder als Sicherheitskomponente eines Produkts verwendet werden,[46] das unter die in Anhang II KI-VO-E spezifizierten Harmonisierungsvorschriften fällt. Kumulativ gibt Art. 6 Abs. 1 lit. b KI-VO-E vor, dass jene EU-Harmonisierungsvorschriften eine Konformitätsbewertung durch Dritte betreffend das Inverkehrbringen oder Inbetriebnehmen des Produkts vorzusehen haben. Die in Anhang II KI-VO-E aufgeführten Harmonisierungsvorschriften gründen sich vorwiegend auf das *New Legislative Framework (NLF)*, einer Sammlung von EU-Rechtsakten, die übergreifende Produktsicherheitsvorschriften ebenso wie gemeinsame Verfahren zur Konformitätsbewertung festlegt.[47]

Nach Art. 6 Abs. 2 KI-VO-E gelten zudem eigenständige KI-Systeme, also solche die nicht bereits Produktsicherheitsvorschriften des *NLF* unterfallen als hochriskant, sofern diese in Anhang III des KI-VO-E benannt sind. Die darin enthaltene Aufzählung inkludiert die übergeordneten Sektoren „Biometrische Identifizierung und Kategorisierung natürlicher Personen“ (Anhang III Nr. 1 KI-VO-E),[48] „Verwaltung und Betrieb kritischer Infrastrukturen“ (Anhang III Nr. 2 KI-VO-E), „Allgemeine und berufliche Bildung“ (Anhang III Nr. 3 KI-VO-E), „Beschäftigung, Personalmanagement und Zugang zur Selbstständigkeit“ (Anhang III Nr. 4 KI-VO-E), „Zugänglichkeit und Inanspruchnahme grundlegender privater und öffentlicher Dienste und Leistungen“ (Anhang III Nr. 5 KI-VO-E), „Strafverfolgung“ (Anhang III Nr. 6 KI-VO-E), „Migration, Asyl und Grenzkontrolle“ (Anhang III Nr. 7 KI-VO-E) sowie „Rechtspflege und demokratische Prozesse“ (Anhang III Nr. 8 KI-VO-E), welche jeweils durch konkretere Anwendungsszenarien der KI-Systeme ausdifferenziert werden. Abschließend ist diese Aufzählung nicht, da gemäß Art. 7 Abs. 1 i.V.m. Art. 73 KI-VO-E seitens der Kommission delegierte Rechtsakte erlassen werden können, um die Aufzählung in Anhang III KI-VO-E zu modifizieren. Diese Option wird in Art. 7 Abs. 1 lit. a KI-VO-E insofern beschränkt, als dass Änderungen des Anhang III KI-VO-E lediglich im Rahmen der in Nr. 1 – 8 festgelegten, übergeordneten Bereiche erfolgen können. Obschon die Konzeption der Festlegung abstrahierter KI-Einsatzsektoren in Kombination mit einer nachträglichen Änderungsoption wegen der zugleich strukturierenden Wirkung und Zukunftstauglichkeit begrüßenswert ist, überzeugt deren Realisierung im KI-VO-E nicht. Die Auswahl übergeordneter Sektoren deckt schon aus aktueller

42 Vgl. *Linardatos*, GPR 2022, 58 (62); *Spindler*, CR 2021, 361 (365); *Petermann/Sauter*, Biometrische Identifikationssysteme, 2002, S. 92.

43 Insgesamt wie hier auch *Linardatos*, GPR 2022, 58 (62); ähnlich *Bomhard/Merkle*, RDi 2021, 276; *Ebers/Hoch/Rosenkranz/Ruschemeier/Steinrötter*, RDi 2021, 528. Kritisch auch *Rostalski/Weiss*, ZfDR 2021, 329 (343 f.).

44 Vgl. *Europäische Kommission,* COM(2021) 206 final – KI-VO-E 2021, S. 4.

45 Der Begriff des sektorbezogenen Ansatzes geht zurück auf *Ebert/Spiecker gen. Döhmann,* NVwZ 2021, 1188 (1190).

46 Legaldefiniert ist der Begriff der Sicherheitskomponente in Art. 3 Nr. 14 KI-VO-E.

47 Vgl. *Spindler,* CR 2021, 361 (366).

48 Nicht nachvollziehbar ist, wieso Art. 3 Nr. 35 KI-VO-E auch die sexuelle oder politische Ausrichtung zu biometrischer Kategorisierung zählt; zu Recht kritisch *Chander/Jakubowska*, EU's AI law needs major changes to prevent discrimination and mass surveillance; abrufbar unter: https://edri.org/our-work/eus-ai-law-needs-major-changes-to-prevent-discrimination-and-mass-surveillance/ (zuletzt abgerufen am 14.11.2022).

Perspektive nicht alle denkbaren hochriskanten Einsatzszenarien von KI-Systemen ab.[49] Nicht nur steht diese Einschränkung einer potentiellen Einbeziehung des – in Anbetracht des hohen Risikopotentials ohnehin schwer nachvollziehbar – exkludierten militärischen KI-Einsatzes in Art. 2 Abs. 3 KI-VO-E entgegen.[50] Es sind zudem auch bereits bekannte Sektoren wie z.B. der Gesundheitssektor in einer Dimension, die über Art. 6 Abs. 1 i.V.m. Anhang II KI-VO-E hinausgeht, der Versicherungs- oder der Immobiliensektor nicht erfasst.[51] Umso mehr Probleme wird es bereiten, dass noch nicht absehbare, übergeordnete Sektoren, in denen sich der Einsatz von KI-Systemen als hochriskant erweist, von der nachträglichen Änderungsmöglichkeit ausgeschlossen sind. Ein derart starrer Sektorenbezug[52] überzeugt auch deswegen nicht, weil Art. 7 Abs. 1 lit. b, Abs. 2 KI-VO-E ohnehin zusätzlich vorschreiben, dass das Gefährdungspotential der Anhang III KI-VO-E ergänzenden KI-Einsatzarten denjenigen, der bereits darin erfassten Arten mindestens vergleichbar sein muss. Warum eine qualitative Vergleichbarkeit nicht auch für die Ergänzung von Sektoren ausreichen kann, erschließt sich nicht.

bb) Folgen einer Klassifizierung als hochriskantes KI-System

Als Konsequenz einer Klassifizierung als Hochrisiko-KI-System werden in den Art. 8 ff. KI-VO-E daran zu stellende Anforderungen geregelt und sodann Pflichten aller an der Verwendung des KI-Systems Beteiligten in Art. 16 ff. KI-VO-E festgeschrieben. Einzelheiten betreffend die Konformitätsbewertung des Systems befinden sich in Art. 40 ff. KI-VO-E. Neben der allgemeinen Auseinandersetzung mit diesen Kategorien erfolgt deren Exemplifizierung an der Identifizierung und Kategorisierung natürlicher Personen anhand biometrischer Merkmale.

(1) Anforderungen (Art. 8 ff. KI-VO-E)

Intendiert wird mit der Pflicht zur Umsetzung der im KI-VO-E formulierten Anforderungen für hochriskante KI-Systeme nach Art. 8 Abs. 2 KI-VO-E, die jeweilige Zweckbestimmung des KI-Systems zu gewährleisten und dem gem. Art. 9 KI-VO-E zu implementierenden Risikomanagementsystem Rechnung zu tragen. Diese Zielsetzung legt den Grund für die Realisierung des von der EU-Kommission verfolgten risikobasierten Regulierungsansatzes, da nur die klar umgrenzte Zwecksetzung des KI-Systems ein treffendes Risikomanagement ermöglicht, welches zugleich spezifisch aus seiner Konstitution erwachsenden Gefahren begegnet ohne die Öffnung für Fortschritt über das erforderliche Maß hinausgehend zu begrenzen. Ob die Anforderungen der Art. 10 ff. KI-VO-E dieses Ziel zu erreichen vermögen, gilt es zu beurteilen. Normiert werden *„Daten und Daten-Governance"*-Vorgaben in Art. 10 KI-VO-E, die *„Technische Dokumentation"* des KI-Systems zum Erfüllungsnachweis der hier zur Sprache stehenden Anforderungen in Art. 11 KI-VO-E, *„Aufzeichnungspflichten"* in Art. 12 KI-VO-E zum Zwecke der Protokollierung des laufenden KI-System-Betriebs, zudem gem. Art. 13 KI-VO-E *„Transparenz und Bereitstellung von Informationen für Nutzer"* um Handlungen und Ergebnisse des KI-Systems für diese nachvollziehbar zu machen. Weiterhin unterstellt Art. 14 KI-VO-E Hochrisiko-KI-Systeme *„menschliche[r] Aufsicht"* und Art. 15 KI-VO-E fordert deren *„Genauigkeit, Robustheit und Cybersicherheit"*. Wenngleich diese Regelungen in weiten Teilen den erarbeiteten Anforderungen der HEG-KI entsprechen,[53] fällt beim Vergleich beider Anforderungsregime auf, dass der KI-VO-E den von der HEG-KI formulierten Katalog der *„Vielfalt, Nichtdiskriminierung und Fairness"* zwar in seinen Erwägungsgründen jedenfalls unter dem Stichwort der Nichtdiskriminierung abdeckt, diesen aber mit keinem Wort im Verordnungsentwurf selbst erwähnt.[54] Zwar dient Art. 10 Abs. 3 KI-VO-E vor allem durch die Vorgabe der Repräsentativität von *„Trainings-, Validierungs- und Testdatensätze[n]"* der Vermeidung diskriminierender Effekte. Sektorabhängig drohen beispielsweise durch den Einsatz hochriskanter KI-Systeme zur biometrischen Identifizierung und Klassifizierung natürlicher Personen (vgl. Art. 6 Abs. 2 i.V.m. Anhang III Nr. 1 KI-VO-E) Diskriminierungsgefahren, denen nicht allein durch repräsentative Datensatzgestaltung zu begegnen ist. Um nur einige Beispiele zu nennen, können in diesem Fall z.B. *automation biases*[55] oder *default-Effekte*[56] ebenso zu Benachteiligungen oder Ungleichbehandlungen führen, wie das Einfließen bewusst oder unbewusst diskriminierender Annahmen programmierender Personen in die Ausgestaltung des KI-Systems.[57] Potentielle Folgen solcher *biases* und weiterer diskriminierender Effekte können im Beispielfall unmittelbar zur schlechteren Performanz des biometrisch identifi-

---

49 So *Ebers/Hoch et al.,* RDi 2021, 532; *Roos/Weitz,* MMR 2021, 844 (845); *Rostalski/Weiss,* ZfDR 2021, 329 (347).

50 Dazu unter Hinweis auf Kritik von NGOs *Horstmann,* ZD-Aktuell 2022, 1182; auf hohes Gefährdungspotential hinweisend *Roos/Weitz,* MMR 2021, 844 (845); Verwunderung bekunden *Valta/Vasel,* ZRP 2021, 142 (143).

51 Vgl. *Ebers/Hoch et al.,* RDi 2021, 532.

52 Diesen Terminus prägen *Rostalski/Weiss,* ZfDR 2021, 329 (347).

53 Bezugnehmend auf den Vorrang menschlicher Aufsicht, technische Robustheit und Sicherheit, Datenqualitätsmanagement, Transparenz, vgl. *HEG*-KI, Guidelines, 2019, S. 17 f.

54 Ebenso schon Europäische Kommission, COM(2020) 65 final – Weißbuch KI, 2020, S. 22, worin bei der Formulierung von Anforderungen zwar explizit auf die HEG-KI verwiesen wird, Nichtdiskriminierung aber nicht als Schlüsselmerkmal, sondern nur im Rahmen allgemeinen Grundrechtsschutzes erwähnt wird.

55 Nach *Cummings,* JOTS 2006, 23 (25) meint automation bias die unhinterfragte Annahme automatisiert generierter Ergebnisse als zutreffend.

56 Zu default-Effekten vgl. *Anti-Bias* v. 7.1.2021, Anti-Bias 2020, abrufbar unter: https://www.anti-bias.eu/anti-bias-strategien/nudges-beispiele/default/ (zuletzt abgerufen am 08.11.22), diese äußern sich in kognitiven Verzerrungen, wonach automatisierte Entscheidungen bevorzugt werden, die kein selbstständiges aktiv werden erfordern.

57 Zu Einflussnahmefaktoren von Softwareentwicklern vgl. nur *Braegelmann/Kaulartz,* Rechtshandbuch AI und Machine Learning 2020, Kapitel 2.2, 34 Rn. 13.

zierenden oder klassifizierenden KI-Systems bei marginalisierten Gruppen führen.[58] Mittelbar droht eine weitreichendere Stigmatisierung, z.B. durch darauf basierenden, fehlerhaft eingeleiteten polizeilichen Ermittlungen bis hin zu Fehlurteilen. Jene Umstände sind nicht mit repräsentativer Datensatzgestaltung zu beheben, sondern bedürften gerade für hochriskante KI-Systeme weiterer Maßnahmen wie der Beteiligung marginalisierter Gruppen am Entwicklungsprozess oder Aufklärungsmaßnahmen über jene Effekte. Augenfällig wird bei der vorausgegangenen Aufzählung, dass Diskriminierungspotentiale letztlich so vielfältig in Erscheinung treten können, wie Maßnahmen, um diesen zu begegnen. Daraus ist jedoch nicht zu folgern, dass allein die aus Art. 9 KI-VO-E- hervorgehende Anforderung, allen denkbar auftretenden Risiken eines KI-Systems angemessen zu begegnen, Diskriminierungsrisiken hinreichend gegenübertritt. Vielmehr sollte die Zielsetzung der Nichtdiskriminierung ausdrücklich in den Anforderungen an hochriskante KI-Systeme aufgegriffen werden, um diese auch für Gestaltungsmaßnahmen vor Inverkehrbringen oder Inbetriebnahme des KI-Systems mit höchster Relevanz zu versehen. Der im KI-VO-E verfolgte Ansatz der Bezugnahme auf Nichtdiskriminierung im Kontext von Gewährleistungen der EU-Grundrechtecharta[59] ist zwar ebenfalls entscheidend, entfaltet aber vorrangig Schutz in Fällen, in denen sich eine Diskriminierung bereits realisiert hat oder jedenfalls das KI-System schon verwendet wird. In diesem Kontext ist hervorzuheben, dass gerade KI-Systemen immanente Spezifika (Undurchsichtigkeit, Komplexität) die Erkennbarkeit bereits aufgetretener Diskriminierungen erschweren, wodurch die Bedeutsamkeit präventiver Maßnahmen wächst.

An diesem Beispiel wird deutlich, dass die im KI-VO-E formulierten Anforderungen teilweise zu unspezifisch sind und insofern einer Nachbesserung bedürfen.[60] Bestmöglich wird der im KI-VO-E intendierte Interessenausgleich zwischen effizientem Schutz vor Risiken und zukunftstauglicher Flexibilität nämlich dann erreicht, wenn Anforderungen an die KI-Regulierung auf Gefahren zugeschnitten sind, die sich aus den technischen Besonderheiten künstlicher Intelligenz ergeben. An anderer Stelle, so z.B. in Art. 10, 14, 15 KI-VO-E wird diesen speziellen Risiken jedenfalls thematisch besser Rechnung getragen. Was hingegen deren inhaltliche Ausdifferenzierung anbelangt, sind auch diese Normen allgemein gehalten.

In Bezug auf das Vorgenannte erweist sich gerade die strukturelle Entscheidung der EU-Kommission, unspezifisch gehaltene Anforderungen in den Art. 8 ff. KI-VO-E zu formulieren und diese mit der weitreichenden Übertragung von Konkretisierungsbefugnissen an private EU-Normungsorganisationen zu kombinieren als nicht unproblematisch. So spricht ErwG 61 S. 1 KI-VO-E von der Normung als Schlüsselrolle zur Ermittlung technischer Lösungen für KI-System-Anwender. Ebenso haben notifizierende Stellen, d.h. solche Stellen die die Konformität hochriskanter KI-Systeme überprüfen, gem. Art. 33 Abs. 11 KI-VO-E an der Arbeit europäischer Normungsorganisationen[61] mitzuwirken. Deren Wirkmacht ist somit einerseits nicht zu unterschätzen und andererseits nicht klar definiert. Wegen der nur begrenzt EU-rechtlich zulässigen Delegation von Entscheidungskompetenzen auf privatrechtliche Organisationen,[62] die vor allem auch auf deren fehlende, demokratische Legitimation zurückzuführen ist,[63] sollte die Rolle solcher Normungsorganisationen jedenfalls eindeutig um- und begrenzt werden.

(2) Pflichten (Art. 16 ff. KI-VO-E)

Der inhaltlichen Befassung mit Pflichten, die mit dem Betrieb hochriskanter KI-Systeme verbunden sind, hat die Auseinandersetzung mit den Adressaten dieser Pflichten vorauszugehen. Die Überschrift des gegenständlichen Kapitels 3 des KI-VO-E richtet sich an Anbieter und Nutzer hochriskanter KI-Systeme. Wie bereits herausgearbeitet unterscheiden sich beide Begrifflichkeiten in ihrem chronologischen Bezugspunkt. Als begrüßenswert erweist sich, speziell für die Pflichten der Anbieter und Nutzer hochriskanter-KI, die Miterfassung des Entwicklungsprozesses des KI-Systems (Art. 3 Nr. 2 KI-VO-E) in zweierlei Hinsicht. Zum einen legt dieser präventiv ausgerichtete Ansatz den Grund für eine verbesserte Überschaubarkeit und Einschätzbarkeit des KI-Systems selbst, indem er den in Art. 16 ff. KI-VO-E normierten Pflichten über den bloßen Verwendungszeitraum hinaus Geltung verschafft. Daraus resultiert u.a., dass die Nachprüfbarkeit der Einhaltung jener Pflichten erleichtert wird. Zum anderen betrifft der Anbieterbegriff in Art. 3 Nr. 2 KI-VO-E zwar nicht per se jeden einzelnen Programmierer, aber die als Entwickler geltende, verantwortliche Person oder Stelle.[64] Da Verletzungen dieser Pflichten gem. Art. 71 Abs. 1, 4 KI-VO-E sanktionsbewehrt sind, besteht also bereits in einem frühen Stadium ein starker „Anreiz", gerichtet an den Verantwortungsträger des Entwicklungsprozesses, zur Einhaltung dieser Pflichten.

Modifiziert werden die Adressaten sog. Anbieterpflichten (Art. 16 – 23 KI-VO-E) in den Art. 24 ff. KI-VO-E. So gelten für Produkthersteller, die unter Rechtsakte fallen, welche in Anhang II Abschnitt A KI-VO-E erfasst sind (Art. 24 KI-VO-E) und für die Fälle des Art. 28 KI-VO-E

58 Nach der "Gender Shades"-Study von *Buolamwini/Gebru,* PMLR 2018, 1 (8) liegt die Fehlerrate biometrischer Gesichtserkennung bei Frauen mit dunkler Haut bei bis zu 34,7 %; die Fehlerrate im Falle weißer Männer liegt stellenweise bei höchstens 0,3 %; diese Erkenntnis stellt keinen Einzelfall dar; vgl. dazu die Zusammenfassung weiterer Studien, die Diskriminierung im Bereich biometrischer Gesichtserkennung zum Gegenstand haben *Najibi* v. 24.10.2020, Racial Discrimination, abrufbar unter: https://sitn.hms.harvard.edu/flash/2020/racial-discrimination-in-face-recognition-technology/ (zuletzt abgerufen am 11.11.22).

59 Vgl. nur Europäische Kommission, COM(2021) 206 final – KI-VO-E 2021, S. 4, 12.

60 So auch *Ebers,* RDi 2021, 588 (590).

61 So z.B. das CEN (Comité Européen de Normalisation), das CENELEC (Comité Européen de Normalisation Électrotechnique) oder das ETSI (European Telecommunications Standards Institute).

62 So schon *EuGH,* Slg. 1958, 16 (44) – Meroni I; Slg. 1956, 57 (81) – Meroni II.

63 Dazu sehr ausführlich *Ebers,* RDi 2021, 588 (593 ff.).

64 Anders gestaltet ist Art. 25 DS-GVO, der produktspezifische Anforderungen nur für Verwender formuliert, wohingegen Hersteller gerade nicht unmittelbar betroffen sind, vgl. *Roos/Weitz,* MMR 2021, 844 (846).

die in Art. 16 – 23 KI-VO-E normierten Pflichten unverändert. Bevollmächtigten für Anbieter, die außerhalb der EU niedergelassen sind (Art. 25 KI-VO-E), Einführern (Art. 26 KI-VO-E), Händlern (Art. 27 KI-VO-E) und Nutzern (Art. 29 KI-VO-E) werden in der sie betreffenden Norm in unterschiedlichem Umfang angepasste Pflichtenregime zugewiesen.[65] Wenngleich diese Zuschreibung von Pflichten auf den ersten Blick unübersichtlich zu sein scheint, ermöglicht sie letztlich eine passgenaue Regelung je tätig werdendem Akteur.

Neben den zuvor erwähnten, günstigen Folgen der Einbeziehung der Entwicklungsphase des KI-Systems, erwächst daraus aber auch eine Schwierigkeit. Aufgrund der untrennbaren Verknüpfung von KI-Systemen mit Daten, besteht im Falle der Eröffnung des Anwendungsbereichs der DS-GVO zu dieser ein Wertungswiderspruch. Wie schon dargestellt, entspricht der *„Verantwortliche"* i.S.v. Art. 4 Nr. 7 DS-GVO, an den sich auch die in Art. 25 DS-GVO normierten Pflichten richten, am ehesten dem Nutzerbegriff aus Art. 3 Nr. 4 KI-VO-E.[66] Richten sich also nach der DS-GVO, in der Terminologie des KI-VO-E gesprochen, Pflichten vorrangig an den Nutzer und betreffen damit auch den Zeitpunkt nach Inverkehrbringen/Inbetriebnahme, so läuft dies der Fokussierung des KI-VO-E auf Anbieterpflichten, die das Entwicklungsstadium vor der Nutzung des KI-Systems betreffen, entgegen. Nutzer hochriskanter KI-Systeme müssen also nach aktuellem KI-VO-E regelmäßig keine eigene Risikobewertung durchführen, die Spezifika künstlicher Intelligenz mitberücksichtigt, obschon gerade diese durch ihre praktischen Anwendungserfahrungen besonders wertvolle Erkenntnisse erwarten können.[67] Abhilfe verschafft insofern auch nicht der Verweis in Art. 29 Abs. 6 KI-VO-E auf die in Art. 35 DS-GVO normierte Datenschutz-Folgenabschätzung, weil diese nicht vergleichbar speziell auf die Risiken automatisierter Verarbeitung enormer Datenmengen ausgerichtet ist, wie es beim KI-Einsatz regelmäßig der Fall ist. Deswegen ist nach hier vertretener Ansicht auch für Nutzer hochriskanter KI-Systeme eine Risikofolgenabschätzung zu implementieren, um gleichlaufend mit der DS-GVO, auch die Nutzerperspektive wirkungsvoll zu berücksichtigen.[68]

Inhaltlich gibt Art. 16 KI-VO-E einen Überblick über alle Pflichten der *Anbieter* hochriskanter KI-Systeme. Mit Ausnahme des Art. 16 lit. a KI-VO-E, der die Einhaltung der zuvor besprochenen Anforderungen normiert und des Art. 16 lit. f, i KI-VO-E, der auf spezielle Konformitätskennzeichnungs- und Registrierungsvorschriften des Kapitel 5 des KI-VO-E verweist, beinhalten die weiteren Anwendungsalternativen des Art. 16 KI-VO-E[69] Verweisungen in die direkt nachfolgenden Vorschriften des gegenständlichen Kapitels. Feststellbar sind qualitativ unterschiedliche Pflichtenkategorien.

Vorgaben betreffen *erstens* die Risikoregulierung des KI-Systems im Sinne des KI-VO-E unmittelbar. So sieht Art. 17 KI-VO-E die Einrichtung eines KI-Qualitätsmanagementsystems vor, Art. 19 KI-VO-E sichert die Durchführung eines Konformitätsbewertungsverfahrens vor Inverkehrbringen oder Inbetriebnahme des KI-Systems ab und Art. 21 KI-VO-E verpflichtet zur Korrektur eines nicht (mehr) den Vorgaben des KI-VO-E entsprechenden KI-Systems. *Zweitens* bestehen Pflichten zur Gewährleistung der Nachprüfbarkeit des KI-Systemverhaltens in Form der Verpflichtung zur technischen Dokumentation (Art. 18 i.V.m Art. 11 KI-VO-E) und in Gestalt der Aufbewahrung automatisch erzeugter Protokolle des hochriskanten KI-Systems (Art. 20 KI-VO-E). Zuletzt sind *drittens* Kooperationspflichten auszumachen, genauer Informationspflichten gem. Art. 22 KI-VO-E über Risiken i.S.d. Art. 65 Abs. 1 KI-VO-E sowie nach Art. 23 KI-VO-E Pflichten zur Zusammenarbeit mit zuständigen, nationalen Behörden.

Eine andere Perspektive, die u.a. eine Aussage über die Eingriffsintensität der Pflichten für Systemanbieter enthält, bietet die Kategorisierung der Pflichten der Art. 16 ff. KI-VO-E in Selbst- und Fremdregulierung.[70] Erstere betrifft die Umsetzung jener Pflichten durch den Anbieter selbst, wohingegen Letztere deren, regelmäßig einschneidendere, Umsetzung durch Dritte mit für den Anbieter verpflichtender Wirkung beinhaltet. Eine gleichmäßige Verteilung beider Kategorien ist indes nicht festzustellen, da Selbstregulierung im Wege der internen Kontrolle der Einhaltung dieser Pflichten den ganz überwiegenden Anteil des KI-VO-E betrifft. Exemplarisch dafür steht die Zusammenschau von Art. 17 KI-VO-E und Art. 21 KI-VO-E, die rein inhaltlich zu überzeugen vermag.[71] So umfasst das in Art. 17 KI-VO-E normierte Qualitätsmanagementsystem den gesamten Lebenszyklus des KI-Systems und bezieht darüber hinaus sogar Qualitätsvorgaben für Daten in Art. 17 Abs. 1 lit. f KI-VO-E ein. Findet eine Abweichung von diesen Vorgaben statt, verpflichtet Art. 21 KI-VO-E zur Vornahme unverzüglicher Korrekturmaßnahmen. Eine erhebliche Einbuße an Durchschlagskraft erfährt dieser umfangreiche und dennoch dynamische Ansatz jedoch dadurch, dass er der Kategorie der Selbstregulierung unterfällt. So bleibt doch zu bezweifeln, dass Anbieter, die das KI-System ebenso wie das Qualitätsmanagementsystem entworfen haben, am besten für dessen kritische Beurteilung geeignet sind. Ohne dabei zwingend von der bewussten Außerachtlassung von Korrekturmaßnahmen auszugehen, ist zu befürchten, dass der Interessenkonflikt zwischen dem Ziel der Inbetriebnahme/Inverkehrbringen bzw. Nutzung des KI-Systems

---

65 Legaldefiniert werden "Bevollmächtigte" in Art. 3 Nr. 5 KI-VO-E, "Einführer" in Art. 3 Nr. 6 KI-VO-E und "Händler" in Art. 3 Nr. 7 KI-VO-E; als Sammelbegriff für alle denkbaren Adressaten fungiert gem. Art. 3 Nr. 8 KI-VO-E der Begriff des "Akteur[s]".

66 So auch *Ebers/Hoch et al.,* RDi 2021, 534; *Ebert/Spiecker gen. Döhmann,* NVwZ 2021, 1188.

67 Vgl. EDPB-EDPS, Joint Opinion on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence, 18.6.2021, S. 11.

68 Vgl. *Ebers/Hoch et al.,* RDi 2021, 534, die konkret von einer KI-Folgenabschätzung sprechen.

69 Konkreter Art. 16 lit. b – e, g, h, j KI-VO-E.

70 Von interner und externer Kontrolle sprechen *Ebert/Spiecker gen. Döhmann,* NVwZ 2021, 1188 (1191). Exkludiert werden an dieser Stelle Sonderregelungen für Kreditinstitute i.S.d. Richtlinie 2013/36/EU, vgl. nur Art. 17 Abs. 3, 18 Abs. 2, 19 Abs. 2, 20 Abs. 2 KI-VO-E.

71 So auch ebd.

einerseits und dessen Rücknahme zum Zwecke der Korrektur andererseits, den unvoreingenommenen Blick für den höchstmöglichen Schutz vor Risiken verstellt. Noch augenfälliger werden die Folgen der Klassifizierung als selbstregulierend in Art. 22 KI-VO-E. Dieser dient dem Schutz der in Art. 65 KI-VO-E verbürgten Rechte, davon umfasst ist nach Art. 65 Abs. 2 KI-VO-E auch der Grundrechtsschutz. Wird ein entsprechender Verstoß festgestellt, kommt nach Art. 65 Abs. 5 KI-VO-E auch das Verbot der Bereitstellung des KI-Systems, dessen Rückruf oder die Entnahme vom Markt in Betracht. Die zur Ermittlung tatsächlich bestehender Risiken seitens der zuständigen Marktüberwachungsbehörde[72] erforderlichen, hinreichenden Gründe, hängen jedoch in Ermangelung der Existenz externer Kontrollen von der Kooperation und Initiative der Anbieter des KI-Systems selbst ab.[73] Unbenommen unbewusster Konsequenzen, die aus Interessenkonflikten resultieren können, bietet diese Konstellation für Anbieter, die Risiken verschleiern wollen, die besten Umstände dies zu tun. An alledem manifestiert sich, dass der interne Kontrollansatz die theoretisch als wirksam einzustufenden Pflichten in letzter Konsequenz erheblich in ihrer Funktionalität beeinträchtigt. Insofern sollte eine Nachbesserung vorgenommen werden, die den Fokus mehr auf Fremdregulierung und externe Kontrollen verschiebt oder jedenfalls den Anteil an Selbst- und Fremdregulierung in Ausgleich bringt.

(3) Konformitätsbewertung (Art. 40 ff. KI-VO-E)
Weder folgt die Konformitätsbewertung hochriskanter KI-Systeme einem einheitlichen Ablauf, noch ist deren Klassifizierung als selbst- oder fremdreguliert uniform. Der Kategorie der Fremdregulierung zuzuordnen ist, für den Fall nicht oder nicht vollständiger Anwendung harmonisierter Vorschriften gem. Art. 40 KI-VO-E, die Konformitätsbewertung biometrischer Systeme zur Klassifizierung und Identifizierung gem. Art. 43 Abs. 1 i.V.m. Anhang VII KI-VO-E durch eine notifizierte Stelle. Eine weitere Sonderregelung gilt gem. Art. 43 Abs. 3 KI-VO-E für KI-Systeme, die unter Anhang II Abschnitt A KI-VO-E fallen. Deren Konformitätsbewertung richtet sich nach den jeweiligen produktsicherheitsrechtlichen Vorschriften, da diese selbst harmonisiert sind – regelmäßig ist das NLF einschlägig[74] – und sie damit ohnehin einer Konformitätsprüfung Dritter unterfallen (vgl. Art. 6 Abs. 1 lit. b KI-VO-E).[75] Abgesehen davon bedürfen gem. Art. 43 Abs. 2 KI-VO-E eigenständige KI-Systeme gem. Anhang III Nr. 2 – 8 KI-VO-E einer Konformitätsbewertung in Gestalt einer selbstregulierenden, internen Kontrolle. Somit ist die Konformitätsbewertungspflicht des Art. 19 KI-VO-E abhängig von der Klassifizierung und der Art des KI-Systems fallspezifisch einer externen oder internen Kontrolle zugänglich.

Das bereits dargestellte Problem der Abhängigkeit von Anbietern des KI-Systems vertieft sich bedeutsam in der Konstellation des Art. 43 Abs. 2 i.V.m. Anhang III Nr. 2 – 8 KI-VO-E. Über die Einhaltung von Pflichten und Anforderungen, sowie die ex-post Überwachung durch die zuständige Behörde hinausgehend, wird auch die Konformitätsbewertung des hochriskanten KI-Systems prä-Inverkehrbringen/Inbetriebnahme in die Hände der Anbieter selbst gelegt.[76] Wenngleich aus der strengeren Behandlung biometrischer Systeme i.S.d. Anhang III Nr. 1 KI-VO-E gem. Art. 43 Abs. I KI-VO-E hervorgeht, dass diese im Verhältnis zu den Sektoren des Anhang III Nr. 2 – 8 KI-VO-E pauschal als riskanter eingestuft werden, ist die de facto nahezu vollständige Verantwortungsabgabe in die Hände der Anbieter, die den gesamten Lebenszyklus des KI-Systems betrifft, nicht angezeigt. Vielmehr sollte als Gegengewicht zur überwiegend selbstregulierend erfolgenden ex-post Überwachung eine vollständige Übertragung der ex-ante Konformitätsbewertung an Dritte, unabhängige Stellen, erwogen werden.[77]

(4) Durchsetzungsmechanismen (Art. 56 ff. KI-VO-E)
Die Durchsetzung des KI-VO-E soll durch das Ineinandergreifen zweier Ebenen gewährleistet werden. Abgesichert wird zum einen auf *nationaler Ebene* die Anwendung sowie Durchführung des KI-VO-E durch eine je Mitgliedstaat zu benennende, nationale Behörde (vgl. Art. 59 KI-VO-E), der in Art. 63 ff. KI-VO-E eine Reihe verschiedener Befugnisse zuteilwerden. Davon umfasst sind nicht nur überwiegend auf die nationale Marktüberwachung fokussierte Kompetenzen, wie der Zugang zu Daten und zur technischen Dokumentation (Art. 64 KI-VO-E), sondern auch mitgliedstaatenübergreifende Regelungen wie z.B. in Art. 65 KI-VO-E, die eine Überprüfung von Maßnahmen anderer Mitgliedstaaten durch die Kommission ermöglichen. Zum anderen erfolgt in Art. 56 KI-VO-E die Einsetzung des *Europäischen Ausschusses für künstliche Intelligenz* (im Folgenden: EAKI) auf *EU-Ebene*, der einen Beitrag zur Kooperation nationaler Behörden leisten und die Umsetzung des KI-VO-E gleichermaßen durch die Kommission, wie durch die nationalen Aufsichtsbehörden unterstützen soll. Dieser Ausschuss setzt sich gem. Art. 57 Abs. 1 KI-VO-E aus Vertretern der nationalen Aufsichtsbehörden gem. Art. 56 KI-VO-E und dem EU-Datenschutzbeauftragten zusammen. Überdies können andere nationale Behörden bei Bedarf eingeladen werden. Wenngleich eine derartige Bündelung nationaler Aufsichtsbehörden auf EU-Ebene sinnvoll ist, sind die Einflussnahmemöglichkeiten des EAKI als sehr gering zu bewerten, weil diesem keine Befugnisse zur Durchsetzung von Maßnahmen zukommen, sondern eine Beschränkung auf unverbindliche Empfehlungen, Stellungnahmen und schriftliche Beiträge (Art. 58 lit. c KI-VO-E) sowie auf die Leistung von Beiträgen und Unterstützungshandlungen (Art. 56 Abs. 2 KI-VO-E). Hinzu kommt, dass mit der EU-Kommission als Vorsitzender des Ausschusses (Art. 57 Abs. 3 KI-VO-E), deren Kompetenz u.a. in der Vorbereitung von Tagesordnung und Aufgaben des

[72] Die Rolle der Marktüberwachungsbehörden vertiefend *Orssich,* EuZW 2022, 254 (260).
[73] Dazu ebenfalls kritisch *Spindler,* CR 2021, 361 (367).
[74] Vgl. *Roos/Weitz,* MMR 2021, 844 (845), die zudem auf den Sonderfall sonstiger harmonisierender Vorschriften i.V.m. Art 84 KI-VO-E näher eingehen.
[75] Vgl. *Rostalski/Weiss,* ZfDR 2021, 329 (347).
[76] Die Begriffe pre-market und post-market prägt *Orssich,* EuZW 2022, 254 (256) im Kontext von Konformitätsbewertung einerseits und Marktüberwachung andererseits.
[77] So im Ergebnis auch EDPB-EDPS, Joint Opinion on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence, 18.6.2021, S. 15.

Ausschusses liegt, ein nicht unerheblicher politischer Einfluss auf die Arbeit des EAKI vorliegt, der ein zusätzliches Abhängigkeitsverhältnis schafft.[78]

Zur verbindlichen Durchsetzung der KI-Verordnung sind somit nationale Aufsichtsbehörden und die Kommission befugt. Konkrete Durchsetzungsmittel können einerseits das KI-System selbst betreffen, indem die Verpflichtung zur Vornahme von Korrekturmaßnahmen in einer bestimmten Frist (Art. 65 Abs. 2 KI-VO-E), das Bereitstellungsverbot des KI-Systems auf dem Markt oder dessen Rücknahme vom Markt (Art. 65 Abs. 5 KI-VO-E) ausgesprochen wird. Daneben kommen auch Sanktionierungen in Form von Geldbußen (Art. 71 Abs. 1 KI-VO-E) in Betracht. Obschon die in Art. 71 Abs. 3 – 5 KI-VO-E vorgesehene Höhe dieser Geldbußen durchaus abschreckend zu sein vermag, verliert dieses Mittel schon dadurch an Durchschlagskraft, dass es Mitgliedstaaten nach Art. 71 Abs. 7 KI-VO-E überlassen bleibt, ob diese Sanktionen auch für Behörden und öffentliche Stellen gelten. Nicht nur geht damit eine Einbuße betreffend den Grad der Harmonisierung einher,[79] noch dazu ist der für Bürger besonders einschneidende, staatliche Einsatz von KI-Systemen jedenfalls optional mit weniger Schutzvorkehrungen versehen.

Hinzu kommt, dass unbenommen der Existenz dieser Durchsetzungsmechanismen aufgrund des hohen Anteils an Selbstregulierung und interner Kontrollen zu bezweifeln ist, in welchen Fällen nationale Marktüberwachungsbehörden oder die EU-Kommission überhaupt auf Verstöße gegen die Verordnung aufmerksam (gemacht) werden. Von einer hinreichenden Gewährleistung der Durchsetzbarkeit der Verordnung ist demnach nicht auszugehen.[80]

cc) Exkurs: Hochriskante KI-Systeme öffentlicher Stellen und Behörden

Mit der Inbetriebnahme und Nutzung hochriskanter KI-Systeme durch öffentliche Stellen und Behörden gehen nicht nur allgemein verringerte Durchsetzungsmöglichkeiten der Verordnung und eine gesteigerte Eingriffsintensität für Bürger einher, sondern in der staatlichen Anwendung intransparenter Technologien manifestiert sich auch das Machtgefälle zwischen Staat und Individuum in besonderem Maße. Diesem Gefälle tritt auch die in der Begründung des KI-VO-E zum Ausdruck kommende Wahrnehmung nicht entgegen, nach der für effektiven Grundrechtsschutz durch Transparenzanforderungen und Rückverfolgbarkeit der KI-Systeme *„im Verbund mit starken Ex-post-Kontrollen gesorgt [sei]“*.[81] Vielmehr führen schwache Durchsetzungsmechanismen, die für öffentliche Stellen und Behörden nochmals aufgeweicht werden (Art. 71 Abs. 7 KI-VO-E) dazu, dass Transparenzpflichten dem dargestellten Durchsetzungsdefizit ebenso wie alle weiteren Anforderungen und Pflichten unterliegen.

Noch größer ist jenes Machtgefälle im Bereich des Einsatzes hochriskanter KI-Systeme zur Strafverfolgung.[82] Dieser ist im KI-VO-E in Art. 6 Abs. 2 i.V.m. Anhang III Nr. 1, 6 KI-VO-E erfasst und begegnet neben den bereits kritisierten Maßnahmen zur Verordnungsdurchsetzung, zudem in Art. 60 Abs. 2 i.V.m. Anhang VIII Nr. 11 KI-VO-E einer weiteren Erleichterung in Gestalt der Begrenzung der Informationsbereitstellung von KI-Systeminformationen in der sog. *„EU-Datenbank für eigenständige Hochrisiko-KI-Systeme“*. Aus Sicht der Strafverfolgungsbehörden liegt auf der Hand, dass die genaue Funktionsweise eines derart verwendeten KI-Systems der Öffentlichkeit nicht ohne weiteres in einer Datenbank frei zugänglich gemacht werden kann. Auch kriminalpolitisch kann sich der Einsatz von KI-Systemen jedenfalls aus Perspektive effektiver Strafverfolgung sowie dem Gedanken der negativen Generalprävention entsprechend,[83] als sinnvoll erweisen. Wird ein solcher Einsatz erwogen, sind aus kriminalpolitischer Perspektive jedoch gleichermaßen soziale Effekte des KI-Einsatzes in Rechnung zu stellen. So können die bereits dargestellten, in KI-Systemen enthaltenen *biases* und weiteren *Diskriminierungsrisiken* durch fehlende Durchsetzbarkeit von Transparenzpflichten zu sozialen Benachteiligungen führen.[84] Abstrakter sind die sozialen Folgen des staatlichen Einsatzes intransparenter und bisweilen für den Staat selbst, in ihrer Funktionsweise nicht nachvollziehbarer Technik, zu berücksichtigen. Drohen können der gesellschaftliche Vertrauensverlust in die Justiz ebenso, wie steigender Überwachungsdruck auf die Gesellschaft, der zur Einschränkung grundrechtlich geschützter Gewährleistungen zu führen vermag. Da diese Überlegungen kriminalpolitisch in Rechnung zu stellen sind, ist auch aus dieser Perspektive die Stärkung von Durchsetzungsmechanismen des KI-VO-E, sowie die Fokussierung auf externe Kontrollen durch unabhängige Stellen als präferabel zu erachten, weil so das Vertrauen in derartige Techniken gestärkt wird. Überdies vermag ein Zugewinn an Transparenz mit einem *Mehr* an Rechtssicherheit einherzugehen.

c) KI-Systeme mit geringem Risiko

Als KI-Systeme mit geringem Risiko sieht die EU-Kommission solche an, die „ein besonderes Risiko in Bezug

78 So auch *Ebers/Hoch*, RDi 2021, 528 (536).

79 Zur unbedingten Erforderlichkeit des Bestehens einheitlicher Durchsetzungsmechanismen auch EDPB-EDPS, Joint Opinion on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence, 18.6.2021, S. 15.

80 Ebenso *Ebers/Hoch*, RDi 2021, 528 (535 f.); *Ebert/Spiecker gen. Döhmann,* NVwZ 2021, 1188 (1193) weisen zudem auf das Fehlen von Beweislastumkehrungen und Kausalitätserleichterungen hin; generelle Schwierigkeiten effektiver Rechtsdurchsetzung beschreiben *Valta/Vasel*, ZRP 2021, 142 (145).

81 Europäische Kommission, COM(2021) 206 final – KI-VO-E 2021, S. 13.

82 Im datenschutzrechtlichen Kontext aber verallgemeinerbar vgl. *Stief,* StV 2017, 470 (475) zur Machtasymmetrie zwischen Strafverfolgungsbehörden und Bürgern.

83 Vgl. dazu nur *Walter,* ZIS 2011, 636 (638 ff.).

84 Zur sozialpolitischen Einbettung der Kriminalpolitik umfassend *Kunz,* NK 2005, 151 (151 f.).

auf Identitätsbetrug oder Täuschung bergen".[85] Die Transparenzpflichten gelten für folgende KI-Systeme:

- solche, die mit Menschen interagieren („Chatbots")[86]
- solche, die Emotionen erkennen oder Menschen biometrisch kategorisieren oder
- solche, die Inhalte erzeugen oder manipulieren („Deepfakes")[87]

aa) KI-Systeme für die Interaktion mit natürlichen Personen

KI-Systeme, die für die Interaktion mit natürlichen Personen bestimmt sind, müssen grundsätzlich so konzipiert und entwickelt werden, dass die natürliche Person davon Kenntnis erlangt, dass es sich beim Gegenüber um ein KI-System handelt. Dadurch soll der Gefahr entgegengewirkt werden, durch Chatbots Stimmungsbilder zu verfälschen und Falschnachrichten zu verbreiten, da durch die Informationslücke über den Gesprächspartner Meinungsbildungsprozesse beeinträchtigt werden können.[88] Für Legal Tech-basierte Anwendungen bedeutet dies, dass der Nutzer eindeutig darüber informiert werden muss, dass er nicht über ein Chat-Programm mit einem juristischen Berater kommuniziert, sondern allein ein KI-System für die Kommunikation verantwortlich ist.[89]

Eine Ausnahme gilt für den Fall, dass aufgrund der Umstände und des Kontextes für den Nutzer offensichtlich ist, dass es sich um ein KI-System handelt. Nicht näher präzisiert wird allerdings, wann ein solcher Fall der Offensichtlichkeit vorliegen soll. Insofern sind Auslegungsschwierigkeiten und Unsicherheiten vorprogrammiert.[90]

Eine weitere Ausnahme ist gem. Art. 52 Abs. 1 S. 2 KI-VO-E für den Bereich der Strafverfolgung vorgesehen. D.h. die Transparenzpflichten müssen von solchen KI-Systemen nicht erfüllt werden, die gesetzlich zur Aufdeckung, Verhütung, Ermittlung und Verfolgung von Straftaten zugelassen sind. Würde man den Strafverfolgungsbehörden hier Transparenzpflichten auferlegen, würde die ganze Ermittlungs- und Präventionsarbeit ad absurdum geführt. Als Rückausnahme ist aber die Einschränkung anzusehen, dass diese Ausnahmevorschrift nicht für Systeme gilt, die der Öffentlichkeit zur Anzeige einer Straftat zur Verfügung stehen. Da es hier um die klassische Form der Chatbots geht, in der natürliche Personen mit einem KI-System in die Kommunikation treten, ist dies aufgrund des Gleichklangs mit den Transparenzpflichten Privater angemessen.

bb) Emotionserkennungssysteme oder Systeme zur biometrischen Kategorisierung

Die Informationspflichten gegenüber betroffenen Personen werden in Art. 52 Abs. 2 KI-VO-E auf Emotionserkennungssysteme oder Systeme zur biometrischen Kategorisierung erweitert. Der Begriff des Emotionserkennungssystems wird in Art. 3 Nr. 34 KI-VO-E legaldefiniert als ein KI-System, das dem Zweck dient, Emotionen oder Absichten natürlicher Personen auf der Grundlage ihrer biometrischen Daten festzustellen oder daraus abzuleiten. Emotionsdaten können auf sehr unterschiedliche Weise generiert werden, beispielsweise durch die Analyse von Texten, durch Gesichtserkennung oder auch durch Spracherkennung.[91] Auch KI-Systeme zur biometrischen Kategorisierung werden in Art. 3 Nr. 35 KI-VO-E legaldefiniert. Danach dienen solche Systeme dem Zweck, natürliche Personen auf der Grundlage ihrer biometrischen Daten bestimmten Kategorien wie Geschlecht, Alter, Haarfarbe, Augenfarbe, Tätowierung, ethnische Herkunft oder sexuelle oder politische Ausrichtung zuzuordnen. Durch die Transparenz, die von solchen KI-Systemen gefordert wird, soll der Betroffene in die Lage versetzt werden, autonom und in Kenntnis der relevanten Parameter zu entscheiden, ob ein derartiges KI-System von ihnen genutzt werden soll oder nicht.[92]

Angesichts der Tiefe des Eingriffs und des hohen Missbrauchspotentials wird kritisiert, dass die Einordnung unter KI-Systeme mit geringem Risiko missglückt ist. Teilweise wird sogar ein generelles Verbot entsprechender Systeme gefordert, wobei für Gesundheits- oder Forschungszwecke eng begrenzte Ausnahmen gelten sollen.[93] Gerade der Einsatz von Emotionserkennungssystemen kann aufgrund des großen Missbrauchspotentials, bspw. emotionale Schwächen zu identifizieren und entsprechende Maßnahmen zur „Stärkung" einzuleiten, erhebliche Auswirkungen auf die Persönlichkeitsentwicklung haben, so dass die Anforderungen an den Einsatz zu überdenken und anzupassen sind.[94]

Hinzu kommt, dass für die Systeme zur biometrischen Kategorisierung wiederum eine Ausnahme für die Zwecke der Strafverfolgung besteht. Die Transparenzpflicht gilt gem. Art. 52 Abs. 2 S. 2 KI-VO-E also nicht für gesetzlich zur Aufdeckung, Verhütung, Ermittlung und Verfolgung

---

85 Europäische Kommission, COM(2021) 206 final – KI-VO-E 2021, S. 39, Erwägungsgrund 70. Kritisch die Literatur, die tw. von KI-Systemen mit mittlerem Risiko ausgeht, vgl. bspw. *Bomhard/Merkle*, RDi 2021, 276 (282). Problematisch ist diese Differenzierung nicht, gelten doch zumindest für diese „Gruppe" KI-Systeme, wie auch immer man sie klassifiziert, die Informations- und Transparenzpflichten gem. Art. 52 KI-VO-E. So weisen Vgl. *Rostalski/Weiss,* ZfDR 2021, 329 (350) zu Recht darauf hin, dass auch Hochrisiko-KI-Systeme im Einzelfall zugleich den Anforderungen aus Art. 52 KI-VO-E unterfallen.

86 Unter einem Chatbot (sog. Gesprächsroboter) versteht man ein textbasiertes, rein virtuelles Dialogsystem, s. *Herrmann*, CR 2022, 350.

87 Bei Deepfakes handelt es sich um „alle Arten synthetischer Medieninhalte (wie Fotos, Audiodateien oder Videos), die mithilfe Künstlicher Intelligenz manipuliert oder sogar komplett erstellt wurden", s. *Schick*, Deepfakes: Wie gefälschte Botschaften im Netz unsere Demokratie gefährden und unsere Leben zerstören können, 2021, S. 24.

88 *Kalbhenn*, ZUM 2021, 663 (669).

89 S. *Engelmann/Brunotte/Lütk*ens, RDi 2021, 317 (321).

90 In diesem Sinne auch *Rostalski/Weiss*, ZfDR 2021, 329 (351); kritisch auch *Engelmann/Brunotte/Lütkens*, RDi 2021, 317 (321).

91 S. *Kalbhenn*, ZUM 2021, 663 (670).

92 *Rostalski/Weiss*, ZfDR 2021, 329 (351).

93 EDPB-EDPS, Joint Opinion on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence, 18.6.2021, S. 15; zur Kritik auch *Bomhard/Merkle*, RDi 2021, 276 (282).

94 So auch *Rostalksi/Weiss*, ZfDR 2021, 329 (352 f.).

von Straftaten zugelassene KI-Systeme, die zur biometrischen Kategorisierung verwendet werden. Dies ist zumindest im Hinblick auf die Parameter der ethnischen Herkunft oder sexuelle oder politische Ausrichtung bedenklich. Beispielsweise hat der deutsche Gesetzgeber bewusst darauf verzichtet, bei der molekulargenetischen Untersuchung nach § 81e StPO Feststellungen zur biogeographischen Herkunft zu treffen.[95]

cc) KI-Systeme, die Inhalte erzeugen oder manipulieren („Deepfakes")

Art. 52 Abs. 3 S. 1 KI-VO-E formuliert bestimmte Transparenzpflichten für ein KI-System, *„das Bild-, Ton- oder Videoinhalte erzeugt oder manipuliert"*, um tatsächlich existierende Personen, bestimmte Gegenstände, Orte, Einrichtungen oder Ereignisse dergestalt nachzuahmen, dass sie einer anderen Person *„fälschlicherweise als echt oder wahrhaftig erscheinen"*. Liegt per Definition ein sog. *Deepfake* vor, ist gem. Art. 52 Abs. 3 S. 1 KI-VO-E offenzulegen, dass eine künstliche Erzeugung oder Manipulation des gegenständlichen Inhalts erfolgt ist. Darüber hinaus normiert Art. 52 Abs. 3 S. 2 KI-VO-E Ausnahmen von dieser Kennzeichnungspflicht, unter anderem im Bereich der Strafverfolgung.

Dass *Deepfakes* generell mit erheblichen Gefahren in gesellschaftlicher sowie privater Hinsicht verbunden sind, lässt sich sowohl anhand konkreter Fälle als auch auf Basis abstrakter Gefahrenszenarien veranschaulichen. So werden Deepfakes jüngst wiederholt im Kontext des Ukraine-Kriegs genutzt, beispielsweise als ein täuschend echt aussehendes Deepfake-Video des ukrainischen Präsidenten *Wolodymyr Selenskyj* auftauchte, in dem er die ukrainische Bevölkerung vermeintlich zur Kapitulation aufruft.[96] Auswirkungen, die bis hin zur Eskalation eines Krieges reichen können, sind damit nicht auszuschließen. Ebenso vermögen Deepfakes gravierende Folgen auf persönlicher Ebene auszulösen, so z.B. im Falle der damals 18 jährigen Studentin *Noelle Martin*, deren Gesicht mittels eines KI-Systems in pornographische Inhalte eingefügt wurde, sodass der Anschein erweckt wurde, sie hätte an der Generierung dieser Inhalte mitgewirkt.[97] Ebenso ist auf abstrakter Ebene nicht auszuschließen, dass z.B. Deepfakes erstellt werden, die unschuldige Personen vermeintlich bei der Begehung einer Straftat darstellen, welche sodann – ebenfalls durch KI-Systeme – nachträglich biometrisch identifiziert werden. Im letzteren Fall wäre künstliche Intelligenz sowohl für die Erzeugung als auch für die Verurteilung einer unschuldigen Person von entscheidender Bedeutung.

Angesichts der massiven Gefahren, die mit der Erzeugung von Deepfakes verbunden sein können, ist nicht nachvollziehbar, diese unter *KI-Systeme mit geringem Risiko* zu subsumieren. Über die exemplarisch veranschaulichten Konstellationen hinausgehend bestehen Risiken für demokratische Prozesse, die in der Natur von Deepfakes als Desinformationsquelle begründet liegen.[98] Daneben droht das Vertrauen in die Technik – eines der selbst erklärten Ziele des KI-VO-E – durch die Erzeugung von Deepfakes erheblichen Schaden zu nehmen.[99] Ein Vertrauensverlust droht ebenso betreffend die Richtigkeit und Echtheit medialer Inhalte, womit ein medialer Vertrauensverlust insgesamt einherzugehen vermag und zwar je mehr, desto genauer gesellschaftlich das weitreichende Potential von Deepfakes an Bekanntheit erlangt.[100] All diesen Risiken vermag die Kennzeichnungspflicht nicht hinreichend abzuhelfen, da in den letzten Jahren bereits eindrücklich zu Tage getreten und wissenschaftlich belegt worden ist, dass sich nachweisliche Falschinformationen trotz deren Korrektur und der gesellschaftlichen Akzeptanz dieser Korrektur dennoch weiter selbst auf solche Personen auswirken, die sich im klaren darüber sind, einer Falschinformation ausgesetzt zu sein (sog. *continued-influence effect)*.[101] Ebenso neigen Menschen nicht selten dazu, Informationen unhinterfragt zu übernehmen, die deren Weltbild bestätigen.[102]

Aus diesen Gründen sollten das Inverkehrbringen/Inbetriebnehmen sowie der Einsatz von KI-Systemen zur Erzeugung von Deepfakes regelmäßig verboten und nur in engen Grenzen erlaubt sein.[103] Dies gilt umso mehr, als dass über Möglichkeiten der sicheren Kennzeichnung von Deepfakes ohnehin noch keine Einigkeit besteht.[104]

III. Gesamtfazit

Die Schaffung einer EU-weit harmonisierten, verbindlichen KI-Verordnung als Rahmenwerk zur Regelung des Inverkehrbringens/Inbetriebnehmens bzw. Nutzens von KI-Systemen ist nicht nur begrüßenswert, sondern auch höchst relevant, um den Umgang mit derartiger, sich rasant fortentwickelnder Technik, entwicklungsoffen und rechtssicher zu gestalten. Ebenfalls erweist sich die Herangehensweise des KI-VO-E in Gestalt der Schaffung ei-

95 In der kriminalpolitischen Diskussion wurde die Aufnahme dieses Merkmals im Vorfeld zur Gesetzesänderung kontrovers diskutiert, vgl. statt vieler *Schneider*, NStZ 2018, 692.

96 Vgl. ZDF, Propaganda im Ukraine Krieg, Fake-Video von Selenskyj im Umlauf; abrufbar unter: https://www.zdf.de/nachrichten/video/panorama-fake-video-selenskyj-100.html (zuletzt abgerufen am 14.11.2022).

97 Vgl. ABC News, The insidious rise of deepfake porn videos — and one woman who won't be silenced; abrufbar unter: https://www.abc.net.au/news/2019-08-30/deepfake-revenge-porn-noelle-martin-story-of-image-based-abuse/11437774 (zuletzt abgerufen am 14.11.2022).

98 Das BMVG spricht sogar von „Desinformation als Mittel der hybriden Kriegsführung", abgerufen unter: https://www.bmvg.de/de/themen/dossiers/weissbuch/gedanken/desinformation-als-mittel-der-hybriden-kriegfuehrung-12286 (zuletzt abgerufen am 14.11.2022); allg. dazu *Rostalksi/Weiss*, ZfDR 2021, 329 (352 f.).

99 So auch *Kalbhenn*, ZUM 2021, 663 (670).

100 Vgl. dazu nur *Chesney/Citron*, CLR 2019, 1785 f.

101 Vgl. *Linardatos,* Das Ende der Verlässlichkeit; abgerufen unter: https://www.lto.de/recht/hintergruende/h/deepfakes-regulierung-europa-eu-schaden-demokratie-manipulation/ (zuletzt abgerufen am 14.11.2022).

102 Ähnlich ebd.

103 So auch *Rostalksi/Weiss*, ZfDR 2021, 329 (352 f.).

104 Vgl. *Shane/Saltz/Leibowicz*, From deepfakes to TikTok filters: How do you label AI content?; abrufbar unter: https://www.niemanlab.org/2021/05/from-deepfakes-to-tiktok-filters-how-do-you-label-ai-content/ (zuletzt abgerufen am 14.11.2022).

nes risikobasierten Ansatzes, der den gesamten Lebenszyklus des KI-Systems erfasst, als sachgerecht. Zu dessen Umsetzung dem Grunde nach geeignet, ist die im Verordnungsentwurf angelegte ex-ante Konformitätsbewertung und ex-post Marktüberwachung, verknüpft mit variierenden Durchsetzungsmechanismen. Demnach ist das implementierte Regulierungsmodell seiner Struktur nach zur Umsetzung des erklärten Ziels der Schaffung eines harmonisierten EU-Rechtsrahmens für KI-Systeme tauglich. Dieser Befund schlägt indes nicht auf die Feststellung einer uneingeschränkten Brauchbarkeit des KI-VO-E in inhaltlicher Hinsicht durch. So liegen in der Ausdifferenzierung des Regulierungsmodells jene Hauptprobleme, die der Eignung des KI-VO-E zur effektiven Regulierung von KI-Systemen entgegenstehen. Im Falle *verbotener KI-Systeme* ist zu kritisieren, dass die Vorschrift des Art. 5 KI-VO-E derart viele Ausnahmen festlegt, dass nicht nur deren intendierte Ausgestaltung als Verbot mit Erlaubnisvorbehalt in Zweifel steht, sondern die mit den Ausnahmeregelungen verbundene Öffnung für länderspezifische Regelungen den Grad der Harmonisierung erheblich verringert. Betreffend *KI-Systeme mit geringem Risiko* erweist sich die konkrete Zuordnung bestimmter Praktiken in diese Kategorie als nicht nachvollziehbar. So ergibt die Risikobewertung KI-basierter *Emotionserkennungssysteme, Systeme zur biometrischen Kategorisierung* oder zur Erstellung von *Deepfakes,* dass damit verbundene Risiken mitnichten gering sind und eher im Bereich des Verbots mit engen Ausnahmen angesiedelt werden sollten.

In besonderem Maße problematisch ist die Kategorie *hochriskanter KI-Systeme*, da diese aufgrund überwiegend intern durchzuführender ex-ante Konformitätsbeurteilungen und ex-post Kontrollen, Akteure, die KI-Systeme einsetzen, ganz regelmäßig nur zur Selbstregulierung verpflichtet. Gehemmt wird dadurch die Durchsetzungsfähigkeit der Anforderungen und Pflichten des KI-VO-E über den gesamten Lebenszyklus des KI-Systems hinweg. So wird durch diesen Ansatz die Ermittelbarkeit, ob Vorgaben des KI-VO-E (nicht) eingehalten worden sind, von Akteuren, die ein eigenes Interesse an der Inbetriebnahme/Inverkehrbringung oder Nutzung des KI-Systems haben, abhängig gemacht. Nicht nur ist im Vergleich zur Beteiligung unabhängiger, externer Stellen an diesem Monitoringprozess die Effektivität des aktuellen Durchsetzungsmodells des KI-VO-E als gemindert einzustufen, sondern je nach Kooperationsbereitschaft der Akteure schwankt zudem der Grad der Durchsetzbarkeit. Dem sind allgemeine, rechtspolitische Bedenken entgegenzuhalten. So beeinträchtigt der selbstregulierende Ansatz die Rechtssicherheit und steht der Intention des KI-VO-E entgegen, der die Schaffung eines Rechtsrahmens vorsieht, welcher die zukunftsoffene Entwicklung von KI-Systemen unterstützt und dennoch *Sicherheit und Vertrauen in die Technik* ermöglicht.

Zugleich bestehen speziell kriminalpolitische Einwände gegen die Ausgestaltung des Einsatzes hochriskanter KI-Systeme zum Zwecke der Strafverfolgung. Hier sind vorrangig soziale Konsequenzen des ermittlungsbehördlichen Einsatzes von KI-Systemen kritisch in Rechnung zu stellen. Nicht unerheblich wirken sich die im KI-VO-E zwar vorgesehenen, aber nur bedingt durchsetzbaren Transparenzvorgaben aus, die das aus dem staatlichen Strafmonopol resultierende Machtgefälle zwischen Staat und Bürger überzustrapazieren vermögen. Ebenso ist der aus dem Einsatz hochkomplexer und so schwer nachvollziehbarer Kriminaltechnik im Bereich der Bild- und Videotechnik, erwachsende Überwachungsdruck zu berücksichtigen. Obschon der staatliche – und konkret ermittlungsbehördliche – Einsatz von KI-Systemen gleichermaßen beispielsweise das bürgerliche Vertrauen in die Kompetenz der Strafverfolgungsorgane zu stärken vermag, ist dies trotzdem nur dann möglich, wenn damit verbundene Vorgaben betreffend die gegenständliche Technik konsequent und gleichförmig umsetzbar sind. Im Ergebnis besteht also trotz der positiv zu bewertenden Grundstruktur des KI-VO-E deutlicher Nachbesserungsbedarf.