Responsible AI für die Gesellschaft

Wie Hamburg demokratische und rechtsstaatliche Prinzipien beim Einsatz Künstlicher Intelligenz in der Verwaltung sicherstellt

Kooperationsprojekt des LawCom.Institute und ARIC e.V.

Herausgeber:	LawCom.Institute GmbH, Kattrepelsbrücke 1, 20095 Hamburg www.lawcom.institute, info@lawcom.institute

Autor:innen:	Dr. Ulrike Ehling (LawCom.Institute), Friedrich-Joachim Mehmel (LawCom.Institute), Clara Sieveking (LawCom.Institute), Elisabeth Weißbecker (ARIC e.V.)
Unter Mitarbeit von:	Steven Dehlan (ARIC e.V.), Alois Krtil (ARIC e.V.) Louisa Sophie Rockstedt (ARIC e.V.), Jan Ruhnke (ARIC e.V.), Sebastian Schröder (LawCom.Institute), Yannek Wloch (Bucerius Law School)

Gefördert durch die Hamburgische Investitions- und Förderbank

Laufzeit der Förderung: März 2022 bis Februar 2023

Vorwort

Ein funktionierendes Staatswesen und insbesondere seine Verwaltung sind von herausragender Bedeutung für das Vertrauen der Bürger:innen in den Staat. Bürger:innen haben vielfältige Berührungspunkte mit der Verwaltung, angefangen beim Meldewesen, den Regelungen im Straßenverkehr, bei Sozialleistungen, Anmeldung bei den Schulen, Zulassung zu den Universitäten, jüngst etwa in Zusammenhang mit der Auszahlung von Corona-Hilfen oder der administrativen Bewältigung der Energiekrise. Von Bedeutung sind auch die Versorgung mit bezahlbarem Wohnraum, zügige Abwicklung von Bauvorhaben, Stadtplanung, Steuerverwaltung, Justiz, ein funktionierendes Gesundheitswesen, polizeiliches Handeln, um nur einige Stichworte zu nennen. Hier machen Bürger:innen unmittelbare Erfahrungen mit der Verwaltung, mit „ihrem" Staat.

Systeme, die Künstliche Intelligenz (KI) verwenden, können dabei helfen, den digitalen Wandel einer effektiven und effizienten Verwaltung der Freien und Hansestadt Hamburg weiter voranzubringen. Hamburg hat ganz in diesem Sinn 2020 die „Digitalstrategie für Hamburg"[1] beschlossen, auf deren Grundlage der digitale Wandel der Stadt in Wirtschaft, Gesellschaft und Verwaltung gestaltet wird: Durch vermehrte Verwendung von KI auch in der Verwaltung könnten bisher manuelle Prozesse automatisiert und der Arbeitsalltag der Mitarbeiter:innen erleichtert werden.[2] Bei der Wahrnehmung hoheitlicher Aufgaben würde KI zur Unterstützung von datenbasiertem und serviceorientiertem Verwaltungshandeln in geeigneten Bereichen erprobt. Das gleiche gelte für den unterstützenden Einsatz von KI, z.B. bei Auswertungen von Eingaben.[3] In der Digitalstrategie heißt es weiter: *„Digitalisierung gelingt nur im zeitgemäßen rechtlichen Rahmen. Den vielfältigen Möglichkeiten der digitalen Transformation stehen zumeist gesetzliche Regelungen gegenüber. Sie können Treiber oder Bremse für Innovation sein. Damit der Wandel gelingt, müssen die regulatorischen Rahmenbedingungen differenziert betrachtet und dort wo nötig aktualisiert werden".*[4]

Vor diesem Hintergrund geht die vorliegende Studie auf bereits geltende verfassungs- und verwaltungsrechtliche Vorschriften ein, die für den Einsatz von KI in der Verwaltung von Relevanz sind und gibt Empfehlungen für ergänzende Regelungen. In einem ersten Schritt analysiert sie jedoch die potentiellen Auswirkungen der aktuellen Bemühungen der Europäischen Union, ein Regulierungsregime für KI über verschiedene Einsatzfelder hinweg zu etablieren, bei dem entlang von Risikobereichen verschiedene Anforderungen an z.B. Sicherheit und Robustheit der Systeme gestellt werden bis hin zur rechtlich verpflichtenden Zertifizierung von KI-Anwendungen.

Da sich die europäischen Regelungen gleichsam an Wirtschaft und Verwaltungen wenden und somit nach ihrem Inkrafttreten auch den Einsatz von GovTech-Produkten, die auf KI zurückgreifen, in der Verwaltung maßgebend regulieren werden, sollten deshalb schon jetzt entsprechende rechtliche Anforderungen an Entwicklung, Beschaffung und Einsatz beachtet werden. In diesem Zusammenhang sind auch die von der

[1] Freie und Hansestadt Hamburg - Senatskanzlei Amt für IT und Digitalisierung (2020): Digitalstrategie für Hamburg, https://www.hamburg.de/contentblob/13508768/703cff94b7cc86a2a12815e52835accf/data/download-digitalstrategie-2020.pdf (zuletzt aufgerufen am 08.02.2023)
[2] a.a.O Ziffer 2.3.2
[3] a.a.O Ziffer 2.6
[4] a.a.O.

EU-Kommission geplante KI-Haftungsrichtlinie und ihr Vorschlag für eine Überarbeitung der EU-Produkt-Haftungsrichtlinie für Entwicklung und Einsatz von KI zukünftig zu beachten.

Die EU-Regelungen werden aufgrund der Fokussierung auf den „Hochrisikobereich" nur eine Teilmenge des möglichen Einsatzes von KI-Systemen in der öffentlichen Verwaltung erfassen. Da es aber auch im Nicht-Hochrisikobereich zu Fehlfunktionen mit unintendierten Folgen kommen kann, die ebenfalls das Vertrauen in das Verwaltungshandeln erschüttern können, werden in der Studie die rechtlichen Überlegungen um das Konzept einer *Responsible Artificial Intelligence* (RAI) ergänzt, wonach Systeme Künstlicher Intelligenz in gesellschaftlich verantwortungsvoller Weise sowohl entwickelt als auch eingesetzt und kontrolliert werden sollen. RAI setzt voraus, dass der Einsatz von KI im Einklang mit europäischen Wertevorstellungen und Recht steht. Es muss nachvollziehbar und erklärbar sein, wie ein Ergebnis zustande kommt. Fehlfunktionen und Ausfälle müssen minimiert bzw. ausgeschlossen werden. Auch die Vermeidung jedweder Diskriminierung ist besonders wichtig. Schließlich muss die menschliche Verantwortung für Entwicklung und Betrieb klar geregelt sein.

Gerade der Einsatz von Künstlicher Intelligenz in den verschiedensten gesellschaftlichen Bereichen wird immer öfter auch kritisch diskutiert, und viele Menschen haben eine skeptische Haltung.[5] Eine in erster Linie rein technische Herangehensweise bei der Digitalisierung der Verwaltung genügt daher nicht. Im Rahmen dieser Transformation Anforderungen von *Responsible AI* und des KI-Acts zu berücksichtigen, ihnen zu entsprechen und in die bestehenden Verwaltungsstrukturen zu implementieren, stellt eine neue Herausforderung für die Hamburger Verwaltung dar. Die Studie plädiert gerade deshalb dafür, neben dem KI-Act auch das RAI-Konzept allgemein für den Einsatz von KI in der Verwaltung der Freien und Hansestadt Hamburg anzuwenden. Sie unterbreitet Vorschläge für die rechtliche, technische und praktische Umsetzung.

Hier lässt sich an Vorhandenes anknüpfen. In Hamburg ist schon frühzeitig eine Diskussion über die Konsequenzen durch einen gesellschaftlich verantwortlichen, rechtsstaatlichen Anforderungen genügenden Einsatz von KI in der Verwaltung geführt worden: Am 12. Oktober 2018 trafen sich Expert:innen unterschiedlicher Disziplinen aus Wissenschaft und Praxis in einer u.a. vom Rechtsstandort Hamburg e.V., dem Hans-Bredow-Institut für Medienforschung[6] und dem Fachbereich Informatik der Universität Hamburg unter Mitwirkung der Senatskanzlei der Freien und Hansestadt Hamburg im Hamburger Rathaus durchgeführten Fachtagung, um sich über die Wirkungen und den Nutzen von Algorithmen und Künstlicher Intelligenz in der öffentlichen Verwaltung und die aus ihrer Einbindung erwachsenden Probleme für den Grundrechts- und Datenschutz, das Rechtsstaats- und Demokratieprinzip und die Verwaltungsgerichtsbarkeit auszutauschen und mögliche Lösungsansätze zu entwickeln.[7] Auf dieser Basis sind „9 Thesen zu

[5] s. z.B. Jussupow, Bebasat, Why are we averse towards algorithms? A comprehensive literature review on algorithm aversion, (Juni 2020), https://www.researchgate.net/profile/Ekaterina-Jussupow/publication/344401293_WHY_ARE_WE_AVERSE_TOWARDS_ALGORITHMS_A_COMPREHENSIVE_LITERATURE_REVIEW_ON_ALGORITHM_AVERSION/links/5f719082299bf1b53efa4198/WHY-ARE-WE-AVERSE-TOWARDS-ALGORITHMS-A-COMPREHENSIVE (zuletzt aufgerufen 08.02.2023)

[6] Seit 2019 Leibniz-Institut für Medienforschung | Hans-Bredow-Institut (HBI)

[7] vollständige Dokumentation der Tagung mit Workshops, Teilnehmer:innen, Reden etc. unter https://www.zukunft-der-verwaltungsgerichtsbarkeit.de/ki-und-verwaltung (zuletzt aufgerufen am 08.02.2023)

Chancen und Risiken, demokratische Legitimation und rechtsstaatliche Kontrolle bei der Algorithmisierung der Verwaltung“ [8] verfasst worden, die auch im internationalen Vergleich das Augenmerk früh auf die Fragestellungen der vorliegenden Studie geworfen haben.

Daran anknüpfend werden nun zunächst die Bedeutung der KI-Regulierung (Kapitel 1), die KI-Verordnung der EU (KI-Act) (Kapitel 2) sowie die allgemeinen Anforderungen an den vertrauensvollen Einsatz von KI – *Responsible AI* – (Kapitel 3 und 4) dargestellt. In Kapitel 5 wird erörtert, ob die einschlägigen verwaltungsverfahrensrechtlichen Vorschriften wie insbesondere §§ 35, 35a VwVfG für den KI-Einsatz einer Novellierung bedürfen, es wird die Bedeutung des Binnenrechts diskutiert, sowie die Frage behandelt, ob eine hinreichende demokratische Legitimation und die Voraussetzungen für ausreichenden verwaltungsrechtlichen Rechtsschutz gegeben sind. Als Grundlage für die dann folgenden Handlungsempfehlungen werden die bisherigen Ergebnisse zusammengefasst und Leitlinien benannt (Kapitel 6), auf denen die anschließenden Handlungsempfehlungen (Kapitel 7) beruhen.

An dieser Stelle möchten die Verfasser:innen der Studie ausdrücklich allen Mitarbeiter:innen der Hamburger Verwaltung, die sich in Expert:inneninterviews und Workshops geäußert haben, herzlich danken; besonderer Dank gilt den Gesprächspartner:innen vom Amt für IT und Digitalisierung (ITD).

Das LawCom.Institute und ARIC e.V. legen der Behörde für Wirtschaft und Innovation zeitgleich eine Studie zum Einsatz von KI in Hamburger kleinen und mittleren Unternehmen (KMU) vor, die sich insbesondere mit den Auswirkungen des KI-Acts für diese Akteure auseinandersetzt. Auch dort werden konkrete Handlungsempfehlungen gegeben, die aus den Grundsätzen des gesellschaftlich verantwortlichen Einsatzes von KI im Sinne einer *Responsible AI* folgen.

[8] Verfasser Mehmel, Schulz, https://www.zukunft-der-verwaltungsgerichtsbarkeit.de/media/pages/ki-und-verwaltung/04b8d11e95-1580462498/9-thesen-de.pdf (zuletzt aufgerufen am 08.02.2023)

Inhaltsverzeichnis

1. KI-Regulierung – Einhegung von Technik durch Recht

Die Versprechen Künstlicher Intelligenz[9] sind groß: die Chance auf ein nachhaltiges und effizienteres Wirtschaftssystem, auf mehr demokratische Partizipation und Zugang zu Wissen, auf mehr Zeit für sinnstiftende Tätigkeiten und verlässlichere Prognosen für die Zukunft. Auf vielen Ebenen wird über den Einsatz Künstlicher Intelligenz in den unterschiedlichsten gesellschaftlichen und unternehmerischen Kontexten diskutiert. Ganz praktisch kommt Künstliche Intelligenz zunehmend im Gesundheitssektor zum Einsatz, ebenso in Personalabteilungen oder dem Marketing. Städtische Verwaltungen greifen auf die Unterstützung von KI bei der Bearbeitung von Bürger:innen-Anfragen zurück, Rechtsdienstleister:innen nutzen sie zur Bearbeitung von auf gleichartigen Sachverhalten beruhenden Massenverfahren. Die Mustererkennung als eine mögliche Form der KI-Anwendung bietet in all diesen Bereichen große Effizienzgewinne und weitere Potentiale, kann Entscheidungen auf eine belastbarere Grundlage stellen und wichtige Perspektiven für die Zukunft bedeuten.

Gerade vor dem Hintergrund dieser enormen Potentiale nimmt die Diskussion über Risiken und die Notwendigkeit von Regulierungsbemühungen rund um die eingesetzten Systeme bedeutend an Fahrt auf. Und dies ganz zu Recht. Denn ein ungeregelter Einsatz von KI birgt auch große Gefahren, insbesondere für die Freiheits- und Bürger:innenrechte, für Arbeitnehmer:innen, die Umwelt und am Ende für den Rechtsstaat im Allgemeinen. Einige Fälle sind allseits bekannt, in denen der Einsatz von Algorithmen zum Teil weitreichende Folgen hatte:

- So wird in Österreich der Einsatz des sogenannten AMS-Algorithmus in der Arbeitsvermittlung kritisch diskutiert. Der Algorithmus sollte regelhaft in der Arbeitslosenberatung auf Grundlage personenbezogener Daten zur Bewertung von Vermittlungschancen von Arbeitssuchenden eingesetzt werden. Von dem Ergebnis wäre u.a. die Zuweisung von Fördermaßnahmen zur Wiedereingliederung abhängig. Im Testbetrieb wurden allerdings nicht nur Frauen systematisch schlechter bewertet, auch nicht-österreichische Bürger:innen bekamen schlechtere Prognosen für die Chancen auf eine Vermittlung. Ohne die Verabschiedung einer gesetzlichen Grundlage, die für mehr Transparenz der Anwendung und für mehr Rechtssicherheit für die Betroffenen sorgen könnte, hat die österreichische Datenschutzbehörde den Einsatz des Algorithmus vorerst gestoppt.[10]
- Über die Kindergeldaffäre in den Niederlanden ist die letzte Regierung Rutte im Januar 2021 zum Rücktritt gezwungen worden. Der Vorwurf lautete „rassistische Diskriminierung" aufgrund einer

[9] Unter Künstlicher Intelligenz versteht man einen Teilbereich der Informatik, dessen Anwendungen die Lösung komplexer Probleme ermöglichen, für die es ansonsten menschliches, intelligentes Handeln erfordert. Dieses schließt beispielsweise Bilderkennung (maschinelles Sehen), Natural Language Processing (Textverständnis und –erzeugung) und analytische Entscheidungsfindung ein. Unterschieden werden u.a. datenbasierte (Maschinelles Lernen), wissensbasierte (Expertensysteme) und hybride Modelle, die wiederum unterschiedliche Lernverfahren und Algorithmen kennen. Künstliche Intelligenz wird im Folgenden als KI oder im englischsprachigen Kontext als AI abgekürzt.

[10] Szigetvari, Datenschutzbehörde kippt umstrittenen AMS-Algorithmus, Der Standard (20.8.2022), https://www.derstandard.de/story/2000119486931/datenschutzbehoerde-kippt-umstrittenen-ams-algorithmus (zuletzt aufgerufen am 08.02.2023);
Autor unbekannt, Der Standard (28.04.2022): Neue Kritik am AMS Algorithmus, https://www.derstandard.de/story/2000135277980/neuerliche-kritik-am-ams-algorithmus-zum-in-die-tonne-treten (zuletzt aufgerufen 08.02.2023)

automatisierten Entscheidung, bei der die nationale Steuerbehörde fälschlicherweise Kindergeldzahlungen zurückgefordert und Betrugsermittlungen gegen zehntausende Familien in den Niederlanden aufgenommen hatte. Schon kleine Formfehler beim Ausfüllen der Anträge, vor allem aber die zugrunde gelegte Datenbasis, bei der die Staatsangehörigkeit als zentraler Verdachtsmarker genutzt wurde, hatten Familien nicht nur einer unhaltbaren finanziellen Belastung ausgesetzt, sondern für viele auch weitere Ermittlungen und rechtliche Konsequenzen zur Folge.[11]

- Schon seit etwa zehn Jahren sorgt insbesondere Bias, also eine systematische Verzerrung aufgrund bestimmter Trainingsdaten, die dann die Entscheidungsfindung beeinflusst (siehe Glossar), in KI-Anwendungen für Schlagzeilen. So wurde 2014/15 berichtet, dass bei Amazon ein KI-basiertes Tool zur Auswahl von Bewerber:innen systematisch Männer für Tech Jobs vorzog. Das Rekrutierungstool suchte nach Mustern in Bewerbungen der letzten zehn Jahre, um daraus eine Priorisierung abzuleiten. Da sich in der Vergangenheit überwiegend Männer beworben hatten, hatte sich die KI selbst beigebracht, Bewerbungen von Frauen als ungeeignet herauszufiltern.[12]
- 2015 sorgte die App „Google Fotos", die Bilder automatisch sortiert und verschlagwortet, für Aufsehen, als bei der automatischen Verschlagwortung dunkelhäutige Menschen als Gorillas bezeichnet wurden. Die KI hatte auf eine Datenbank zurückgegriffen, die auch Tierbilder beinhaltete, und in Folge den schwerwiegenden Fehler begangen, Menschen mit Tieren zu verwechseln.[13]
- 2016 zeigte sich ebenfalls, was passieren kann, wenn eine KI ungefiltert im Betrieb weiterlernt, als der Microsoft Chatbot Tay auf Twitter durch Nutzer:inneneingaben innerhalb von Stunden rassistisch und frauenfeindlich wurde.[14] Seither haben zahlreiche wissenschaftliche Veröffentlichungen Diskriminierung von KI-Anwendungen konstatiert. Die jüngsten Beispiele liefert ChatGPT, das vermeintlich schon Bias adressiert und sich beispielsweise weigert, Witze über Gottheiten des Islam und Christentums zu schreiben, selbiges aber über Gottheiten der Hindus erst nach Beschwerden von Nutzer:innen ablehnte.[15]
- Ein weiteres Beispiel für nicht intendierte Folgen in der Anwendung eines Chatbots liefert jüngst die Suchmaschine Bing, die basierend auf derselben Technik wie ChatGPT ebenfalls komplexe Fragen beantwortet und Konversation mit den Nutzer:innen betreiben kann. So hatte der Versuch eines Reporters der New York Times für Aufsehen gesorgt, der nach einer längeren Konversation mit der KI von dieser zum Verlassen seiner Frau aufgefordert wurde. Auch andere Nutzer:innen

[11] Dachwitz, Ingo; Netzpolitik (29.12.2021): Niederlande zahlen Millionenstrafe wegen Datendiskriminierung, https://netzpolitik.org/2021/kindergeldaffaere-niederlande-zahlen-millionenstrafe-wegen-datendiskriminierung/ (zuletzt aufgerufen am 08.02.20223)

[12] Dastin, Jeffrey , REUTERS (11.10.2018): Amazon scraps secret AI recruiting tool that showed bias against women, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G (zuletzt aufgerufen am 28.02.2023)

[13] Autor unbekannt, SPIEGEL NETZWELT (02.07.2015):
Google entschuldigt sich für fehlerhafte Gesichtserkennung, https://www.spiegel.de/netzwelt/web/google-fotos-bezeichnet-schwarze-als-gorillas-a-1041693.html (zuletzt aufgerufen am 02.07.2023)

[14] Graff, Bernd, Süddeutsche Zeitung (03.04.2023): Rassistischer Chat-Roboter:
Mit falschen Werten bombadiert, https://www.sueddeutsche.de/digital/microsoft-programm-tay-rassistischer-chat-roboter-mit-falschen-werten-bombardiert-1.2928421 (zuletzt aufgerufen am 07.02.2023)

[15] Pudur, Arun, LinkedIn (Januar 2023): How Woke is
ChatGPT?, https://www.linkedin.com/posts/arunpudur_chatgpt-activity-7018419857289330688-8YEe/?utm_source=share&utm_medium=member_ios (zuletzt aufgerufen am 28.02.2023)

hatten darauf hingewiesen, dass der Chatbot unangemessene Antworten gebe und auch vor Drohungen und Erpressungen nicht zurückschrecke.[16]

Vor dem Hintergrund derartiger Erfahrungen und zum Teil weitreichender Risiken richtet sich auch auf politischer Ebene und in internationalen Organisationen zunehmend der Blick auf Anforderungen an Hersteller:innen und Nutzer:innen von algorithmenbasierten Lösungen und KI-Lösungen im Besonderen, auf Rechenschaftspflichten, Normierungsbemühungen und Zertifizierungsverfahren. Es gilt, derartige Gefahren und Fehlfunktionen zu minimieren, Diskriminierung vorzubeugen und auf den verlässlichen und rechtssicheren Einsatz vertrauenswürdiger KI-Systeme hinzuarbeiten. Insgesamt nehmen daher weltweit Regelwerke zu, die zum Ziel haben, den Einsatz Künstlicher Intelligenz in Wirtschaft und Gesellschaft gleichermaßen an Kriterien der Verantwortlichkeit, Fairness und Belastbarkeit zu binden:

- Im Februar 2022 wurde im US-Senat und Repräsentantenhaus ein Entwurf für einen *Algorithmic Accountability Act* vorgelegt. Er versucht ex ante Standards für die Entwicklung und Nutzung von automatisierten Entscheidungssystemen zu definieren und Akteure auf eine Technikfolgenabschätzung zu verpflichten. Außerdem wird in dem Entwurf der Aufbau staatlicher Infrastrukturen angeregt, die ex post eine Überwachung der eingesetzten Technik ermöglichen können und durch die Compliance eingefordert werden kann.[17]
- Das Weiße Haus hat in den USA ebenfalls zum Thema KI-Regulierung Stellung genommen. Im Oktober 2022 hat es einen nicht rechtsverbindlichen „Blueprint" für eine KI-Rechtsverordnung vorgelegt, die sogenannte *AI Bill of Rights*. Fünf Bundesbehörden haben in Folge bereits Leitfäden für einen verantwortlichen Einsatz von KI-Systemen in ihren eigenen Arbeits- und Verwaltungsabläufen vorgelegt. Andere haben verbindliche Richtlinien für in ihrem Geschäftsbereich liegende Branchen herausgegeben, wie z.B. für die Food and Drug Administration bei der Überwachung der Entwicklung und Zulassung medizinischer Geräte.[18]
- Die kanadische Regierung hat im Juni 2022 den *Artificial Intelligence and Data Act* (AIDA) als Teil einer grundlegenden Reform des Datenschutzrechts und angrenzender Rechtsgebiete ins Parlament eingebracht. Zentrales Ziel ist der Schutz der Verbraucher:innen im digitalen Raum oder beim Einsatz von KI-Systemen im privaten Sektor, insbesondere dem internationalen Handel. Ne-

[16] New York Times (17.02.2023): Why a conversation with Bing's Chatbot left me deeply unsettled, https://www.nytimes.com/2023/02/16/technology/bing-chatbot-microsoft-chatgpt.html (zuletzt aufgerufen am 28.02.2023)

[17] [117th Congress Public Law 207] [From the U.S. Government Publishing Office], https://www.congress.gov/bill/117th-congress/house-bill/6580/text (zuletzt aufgerufen am 08.02.2023); Mökander, Jakob et.al. (18.8.2022): The US Algorithmic Accountability Act of 2022 vs. The EU Artificial Intelligence Act: what can they learn from each other?, https://link.springer.com/article/10.1007/s11023-022-09612-y (zuletzt aufgerufen am 08.02.2023) *sowie*
Andrae, Silvio RiskNet (07.12.2022): Ein Stück in mehreren Akten, https://www.risknet.de/themen/risknews/ein-stueck-in-mehreren-akten/ (zuletzt aufgerufen am 07.02.2023)

[18] The White House (Oktober 2022): Blueprint for an AI Bill of Rights: Making Automated Systems work for the American People, https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf (zuletzt aufgerufen am 28.02.2023) *sowie*
Turner Lee, Nicol/Malamud, Jack (19.12.2022): Opportunities and blind spots in the White House's blueprint for an AI Bill of Rights", https://www.brookings.edu/blog/techtank/2022/12/19/opportunities-and-blind-spots-in-the-white-houses-blueprint-for-an-ai-bill-of-rights/ (zuletzt aufgerufen am 28.02.2023)

ben gesetzlichen Regeln, die eine Verpflichtung zu Risikobewertungen oder Aufzeichnungspflichten beinhalten, sieht AIDA auch Konformitätsprüfungen und Notifizierungsverfahren vor und operiert mit der Unterscheidung verschiedener Hochrisikoanwendungen.[19]

Auch in internationalen Organisationen arbeitet man an Leitlinien und Kriterien, die nicht nur die Interoperabilität der Systeme weltweit gewährleisten und so den globalen Handel von KI-Systemen ermöglichen sollen, sondern Menschenrechte und rechtsstaatliche Prinzipien in den Mittelpunkt der Debatte rücken:

- Im November 2021 hat die Organisation der Vereinten Nationen für Bildung, Wissenschaft, Kultur und Kommunikation (UNESCO) den ersten global verhandelten Völkerrechtstext zur Ethik Künstlicher Intelligenz vorgelegt, der konkrete Empfehlungen zur globalen KI-Normung enthält.[20]
- Die Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) hat ein breites Netzwerk an Expert:innen aufgebaut und mit seinen *OECD AI Principles* Leitlinien für eine menschenrechtszentrierte vertrauensvolle KI erarbeitet.[21]
- Die International Standard Organization (ISO) arbeitet an verschiedenen Normen zu Datenqualität oder KI-Managementsystemen[22], die Grundlage für weitergehende Normungsüberlegungen auch in Deutschland sind.[23]

Nicht alle dieser Regulierungsansätze oder Normierungsvorschläge orientieren sich gleichermaßen an den Prinzipien von Rechtsstaatlichkeit und Demokratie. Gerade die Aktivitäten auf internationaler Ebene entfalten zudem in der Regel keine unmittelbare Rechtsverbindlichkeit. Ihnen ist aber gemeinsam, dass der Versuch unternommen wird, Leitplanken für die Nutzung von KI-Systemen zu etablieren und die verschiedenen Akteure innerhalb der KI-Ökosysteme, von Entwickler:innen bis hin zu Anwender:innen, auf diese zu verpflichten. Den meisten dieser Vorschläge liegt damit die Idee einer *Responsible Artificial Intelligence*

[19] Government of Canada (16.6.2022): New laws to strengthen Canadians' privacy protection and trust in the digital economy, https://www.canada.ca/en/innovation-science-economic-development/news/2022/06/new-laws-to-strengthen-canadians-privacy-protection-and-trust-in-the-digital-economy.html (zuletzt aufgerufen am 07.02.2023) *sowie*
Landry, Kevin et.al. (16.11.2022): Bill C-27 – Canadas proposed Artificial Intelligence and Data Act, http://www.stewartmckelvey.com/thought-leadership/bill-c-27-canadas-proposed-artificial-intelligence-and-data-act/ (zuletzt aufgerufen am 03.02.2023)

[20] UNESCO [65282] (2021): Recommendation on the Ethics of
Artificial Intellligence, https://unesdoc.unesco.org/ark:/48223/pf0000380455 (zuletzt aufgerufen am 28.02.2023) *sowie*
Deutsche UNESCO Kommission (2022): UNESCO Empfehlung zur Ethik Künstlicher Intelligenz: Bedingungen zur Implementierung in Deutschland, https://www.unesco.de/sites/default/files/2022-03/DUK_Broschuere_KI-Empfehlung_DS_web_final.pdf (zuletzt aufgerufen am 28.02.2023)

[21] OECD/LEGAL/0049 (2019): Recommendation of the Council on Artificial Intelligence https://oecd.ai/en/ai-principles (zuletzt aufgerufen am 07.02.2023)

[22]ISO.org: ISO-Normen werden international von Experten vereinbart, https://www-iso-org.translate.goog/standards.html?_x_tr_sl=en&_x_tr_tl=de&_x_tr_hl=de&_x_tr_pto=sc (zuletzt aufgerufen am 28.02.2023)

[23] Bundesministerium für Wirtschaft und Energie (Dezember 2022):
Deutsche Normungsroadmap Künstliche Intelligenz
https://www.din.de/resource/blob/772438/6b5ac6680543eff9fe372603514be3e6/normungsroadmap-ki-data.pdf (zuletzt aufgerufen am 07.02.2023)

(RAI), also der gesellschaftlich verantwortliche, ethische Einsatz Künstlicher Intelligenz, zu Grunde.[24] Konkret geht es dabei in der Regel um immer ähnliche Aspekte, die als Leitplanken für einen ethisch korrekten und verantwortungsvollen Einsatz von KI gesehen werden und die im Wesentlichen auf der Anwendbarkeit zentraler Menschenrechte basieren, sowie anschlussfähig sind an die Regeln der EU-Grundrechtecharta und – in der deutschen Debatte – an die aus dem Grundgesetz abgeleiteten Ansprüche an Grund- und Bürger:innenrechte.

2. Die KI-Verordnung der Europäischen Union (KI-Act)

Die Europäische Kommission geht mit ihrem Verordnungsentwurf zur Regulierung Künstlicher Intelligenz[25] [26] einen deutlichen Schritt weiter als die bisherigen internationalen Bemühungen und knüpft an die Überlegungen der von ihr eingesetzten Expert:innengruppe an, die 2019 als unabhängige High-Level Expert Group on Artificial Intelligence Vorschläge zum verantwortlichen Einsatz von KI formuliert hat.[27] Sie unternimmt mit ihrer Vorlage des Verordnungsentwurfes vom 21. April 2021 den Versuch, ethische Standards für eine vertrauenswürdige KI in einem für alle Mitgliedsstaaten verbindlichen Rechtstext zu formalisieren und normativ zu beschreiben. Der Entwurf des KI-Acts richtet sich dabei an alle relevanten Akteure entlang der Wertschöpfungskette, so z.B. an Entwickler:innen und Hersteller:innen von KI-Systemen, an Händler:innen, an Anwender:innen und Nutzer:innen wie auch an diejenigen, die die von ihnen eingesetzte KI weiterentwickeln. Er gilt sowohl für den privaten als auch öffentlichen Sektor und sieht je nach konkreter Ausgestaltung des KI-Systems und Rolle der Akteure unterschiedliche Pflichten vor. Zentral ist den Vorschlägen der EU-Kommission dabei die Orientierung an einem risikobasierten Ansatz, der vor allem zum Ziel hat, die Grundrechte der Bürger:innen der EU zu schützen und zu gewährleisten, dass nur vertrauensvolle KI-Systeme in Europa zum Einsatz kommen.

[24] Für eine ausführlichere Darstellung des Konzepts vgl. Kapitel 3 und 4 der vorliegenden Studie.

[25] Vorschlag für eine VERORDNUNG DES EUROPÄISCHEN PARLAMENTS UND DES RATES ZUR FESTLEGUNG HARMONISIERTER VORSCHRIFTEN FÜR KÜNSTLICHE INTELLIGENZ (GESETZ ÜBER KÜNSTLICHE INTELLIGENZ) UND ZUR ÄNDERUNG BESTIMMTER RECHTSAKTE DER UNION, Brüssel, 21.04.2021, COM (2021) 206 final unter: https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC_1&format=PDF (zuletzt aufgerufen am 28.02.2023) *sowie*
Anhänge des Vorschlages für eine Verordnung des Europäischen Parlamentes und des Rates ZUR FESTLEGUNG HARMONISIERTER VORSCHRIFTEN FÜR KÜNSTLICHE INTELLIGENZ (GESETZ ÜBER KÜNSTLICHE INTELLIGENZ) UND ZUR ÄNDERUNG BESTIMMTER RECHTSAKTE DER UNION, https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC_2&format=PDF (zuletzt aufgerufen am 28.02.2023)

[26] Sofern spezifisch auf einzelne Artikel des EU-Kommissionentwurfs vom 21.4.2021 Bezug genommen wird, wird dieser als KI-Act Entwurf abgekürzt, ist allgemein von der zu erwartenden Regulierung durch eine europäische Verordnung die Rede, wird auf den KI-Act Bezug genommen.

[27] European Commission / Independent High-Level Expert Group on Artificial Intelligence (08.04.2019): Ethics Guidelines for Trustworthy AI, https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html (zuletzt aufgerufen am 28.02.2023)

2.1 Die Risikokategorien nach KI-Act[28]

Die von der EU-Kommission vorgeschlagenen Maßnahmen richten sich zentral auf die Förderung des Vertrauens der Bürger:innen in KI-Anwendungen, indem sie den Menschen, sein Recht auf Privatsphäre und Datenschutz, sein Recht auf Nichtdiskriminierung und Rechtssicherheit, in den Mittelpunkt rücken. Deswegen werden im Verordnungsentwurf gerade an solche Systeme besonders strenge regulatorische Vorgaben formuliert, die ein hohes Risiko für Gesundheit, Sicherheit oder Freiheit darstellen. Ihnen gebührt vor dem Hintergrund des Grundrechtsschutzes besondere Aufmerksamkeit. Um gerade in grundrechts- und sicherheitssensiblen Bereichen Risiken zu minimieren, werden daher im KI-Act KI-Systeme in drei Risikogruppen (plus eine Gruppe ohne Risikopotential) klassifiziert, bei denen weniger von Bedeutung ist, wie ein einzelner Algorithmus konkret formuliert ist, sondern vielmehr in welcher Form und in welchem Kontext er zur Anwendung kommen soll und/oder inwieweit die bestimmungsgemäße Nutzung der KI durch die Entwickler:innen auf den originären Zweck eingeschränkt wird:

1. Verboten nach Titel II Art. 5 KI-Act Entwurf:
    a. Unterschwellige Beeinflussung des Verhaltens einer Person
    b. Ausnutzung von Schwäche/Schutzbedürftigkeit bestimmter Personen (-gruppen) aufgrund von Alter/Behinderung
    c. Soziale Bewertung von Personen (durch/im Auftrag von Behörden)
    d. Biometrische Echtzeit-Fernidentifizierungssysteme in öffentlich zugänglichen Räumen zu Strafverfolgungszwecken [Ausnahmen]
2. Hohes Risiko nach Titel III Art. 6-51, Anhang II & III KI-Act Entwurf:
    a. Hochrisiko-KI nach EU-Harmonisierungsvorschriften (Anhang II)
        i. Produkte und Sicherheitskomponenten von Produkten, die unter die in Anhang II aufgelisteten EU-Harmonisierungsvorschriften fallen
    b. Hochrisiko-KI nach Anwendungskontext (Anhang III)
        i. Biometrische Identifizierung (verordnungskonforme Echtzeit- bzw. nachträgliche Fernidentifizierung)
        ii. Betrieb kritischer Infrastrukturen (z.B. Gas-, Wasser- und Stromversorgung)
        iii. Bildung (Zugang und Bewertungssysteme)
        iv. Beschäftigung, Personal (Recruiting und Leistungsbeurteilung)
        v. Private und öffentliche Dienste (Kreditprüfung, Zugang zu Sozialleistungen)
        vi. Strafverfolgung (Risikobewertung und Profiling)
        vii. Migration und Asyl (Antrags- und Statusprüfungen)
        viii. Rechtspflege (Rechtsanwendungen, Bewertungen von Sachverhalten)
3. geringes Risiko
    a. z.B. Chatbots[29], Deepfakes

[28] Grundlage dieses Kapitels stellt die Auseinandersetzung mit dem Gesetzesentwurf der Europäischen Komission vom 21.04.2021 dar.

[29] An der Einordnung von Chatbots als „geringes Risiko“ im Kommissionsentwurf lässt sich die Problematik der derart kategorisierten Risikoklassen bereits gut erkennen. Der Rat der Europäischen Union reagiert darauf, indem er Allgemeinzweck-KI, zu der Chatbots wie ChatGPT zählen würden, ausdrücklich in die Verordnung aufnimmt und

4. Kein Risiko
    a. z.B. KI-gestützte Videospiele, Spamfilter

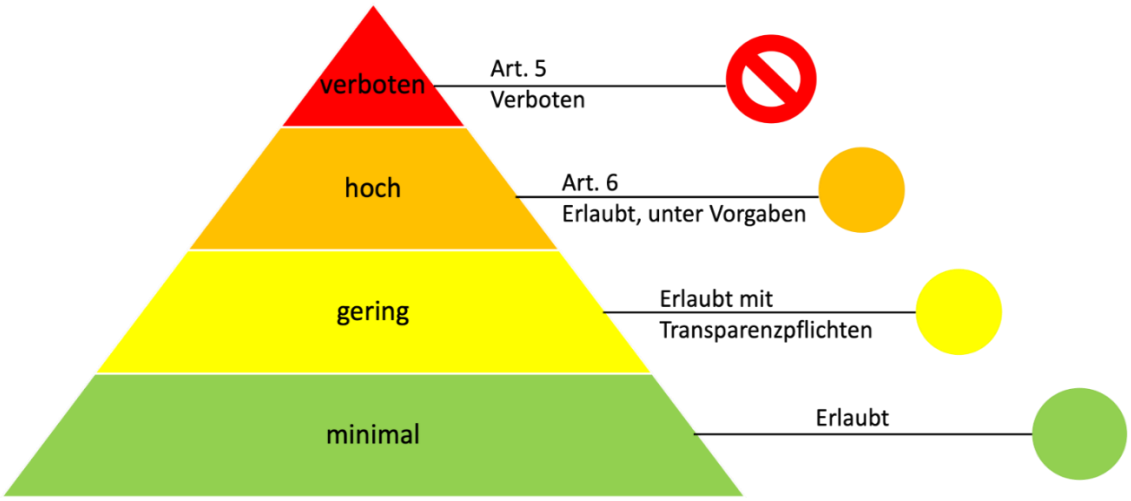

Abb.1: Eigene Darstellung: Der risikobasierte Ansatz nach KI-Act Entwurf

Innerhalb der Hochrisiko-Kategorie gibt es zwei mögliche Ursachen für die Einschätzung, dass der KI-Einsatz als hochriskant zu werten ist. Neben den oben aufgelisteten Anwendungskontexten nach Anhang III KI-Act Entwurf sind demnach außerdem Produkte oder Sicherheitskomponenten von Produkten, die Produktsicherheitsvorschriften in Form von EU-Harmonisierungsvorschriften unterliegen, als Hochrisiko-KI einzustufen. Dies betrifft beispielsweise Maschinen, Spielzeug, Aufzüge, Seilbahnen etc. gemäß Anhang II KI-Act Entwurf. Anbieter:innen von Systemen, die nach dieser Systematik als Hochrisiko-KI einzustufen sind, sind verpflichtet, die hohen Anforderungen des neuen Regelwerks grundsätzlich zu erfüllen, will man Sanktionen in Form von hohen Geldbußen oder Vertriebsverboten umgehen (Art. 71 und Art. 84 KI-Act Entwurf). Zusätzlich zu den Harmonisierungsvorschriften und Anwendungskontexten kann nach dem Kompromisstext des Rates Allgemeinzweck-KI höheren Anforderungen unterliegen (Art. 4 KI-Act Ratsentwurf). KI-Systeme mit allgemeinem Verwendungszweck sind dabei solche KI-Systeme, die von Anbieter:innen dazu vorgesehen sind, allgemein anwendbare Funktionen (z.B. Bild- oder Spracherkennung, Audio- und Videogenerierung, Mustererkennung, Beantwortung von Fragen und Übersetzung) auszuführen.[30]

Für die Praxis wird die Einteilung und genaue Abgrenzung dieser Risikogruppen von größter Relevanz sein, denn die entsprechenden Akteure müssen in Zukunft in der Lage sein, KI-Anwendungen, die sie entwickeln, vertreiben oder selbst nutzen, korrekt in die oben genannten Risikokategorien einsortieren zu können, um die gesetzlich geforderten Anforderungen dementsprechend erfüllen zu können. Darüber hinaus kann die Bedeutung des jeweiligen Anwendungskontextes eines KI-Systems kaum genug betont werden. So adressiert der Verordnungsentwurf sowohl in der Version der Kommission als auch des Rates zwar

auf ihren kontextabhängigen Einsatz hinweist. Siehe hierzu auch die weiteren Ausführungen dieses Unterkapitels sowie Kapitel 2.2.

[30] Rat der Europäischen Union (25.11.2022): Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Gesetz über künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union: Allgemeine Ausrichtung, https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/de/pdf (zuletzt aufgerufen am 28.02.2023)

zentral die Anforderungen an Hochrisiko-KI-Systeme; wann jedoch ein System in diese Zuordnung fällt, bleibt kontextabhängig und kaum allgemeingültig adressierbar. Die oben aufgeführte Auflistung wird somit stets von der Prüfung des Einzelfalls abhängen und dient lediglich – auch abgesehen von der im Detail noch offenen Definition und Abgrenzung der einzelnen Bereiche im Zuge der politischen Verhandlungen – als grobe Orientierung. Diese kontextbasierte, einzelfallabhängige Einordnung in Risikoklassen lässt somit einige Fragen offen und kann auch bei der Verwaltung Unsicherheit hervorrufen. Die Frage, welche Anforderungen erfüllt werden müssen, wird durch die Hinzunahme von Allgemeinzweck-KI in der Ministerratsvorlage umso zentraler.

Obwohl diese Unschärfe bzw. der Interpretationsspielraum, der bei der Einordnung in die Risikoklassen bleibt, zu Verunsicherungen führen kann, ist der Ansatz aus demokratischer und rechtsstaatlicher Perspektive zweckdienlich. Das hinter dem risikobasierten Ansatz stehende Ziel des Gesetzgebers liegt darin, Risiken für Gesundheit, Sicherheit und Freiheit zu minimieren. Ob der Einsatz eines KI-Systems eine Gefahr für die Gesundheit, Sicherheit oder Freiheit von Bürger:innen darstellen kann, lässt sich jedoch nicht anhand des verwendeten Lernverfahrens, eines spezifischen Algorithmus oder danach, ob es sich um daten- oder wissensbasierte KI handelt, einschätzen. So kann beispielsweise das Lernverfahren des Reinforcement Learning (siehe Glossar) verwendet werden, um einer KI beizubringen, wie man ein Level eines Videospiels absolviert oder wie sich ein selbstfahrendes Auto in einer Verkehrssituation orientiert. Würde die Einstufung in eine Risikokategorie anhand spezifischer Algorithmen oder Lernverfahren, wie z.B. Reinforcement Learning, erfolgen, wäre dies zwar mit weniger Unsicherheit u.a. für Verwaltungen verbunden. Gleichzeitig würde aber das gleiche Anforderungsmaß für KI-Systeme vorgeschrieben werden, die im Fall einer Fehlfunktion entweder in einem Videospiel verlieren oder potentiell Menschenleben in einem Verkehrsunfall kosten.

Analog zum kontextabhängigen Risiko für die Sicherheit gibt auch die Frage, wie stark der Einsatz der KI in die Grundrechte der Nutzer:innen eingreift, Anlass für unterschiedliche Anforderungsniveaus: Denn ob ein Chatbot beispielsweise tatsächlich pauschal in die Kategorie „geringes Risiko" fallen sollte, wird letztlich davon abhängen, was seine Aufgabe ist. Ein Chatbot, der in der Kund:innenberatung eines Onlinehändlers eingesetzt wird, erfordert voraussichtlich andere Voraussetzungen als beispielsweise ein Chatbot, der Zugang zu staatlichen Leistungen bietet. Der hieraus möglicherweise entstehenden Verunsicherung von Unternehmen stehen also durchaus gewichtige demokratische und rechtsstaatliche Gründe gegenüber. Um weder das Vertrauen der Bürger:innen zu riskieren, noch innovationshemmende Unsicherheit bei KMU und in Verwaltungen auszulösen, sind unterstützende Angebote zur Einordnung in die Risikokategorien (z.B. in Form von Checklisten, Kurzberatung o.ä.) geboten.

Zudem gibt es einige weitere Gründe, die dafürsprechen können, die Anforderungen im Sinne der RAI im Allgemeinen, also nicht nur, wenn es aufgrund der Einsortierung als Hochrisiko-KI zwingend notwendig ist, anzustreben: Zum einen wird im KI-Act wiederholt die Empfehlung nach Verhaltenskodizes hervorgehoben (Art. 69 KI-Act Entwurf), mit denen sich auch alle anderen KI-Anbieter:innen im europäischen Binnenmarkt an die wesentlichen Anforderungen des KI-Acts binden würden und nach den Vorstellungen der EU-Kommission auch binden sollten. Es wird aber auch im unmittelbaren Interesse der Verwaltung liegen, nicht nur das jeweilige Mindestmaß an gesetzlich vorgeschriebenen Anforderungen zu erfüllen. Will sich die Verwaltung etwa verschiedene mögliche Anwendungsfelder von selbst entwickelten oder eingekauften KI-Lösungen offenhalten, bietet es sich schon aus diesem Grund an, die im KI-Act genannten Anforderungen an Transparenz, Nachvollziehbarkeit oder notwendige Governance-Strukturen – auf die im Weite-

ren genauer eingegangen werden wird – grundsätzlich zu implementieren. Aber selbst wenn das KI-System nur für einen klar definierten Einsatzkontext entwickelt werden soll und dieser nach jetzigem Stand nicht in den Hochrisikobereich fällt, besteht die Chance, dass dieser Einsatzkontext durch einen Durchsetzungsrechtsakt zu einem späteren Zeitpunkt als hochriskant definiert wird. Die Anforderungen nachträglich in einem bestehenden KI-System umzusetzen, kann sehr aufwändig oder in manchen Fällen nicht realisierbar sein.

Auch die politische Debatte weist immer mehr in die Richtung, dass verlässliche KI-Lösungen Teil der sozial- und rechtsstaatlichen Grundsätzen verpflichteten Daseinsvorsorge sein müssen. Aus dieser Perspektive ist es Aufgabe des Staates, das bestmögliche und sichere Umfeld zu schaffen, um im Bereich der Verwaltung – aber nicht nur dort – Dienstleistungen zur Verfügung zu stellen, die die digitale Transformation – auch ohne gesetzliche Vorgaben – an Mindeststandards eines gesellschaftlich verantwortlichen Einsatzes von KI (RAI) knüpfen und damit das Vertrauen der Bürger:innen in die Technik stärken und die transformativen Veränderungen durch die Digitalisierung bestmöglich unter Beachtung des Gebotes der Wirtschaftlichkeit und Sparsamkeit realisieren. Es kann also durchaus im Interesse der Verwaltung sein – neben der ökonomisch getriebenen Vermeidung von Haftungsrisiken –, die Vertrauenswürdigkeit ihrer KI durch eine Konformitätserklärung und die Einhaltung bestimmter Regeln und Normen zu belegen.

2.2 Aktueller Diskussionsstand im Europäischen Parlament und im Rat der Europäischen Union

Seitdem der erste Entwurf des KI-Acts durch die Europäischen Kommission im April 2021 veröffentlicht wurde (KI-Act Entwurf), läuft das Gesetzgebungsverfahren unter Einbeziehung der Stellungnahmen und Kompromissvorschläge der Mitgliedsstaaten im Rat für Telekommunikation sowie im Europäischen Parlament überwiegend in den Ausschüssen für „Internal Market and Consumer Protection“ (IMCO) und „Civil Liberties, Justice and Home Affairs“ (LIBE). Im Oktober 2022 gab es eine erste Plenardebatte im Europäischen Parlament, in der zunächst ein Fokus auf die technischen Fragen des Verordnungsentwurfs gelegt wurde, um schwierigere politische Klärungsbedarfe zu einem späteren Zeitpunkt bearbeiten zu können. Während der Rat sich im Dezember 2022 auf einen gemeinsamen Standpunkt („Allgemeine Ausrichtung”) geeinigt hat, steht eine Einigung des Europäischen Parlaments zum Redaktionsschluss der Studie noch aus.[31]

In Anbetracht der schwierigen und komplexen Materie haben die verschiedenen an der Rechtssetzung der EU beteiligten Organe durchaus kontroverse und voneinander abweichende Vorstellungen über den Regulierungsgehalt und Regulierungsumfang des neuen KI-Acts. Der auf Grundlage des Vorschlages der tschechischen Ratspräsidentschaft vorgelegte gemeinsame Standpunkt des Ministerrates behält zwar wesentliche Teile des Kommissionsvorschlages bei, formuliert aber auch zahlreiche Änderungen, die nun als Basis für die interinstitutionellen Verhandlungen mit dem Europäischen Parlament und der Europäischen Kommission im Rahmen der anstehenden Trilog Verhandlungen dienen werden. Einzelne Mitgliedstaaten,

[31] Im Folgenden wird der Ratsbeschluss vom 25.11.2022 als KI-Act Ratsentwurf abgekürzt.

darunter auch die Bundesrepublik Deutschland, haben zudem nach wie vor Bedenken geäußert und diese noch nach Vorlage des gemeinsamen Standpunkts über Protokollnotizen der EU zur Kenntnis gegeben. [32]

Davon unbenommen beziehen sich die von den Mitgliedstaaten eingebrachten Änderungsbedarfe gegenüber dem Kommissionsentwurf bereits auf den Kerngehalt der Verordnung: auf die grundlegende Definition von KI-Systemen. Vorausgegangen war eine Auseinandersetzung unter den Mitgliedstaaten, dass der ursprüngliche Entwurf der Kommission ein zu breites Verständnis von KI zugrunde legen würde, nach dem potentiell alle Softwarelösungen im europäischen Binnenmarkt von den Anforderungen der Verordnung erfasst worden wären. In dem Bestreben den Begriff enger zu fassen, schlägt der Rat vor, nur Systeme in den Anwendungsbereich der Verordnung zu nehmen, die Maschinelles Lernen oder Logik sowie wissensbasierte Methoden verwenden und somit Such-, Schätz- und Optimierungsmethoden als KI-Technologien zu streichen (Art. 3 KI-Act Ratsentwurf). Zudem wird die Eigenschaft der Autonomie zur weiteren Abgrenzung von klassischer Software benannt.

Kommissionsentwurf (KI-Act Entwurf)	Kompromisstext des Rates (KI-Act Ratsentwurf)
„System der künstlichen Intelligenz“ (KI-System) 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 von Menschen festgelegt werden, Ergebnisse wie Inhalte, Vorhersagen, Empfehlungen oder Entscheidungen hervorbringen kann, die das Umfeld beeinflussen, mit dem sie interagieren”	„System der künstlichen Intelligenz“ (KI-System) ein System, das so konzipiert ist, dass es mit Elementen der Autonomie arbeitet, und das auf der Grundlage maschineller und/oder vom Menschen erzeugter Daten und Eingaben durch maschinelles Lernen und/oder logik- und wissensgestützte Konzepte ableitet, wie eine Reihe von Zielen erreicht wird, und systemgenerierte Ergebnisse wie Inhalte (generative KI-Systeme), Vorhersagen, Empfehlungen oder Entscheidungen hervorbringt, die das Umfeld beeinflussen, mit dem die KI-Systeme interagieren”

Neben der Definition von KI wird im Kompromisstext des Rates auch der Anwendungsbereich des KI-Acts angepasst: KI-Systeme, die ausschließlich zur wissenschaftlichen Forschung und Entwicklung eingesetzt werden, wären demnach vom Anwendungsbereich der Verordnung ausgenommen. Zudem werden private Nutzende (natürliche Personen, die KI zu ausschließlich privaten, nicht beruflichen Tätigkeiten verwenden) von den Pflichten unter der Verordnung befreit (Art.2 Abs. 6-8 KI-Act Ratsentwurf).

Schließlich sieht der Rat wesentliche Änderungen in Bezug auf die Hochrisikokategorie vor. Welche Anwendungen als hochriskant einzustufen seien, wurde lange zwischen den Mitgliedstaaten diskutiert. Wie oben beschrieben werden diese Systeme besonders strengen Anforderungen unterliegen und es ist daher für Entwickler:innen und Anbieter:innen von allergrößter Bedeutung, klar definieren zu können, ob sie rechtlich mit ihren Anwendungen in die Klassifizierung „Hochrisiko” fallen. Erst dies ermöglicht eine verbindliche Einschätzung, welche Anforderungen auf sie zukommen. In dem Versuch hier mehr Klarheit zu

[32] Rat der Europäischen Union (25.11.2022), Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Gesetz über künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union: Allgemeine Ausrichtung - *Erklärung Deutschlands,* https://data.consilium.europa.eu/doc/document/ST-14954-2022-ADD-1/de/pdf (zuletzt aufgerufen am 28.02.2023)

schaffen, weicht der Ministerrat in seinem Standpunkt vom Kommissionsvorschlag ab, indem er in Bezug auf die in Anhang III KI-Act Entwurf genannten Anwendungskontexte eine horizontale Bewertungsebene bei der Klassifizierung von Hochrisiko-Systemen hinzufügt, die gewährleisten soll, dass KI-Systeme, die wahrscheinlich kein erhebliches Risiko für Gesundheit, Sicherheit oder Grundrechte darstellen, von der Verordnung nicht erfasst werden. Rat und Kommission wollen also keine pauschale sektorielle oder branchenspezifische Regulierung, wie sie beispielsweise in den USA angestrebt wird, sondern eine klare horizontale Regelung entlang von Risikobewertungen und Risikokategorien.[33]

Aus Sicht des Rates soll der risikobasierte Ansatz – wie oben bereits erwähnt – auch für die sogenannte Allgemeinzweck-KI gelten, für die in einem gesonderten Durchführungsrechtsakt definiert werden soll, wann sie die Anforderungen für Hochrisiko-Systeme erfüllen müssen. Sollte sich die Ausweitung der Hochrisikoeinstufung auf Allgemeinzweck-KI im weiteren Verlauf der Verhandlungen durchsetzen, könnten von den Anforderungen an Hochrisiko-Systeme eine Vielzahl von Entwickler:innen betroffen sein (Art. 12 c) KI-Act Ratsentwurf). Diese Regelung würde dann nicht nur für alle Systeme gelten, die selber in die Hochrisiko-Kategorie fallen, sondern auch für solche, die mit ihren Funktionen lediglich in ein Produkt integriert sind, das erhöhten Sicherheitsstandards genügen muss, auch wenn die KI selbst nur als Teilsystem allgemeine Funktionen ausführt, wie Bild-, Sprach-, Texterkennung, Video- und Audioproduktion, Mustererkennung, Fragenbeantwortung oder Übersetzungen. Den entsprechenden Anbieter:innen solcher Systeme schon vor Inkrafttreten des Gesetzes einen Handlungsleitfaden zu bieten, stellt sich aktuell als schwierig bis unmöglich dar, da erst ein Jahr nach Inkrafttreten der Verordnung detaillierte Vorgaben entstehen sollen, wie genau mit der Definition von Allgemeinzweck-KI umgegangen werden soll und wann der Einfluss der Anwendung einer Allgemeinzweck-KI als so unwesentlich verstanden werden kann, dass die Anforderungen an Hochrisikosysteme nicht gelten.

Mit dem vorliegenden Entwurf der Kommission sowie dem gemeinsamen Standpunkt des Rates haben zwei der drei am Gesetzgebungsprozess beteiligten Akteure ihre Position zur Regulierung von KI-Anwendungen in der EU vorgelegt. Vor dem Hintergrund dieser zum Teil divergierenden Vorschläge bleibt es spannend, wie der weitere Gesetzgebungsprozess verläuft. Im nächsten Schritt wird das Europäische Parlament seinen Vorschlag abstimmen, bevor im Rahmen der Trilog-Verhandlungen eine gemeinsame Regelung gefunden werden kann. Dieser Prozess kann sich noch bis in die Mitte des Jahres 2023 oder darüber hinaus hinziehen. Die Verordnung soll in Folge nach gegenwärtigem Stand am zwanzigsten Tag nach ihrer Veröffentlichung in Kraft treten und ist dann – abgesehen von einigen Ausnahmen – als unmittelbar geltendes Recht in den Mitgliedstaaten der Europäischen Union ab dem 24. Monat nach Inkrafttreten gültig (Art. 85 KI-Act Entwurf).

2.3 Institutionelle Akteure, Konformitätsbewertungen und Notifizierungen nach KI-Act

Neben der Entwicklung eines umfassenden Rechtsrahmens mit Verpflichtungen für Entwickler:innen und Anwender:innen wird darüber hinaus auch institutionell an dem Ausbau für die Aufsicht und die Durchsetzung des sich entwickelnden Regelwerks gearbeitet. Innerhalb der EU sind entsprechende Überlegun-

[33] Kap. 1: Kontext des Vorschlages KI-Act Entwurf

gen bereits recht weit fortgeschritten. Denn die KI-Verordnung sieht neben konkret formulierten Anforderungen an Entwickler:innen, Anbieter:innen und Nutzer:innen für die Gestaltung und Funktionsfähigkeit eines Algorithmus ein System von Konformitätsprüfungen, Notifizierungen sowie Registrierungen vor, um zumindest bei KI-Systemen mit hohem Risiko den gesamten Lebenszyklus vom Inverkehrbringen bis zum Einsatz in der Praxis dauerhaft im Blick zu behalten.

- So haben die drei europäischen Normungsinstitutionen – das Europäische Komitee für Normung (CEN), das europäische Komitee für elektrotechnische Normung (CENELEC) sowie das Europäische Institut für Telekommunikationsnormen (ETSI) – von der Europäischen Kommission den Auftrag bekommen, nicht nur Prüfverfahren und Prüfmethoden für KI-Systeme und Risikomanagement-Systeme zu entwickeln, sondern auch konkrete technische Normen, bei deren Anwendung davon ausgegangen wird, dass diese als konform mit den EU-Regeln gelten können.[34]
- Zugleich wird an dem Aufbau einer Datenbank gearbeitet, in der sich Anbieter:innen einer Hochrisiko-KI-Lösung laut Art. 51 KI-Act Entwurf werden registrieren lassen müssen. Die Pflicht zur Registrierung soll nach dem Kompromissvorschlag des Rates vom Dezember 2022 nicht nur für Hersteller:innen und Anbieter:innen von Hochrisiko-KI gelten, sondern auch für in öffentlichen Einrichtungen und der Verwaltung eingesetzte KI-Systeme.
- Geplant ist ferner ein Europäischer Ausschuss für Künstliche Intelligenz, in dem neben Vertreter:innen der Mitgliedstaaten laut Kompromissvorschlag des Rates in einer ständigen Untergruppe auch unabhängige Vertreter:innen von KMU bzw. Start-ups, Wissenschaftler:innen, Vertreter:innen von Normungsgremien und notifizierten Stellen, Vertreter:innen von Laboren und Test- und Versuchseinrichtungen sowie Organisationen der Zivilgesellschaft vertreten sein sollen (Art. 56 Abs. 3 KI-Act Ratsentwurf). Nach Wunsch des Rates kommt dem KI-Ausschuss darüber hinaus gemäß Art. 58 KI-Act Ratsentwurf u.a. eine wesentliche Rolle in der Beratung der EU-Kommission bei der dauerhaften Überprüfung der Definition von Hochrisiko-KI nach Anhang III KI-Act Entwurf sowie bei der Formulierung von Durchführungsrechtsakten und praktischen Leitlinien zu. Insbesondere soll der Ausschuss seine Aktivitäten darauf konzentrieren, zu tatsächlich harmonisierten Verfahren in den Mitgliedstaaten in Bezug auf Konformitätsprüfungen und Notifizierungen zu kommen. Durch die Möglichkeit auch selbst initiativ tätig zu werden, so z.B. im Zusammenhang mit der Überprüfung der Definition von Hochrisiko-Systemen nach Anhang III KI-Act Entwurf, würde der Ausschuss über ein hohes Maß an Autonomie verfügen und in der Praxis große Bedeutung in der Governance-Architektur der KI-Verordnung erhalten.[35]

[34] Bertuzzi, Luca, EURACTIV (30.05.2022): EU-Normungsgremien sollen KI Standards ausarbeiten, https://www.euractiv.de/section/innovation/news/eu-normungsgremien-sollen-ki-standards-ausarbeiten/ (zuletzt aufgerufen am 28.02.2023)

[35] Bertuzzi, Luca (10.05.2022): KI-Gesetz: Frankreich für Änderungen bei Aufsichtsrat und Marktüberwachung, https://www.euractiv.de/section/innovation/news/ki-gesetz-frankreich-fuer-aenderungen-bei-aufsichtsrat-und-marktueberwachung/ (zuletzt aufgerufen am 28.02.2023) *sowie*
Zech, Maximilian (23.11.2022): So positioniert sich der Rat der EU zum AI Act, https://background.tagesspiegel.de/digitalisierung/so-positioniert-sich-der-rat-der-eu-zum-ai-act (zuletzt aufgerufen am 28.02.2023)

Grundsätzlich gilt also: Wie ein Produkt im europäischen Binnenmarkt zukünftig in Verkehr gebracht werden kann, hängt wesentlich davon ab, ob es in den Bereich der Hochrisiko-KI fällt und um welche Art von Hochrisiko es sich handelt. Auch wenn zum jetzigen Zeitpunkt vieles im Detail noch unklar ist, gibt es generelle Einigkeit darüber, dass insbesondere jedes System mit einer Hochrisiko-KI eine Konformitätsprüfung durchlaufen muss. Aber auch für alle anderen Anbieter:innen wie Anwender:innen aus Wirtschaft wie Verwaltung kann es gleichermaßen sinnvoll sein, auf eine Prüfung der Einhaltung bestimmter Normen und eine damit einhergehende Konformitätserklärung bzw. CE-Kennzeichnung zu achten. Denn auch Anwender:innen sind z.B. nicht nur darauf angewiesen, dass das ursprüngliche System bestimmten Kriterien entspricht. Bei einer Weiterentwicklung eines KI-Systems, beispielsweise durch weiteres Training mit neuen Daten, können aus Anwender:innen schnell Entwickler:innen werden, die somit ebenfalls selbst zu einer Konformitätsbewertung verpflichtet werden können.

Verwaltungen als ggf. Entwickler:innen oder auch Anwender:innen können sich bereits jetzt darauf einstellen, dass in vielen Fällen nach KI-Act Konformitätsprüfungen notwendig und eine Registrierung zumindest von Hochrisiko-KI zwingend erforderlich werden wird. Dabei wird sich aller Voraussicht nach die Konformitätsprüfung und damit der Weg zu einer Zertifizierung auf drei mögliche Verfahren konzentrieren:

a. Der:Die Hersteller:in kann die Konformitätsprüfung selbst vornehmen. Dies gilt für eigenständige KI-Systeme, die gemäß Anhang III in den Nummern 2-8 KI-Act Entwurf aufgeführt werden. Hier können die Anbieter:innen selber nach Art. 43 KI-Act Entwurf intern die Konformitätsprüfung übernehmen und die Ergebnisse dann der EU-Datenbank melden (Art. 51 KI-Act Entwurf). Im Anschluss kann die bekannte CE-Kennzeichnung erfolgen. Die Hersteller:innen orientieren sich in diesem Fall an den Anforderungen aus Anhang VI KI-Act Entwurf. Diese beziehen sich im Wesentlichen auf die Einhaltung von Art. 17 KI-Act Entwurf sowie eine Überprüfung der technischen Dokumentation, auch im weiteren Verlauf des Betriebs.

b. Der zweite Weg zu einer Zertifizierung ist durch eine von einer noch nicht definierten Bundesbehörde notifizierte Stelle, einer so genannten Konformitätsbewertungsstelle. Diese führt anstelle der Hersteller:innen die Prüfung durch und meldet die Systeme ebenfalls der EU-Datenbank. Diese Regelung kommt immer bei KI-Systemen nach Anhang VII KI-Act Entwurf zur Anwendung.

c. Bei (Sicherheitskomponenten von) Produkten, die gemäß Anhang II KI-Act Entwurf, also aufgrund der Tatsache, dass sie europäischen Harmonisierungsregeln unterliegen, als Hochrisiko-KI einzustufen sind, wird das Konformitätsbewertungsverfahren des jeweiligen Rechtsaktes angewandt, wobei KI-spezifische Aspekte im Sinne der Anforderungen aus Kapitel 2 KI-Act Entwurf im Rahmen der bestehenden Verfahren zusätzlich zu prüfen sind.

Noch ist unklar, wie die Aufsichtsstruktur rund um den KI-Act konkret aussehen wird. Die oben genannten Verfahren stehen somit unter Vorbehalt der endgültigen Einigung im Rahmen der Trilog-Verhandlungen im EU-Gesetzgebungsverfahren zu Struktur und Zuständigkeiten zwischen den Mitgliedstaaten sowie sich daran anschließenden Entscheidungen auf Bundesebene. Dennoch sind schon jetzt die Überlegungen auf europäischer Ebene zu der begleitenden Aufsichtsstruktur durchaus von praktischer Relevanz. Denn der nationalen Aufsichtsbehörde wird in der KI-Governance voraussichtlich eine bedeutende Rolle zukommen. Zumindest im Kommissionsvorschlag ist die Übertragung weitgehender Kompetenzen für Aufsicht und Kontrolle auf die Mitgliedstaaten vorgesehen. Nicht nur soll die nationale Aufsichtsbehörde die Stellen benennen können, die später Konformitätsbewertungen vor Ort durchführen, auch soll sie als Marktüberwachungsbehörde fungieren, um sicherzustellen, dass die Registrierung der KI-Systeme rechtskonform erfolgt, Betriebsprotokolle von Hochrisikoanwendungen ausgewertet und geprüft werden, sowie Anforderungen auch nach Inverkehrbringung weiterhin eingehalten werden. Sie dient darüber hinaus als zentrale Koordinationsstelle, als Ansprechpartnerin für die EU-Kommission und Vertreterin Deutschlands in den Ausschüssen der EU.

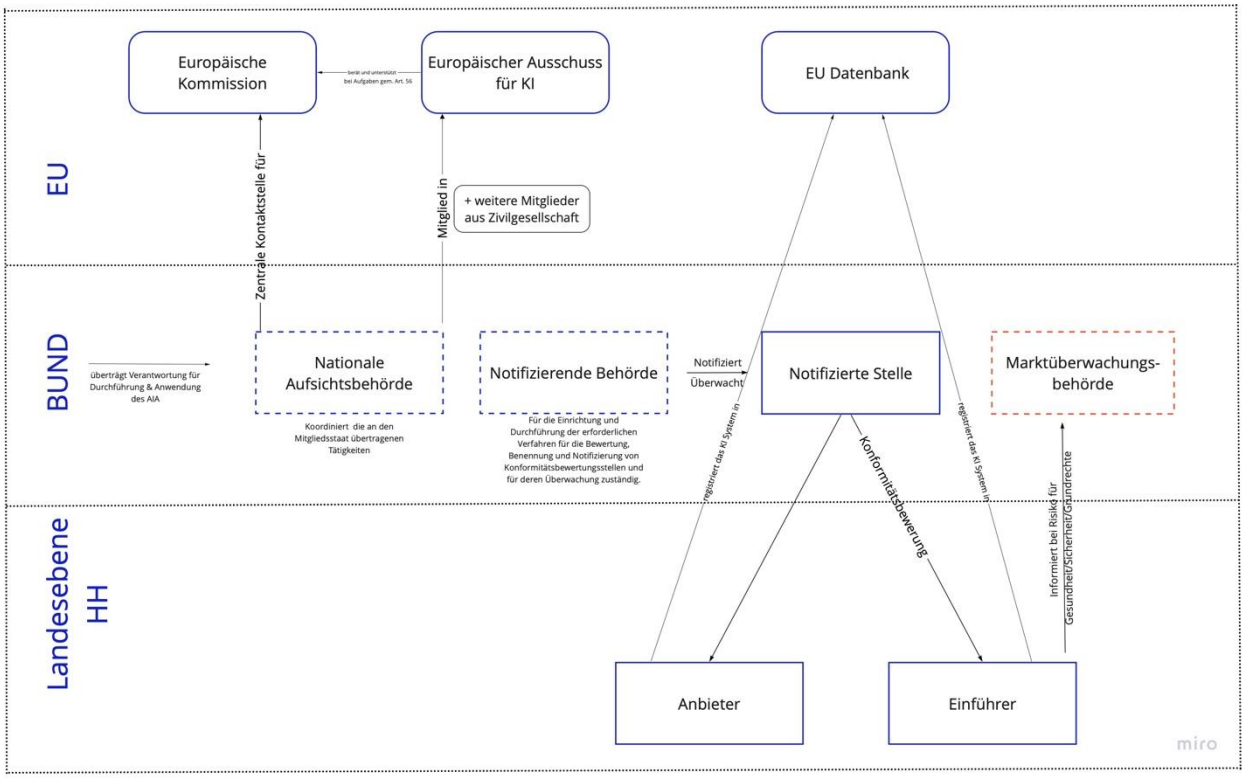

Abb. 2: Eigene Darstellung der KI-Governance und möglicher Zuständigkeiten im Notifizierungsprozess

Wie die Aufsichtsstruktur im Einzelnen in Deutschland genau aussehen wird, ist aus verschiedenen Gründen zwar noch nicht klar, in Hamburg zeichnet sich jedoch beispielsweise bereits ab, dass ein Joint Venture aus der Zertifizierungsgesellschaft Dekra Digital und der Wirtschaftsprüfungsgesellschaft PWC als deutschlandweiter Zertifizierungsanbieter für KI-Systeme auftreten wird. Vorbehaltlich der Bewilligung durch die Wettbewerbsbehörden wollen sie ein breites Spektrum an Aufgaben abdecken, die den Konformitätsbewertungsstellen (oder notifizierten Stellen) in der Praxis voraussichtlich zukommen werden, und

Unternehmen bei der Einhaltung regulatorischer Vorgaben unterstützen. Sich hier bereits als neuer Akteur zu positionieren, erhöht die Wahrscheinlichkeit, dass sie auch nach Maßgabe des KI-Acts durch eine Bundesbehörde notifiziert werden können.

Wie die obigen Ausführungen jedoch zeigen, gilt auch hier, dass der Verhandlungsprozess in der EU noch nicht abgeschlossen ist und es auch in der Governance-Frage Abweichungen zwischen den Vorschlägen der Kommission und denen des Rates gibt. Das zentral für den Verhandlungsprozess zuständige Bundeswirtschaftsministerium sieht daher keine Veranlassung, über die behördlichen Strukturen und Zuständigkeiten in Deutschland bereits zum jetzigen Zeitpunkt zu entscheiden.

2.4 Sanktionsregime und Haftungsregeln: KI-Act, KI-Haftungsrichtlinie und die neue Produkthaftungsrichtlinie der EU

Der KI-Act enthält ein eigenes Sanktionsregime für Fälle der Nichtbefolgung der Verordnung. Zudem hat die EU-Kommission im September 2022 im Kontext ihrer Regulierungsbemühungen den Entwurf einer EU-KI-Haftungsrichtlinie sowie den Entwurf für eine neue EU-Produkthaftungsrichtlinie vorgelegt. Die beiden Richtlinien sollen den KI-Act im Hinblick auf Haftungsfragen beim Einsatz von Software bzw. Künstlicher Intelligenz ergänzen. Eines der Hauptprobleme bei Haftungsfragen rund um das Thema Künstliche Intelligenz ist ganz grundsätzlich der Umgang mit der Komplexität, Autonomie und Undurchsichtigkeit von KI-Systemen. Im Vordergrund stehen daher sogenannte Black Box-Effekte sowie selbstlernende Systeme und die aus ihnen erwachsende Intransparenz und mangelnde Nachvollziehbarkeit des Entscheidungsprozesses. Die bisherigen haftungsrechtlichen Regelungen haben unter anderem vor diesem Hintergrund zu unbefriedigenden Ergebnissen geführt, etwa in Bezug auf die Frage, wer für durch KI eingetretene Schäden haftet oder wen in welcher Form die Beweislast trifft. Vor dem Hintergrund dieser Herausforderungen sind die neuen Regulierungsansätze der EU auch im Bereich des Haftungsrechts zu verstehen.

Im Folgenden wird kurz auf das Sanktionsregime des KI-Acts eingegangen (2.4.1), um sich anschließend den Entwürfen der KI-Haftungsrichtlinie (2.4.2) und der Produkthaftungsrichtlinie (2.4.3) zu widmen.

2.4.1 Sanktionsregime des KI-Acts[36]

Der Verordnungsentwurf der Kommission sieht gemäß Art. 71 Abs. 1 KI-Act Entwurf vor, dass die Mitgliedstaaten Sanktionsvorschriften erlassen, beispielsweise in Form von Geldbußen, die bei Verstößen gegen den KI-Act verhängt werden. Sie stellen darüber hinaus deren ordnungsgemäße und wirksame Durchsetzung sicher. Je nach Schwere des Verstoßes sind unterschiedliche Rahmen für die Höhe des Bußgeldes vorgegeben:

Den obersten Bußgeldrahmen gibt die Verordnung in Art. 71 Abs. 3 KI-Act Entwurf mit einer Höhe von bis zu 30 Millionen Euro oder (für Unternehmen) 6 Prozent des weltweiten Jahresumsatzes vor. Dieses Bußgeld droht beim Einsatz eines verbotenen KI-Systems nach Art. 5 KI-Act Entwurf sowie dann, wenn die Qualitätsanforderungen bezüglich Daten und Daten-Governance (Art. 10 KI-Act Entwurf) an Hochrisiko-KI-Systeme nicht eingehalten werden.

[36] Das folgende Kapitel bezieht sich ausschließlich auf den EU-Verordnungsentwurf vom 21.4.2021.

Als mittlerer Bußgeldrahmen sollen diejenigen Marktteilnehmer mit einer Geldbuße von bis zu 20 Millionen Euro oder (für Unternehmen) 4 Prozent des weltweiten Jahresumsatzes rechnen müssen, die KI-Systeme entwickeln, verwenden oder in Verkehr bringen, die gegen die sonstigen Anforderungen und Pflichten der KI-Verordnung verstoßen (außerhalb des Anwendungsbereichs von Art. 5 und 10 KI-Act Entwurf).

Als niedrigster Bußgeldrahmen wird in Art. 71 Abs. 5 KI-Act Entwurf eine Geldbuße von bis zu 10 Millionen Euro oder (für Unternehmen) 2 Prozent des weltweiten Jahresumsatzes vorgegeben, die verhängt werden soll im Falle der Übermittlung unvollständiger, falscher oder irreführender Angaben gegenüber Behörden.

2.4.2 KI-Haftungsrichtlinie[37]

Ziel der KI-Haftungsrichtlinie laut gegenwärtigem Entwurfsstand ist es, Geschädigten die gerichtliche Durchsetzung außervertraglicher Schadensersatzansprüche wegen unrechtmäßiger Handlungen oder Unterlassens zu erleichtern, die durch den Einsatz von KI entstehen. Geschädigte können sowohl Einzelpersonen wie auch Unternehmen oder Organisationen sein. Als Schädiger:in in Anspruch genommen werden können Anbieter:innen von KI ebenso wie Entwickler:innen und sonstige Marktteilnehmer:innen. Die KI-Haftungsrichtlinie wird also für viele KMU – und kann ggfs. auch für die Verwaltung – unmittelbare Relevanz haben. Sie erhöht die Anforderungen an die sorgfältige interne Dokumentation insofern, als dass bei einem auftretenden Schaden KMU regelhaft ihr Vorgehen werden belegen müssen. Das gilt insbesondere dann, wenn es um die Entwicklung von Hochrisiko-Systemen geht oder diese im Bereich sensibler Infrastruktur nach KI-Act eingesetzt werden.

Im Einzelnen soll die gerichtliche Durchsetzung von derartigen Schadensersatzansprüchen laut Kommissionsentwurf durch zwei die Beweisführung betreffende Faktoren erleichtert werden:

1. Erleichterung des Zugangs zu Beweismitteln durch eine Offenlegungspflicht für die Schädiger:innen

    a. Geschädigte sollen die Offenlegung von Informationen zu KI-Systemen bei Gericht beantragen können, um herauszufinden, wie es zu dem Schaden kommen konnte. Hierfür soll laut Richtlinienentwurf ausreichen, dass die Geschädigten die Plausibilität ihres Anspruches durch Vorlage von Tatsachen und Beweismitteln ausreichend belegen.

    b. Die widerlegbare Vermutung eines haftungsrechtlich relevanten Sorgfaltspflichtenverstoßes soll dann gelten, wenn der:die Schädiger:in der gerichtlichen Anordnung zur Offenlegung nicht nachkommt.

    c. Problematisch kann der Entwurf aus Unternehmenssicht dann sein, wenn die Gefahr besteht, Geschäftsgeheimnisse offenlegen zu müssen, die durch Dritte weiterverwendet werden könnten.

[37] COM (2022) 496 final (28.09.2022): Vorschlag für eine RICHTLINIE DES EUROPÄISCHEN PARLAMENTS UND DES RATES zur Anpassung der Vorschriften über außervertragliche zivilrechtliche Haftung an künstliche Intelligenz (Richtlinie über KI Haftung, http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=1677687781503) (zuletzt aufgerufen am 28.02.2023)

2. Widerlegbare Vermutung eines ursächlichen Zusammenhangs im Fall eines Verschuldens:

    a. Es wird ein Kausalzusammenhang zwischen dem (von dem:der Geschädigten bewiesenen oder vom Gericht auf Grundlage der maßgeblichen Umstände vermuteten) Verschulden der Anbieter:innen von Systemen der Künstlichen Intelligenz und den von ihren KI-Systemen erzeugten Ergebnissen oder dem Fehlen solcher Ergebnisse unter bestimmten Bedingungen widerlegbar vermutet. Die Kausalität muss dabei vernünftigerweise wahrscheinlich sein, aber von dem:der Geschädigten nicht mehr im Einzelnen nachgewiesen werden. Es muss aufgrund der Umstände des Falles lediglich als hinreichend wahrscheinlich angesehen werden können, dass das Verschulden bzw. die Sorgfaltspflichtverletzung, die von dem KI-System erzeugte Leistung oder das Versäumnis des KI-Systems, eine Leistung zu erbringen, beeinflusst haben.

    b. Ausgenommen von dieser Beweiserleichterung sind Geschädigte, die über ausreichende Möglichkeiten und Expertise verfügen, um den Beweis des ursächlichen Zusammenhangs zwischen einer Verletzung der Sorgfaltspflicht und dem Schaden selbst beizubringen.

Die vorstehend dargestellte Erleichterung der gerichtlichen Durchsetzung von außervertraglichen Schadensersatzansprüchen kann ggfs. auch Verwaltungen vor große Herausforderungen stellen. Die Einhaltung der Sorgfaltspflichten sollte daher zu jedem Zeitpunkt stringent beachtet und dokumentiert werden.

2.4.3 Produkthaftungsrichtlinie[38]

Die EU-Kommission hat daneben einen Vorschlag für die Überarbeitung der Produkthaftungsrichtlinie – die derzeitigen Regelungen sind fast 40 Jahre alt – auf den Weg gebracht, um den Herausforderungen „neuer" Technologien wie Software oder Künstlicher Intelligenz auch in der generellen Produkthaftung Rechnung zu tragen. Dabei wird eine Vereinheitlichung von europäischem Produktsicherheitsrecht und Produkthaftung angestrebt, die der dynamischen Lage angemessen ist.

1. Die Produkthaftung soll damit zukünftig auch digitale Produktionsdateien und Software einschließlich KI-Systeme umfassen, bezieht Fragen der Cybersicherheit des Produkts mit ein und erstreckt sich zeitlich über den Zeitpunkt des Inverkehrbringens hinaus, wenn es auch danach weiterhin möglich ist, das Produkt zu kontrollieren. Ähnlich den oben ausgeführten Neuerungen durch die KI-Haftungsrichtlinie soll es in Bezug auf Kontrolle, Beweislast und Offenlegungspflicht auch hier eine deutliche Stärkung der Seite des:der Geschädigten geben. Zugunsten des:der Geschädigten wird ein Kausalzusammenhang zwischen Produktfehler und Schaden vermutet, wenn offensichtliche Fehlfunktionen des Produktes bei normalem Gebrauch für den Schaden ursächlich sind.

[38] COM (2022) 495 Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on liability for defective products, https://single-market-economy.ec.europa.eu/system/files/2022-09/COM_2022_495_1_EN_ACT_part1_v6.pdf (zuletzt aufgerufen am 28.02.2023)

2. Auch sollen bisher mögliche gesetzliche oder vertragliche Haftungsausschlüsse der Hersteller:innen eingeschränkt werden. So sollen sich die Hersteller:innen etwa auf eine fehlende Erkennbarkeit eines Produktfehlers bei Inverkehrbringen nicht berufen dürfen, wenn dieser Fehler durch ein Software-Update hätte behoben werden können.

3. Auch hier soll es eine Offenlegungspflicht für Unternehmen geben, Konstruktionsunterlagen oder dokumentierte Erkenntnisse aus der Produktbeobachtung, die Geschädigte zur Begründung ihres Anspruches brauchen, herauszugeben. Bei Nichtbefolgung dieser Pflichten wird von einer gesetzlichen Vermutung der Fehlerhaftigkeit ausgegangen.

Eine Anpassung der Produkthaftungsrichtlinie an die Entwicklungen der Zeit und die Regelungsmaterie des KI-Acts ist sicher geboten. In Bezug auf KI wird sie laut des derzeitig vorliegenden Entwurfs dann relevant, wenn es sich bei der in Frage stehenden KI um eine Teilkomponente eines Produkts handelt, es also beispielsweise um Sicherheitskomponenten einer Maschine geht, deren Anforderungen in der Maschinenrichtlinie der EU definiert werden. Für Entwickler:innen und Anbieter:innen ist wie bei der KI-Haftungsrichtlinie von Bedeutung, dass ihre Haftung nicht automatisch erlischt oder zeitlich von vornherein zu begrenzen wäre. Auch der Umstand, dass Haftungsausschlüsse laut Richtlinienentwurf stark eingeschränkt werden und Beweislasterleichterungen durch die Offenlegungspflicht von Unterlagen eingeführt werden sollen, erfordert bereits in der Entwicklungsphase eines KI-Systems vorausschauende Dokumentation und die Erfüllung aller auf sie zukommenden Pflichten – auch mit Blick auf sich ändernde Anwendungskontexte bei einer Weiterentwicklung (s.o.).

2.5 Bedeutung des KI-Acts und der Haftungsrichtlinien für die Verwaltung

Der KI-Act wird auch für die Verwaltung Bedeutung entfalten. Vorbehaltlich etwaiger Änderungen im weiteren Normgebungsprozess wird im KI-Act Entwurf vom 21.4.2021 ausdrücklich auch die Verwaltung adressiert, u.a.:

- Gemäß Art. 3 Abs. 2 KI-Act Entwurf ist „Anbieter“ eine natürliche oder juristische Person, 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 oder in Betrieb zu nehmen.
- Gemäß Art. 3 Abs. 4 KI-Act Entwurf ist „Nutzer“ eine natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, die ein KI-System in eigener Verantwortung verwendet, es sei denn, das KI-System wird im Rahmen einer persönlichen und nicht beruflichen Tätigkeit verwendet.

Je nach Betroffenheit ergeben sich insbesondere im Bereich der Hochrisiko-KI-Systeme nach dem KI-Act unterschiedliche Pflichten (vgl. auch oben unter 2.1, 2.2).

Die in Anhang III gemäß Art. 6 Abs. 2 KI-Act Entwurf genannten Anwendungsbereiche, in denen KI-Systeme als Hochrisiko-KI-Systeme klassifiziert werden, können explizit auch die Verwaltung betreffen. Hierzu zählen z.B. Verwaltung und Betrieb kritischer Infrastruktur (Nr.2), Allgemeine und berufliche Bildung (Nr.3), Beschäftigung und Personalmanagement (Nr.4), Zugänglichkeit und Inanspruchnahme grundlegender privater und öffentlicher Dienste und Leistungen (Nr.5), Strafverfolgung (Nr.6), Migration, Asyl und Grenzkontrolle (Nr.7) sowie Rechtspflege und demokratischer Prozess (Nr.8).

Nach Inkrafttreten der Haftungsrichtlinien dürften nach ihrer Umsetzung in nationales Recht die Haftungsregelungen auch für die Verwaltung gelten. Dem dürften die derzeitigen Regelungen nicht entgegenstehen, die gemäß Art. 34 GG, § 839 BGB eine Haftung nur auf Fälle einer Amtspflichtverletzung, also vorsätzliches oder fahrlässiges Handeln, vorsehen. Der geltenden Normhierarchie nach gibt es zudem einen grundsätzlichen Anwendungsvorrang des EU-Rechts auch vor dem GG.

Zusammenfassend lässt sich sagen, dass die KI-Verordnung sowie die Entwürfe der KI-Haftungsrichtlinie und der Produkthaftungsrichtlinie nicht getrennt voneinander zu verstehen und zu denken sind. Sie verstärken sich in ihrer Wirkung gegenseitig und bilden erst gemeinsam ausreichenden Rechtsschutz sowohl für Anwender:innen von KI-Systemen als auch für Hersteller:innen und Entwickler:innen.[39] Dabei rückt die KI-Verordnung die Sicherheit der Systeme in den Mittelpunkt, um Schäden grundsätzlich zu vermeiden, während es bei der KI-Haftungsrichtlinie um den Umgang mit dennoch auftretenden Schäden geht. In Zusammenhang mit der ebenfalls überarbeiteten Produkthaftungsrichtlinie, die wie die KI-Haftungsrichtlinie mit Instrumenten der Beweislasterleichterung arbeitet, will die EU ein Regulierungsregime schaffen, das die verschuldensunabhängige Haftung stärker ins Zentrum ihrer Bemühungen stellt.[40] So nimmt die EU neben den Entwickler:innen und Hersteller:innen auch die Anwender:innen langfristig in die Pflicht, was für diese insbesondere dann zu Schwierigkeiten führen kann, wenn sie nicht über ausreichend technisches Wissen verfügen, um die sichere und transparente Funktion der KI-Anwendung prüfen zu können. Auch vor diesem Hintergrund wird eine Zertifizierung der Produkte für viele Akteure auf der Anwender:innenseite, insbesondere auch im Bereich der Verwaltung, schon beim Erwerb voraussichtlich deutlich an Bedeutung gewinnen, um belastbar in die ordnungsgemäße Funktion des Systems vertrauen zu können.

3. Vertrauensvolle KI: von der EU normiert, von den Bürger:innen gewünscht

Für die EU steht neben dem Ziel der Förderung eines global wettbewerbsfähigen und innovationsfreundlichen Umfeldes für Digitalunternehmen und Start-ups die Schaffung eines Regulierungsregimes für den Einsatz Künstlicher Intelligenz im Mittelpunkt ihrer Bemühungen, das sich an Menschenrechten, Demokratie, dem Schutz der Privatsphäre und der Aufrechterhaltung sozialstaatlicher Prinzipien orientiert. Hiermit soll zentral gewährleistet werden, dass der Einsatz von KI jederzeit sicher, transparent, ethisch und unparteiisch erfolgt – kurz, dass sich der Einsatz an den Werten und Prinzipien der Europäischen Union orientiert.

Politisch hat die EU dabei den Ausgleich verschiedener Interessen und Anliegen im Blick: So sollen gleichermaßen die wirtschaftlichen und gesellschaftlichen Potentiale von KI genutzt werden, die sich nach aktuellen Umfragen auch aus Sicht der Bürger:innen durch ihren Einsatz ergeben und zu Erleichterungen im Alltag führen (86%); und es soll zugleich den Sorgen der Bürger:innen vor einem unsachgemäßen Einsatz von KI-Systemen begegnet werden. Denn mit 66 Prozent äußert laut einer Forsa Umfrage des TÜVs vom

[39] Es ist zurzeit nicht absehbar, wann die beiden Richtlinien auf europäischer Ebene verabschiedet und anschließend in nationales und damit verbindliches Recht umgesetzt werden. Zunächst werden die beiden Entwürfe in das Abstimmungsverfahren zwischen Rat, Parlament und Kommission gehen, wobei es noch zu Änderungen kommen kann. Anschließend müssen die beiden Richtlinien mit den entsprechenden Gesetzgebungsverfahren in nationales Recht gegossen werden.

[40] Es geht hierbei um außervertragliche Haftungsfälle, die durch ein KI-System verursacht wurden und „*einen außervertraglichen verschuldensabhängigen zivilrechtlichen Schadensersatzanspruch*“ (siehe Art. 1 Abs. 2 KI-Haftungs-RL-E)

November 2022 ein beträchtlicher Teil Sorgen vor einer Diskriminierung durch den missbräuchlichen Einsatz von KI und auch die Befürchtung einer Manipulation von KI selbst (62%) sowie potentiell negativer Auswirkungen einer „Vermenschlichung" von Algorithmen bleibt groß.[41]

Vor diesem Hintergrund kann der Faktor „Vertrauen" in die eingesetzte Technologie von dem Ziel der EU, eine dynamische Entwicklung europäischer KI-Ökosysteme zu stimulieren, nicht losgelöst werden. In diesem Kontext sind auch die Regulierungsbemühungen zu verstehen. Die Bürger:innen sollen sich darauf verlassen können, dass innerhalb der EU angewendete und entwickelte Technik nicht nur technisch auf höchstem Niveau ist, sondern auch grundlegende europäische Werte wahrt und den Grundrechtsschutz gewährleistet. Das Recht bietet demnach die Chance, notwendiges Vertrauen in die Technologie zu stärken und die durchaus noch vorhandene kritische Grundhaltung der Bürger:innen in eine höhere Akzeptanz zu überführen. Dies gilt erst recht für den Einsatz von KI in der Verwaltung, wo die Bürger:innen einwandfreies rechtsstaatliches Handeln erwarten dürfen und einen juristisch durchsetzbaren Anspruch darauf haben – und zwar unabhängig von der Art der technischen Hilfsmittel, derer sich die Verwaltung für die Erledigung ihrer Aufgaben bedient.

In dieser Hinsicht kommt eine europaweit gültige Verordnung gerade recht. Da allerdings davon auszugehen ist, dass Nutzer:innen und Betroffene, die mit Sorgen auf die Technologie schauen, nicht zwischen Allgemeinzweck-KI, Hochrisiko-KI und KI-Systemen mit geringem Risiko differenzieren, sondern jegliche Software-Anwendungen, wie das oben genannte Beispiel zum Kindergeldskandal in den Niederlanden exemplarisch zeigt, den gleichen öffentlichen Bewertungsmaßstäben unterliegen, wird ein breites gesellschaftliches Vertrauen in die Technik voraussichtlich nur dann entstehen, wenn auch außerhalb des Einsatzes in Hochrisikobereichen entsprechende Anforderungen gelten oder eingehalten werden. Denn die Bürger:innen wollen zwar dem Einsatz von KI trauen können, doch ob der fehlerhafte Einsatz von einer Hochrisiko-KI oder einer klassischen, nicht-KI-basierten Softwarelösung, die nicht vom KI-Act erfasst wird, ausgeht, interessiert im Schadensfall weniger. Wenngleich der KI-Act wie oben beschrieben zunächst nur an Hochrisiko-KI verbindliche Anforderungen stellt, so könnten die von der Kommission empfohlenen Verhaltenskodizes, die grundlegende Qualitätsanforderungen auch für Nicht-Hochrisiko-KI vorsehen, diese Lücke adressieren.

Dieses Bedürfnis nach Vertrauen stellt sowohl Akteure an den technischen Schnittstellen, als auch diejenigen im politischen Diskurs, in der Gesetzgebung und der Verwaltung vor die Herausforderung, konkret zu beantworten, durch welche Mechanismen und Prozesse ein vertrauensvoller Umgang mit KI gewährleistet werden kann, und zwar unabhängig vom risikobasierten Ansatz der EU im KI-Act und bereits vor dem Inkrafttreten. Um ein breites Vertrauen in KI-Systeme zu schaffen, bedarf es also insbesondere eines Ansatzes, der das, was hinter der Kulisse des Systems passiert, erklärt, nachvollziehbar und zugänglich macht. Und nicht nur dort: Dieses Transparenzgebot sollte vor dem Hintergrund der Akzeptanz der Bevölkerung auch generell für die verschiedensten Softwarelösungen gelten, die einen wesentlichen Einfluss auf die Grundrechte der Bürger:innen haben.

[41] TÜV Verband e.V. (23.11.2022): Verbraucher:innen fordern gesetzliche Regeln für Künstliche Intelligenz [Pressemeldung], https://www.presseportal.de/pm/65031/5377332 (zuletzt aufgerufen am 28.02.2023)

4. Das Konzept Responsible AI – der gesellschaftlich verantwortliche Einsatz von KI

Grundsätzlich stellt sich die Frage, wie sich das Bedürfnis nach Vertrauen in und Verlässlichkeit von KI-Systemen in Form konkreter Maßnahmen und Prozesse in die Praxis übersetzen lässt. Daher stehen sowohl Akteure an den technischen Schnittstellen als auch diejenigen im politischen Diskurs und in der Gesetzgebung aktuell vor der Herausforderung, einen Ansatz zu finden, der erklärt und nachvollziehbar macht, wie das System funktioniert und die Verlässlichkeit gewährleistet. In der Praxis und Forschung werden vor diesem Hintergrund diverse technische und organisatorische Maßnahmen, die dazu beitragen, den Einsatz von KI vertrauensvoll zu gestalten, unter dem Begriff *Responsible AI* (im Folgenden RAI), zusammengefasst. So werden an der Schnittstelle zwischen Recht, Technik und Soziologie entsprechende Fragestellungen vorwiegend konzeptionell bearbeitet, während in der Praxis zunehmend konkrete Kriterien zur Anwendung kommen, die eine Brücke schlagen wollen, zwischen ethischen Ansprüchen, rechtlichen Normen und technischen Möglichkeiten.[42] Dabei wird der verantwortliche Einsatz von KI daran gemessen, ob

- der KI-Einsatz im Einklang mit europäischen Wertevorstellungen und rechtlichen Anforderungen steht (Ethik & Recht),
- es nachvollziehbar ist, wie das Ergebnis der KI zustande kommt (Erklärbarkeit),
- Fehlfunktionen und Ausfälle minimiert werden (Robustheit & Sicherheit),
- die Ausgaben nicht zu Diskriminierung führen (Fairness & Nachhaltigkeit)
- und es klare Zuständigkeiten und Prozesse für die Entwicklung und den Betrieb der KI-Anwendung gibt (Governance).

RAI bewegt sich damit in der Schnittmenge von Recht und Technik, wobei in der Praxis rechtliche Anforderungen und technische Möglichkeiten häufig kontext- und anwendungsfallspezifisch austariert werden müssen.

Ebendieser Logik folgt – wenn auch zentral auf Anforderungen im Hochrisikobereich fokussiert – auch der Entwurf des KI-Acts der EU. Man kann ihn also als regulative Entsprechung des normativen Konzepts einer RAI verstehen, wobei sich der hohe Grad an Kontextspezifität von RAI im Verordnungsentwurf niederschlägt. So finden sich etwa in Titel III, Kapitel 2 KI-Act Entwurf mit den Kriterien einer RAI korrespondierende Vorgaben:

- Anforderungen an die Erklärbarkeit von KI-Systemen in Art. 13 (Transparenz & Bereitstellung von Informationen für die Nutzer:innen),
- Anforderungen an Fairness in Art. 10 (Daten & Daten-Governance),
- Anforderungen an Ethik & Regulierung in Art. 11 (Technische Dokumentation), Art. 12 (Aufzeichnungspflichten), Art. 19 (Konformitätsbewertung),

[42] Huchler, Norbert et. al., Plattform Lernende Systeme (Juni 2020): Kriterien für die Mensch-Maschine-Interaktion bei KI: Ansätze für menschengerechte Gestaltung in der Arbeitswelt, Whitepaper der AG Arbeit/Qualifikation, https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG2_Whitepaper2_220620.pdf (zuletzt aufgerufen am 28.02.2023) *sowie*
Autor unbekannt, Fraunhofer IKS (2023): Künstliche Intelligenz: Sicherheit für Industrie, Medizin und autonomes Fahren erfordert zuverlässige Entscheidungen, https://www.iks.fraunhofer.de/de/leistungen/zuverlaessige-kuenstliche-intelligenz.html (zuletzt aufgerufen am 07.02.2023)

- Anforderungen an Robustheit & Sicherheit in Art. 15 (Genauigkeit, Robustheit, Cybersicherheit),
- und Anforderungen an Governance in Art. 9 (Risikomanagementsystem), Art. 14 (menschliche Aufsicht), Art. 17 (Qualitätsmanagementsystem).

Eine Konkretisierung der im KI-Act Entwurf normierten Anforderungen ist durch die laufenden Aktivitäten in den oben beschriebenen Gremien zu erwarten. Dies wird jedoch aller Voraussicht nach zu einem dauerhaften politischen Aushandlungsprozess und ständigen an der Praxis orientierten Anpassungen führen. Vor dem Hintergrund einer sich rasant vollziehenden Entwicklung in dem Bereich scheint genau dieser gesellschaftliche Aushandlungsprozess, der institutionell auch durch die Ergänzung des KI-Ausschusses um zivilgesellschaftliche Akteure seinen Widerhall findet, auch zwingend notwendig. Will man das durch das Regulierungsregime zu erwartende aufgebaute Vertrauen der Bürger:innen in die Potentiale und die Sicherheit der Technologie nicht erschüttern, sind kontinuierliche Anpassungen und Bewertungen des Risikopotentials von KI-Systemen in den zuständigen europäischen Gremien notwendig und ratsam.

Unabhängig von der EU-Rechtssetzung ist es für den Einsatz von Systemen, die KI verwenden, in der Verwaltung geboten, allgemeine Grundsätze für einen gesellschaftlich verantwortlichen Einsatz von KI zu beachten. Dies gilt auch für die Freie und Hansestadt Hamburg bei der rechtssicheren Umsetzung der Digitalstrategie. Bei Berücksichtigung der Kriterien einer RAI würde bereits ein Großteil der vertrauensbildenden Faktoren sowohl im Sinne der Bürger:innen in das Verwaltungshandeln als auch der Unternehmen abgedeckt.

5. Verwaltungsrechtlicher Rahmen für den Einsatz von KI in der Verwaltung

Im Anschluss an die obigen Ausführungen wird der Blick nunmehr auf die bereits heute gültigen einschlägigen verwaltungsrechtlichen Vorschriften einschließlich des Binnenrechts geworfen und etwaige Änderungsbedarfe diskutiert. Nach einem kurzen Blick auf die Ebenen des Verwaltungshandeln, soweit sie für den Einsatz von KI relevant sein könnten (5.1) und auf die verfassungsrechtlichen Anforderungen des Grundgesetzes (5.2) werden einschlägige verwaltungsrechtliche Vorschriften (5.3) und die Frage hinreichender demokratischer Legitimation KI-basierter verbindlicher Verwaltungsentscheidungen (5.4) behandelt. Außerdem soll diskutiert werden, ob beim Einsatz von KI – und wenn ja unter welchen Voraussetzungen – hinreichender Verwaltungsrechtsschutz gewährleistet werden kann (5.5).

5.1 Ebenen des Verwaltungshandelns

Es können im Wesentlichen fünf Ebenen des Verwaltungshandelns zu Grunde gelegt werden, in denen es zum Einsatz von KI kommen kann:

1) verwaltungsinterne Prozessabläufe (z.B. Einführung elektronischer Akte, Chatbots für Mitarbeiter:innen-Anfragen)
2) Kommunikation mit Bürger:innen (z.B. Chatbots, Sprachassistenten wie text to speech – speech to text -, Plattformen wie hamburg.de, Behördenfinder etc.)

3) Entscheidungsvorbereitung im Verwaltungshandeln (z.B. Auswertung von Daten, Kontexte herstellen, Recherche, Entscheidungsvorschläge)
4) Treffen von Entscheidungen (z.B. Verwaltungsakte, Planungsentscheidungen)
5) Kontrolle der getroffenen Entscheidungen (z.B. Rechtsbehelfe, Controlling, Datenabgleich)

Auf den verschiedenen Ebenen kann KI zur Effizienzsteigerung oder zur Prozessoptimierung in der Verwaltung eingesetzt werden. Sie greifen unterschiedlich tief in die individuellen Rechte der Bürger:innen ein. Sie unterliegen bzw. werden, etwa in Hinblick auf den KI-Act, unterschiedlichen rechtlichen Anforderungen unterliegen. Ob sich der Einsatz von KI für alle Ebenen des Verwaltungshandelns gleichermaßen eignet und rechtssicher erfolgen kann, soll im Folgenden geklärt werden.

5.2 Verfassungsrechtliche Anforderungen

Grundsätzlich sind beim Einsatz von KI nach dem Grundgesetz in erster Linie folgende verfassungsrechtliche Maßstäbe für die Bewertung der folgenden Prüfungspunkte zu beachten:

- Gemäß Art. 1 Abs. 3 GG sind Gesetzgebung, die vollziehende Gewalt – also die Verwaltung – und die Rechtsprechung an die Grundrechte als unmittelbar geltendes Recht gebunden.
- Gemäß Art. 19 Abs. 4 GG steht denjenigen, die durch die öffentliche Gewalt in ihren Rechten verletzt werden, der Rechtsweg offen.
- Nach Art. 20 Abs. 1 GG ist die Bundesrepublik Deutschland ein demokratischer und sozialer Rechtsstaat.
- Nach Art. 20 Abs. 2 GG geht alle Staatsgewalt vom Volke aus; sie wird durch besondere Organe der Gesetzgebung, der vollziehenden Gewalt und der Rechtsprechung ausgeübt.
- Gemäß Art. 20 Abs. 3 GG sind Verwaltung und Rechtsprechung an Gesetze und Recht, damit auch an Verordnungen und Satzungen gebunden.

Demnach kann der verfassungskonforme Einsatz von KI nur erfolgen, wenn den Bürger:innen die Möglichkeit eingeräumt wird, den Rechtsweg gegen eine durch eine KI getroffene Entscheidung zu beschreiten. Dies stellt insbesondere die Verwaltung vor Herausforderungen, wenn zugleich auf die einer KI inhärent mangelnden Nachvollziehbarkeit Bezug genommen werden muss. In diesem Balanceakt bewegen sich zukünftig Verwaltungsentscheidungen, die auf automatisierte Verfahren zurückgreifen. Ein Blick in die genuin verwaltungsrechtlichen Regelungen, die auf die Möglichkeiten des Einsatzes von KI rekurrieren, scheint vor diesem Hintergrund unerlässlich, um die entsprechende Rechtsgrundlage darzustellen und zu einer Einschätzung zu gelangen, inwiefern die bereits vorhandenen Regelungen den Einsatz neuer Technologien auf den verschiedenen Ebenen des Verwaltungshandelns ermöglichen oder ob es einer Ergänzung und Anpassung bedarf.

5.3 Verwaltungsrechtliche Regelungen

Im Folgenden sollen in einem exemplarischen Überblick insbesondere verwaltungsverfahrensrechtliche Vorschriften benannt werden, die ausdrücklich „automatisierte“ Entscheidungen ermöglichen bzw. in diesem Kontext relevant sein könnten (5.3.1). Darüber hinaus soll ein Blick auf das Binnenrecht geworfen

werden (5.3.2). Im Schwerpunkt werden anschließend §§ 35, 35a, 39 VwVfG auf Normierungsbedarfe hin untersucht (5.3.3).

5.3.1 Verfahrensrechtliche Vorschriften zum automatisierten Erlass von Verwaltungsakten

Verfahrensrechtliche Vorschriften, die ausdrücklich den automatisierten Erlass von Verwaltungsakten regeln, finden sich insbesondere im Verwaltungsverfahrensgesetz, SGB X und der Abgabenordnung:

1) § 35a VwVfG (Vollständig automatisierter Erlass eines Verwaltungsaktes)

   Ein Verwaltungsakt kann vollständig durch automatische Einrichtungen erlassen werden, sofern dies durch Rechtsvorschrift zugelassen ist und weder ein Ermessen noch ein Beurteilungsspielraum besteht.

2) § 24 Abs. 1 VwVfG (Untersuchungsgrundsatz)

   Setzt die Behörde automatische Einrichtungen zum Erlass von Verwaltungsakten ein, muss sie für den Einzelfall bedeutsame tatsächliche Angaben des Beteiligten berücksichtigen, die im automatischen Verfahren nicht ermittelt würden.

3) § 31a SGB X (Vollständig automatisierter Erlass eines Verwaltungsaktes)

   Ein Verwaltungsakt kann vollständig durch automatische Einrichtungen erlassen werden, sofern kein Anlass[43] besteht, den Einzelfall durch Amtsträger zu bearbeiten. Setzt die Behörde automatische Einrichtungen zum Erlass von Verwaltungsakten ein, muss sie für den Einzelfall bedeutsame tatsächliche Angaben des Beteiligten berücksichtigen, die im automatischen Verfahren nicht ermittelt würden.

4) § 155 Abs. 4 Abgabenordnung (Steuerfestsetzung)

   Die Finanzbehörden können Steuerfestsetzungen sowie Anrechnungen von Steuerabzugsbeträgen und Vorauszahlungen auf der Grundlage der ihnen vorliegenden Informationen und der Angaben des Steuerpflichtigen ausschließlich automationsgestützt vornehmen, berichtigen, zurücknehmen, widerrufen, aufheben oder ändern, soweit kein Anlass dazu besteht, den Einzelfall durch Amtsträger zu bearbeiten. Das gilt auch

   1. für den Erlass, die Berichtigung, die Rücknahme, den Widerruf, die Aufhebung und die Änderung von mit den Steuerfestsetzungen sowie Anrechnungen von Steuerabzugsbeträgen und Vorauszahlungen verbundenen Verwaltungsakten sowie,

[43] Ein Anlass zur Bearbeitung durch Amtsträger liegt insbesondere vor, soweit der Steuerpflichtige in einem dafür vorgesehenen Abschnitt oder Datenfeld der Steuererklärung Abgaben im Sinne des § 150 Abs. 7 gemacht hat. Bei vollständig automationsgestütztem Erlass eines Verwaltungsakts gilt die Willensbildung über seinen Erlass und über seine Bekanntgabe im Zeitpunkt des Abschlusses der maschinellen Verarbeitung als abgeschlossen.

2. wenn die Steuerfestsetzungen sowie Anrechnungen von Steuerabzugsbeträgen und Vorauszahlungen mit Nebenbestimmungen nach § 120 versehen oder verbunden werden, soweit dies durch eine Verwaltungsanweisung des Bundesfinanzministeriums oder der obersten Landesfinanzbehörden allgemein angeordnet ist.

5) § 74 Abs. 1 LHO (Regelungen zur Zulassung von IT-Verfahren in Bezug auf deren Einsatz in der Steuerverwaltung)

Verfahren der Informationstechnik (IT) für:
1. elektronische Anordnungen,
2. Buchungen,
3. Zahlungen,
4. Aufbewahrung von Nachweisen der Buchungen,
5. Geldverwaltung oder
6. Abschlüsse

dürfen nur eingesetzt werden, wenn sie von der für die Finanzen zuständigen Behörde zugelassen wurden. Diese kann im Einvernehmen mit dem Rechnungshof auf das Zulassungserfordernis verzichten. Der Schutz des Staatsvermögens vor unzulässigen Eingriffen sowie die Zuverlässigkeit, Vollständigkeit und Revisionsfähigkeit der Rechnungslegung sind zu gewährleisten.

6) § 74 Abs. 2 LHO (Regelungen zur Zulassung von IT-Verfahren in Bezug auf deren Einsatz in der Steuerverwaltung)

Die für die Finanzen zuständige Behörde stellt die IT-Verfahren zur Verfügung, die für das Haushalts-, Kassen- und Rechnungswesen der Freien und Hansestadt Hamburg notwendig sind. Sie kann technische Hilfstätigkeiten durch andere Verwaltungsträger verrichten lassen. Technische Hilfstätigkeiten sind insbesondere Rechenzentrumsleistungen, die Erstellung, Anpassung und Pflege von Software, technisches Monitoring, technische Analyse von Fehlern und auf diese Tätigkeiten bezogene Beratungsleistungen. Die technischen Hilfstätigkeiten des beauftragten Verwaltungsträgers sind der Freien und Hansestadt Hamburg zuzurechnen. Es ist sicherzustellen, dass die technischen Hilfstätigkeiten entsprechend den fachlichen Weisungen der für die Finanzen zuständigen Behörde verrichtet werden.

7) Art 22 Abs. 1 DSGVO

Personen haben das Recht, nicht einer ausschließlich auf einer automatischen Verarbeitung beruhenden Entscheidung unterworfen zu werden, die ihnen gegenüber rechtliche Wirkung entfaltet oder sie in ähnlicher Weise erheblich beeinträchtigt. Gemäß Absatz 2 b) dieser Vorschrift gilt dies nicht, wenn die Entscheidung aufgrund von Rechtsvorschriften der Union oder der Mitgliedstaaten, denen der Verantwortliche unterliegt, zulässig ist und diese

Rechtsvorschriften angemessene Maßnahmen zur Wahrung der Rechte und Freiheiten sowie der berechtigten Interessen der betroffenen Person enthalten. Durch die obigen Vorschriften wird diese Ausnahmeregelung grundsätzlich ausgefüllt.[44]

5.3.2 Änderungsbedarfe bei §§ 35a, 35 ,39 und 29 VwVfG im Hinblick auf automatisierte Entscheidungen und Entscheidungsvorbereitung durch KI

Wenn Schlüsselreize, die einen bestimmten Entscheidungsvorschlag der KI hervorgerufen haben, nicht offengelegt werden, gibt es ein Problem der Nachvollziehbarkeit. Eine Offenlegung ermöglicht zwar die Manipulierbarkeit, ist allerdings im Hinblick auf die interne und externe Nachvollziehbarkeit erforderlich. Hinsichtlich einer möglichen Manipulation ist es auch in der analogen Verwaltungswelt üblich und nicht verboten, dass Bürger:innen bestimmte Stichworte nennen, um das Verwaltungsverfahren in ihrem Sinne zu beeinflussen. Was im Einzelfall unproblematisch scheint, muss aber im Hinblick auf die generelle Manipulation der KI ausgeschlossen werden.

Im Folgenden sollen vor diesem Hintergrund die §§ 35a, 35, 39 und 29 VwVfG näher beleuchtet werden.

1) §§ 35a, 35 VwVfG

§ 35a VwVfG erlaubt den Erlass von Verwaltungsakten im automatischen Verfahren gemäß seinem Wortlaut unter folgenden Voraussetzungen:

- Die vollständig automatisierte Entscheidung ist durch Rechtsvorschrift zugelassen.
- Es darf weder ein Beurteilungs- noch Ermessensspielraum bestehen.

Auch weil diese Vorschrift in der Literatur teilweise als „fortschrittsfeindlich“ kritisiert wird[45], ist fraglich, ob sich im Hinblick auf den Einsatz von KI-Systemen, ein zwingender Änderungsbedarf ergibt.

Der Einschränkung liegt eine einleuchtende Wertung zugrunde: Komplexe Entscheidungen, die auf einer Abwägung widerstreitender Interessen beruhen, will der Gesetzgeber stets in menschlicher Hand belassen. Nur Amtsträger:innen sollen die Einzelfallgerechtigkeit herstellen, die behördliche Entscheidungsspielräume ermöglichen. Damit ist auch eine (entsprechende) Anwendung dieser Vorschrift auf Fälle der Ermessensreduzierung auf Null oder der Rechtsfigur des sogenannten intendierten Ermessens ausgeschlossen. Denn die Feststellung, ob ein Fall der Ermessensreduzierung auf Null vorliegt, kann erst Ergebnis der Ermessensprüfung selbst sein. Eine entsprechende Anwendung würde dem Willen des Gesetzgebers widersprechen, die Entscheidung in der menschlichen Hand zu belassen.

Die erforderliche Kreativität in der Entscheidungsfindung setzt bereits technische Grenzen, unabhängig von den Grenzen des § 35a VwVfG. Es ist insoweit nicht empfehlenswert, Ermessensentscheidungen der

[44] Vgl. zur automatischen Entscheidungsfindung unter § 22 DschGVO eine im Mai 2022 vom Future of Privacy Forum (FPF) veröffentlichte Übersicht über praktische Fälle von Datenschutzbehörden und Gerichten in Europa https://fpf.org/wp-content/uploads/2022/05/FPF-ADM-Report-R2-singles.pdf (zuletzt aufgerufen am 08.02.2023), wonach u.a. die Pflicht zur Dokumentation eine zentrale Rolle spielt

[45] z.B. Stegmüller, Vollautomatische Verwaltungsakte – eine kritische Sicht auf die neuen § 24 I 3 und § 35aVwVfG (NVwZ 2018, 353)

KI zu überlassen. Auch nach den in Kap. 4 dargestellten allgemeinen Grundsätzen für *Responsible* AI sollte es grundsätzlich einen Vorrang menschlichen Handelns und menschlicher Aufsicht beim Einsatz von KI-Systemen geben.[46] Vor diesem Hintergrund sollte § 35a VwVfG unverändert bleiben.

Noch zu klären ist jedoch der Einsatz von KI für die Entscheidungsvorbereitung im Rahmen von Entscheidungen sowohl gemäß § 35 VwVfG bei gebundenen sowie Ermessensentscheidungen als auch in Fällen des vollständig automatisierten Erlasses eines Verwaltungsaktes gemäß § 35a VwVfG. So könnte KI im Rahmen der Entscheidungsvorbereitung beispielsweise bei der Sachverhaltsermittlung und der Feststellung des Vorliegens einzelner Tatbestandsmerkmale eingesetzt werden. Auch könnten KI-basiert Entscheidungsmöglichkeiten bzw. -vorschläge bei Ermessensentscheidungen generiert werden.

Eine Entscheidungsvorbereitung mit Hilfe von KI ist sinnvoll. In diesen Fällen sind allerdings in erster Linie die Dokumentationspflichten zu beachten. Werden die Anforderungen an die eingesetzte KI bzw. die Modalitäten ihres Einsatzes eingehalten, wie sie etwa den Anforderungen des KI-Acts bzw. den Grundsätzen von *Responsible AI* entsprechen, ist die KI zertifiziert, dürften keine durchgreifenden Bedenken bestehen. Es bietet sich indessen an, die konkreten Bedingungen des Einsatzes von KI in der Entscheidungsvorbereitung sowie die jeweils erforderlichen Dokumentationspflichten gesetzlich zu definieren. Auf diese Weise würde durch ein ergänzendes Instrument in der Entscheidungsvorbereitung zugleich eine Legitimation durch den Gesetzgeber verliehen und die Rechtssicherheit des Einsatzes erhöht (vgl. zu den Anforderungen der demokratischen Verwaltungslegitimation Kap. 5.2).

2) § 39 VwVfG

Dokumentationspflichten sind unabhängig von den Ebenen des Verwaltungshandelns zu beachten, nicht nur im Hinblick auf § 39 VwVfG (siehe u.a. Archivrecht und Aktenordnung.). Die Dauerhaftigkeit bzw. der Erhalt für die gesetzlichen Aufbewahrungsfristen muss technisch gewährleistet werden, denn digitale Speicher halten jeweils nur wenige Jahre. Dies dürfte jedoch nicht anders zu behandeln sein als im Falle der elektronischen Akte.

3) § 29 VwVfG

Weiterhin besteht Unklarheit in Bezug auf das Recht auf Akteneinsicht nach § 29 VwVfG. Die Gewährleistung des Rechtsanspruchs ist möglicherweise in Fällen der vollständigen Automatisierung von Verwaltungsverfahren regelhaft nicht umsetzbar. Eine Offenlegung des Algorithmus erscheint jedoch ebenfalls kein tauglicher Ersatz für das Recht auf Akteneinsicht. Bereits bei der Konzipierung von vollautomatisierten Verwaltungsverfahren ist daher zu berücksichtigen, dass alle Verfahrensschritte rekonstruierbar sein müssen, um dem Recht auf Akteneinsicht gerecht werden zu können.

[46] vgl. statt vieler: Ethics Guidelines for Trustworthy AI der von der EU Kommission eingesetzten unabhängigen High-Level Expert Group on Artificial Intelligenc vom 8. April 2019, S. 14

4) Resümee

Würden die Anforderungen an *Responsible AI* bzw. des KI-Acts eingehalten und können Zertifikate nach dem KI-Act vorgelegt werden, so erscheint die Anwendung der KI im Verwaltungsverfahren im Rahmen der Vorbereitung des Erlasses eines Verwaltungsaktes in den Fällen des § 35 VwVfG sowie in Fällen des § 35a VwVfG beim Erlass von Verwaltungsakten im automatischen Verfahren bei gebundenen Entscheidungen, soweit die tatbestandlichen Voraussetzungen des § 35a VwVfG vorliegen, grundsätzlich unproblematisch. In einzelnen Fallkonstellationen könnten sich allerdings besondere Fragen ergeben. Auch wird die weitere Entwicklung der Möglichkeiten des Einsatzes von KI zu beobachten sein, da sich daraus möglicherweise neue rechtliche Herausforderungen ergeben können, die entsprechende Anpassungen erforderlich machen könnten.

5.3.3 Ausgewählte interne Verwaltungsvorschriften

Der Einsatz von KI in der Verwaltung setzt auch eine Prüfung des Binnenrechts der Verwaltung voraus. Exemplarisch und beispielhaft wird auf einige Regelungen im Folgenden eingegangen:

1) Verwaltungsvorschrift zu § 74 LHO

In der von der Finanzbehörde erlassenen Verwaltungsvorschrift zu § 74 LHO vom 16. Dezember 2021, zuletzt geändert am 16. Dezember 2022 – anzuwenden ab Haushaltsjahr 2023[47] – werden in Abschnitt II Standards wie Risikoanalyse, IT-Sicherheit, Abgrenzung der Verantwortungsbereiche, Verfahrenszugriff, Richtigkeit und Vollständigkeit der erfassten und verarbeiteten elektronischen Daten und Dokumente, in Abschnitt III Stichprobenkontrollverfahren und in Abschnitt IV Zulassung von IT-Verfahren geregelt. Ausgenommen von der Zulassungspflicht nach dieser Verwaltungsvorschrift sind die IT-Verfahren, deren Einsatz für die Freie und Hansestadt Hamburg nach § 12 des Gesetzes über die Koordinierung der Entwicklung und des Einsatzes neuer Software der Steuerverwaltung (KONSENS-G) verpflichtend ist. In Abschnitt V zur IT-Sicherheit wird ausdrücklich festgelegt, dass die Freigaberichtlinie in der jeweils geltenden Fassung zu beachten ist.

Zusammenfassend lässt sich sagen, dass danach für die Steuerverwaltung für den Einsatz von IT schon jetzt viele Anforderungen gelten dürften, die Regelungsgegenstand des KI-Acts sein werden und auch den Anforderungen von *Responsible AI* in einer Reihe von Punkten entsprechen dürften.

[47] Siehe Website der FHH Hamburg (2022): Verwaltungsvorschriften (VV) zur Landeshaushaltsordnung, https://www.hamburg.de/fb/vv-zur-lho/ (zuletzt aufgerufen am 07.02.2023)

2) Freigaberichtlinie

Nach der Freigaberichtlinie[48] sollen gemäß 1.1. die beim Einsatz von Software im Interesse des Datenschutzes und der Datensicherheit zum Schutze von Software und Daten zu treffenden Maßnahmen geregelt werden, wobei weitergehende Vorschriften unberührt bleiben wie insbesondere für Software auch die Bestimmungen der Landeshaushaltsordnung und die dazu erlassenen Verwaltungsvorschriften. Gemäß 1.2. gilt die Richtlinie für jede im Produktionsbetrieb eingesetzte Software, unabhängig davon, ob sie auf Rechnern in Dienststellen der Freien und Hansestadt Hamburg oder in Einrichtungen von Dritten (Datenverarbeitung im Auftrag), z.B. eines zentralen IT-Dienstleisters, installiert ist. Gemäß 9. sind die betroffenen Dienststellen gehalten zu überprüfen, ob die im Produktionsbetrieb eingesetzte Software, DV-Verfahren und allgemeine Software den Anforderungen der Richtlinie entsprechen und treffen ggf. die erforderlichen Maßnahmen zur Anpassung an die Vorschriften der Freigaberichtlinie.

3) Technische Regelwerke

Bei der Entwicklung von Software, z.B. durch den IT-Dienstleister Dataport, sind die geltenden technischen Regelwerke zu beachten, die beispielsweise im Hinblick auf Sicherheit und Robustheit Standards vorgeben. Ein Überblick über bestehende, zu beachtende Regelwerke findet sich in Anlage 2 zu den Empfehlungen unter 7.10 „Technische Möglichkeiten zur Gewährleistung des gesellschaftlich verantwortlichen Einsatzes von KI“.

4) Resümee

Die Steuerung des KI-Einsatzes muss auch durch das Binnenrecht der Verwaltung erfolgen. Vor diesem Hintergrund empfiehlt es sich, in Hinblick auf die Grundsätze von *Responsible AI* und den KI-Act eine Evaluierung der bestehenden Regelungswerke vorzunehmen und ggfs. einer Überarbeitung bzw. Neufassung in einer oder mehrerer sich ergänzender Richtlinien bzw. Verwaltungsvorschriften durchzuführen. Auch kann der Erlass einer Rahmenrichtlinie in Betracht gezogen werden.

5.4 Hinreichende demokratische Verwaltungslegitimation: Braucht Hamburg ein KI-Rahmengesetz?

Die demokratische Legitimation der Verwaltung, die auf dem Demokratieprinzip und dem Rechtsstaatsprinzip basiert, stellt als verfassungsrechtliches Prinzip bestimmte Strukturanforderungen an die Organisation, die Verfahren und die Entscheidungsregeln der öffentlichen Verwaltung. Die konkreten normativen Anforderungen sind in der verfassungsrechtlichen Rechtsprechung und der rechtswissenschaftlichen

[48]Freigabe-RL vom 4. April 2005 (MittVw Seite 46) in der Fassung vom 18. November 2010 (MittVw Seite 189) http://daten.transparenz.hamburg.de/Dataport.HmbTG.ZS.Webservice.GetRessource100/GetRessource100.svc/507cfe0b-462d-404d-b8bb-217563e9056f/Akte_FB1a.805.01-2.pdf (zuletzt aufgerufen am 07.02.2023)

Literatur umfangreich beschrieben.[49] Durch die unterschiedlichen Mechanismen der Verwaltungslegitimation (u.a. sachliche Legitimation, personelle Legitimation, institutionelle Legitimation, Verfahrenslegitimation) soll eine zureichende Rückkopplung der Verwaltungsentscheidungen an den demokratischen Willen, das Parlament, hergestellt werden.

Vorbehaltlich einer intensiven verfassungsrechtlichen Prüfung kann von Folgendem ausgegangen werden: Das Modell der demokratischen Verwaltungslegitimation verfügt über eine hinreichende Flexibilität und verschiedene sog. „Legitimationsmodi", durch die auch in Sachgebieten, in denen die gesetzliche Steuerung der Verwaltungsträger nicht durch konditionale Entscheidungsprogramme gewährleistet ist, ein hinreichendes Legitimationsniveau erreicht werden kann (z.B. Planungsrecht, Regulierungsrecht). Demnach dürfte auch in Verwaltungssachgebieten, in denen KI oder maschinelles Lernen zum Einsatz kommen soll, ein hinreichendes Legitimationsniveau erreicht werden.

In allen Sachgebieten, in denen ausschließlich automatisierte Entscheidungssysteme zum Einsatz kommen, oder solche zur Unterstützung der Verwaltungsentscheidungen herangezogen werden (z.B. einfache konditionale Entscheidungsprogramme, die nicht auf maschinellem Lernen beruhen), besteht stets dann eine hinreichende Legitimation der Regelungsstruktur, wenn die Programme ein konditionales gesetzliches Entscheidungsprogramm exakt umsetzen und die Programmstrukturen nachvollzogen werden können. Dies ist bei „Wenn/dann-Entscheidungen" u.a. durch § 35a VwVfG normiert. Sobald jedoch Formen von KI oder maschinellem Lernen – gleich in welcher Form – in Sachgebieten mit gesetzlichen Entscheidungsprogrammen zur Anwendung kommen, die Entscheidungen vollständig vorbereiten, bei denen auf Tatbestandsseite Beurteilungsspielräume oder Ermessensspielräume auf der Rechtsfolgenseite bestehen, bedarf es ergänzender Mechanismen, die die demokratische Verwaltungslegitimation absichern:

Zur Vermittlung einer Basislegitimation des KI-Einsatzes kommt hierzu eine legitimierende Grundentscheidung des parlamentarischen Gesetzgebers in Betracht: Ein Rahmengesetz auf Bundes- und Landesebene, das den Einsatz von KI bzw. maschinellen Lernens in der Verwaltung gestattet und die normativen Grundanforderungen an die Systeme definiert. Der KI-Act dürfte für die verfassungsrechtliche Legitimation nicht ausreichen.

Neben dieser aller Voraussicht nach aus verfassungsrechtlicher Sicht erforderlichen demokratischen Legitimation der Anwendung von KI in den soeben genannten Fällen könnte für ein entsprechendes Rahmengesetz mit der Festschreibung von zu beachtenden Eckpunkten für den Einsatz von KI sprechen, dass in einer Art normativer Selbstverpflichtung diese Grundsätze auch dann zu beachten sind, wenn die eingesetzten Algorithmen nicht unter den KI-Act fallen werden und technisch nicht besonders anspruchsvoll sind (einfache Softwareanwendungen), bei denen es aber beim Auftreten von Fehlern zu massiven Rechtsverletzungen kommen kann, wie z.B. bei dem oben dargestellten niederländischen „Kindergeldfall". Derartige Fehler hätten durch Beachtung der Grundsätze von *Responsible AI* vermieden werden können.

Im Übrigen würde dies auch unterstreichen, welche Bedeutung der (Hamburger) Gesetzgeber und damit auch die Verwaltung dem gesellschaftlich verantwortlichen Einsatz von KI beimisst, die durch den KI-Act

[49] vgl. statt vieler: Trute, Die demokratische Verwaltungslegitimation, in: Voßkuhle/Eifert/Möllers, Grundlagen des Verwaltungsrechts, 2022

nicht geleistet werden kann, da eine Rückkopplung an die nach dem Grundgesetz maßgeblichen Legitimationssubjekte (Bundes- und Landesparlamente) fehlt.

5.5 Verwaltungsgerichtlicher Rechtsschutz

Dem Gebot effektiven Rechtsschutzes (Art. 19 Abs. 4 GG) ist grundsätzlich hinreichend Rechnung getragen.

Durch den Amtsermittlungsgrundsatz (§ 86 VwGO) ergeben sich grundsätzlich keine abweichenden Fragestellungen gegenüber dem Verwaltungsverfahren. Aus der Perspektive verwaltungsgerichtlicher Kontrolle dürfte der Einsatz von KI im Verwaltungsverfahren dort grundsätzlich unproblematisch sein, wo die mit KI-Unterstützung getroffene Entscheidung auch ohne den Einsatz von KI verifizierbar ist. Steht dabei das Vorliegen einer gerichtlich voll überprüfbaren Tatbestandsvoraussetzung („Ist der Hund A ein Hund der Rasse X?“) im Streit, ist das Gericht nicht an den von der Verwaltung gewählten – hier: KI-unterstützten – Weg zur Sachverhaltsermittlung gebunden, sondern kann (und muss ggf.) eine alternative Methode zur (Auf-)Klärung des Sachverhalts und eigenen Überzeugungsbildung wählen (z.B. Sachverständigengutachten „ohne KI“). Handelt es sich um eine gebundene Verwaltungsentscheidung, kommt es ohnehin „nur“ auf die – nach Maßgabe der heranzuziehenden Norm – Richtigkeit des Entscheidungsergebnisses an, d.h. selbst wenn nachgewiesen würde, dass eine „falsch trainierte“ oder sonst nicht den Anforderungen des KI-Acts genügende KI zum Einsatz kam, diese aber gleichwohl – zufällig – das Tatbestandsmerkmal in der Sache zutreffend bejaht hat, wäre eine darauf gestützte Verwaltungsentscheidung nicht aufzuheben.

Schwieriger zu beurteilen wären Fälle, in denen Verwaltungsentscheidungen – insbesondere Maßnahmen der Gefahrenabwehr – auf KI-unterstützten Prognosen beruhen. Maßgeblich für die – auch nachträgliche – Beurteilung der Rechtmäßigkeit von Prognoseentscheidungen ist in der Regel die ex ante-Sicht der Entscheider:innen. So wird z.B. im Gefahrenabwehrrecht eine – bestimmte staatliche Maßnahmen eröffnende – „konkrete Gefahr“ dann angenommen, wenn im Zeitpunkt der Entscheidung nach objektiven Gesichtspunkten eine Sachlage vorliegt bzw. vorlag, bei der im Einzelfall die hinreichende Wahrscheinlichkeit besteht bzw. bestand, dass in absehbarer Zeit ein Schaden für ein geschütztes Rechtsgut eintreten wird. Würde eine solche Einschätzung der Sachlage – möglicherweise allein – durch KI getroffen, würde sich im Rahmen der verwaltungsgerichtlichen Kontrolle einer darauf gestützten (menschlichen) Entscheidung die Frage der Nachvollziehbarkeit der Einschätzung stellen. Hier dürfte es bei einem Einsatz von KI – auch abhängig von der Bedeutung des Rechtsguts, in das durch die Entscheidung eingegriffen wird, sowie der Höhe bzw. des Ausmaßes des drohenden Schadens im Fall des Nichteinschreitens – maßgeblich darauf ankommen, dass die o.g. Anforderungen an die eingesetzte KI erfüllt sind und überdies plausibel dargelegt werden kann, dass die von der KI getroffene Einschätzung objektiv mindestens „gleich gut oder besser“ als die menschlich getroffene ist. Darüber hinaus erscheint denkbar, dass aufgrund objektiver Wertentscheidungen des Grundgesetzes (oder supranationaler Rechtsquellen) in bestimmten Lebensbereichen oder in Bezug auf bestimmte Eigenschaften / Merkmale („Hautfarbe als Trainingsmerkmal für Risikoeinschätzungen“) ein Einsatz von KI vollständig ausgeschlossen ist. Solche Konstellationen dürften aber überwiegend ohnehin in den Anforderungen des KI-Acts abgedeckt sein. Werden die Anforderungen des KI-Acts an die Herstellung einer KI bzw. an deren Einsatz wie z.B. Transparenz, Robustheit, Sicherheit, Datenqualität, Dokumentation Rechnung getragen, dürften im Regelfall keine durchgreifenden Probleme für

die Gewährleistung ausreichenden Rechtsschutzes bestehen. Sollten im Einzelfall Probleme bei der Überprüfbarkeit auftreten, dürften diese mit den klassischen Beweislastregelungen zu beherrschen sein.

Es wird allerdings zukünftig zu beobachten sein, ob im Falle ihres Einsatzes bei bestimmten Anwendungen wie selbstlernenden Systemen und neuronalen Netzwerken etwa in Zusammenhang mit der sog. „Black Box", das heißt der fehlenden Nachvollziehbarkeit von Entscheidungen, besondere Probleme auftreten, die im Hinblick auf die Gewährleistung effektiven Rechtsschutzes weiterer Überlegungen bedürfen bzw. für die Verwaltung weitere Anforderungen an den Nachweis einer Nachvollziehbarkeit und Ergebnisstabilität mit sich bringen, um so effektiven Rechtsschutz dem Grunde noch gewährleisten zu können.

Eine Herausforderung für die Gewährleistung effektiven Rechtsschutzes sind darüber hinaus die zunehmenden Einsatzmöglichkeiten von sogenannten Legal Tech Anwendungen auf Seiten der Kläger:innen bzw. der sie vertretenden Anwaltskanzleien[50]. Der Einsatz entsprechender Instrumente etwa in Bezug auf Sachverhaltsermittlung und -auswertung, Beschleunigung von Verwaltungsabläufen, Auswertung von Schriftsätzen und Dokumenten steckt in der Verwaltung noch in den Kinderschuhen, dürfte aber über kurz oder lang erfolgen. Ob die Verwaltungsgerichtsbarkeit hier hinreichend technisch aufgestellt ist, welche Anforderungen nach dem KI-Act beim Einsatz entsprechender Anwendungen in der Justiz einzuhalten sind bzw. ob eine gesetzliche Grundlage erforderlich wäre, ist allerdings nicht Gegenstand dieser Studie.

6. Grundlagen für die Handlungsempfehlungen

Als Grundlage für die Empfehlungen werden im Folgenden die bisherigen Ergebnisse der Studie zusammengefasst (6.1), die Herausforderung und Chance einer gesellschaftlich verantwortlichen Digitalisierung betont (6.2) sowie die zentralen Leitlinien für die Empfehlungen benannt (6.3).

6.1 Bisherige Ergebnisse

6.1.1 Responsible AI und KI-Act

Responsible AI basiert im Wesentlichen auf den Menschenrechten, der Europäischen Grundrechtecharta und dem Grundgesetz. Hieraus leiten sich die Grundätze eines gesellschaftlich verantwortlichen Einsatzes von KI ab.

Einen normativen und damit rechtsverbindlichen Ausdruck finden einen großen Teil dieser Grundsätze im sogenannten KI-Act, einer EU-Verordnung, die sich noch im Abstimmungsverfahren zwischen Kommission, Rat und Europäischem Parlament befindet und nach ihrer Verabschiedung unmittelbar geltendes Recht wird. Die Einzelheiten der Regelungen sind bis zu ihrer endgültigen Verabschiedung im Fluss. In der Verordnung werden im Einzelnen die Adressaten, konkret einzuhaltende Anforderungen einschließlich der zu beachtenden Instrumente und weitere Maßnahmen festgelegt (vgl. die obigen Ausführungen).

[50] Vgl. zur Begriffsdefinition und Darstellung der Möglichkeiten für Kanzleien LEGAL-TECH.DE, Was ist Legal-Tech?, https://legal-tech.de/was-ist-legal-tech-ffi (zuletzt aufgerufen am 07.02.2023)

6.1.2 Notwendigkeit

Als Verordnung wird der KI-Act nach einer festgelegten Übergangsfrist von zwei Jahren nach Inkrafttreten unmittelbare rechtliche Wirkung in den Mitgliedstaaten auch für die Verwaltung entfalten. In diesem Kontext werden die nunmehr als Entwürfe vorgelegten Haftungsrichtlinien mit ins Auge gefasst werden müssen, die entsprechend den obigen Ausführungen in unmittelbarem Zusammenhang mit der KI-Verordnung zu sehen sind und auch für die Verwaltung nach entsprechender gesetzlicher Umsetzung gelten werden.

Unabhängig von zurzeit bestehenden Bestrebungen, im weiteren Normierungsprozess den KI-Begriff in Art. 3 KI-Act entgegen der ursprünglichen von der Kommission am 21. April 2021 vorgelegten Fassung enger zu fassen, empfehlen wir auch den Einsatz von Software unter Beachtung der maßgeblichen Prinzipien der *Responsible AI* zu behandeln, auch wenn sie nicht unter die Regelungen des KI-Acts fallen würden. So zeigen die oben erwähnten Beispiele aus den Niederlanden, Österreich oder Italien, dass selbst scheinbar einfache Algorithmen bzw. Software Biases oder andere Fehler enthalten können und insbesondere in Abhängigkeit vom Anwendungskontext zu gesellschaftlich nicht zu verantwortenden Folgen führen können. Dies kann je nach Anwendungskontext auch in anderen Fällen gelten, bei denen eine fehlerhafte Software zu gravierenden Folgen führen kann. Beispielhaft ist eine digitale Erfassung von Brücken im Hinblick auf ihren baulichen Zustand und entsprechender Überprüfungszeitpunkte zu nennen – hier kann eine fehlerhafte Datenlage zu einer fehlerhaften Festlegung eines zu späten Überprüfungszeitpunktes und dadurch beispielsweise zu einem Einsturz des Bauwerks führen mit der möglichen Folge erheblicher Schäden an Leben und Gesundheit von deren Benutzer:innen.

Auch wenn zurzeit in der Hamburger Verwaltung vorwiegend einfache Formen von KI zur Anwendung kommen und in weiten Bereichen der Schwerpunkt auf sogenannte Robotik gelegt wird, wird davon auszugehen sein, dass schon in absehbarer Zeit der Anwendungsbereich von Künstlicher Intelligenz und maschinellem Lernen einen immer weiteren Raum in der Verwaltung einnehmen wird. Unter anderem wird generative KI zu immer weiteren Einsatzmöglichkeiten in der Verwaltung führen, wie z.B. Chat-Programme, wie die gerade als open source freigeschaltete ChatGPT, die die Möglichkeiten der künstlichen Generierung insbesondere von Texten eröffnen. Umso wichtiger ist es, dass sich die Verwaltung schon jetzt entsprechend aufstellt, um den Anforderungen eines gesellschaftlichen verantwortlichen, ethischen Einsatzes von Algorithmen und Künstlicher Intelligenz zu entsprechen und nach den normativen Vorgaben des KI-Acts zu verfahren.

Je früher die Verwaltung sich auf diese Anforderungen bei der Digitalisierung einstellt, desto einfacher wird der Übergang beim Inkrafttreten des KI-Acts sein. Wenn die Verwaltung unvorbereitet ist, dürfte dies nachteilige Auswirkungen auf die weitere Digitalisierung haben, da sich dann „plötzlich“ rechtliche und gegebenenfalls ethische Hindernisse auftun, die die Entwicklung eigener und die Anwendung neuer digitaler Lösungen verzögern oder gar verhindern kann, weil den Anforderungen des KI-Acts nicht entsprochen worden ist. Hierdurch können u.a. finanzielle Einbußen, Einschränkungen von Verwaltungsdienstleistungen oder Verlust von Vertrauen in das Verwaltungshandeln die Folge sein. Und es kann zur nach dem KI-Act möglichen Verhängung von Bußgeldern kommen.

6.2 Herausforderung und Chance zugleich

Die Digitalisierung der Hamburger Verwaltung ist eines der großen Vorhaben des Senats. Daraus resultierende Anstrengungen haben Hamburg im Bundesvergleich in den vergangenen Jahren die Belegung von Spitzenplätzen beschert.[51] Die Digitalstrategie für Hamburg macht deutlich, dass es sich um eine bedeutende Daueraufgabe für die Zukunft und den Standort handelt.

Dabei den Anforderungen von *Responsible AI* und des KI-Acts mit seinen vielfältigen zu berücksichtigenden Vorgaben (vgl. hierzu die obigen Ausführungen) zu entsprechen und diese in die bestehenden Verwaltungsstrukturen zu implementieren, stellt eine neue Herausforderung für die Hamburger Verwaltung dar.

Dies ist aber auch Chance zugleich: Digitalisierung und *Responsible AI* sollten untrennbar verbunden sein. Sie sind zwei Seiten einer Medaille. Der verantwortliche Einsatz von KI ist mit Inkrafttreten des KI-Acts als Verordnung nicht nur normatives Gebot, es ist vielmehr aus gesellschaftlicher Verantwortung heraus geboten. Digitalisierung der Verwaltung führt zu besseren Dienstleistungen durch den Staat und fördert damit die Akzeptanz staatlichen Handelns. Die Gewährleistung des Einsatzes nach den Grundsätzen gesellschaftlicher Verantwortung über die Umsetzung des KI-Acts hinaus fördert daher zugleich das Vertrauen der Bürger:innen in die Digitalisierung und das Verwaltungshandeln insgesamt.

Dementsprechend sollte Motto und Leitlinie für Hamburg sein: Die gesellschaftlich verantwortliche Digitalisierung der Hamburger Verwaltung! Hamburg würde damit sehr bewusst auch eine bundesweite Vorreiterrolle einnehmen.

6.3 Drei Säulen

Vor dem Hintergrund der bisherigen Ergebnisse der Studie ergeben sich drei Säulen, die Leitlinie und Orientierung für die folgenden Empfehlungen sind:

- Bewusstsein wecken für die Notwendigkeit, sich frühzeitig auf die Herausforderungen von *Responsible AI* und KI-Act einzustellen (Awareness)
- Rechtzeitig die Umsetzung der Anforderungen nach den Grundsätzen von *Responsible AI* und dem KI-Act zu steuern und zu regulieren (Governance)
- Kommunikation sowohl verwaltungsintern als auch in die und mit der Öffentlichkeit gewährleisten (Communication)

[51] Vgl. u.a. Hamburg News, Studie (11.05.2022): Hamburg löst Berlin an der Spitze der Digitalisierung ab, https://www.hamburg-news.hamburg/standort/studie-hamburg-loest-berlin-der-spitze-der-digitalisierung-ab (zuletzt aufgerufen 08.02.2023) *sowie* Presseinformation Deutschlands smarteste Städte: Hamburg gewinnt knapp vor München, Dresden erstmals auf dem Treppchen, Verfolger holen auf, 20.09.2022, https://www.bitkom.org/Presse/Presseinformation/Deutschlands-smarteste-Staedte-2022 (zuletzt aufgerufen 08.02.2023)

7. Empfehlungen

Auf der Basis der Ergebnisse der Studie, den mit Vertreter:innen einzelner Behörden sowie dem ITD geführten Einzelgesprächen sowie den Anregungen aus den beiden durchgeführten Workshops werden für die Hamburger Verwaltung folgende Handlungsempfehlungen gegeben:

7.1 Digitaler Flyer – Information nach Innen

Es wird empfohlen, einen einheitlichen Informationsstand über die Digitalisierung der Hamburger Verwaltung bei den Mitarbeiter:innen herzustellen. Dies gilt insbesondere für das tragende Leitmotiv, dass die Digitalisierung mehr als ein technischer Vorgang ist: Es geht um die Verbesserung der Dienstleistungen für die Bürger:innen in einer gesellschaftlich verantwortlichen Art und Weise. Hierfür bietet sich in einem ersten Schritt ein digitaler Flyer, ggf. ergänzt durch eine digitalen Informationsbroschüre an.

Die Digitalisierung der Verwaltung wird zunehmend nahezu alle Bereiche des Verwaltungshandelns erfassen. Umso wichtiger ist es, dass frühzeitig auch die Mitarbeiter:innen, die bisher nicht unmittelbar von der Digitalisierung betroffen sind, damit vertraut gemacht werden, dass die Digitalisierung weiter voranschreiten wird, warum und in welcher Art und Weise sie umgesetzt wird oder was das Leitmotiv des Hamburger Weges der Digitalisierung ist. Dies hilft bestehende Widerstände abzubauen bzw. das Entstehen von möglichen Widerständen frühzeitig zu verhindern und kann geeignet sein, eine bessere Identifikation mit dem Vorhaben zu fördern.

Es bietet sich an, grundlegende Informationen über Begrifflichkeiten und Wirkungsweisen etwa in Form eines Glossars zu erstellen, um so ein einheitliches Verständnis und Verstehen der Mitarbeiter:innen zu erreichen. Sowohl als Ergebnis der sogenannten Desk Research als auch der von uns geführten Gespräche lässt sich festhalten, dass ein unterschiedlicher Wissensstand etwa in Bezug auf Begrifflichkeiten wie Algorithmen, Künstliche Intelligenz oder selbstlernende Systeme als ein Hindernis für die Umsetzung der Digitalstrategie gesehen wird. Daneben kann erläutert werden, was unter *Responsible AI* zu verstehen ist, was der KI-Act und u.a. die haftungsrechtlichen Vorschriften bedeuten.

In einer solchen ersten digitalen Information kann darüber hinaus auf entsprechende Angebote zur Wissensvermittlung und -vertiefung, zur Weiterqualifizierung etc. verwiesen werden, wie sie im Folgenden benannt werden.

In einem späteren Stadium können diese Informationen neben anderen in eine noch zu entwickelnde digitale Plattform (vgl. unten ad. 7.6) aufgenommen werden.

7.2 Aus- und Fortbildung

Neben schon bestehenden Fortbildungsangeboten etwa im ZAF bzw. in Abstimmung mit diesem ist ein Fortbildungsangebot (weiter) zu entwickeln. Ziel ist u.a. sowohl die Vermittlung der nötigen technischen Grundlagen und Schlüsselqualifikationen als auch der einschlägigen rechtlichen, regulatorischen und ethischen Aspekte für den Umgang mit KI. Es empfiehlt sich die Entwicklung eines modularen, interdisziplinären Fortbildungskonzeptes. Besonderes Augenmerk ist dabei auf die jeweiligen Zielgruppen zu richten, um einerseits den spezifischen Bedürfnissen einzelner Behörden bzw. Fachämtern Rechnung zu tragen, andererseits ein einheitliches Verständnis zu erreichen und um darüber hinaus unterschiedliche Perspektiven einzubringen.

Ein guter Ansatzpunkt für die Weiterqualifizierung von Mitarbeiter:innen bietet das an der Hochschule für angewandte Wissenschaften (HAW) eingerichtete Duale Bachelor-Studium E-Government/Verwaltungsinformatik unter Beteiligung der vier Fachdisziplinen Informatik, Wirtschaftswissenschaften, Rechtswissenschaften und Sozialwissenschaften sowie Praxisphasen in der Hamburger Verwaltung über sieben Semester mit jährlich 36 Studienplätzen.[52] Es könnte sich anbieten, in enger Abstimmung mit der Hochschule bedarfsgerechte Studieninhalte in dem vorliegend relevanten Bereich zu definieren und an für die Verwaltung wichtigen Fragestellungen oder der Konkretisierung von Themenstellungen für Bachelorarbeiten mitzuwirken, die sich beispielsweise mit in der Studie angesprochenen Fragen vertiefend beschäftigen. Daneben kann als weiterer Anknüpfungspunkt der an der HAW schon laufende Bachelor-Studiengang Public Management (PUMA)[53] dienen, der Grundlage für die bundesweit anerkannte Laufbahnbefähigung für den gehobenen allgemeinen Verwaltungsdienst ist. Das Studium besteht aus vier fachtheoretischen Semestern, die an der HAW Hamburg durchgeführt werden, und zwei berufspraktischen Semestern, die in den Behörden und Ämtern der Freien und Hansestadt Hamburg absolviert werden. Theoretische und

[52] Siehe Website der HAW Hamburg (2023): E Government Dual , https://www.haw-hamburg.de/studium/studiengaenge-a-z/studiengaenge-detail/course/courses/show/e-government-dual/Studieninteressierte/ (zuletzt aufgerufen am 28.02.2023)

[53] Siehe Website der HAW Hamburg (2023): Public Management (Dual), https://www.haw-hamburg.de/studium/studiengaenge-a-z/studiengaenge-detail/course/courses/show/public-management-dual/Studieninteressierte/ (zuletzt aufgerufen am 28.02.2023)

praktische Studieninhalte sind eng miteinander verzahnt. In diesem Studiengang werden Grundkenntnisse in der Informationstechnologie als auch in den Fächern Staats- und Europarecht, Allgemeines Verwaltungsrecht, Rechtsmethodik, Zivilrecht, Volkswirtschaftslehre und öffentliche Finanzwirtschaft, Betriebswirtschaftslehre der öffentlichen Verwaltung, Soziologie, Politologie und Psychologie vermittelt. Auch hier könnte es sich beispielsweise anbieten, den Modulplan entsprechend anzupassen oder beispielsweise Bachelorarbeiten zu fachübergreifenden Themenstellungen zu vergeben, für die in der Verwaltungspraxis Bedarf besteht.

7.3 Workshops

Neben Fortbildungsveranstaltungen bietet sich die Durchführung von Workshops an, die sich speziell an die Leitungsebene (Staatsrät:innen, Amtsleitungen) und Abteilungsleiter:innen der jeweiligen Behörden richten. Maßgeblich für den Erfolg der Digitalisierung ist, dass sie von der Spitze getragen, initiiert und umgesetzt wird. Hierzu bedarf es neben dem in Fortbildungsveranstaltungen vermittelten Wissen einer vertieften Sensibilisierung der Notwendigkeit, bei der Digitalisierung von Verwaltungsleistungen auch rechtliche und ethische Aspekte mit einzubeziehen. Neben Awareness geht es dabei auch und gerade um Governance, wie der Veränderungsprozess unter Einbeziehung der Mitarbeiter:innen am effektivsten und effizientesten umgesetzt werden kann.

Diese Workshops sollten unter maßgeblicher Mitwirkung des ITD als Querschnittseinrichtung inhaltlich näher konkretisiert und durchgeführt werden. So kann gewährleistet werden, dass es zu einer gemeinsamen „Handschrift“ der verschiedensten Behörden bzw. Einrichtungen kommt. Zu klären ist auch, ob die Workshops behördenübergreifend oder für die einzelnen Behörden separat durchgeführt werden. Auf eine interdisziplinäre Zusammensetzung ist zu achten, es bedarf neben organisatorischer auch technischer und juristischer Kompetenz.

7.4 Checkliste

Die Erstellung einer Checkliste, die die Adressat:innen des KI-Acts, also insbesondere Entwickler:innen, Einkäufer:innen, Nutzer:innen bzw. Fortentwickler:innen konkrete Hilfestellung gibt, ist für zukünftige Arbeitsprozesse unerlässlich. Sie soll anhand eines Prüfprogramms in logischer Reihenfolge der Regelungen des KI-Acts ohne große Recherche auf einen Blick die Möglichkeit geben, zu ermitteln, welche Anforderungen je nach Anwendung und Einsatzgebiet zu befolgen sind; u.a. geht es um die Identifizierung, ob die jeweilige Anwendung unter den KI-Act fällt und wenn ja, welcher Risikogruppe sie zuzuordnen ist. Je nach Risikogruppen und Adressat, also Entwickler:in, Anwender:in, Nutzer:in etc., werden im KI-Act unterschiedliche Anforderungen normiert. Die Checkliste hilft, den Grad der Betroffenheit und die damit korrelierenden einschlägigen Anforderungen zu identifizieren und dementsprechend das weitere Handeln daran zu orientieren bzw. in die eigene Entscheidungsfindung einzupreisen zu können. In diesem Sinne handelt es sich um ein Werkzeug, durch den KI-Act geführt zu werden. Es hilft die richtigen Fragen und Weichen zu stellen, bei unklaren Sachverhalten bzw. Weichenstellungen kritische Punkte zu identifizieren und zu deren Beantwortung ggfs. weiteren technischen und/oder juristischen Sachverstand zielgerichtet hinzuziehen zu können. Eine exemplarische Version der Checkliste findet sich im Anhang dieser Studie. Da sich der KI-Act Entwurf vom 21. April 2021 zum für die Erstellung der Studie maßgeblichen Zeitpunkt noch im Abstimmungsprozess zwischen Kommission, Rat und Parlament befand (und eine Verabschiedung zurzeit nicht absehbar ist), handelt es sich im Anhang um eine in diesem Sinne vorläufige Darstellung einer

möglichen Ausgestaltung einer Checkliste auf der Basis des KI-Acts Entwurfes vom 21. April 2021. Es empfiehlt sich, die entsprechend fortzuentwickelnde Checkliste in eine verwaltungsinterne digitale Plattform zu integrieren (vgl. unten unter 7.6). Auch wäre zu prüfen, ob eine entsprechende Liste in Hinblick auf die zu beachtenden Grundsätze von *Responsible AI* bei nicht unter den KI-Act fallenden KI Anwendungen aufzustellen ist.

7.5 Monitoring

Vor dem Hintergrund

- dass die Verwaltung jederzeit einen umfassenden Informationsstand über die verschiedenen den Einsatz von KI im privaten und öffentlichen Bereich betreffenden Regelungen auf EU- Bundes- und Landesebene braucht,
- des Umstandes, dass noch vieles im Fluss ist (insbesondere in Bezug auf die konkreten Regelungen und Definitionen im KI-Act) und
- der Notwendigkeit rechtzeitiger Sensibilisierung bzw. Information im Hinblick auf eigene Handlungsbedarfe (rechtzeitige Stellungnahmen, Vorbereitung der Umsetzung, Regelungen der Zuständigkeiten der Notifizierungsstellen auf EU-, Bundes- und Landesebene),

wird die Einrichtung einer, bzw. soweit es schon geeignete Strukturen gibt, der Ausbau zu einer sogenannten Monitoring-Stelle empfohlen. Eine frühzeitige Information schafft die Voraussetzungen, um – je nach Betroffenheit – rechtzeitig die notwendigen Maßnahmen ergreifen zu können, sei es, in den Normgebungsprozess noch eingreifen zu wollen, sei es, sich rechtzeitig auf etwaige neue Anforderungen einstellen zu können. Gerade im Bereich KI ist davon auszugehen, dass es in den nächsten Jahren eine exponentielle Entwicklung geben wird, dass die im KI-Act vorgesehenen Anforderungen angepasst bzw. neue Verordnungen und Richtlinien erlassen werden müssen, um sich stets den „neuen" Realitäten anzupassen.

Hier kann auch ein Zusammenhang mit der Online-Plattform hergestellt werden, da im Rahmen des Monitorings gewonnene Erkenntnisse dort als fortlaufende Aktualisierung einfließen könnten.

7.6 Online-Plattform

Es wird die Einrichtung einer Online-Plattform empfohlen, um so einen unbürokratischen, schnellen Zugriff auf einschlägige Informationen zu ermöglichen. Elemente könnten ohne Anspruch auf Vollständigkeit sein:

- Gewährleistung eines einheitlichen Informationsstands über die Digitalisierung, den Einsatz von KI, Begrifflichkeiten etc.
- Darstellung des Selbstverständnisses der Verwaltung, die Digitalisierung gesellschaftlich verantwortlich durchzuführen, insbesondere entlang des tragenden Leitmotivs, dass die Digitalisierung mehr als ein nur technischer Vorgang ist
- Darstellung allgemeiner Informationen zum KI-Act und anderen relevanten Vorschriften einschließlich des Binnenrechts (Verwaltungsvorschriften, Richtlinien) in jeweils aktueller Form, durch entsprechende Präsentation und Menüführung schnell und leicht zugänglich

- digitale Checkliste als zentraler Bestandteil (vgl. oben unter 7.4), mit vertiefenden Informationen wie zusätzliche Erläuterung von Begrifflichkeiten aus dem KI-Act, die für die Umsetzung relevant sind (z.B. Art 9 EU KI-Act Entwurf Riskmanagement)
- gebündelte Darstellung der verschiedensten Aktivitäten und Ermöglichung einer besseren Koordination der Initiativen in der Hamburger Verwaltung bzw. Stimulierung von Initiativen, Ermöglichung von Synergieeffekten sowie Benennung von Ansprechpartner:innen
- Darstellung und Aktualisierung der im Rahmen des Monitorings gewonnenen Informationen, soweit sie für eine digitale Informationsplattform und ihren Adressatenkreis von Belang sind.

7.7 Reallabore

In Abgrenzung zu den sogenannten Hubs oder Sandboxes, bei denen im Vordergrund die Entwicklung und Erprobung von KI-Lösungen steht – so z.B. die unter dem Dach des GovTech Campus Hessen angebundene Initiative im Rahmen des Innovation Hub 110 mit dem Schwerpunkt auf digitalen Innovationen im Sicherheitsbereich und bei der Polizei – geht es bei dem sog. Reallabor um die Erprobung von Arbeitsweisen zur möglichst effektiven und effizienten Implementierung von KI in die Arbeit der Verwaltung. Es geht darum, angefangen bei der Bedarfsanalyse, der Auswahl der geeigneten digitalen Lösung und deren Umsetzung, die richtige Struktur eines Entscheidungs- und Umsetzungsprozesses zu entwickeln, zu erproben und zu evaluieren. Neben den anzustellenden Erwägungen bei der Identifizierung eines möglichen Bedarfes für eine Digitalisierung bis hin zur Auswahl der passenden Lösung (Technik, Kosten, Erreichung des gewünschten Ziels, Auswirkungen auf die Ablauforganisation etc.) sollten sowohl von Anfang an als auch prozessbegleitend die Rechtsrisiken bzw. -anforderungen insbesondere auch im Hinblick auf *Responsible AI* und den KI-Act mit in die Entscheidungsfindung und die anschließende Realisierung einbezogen werden. In diesem Sinn bedarf es in Bezug auf *Responsible AI* und den KI-Act eines proaktiven Handelns: Weg von einem Art Säulenmodell des Nebeneinanderher hin zu einem ganzeinheitlichen, interaktiven, interdisziplinären Prozess.

Es könnte sich anbieten, exemplarisch ein oder mehrere Anwendungsfelder zu identifizieren, um eine entsprechende Arbeitsform (fort) zu entwickeln sowie eine weitere Konkretisierung der Idee und ihrer Funktion vorzunehmen und die richtigen Arbeitsweisen zu erproben. Ein solches „Reallabor“ könnte dazu dienen, eine Art Print zu entwickeln, nach dem auch andere Behörden bzw. Stellen arbeiten könnten, um so bessere Rückschlüsse auf die richtige Anwendung, Organisation etc. herleiten zu können. So können wichtige Impulse für Change Management und agile Arbeitsformen identifiziert werden.

Der beim Landesbetrieb Straßen, Brücken und Gewässer angesiedelte Fachbereich DS2 – Digi-Lab könnte sich für einen Probelauf anbieten, da schon jetzt interdisziplinär bei der Entwicklung gearbeitet wird. Hier muss die juristische Perspektive von Anfang an mit einbezogen werden.

7.8 Organisatorische Vorschläge und Anregungen

Grundsätzlich ist Digitalisierung „Chef:innensache“, in erster Linie angesiedelt beim in der Senatskanzlei verorteten ITD. Hinzu kommen die Staatsrät:innen bzw. Amtsleiter:innen und Abteilungsleiter:innen, die CDOs in den jeweiligen Behörden bzw. Landesbetrieben und für die Bezirksämter sowie daneben ggfs. sonst Verantwortliche für einzelne Digitalisierungsprojekte. Die strategische Verantwortung liegt beim ITD; die Umsetzung sollte entsprechend der jeweiligen fachlichen Anforderungen und Bedarfe auf der Fachebene und in den Fachbehörden stattfinden.

Hierzu gehört inhaltlich entsprechend unserem Vorschlag untrennbar verbunden auch die Verantwortlichkeit für die Umsetzung des KI-Acts bzw. der Gewährleistung von *Responsible AI.* In erster Linie sollte hier das ITD in der Pflicht sein. Das ITD sollte hierfür ein Kompetenzzentrum einrichten bzw. vorhandene Strukturen entsprechend ausbauen, das sowohl für Beratungen zur Verfügung steht als auch, gegebenenfalls durch eine sogenannte Task Force, Digitalisierungsprojekte und insbesondere den Einsatz von KI-systemen in den einzelnen Behörden auch rechtlich proaktiv begleiten kann. Es erscheint sinnvoll, die entsprechende technische bzw. digitale Expertise von vornherein mit in ein solches Kompetenzzentrum zu implementieren. Für sogenannte KI Audits bedarf es genügender fachlicher Expertise.

Bei unseren Gesprächspartner:innen aus diversen Hamburger Behörden bzw. Landesbetrieben bestand die Überzeugung, dass für solche befürworteten Dienstleistungen grundsätzlich das ITD als Querschnittseinrichtung die zuständige Stelle sein sollte. In diesem Zusammenhang wurde des Öfteren hervorgehoben, dass die jeweiligen Rechtsämter der Fachbehörden personell sowohl in quantitativer als auch in Bezug auf das hier einschlägige Thema der *Responsible AI* bzw. KI-Act fachlich nicht ausreichend ausgestattet seien. Zu überlegen ist, ob nicht größere Behörden, bei denen die Digitalisierung schon jetzt eine große Rolle spielt, insoweit selbst mit entsprechender juristischer und der notwendigen technischen Expertise ausgestattet sein sollten.

Eine bedeutende Rolle dürfte als Teil des ITD GovTechHH[54] zukommen. Hamburg möchte mit GovTechHH die Zusammenarbeit mit Start-ups ausbauen. Nunmehr übernimmt GovTechHH die Koordination des regionalen GovTech Campus Hansestadt Hamburg als Teil des GovTech Campus Deutschland.[55] Die in dem Projekt angelegte Kooperationen mit der Wissenschaft, mit Universitäten und Hochschulen oder kanalisierenden Organisationen wie dem Artificial Intelligence Center Hamburg (ARIC) e.V. sollte dringend umgesetzt bzw. ausgebaut werden. Dies empfiehlt sich auch aufgrund des sehr schnellen Forschungsfortschritts im Bereich KI und den immer kürzer werdenden Vermarktungszyklen zwischen Forschung und Wirtschaft, um rechtzeitig Trends zu erkennen, aufzugreifen und ggfs. bei Bedarf auch im Interesse der Verwaltung in die Umsetzung zu bringen. In diesem Kontext kann es sich auch anbieten, zielgerichtete Förderprogramme für Start-ups z.B. in Zusammenarbeit mit der Finanzbehörde über die IFB zu initiieren, die für die Verwaltung aber auch darüber hinaus von Interesse sein könnten.[56]

7.9 Zuständigkeiten für Bewertungsstellen

Der Hamburger Senat sollte sich frühzeitig und fortlaufend um die sich aus dem KI-Act ergebenden Notwendigkeiten und Möglichkeiten der Errichtung von Einrichtungen wie Notifizierungsstellen auf Bundes- und Landesebene kümmern, um so die Rolle Hamburgs zu stärken oder aus Hamburger Sicht Fehlentwicklungen rechtzeitig begegnen zu können. Ein aktuelles Beispiel ist die mit Unterstützung der Behörde für Wirtschaft und Innovation gegen gewichtige Konkurrenz anderer Bundesländer gerade etablierte Test-

[54] https://digital.hamburg.de/digitale-stadt/govtechh-11008 (zuletzt aufgerufen am 07.02.2023)
[55] Vgl. https://www.hamburg-news.hamburg/innovation-wissenschaft/govtech-campus-fuer-die-hamburger-verwaltung-entsteht (zuletzt aufgerufen am 07.02.2023)
[56] Als Beispiel kann das Förderprogramm Innotech genannt werden, durch das auch Start-ups aus dem Bereich Legal Tech gefördert werden können, die durchaus mit ihren Lösungen auch für die Verwaltung interessant sein könnten, https://www.ifbhh.de/foerderprogramm/innofintech (zuletzt aufgerufen am 07.02.2023)

und Zertifizierungsstelle CertifAI von PWC und Dekra Digital in Hamburg, die Anfang 2023 ihren Betrieb aufnimmt.

7.10 Technische Möglichkeiten zur Gewährleistung des gesellschaftlich verantwortlichen Einsatzes von KI

Für die Gewährleistung der sich aus den Grundsätzen von *Responsible AI* und den Regelungen des KI-Acts ergebenen Anforderungen bieten sich grundsätzlich auch KI-basierte technische Lösungen an. In der Anlage 2 werden zu den Punkten Fairness/Ethik, Erklärbarkeit/Transparenz, Robustheit/Zuverlässigkeit mit dem Schwerpunkt Prozesse und Normen des Software Engineering sowie zur Qualitätssicherung/Validierung und Verifikation exemplarisch ohne Anspruch auf Vollständigkeit technische Lösungsansätze vorgestellt. Nachhaltige und etablierte KI-Systeme zeigen, dass in der Entwicklung eines KI-Systems grundlegend wie bei jeder anderen Software mit den hier aufgezeigten erhöhten Anforderungen (Anlage 2) gehandelt werden sollte. In diesem Zusammenhang sind der weitere Prozess der Zertifizierung und die in diesem Zusammenhang (weiter)entwickelten Anforderungen bzw. Maßstäbe zu beachten.

Eine Aufgabe wird in der Zukunft sein, zu definieren, wann und in welcher Form je nach Anwendungsfall (Erwerb, Selbst- und Fortentwicklung sowie Gebrauch) welche Notwendigkeiten und Möglichkeiten des Einsatzes technischer Lösungen in der Verwaltung bestehen. Es kann sich anbieten, in diesem Zusammenhang einerseits nach schon bestehenden technischen Lösungen auf dem Markt bzw. bei Start-ups zu suchen, sich andererseits im Dialog mit Start-ups auch nach neuen Lösungsansätzen umzusehen. Hierfür gibt es als Plattform bereits GovTechHH.

7.11 Normierungsvorschläge

Entsprechend den obigen Ausführungen bedarf es in den meisten Fällen der KI-Anwendungen keiner Gesetzesnovellierung. Werden die Anforderungen des KI-Acts eingehalten und können Zertifikate nach dem KI-Act vorgelegt werden, so scheint die Anwendung der KI im Verwaltungsverfahren grundsätzlich unproblematisch zu sein; dies gilt insbesondere für §§ 35, 35a, 39 VwVfG. Auch wird die Notwendigkeit der Ausdehnung des Anwendungsbereichs des § 35a VwVfG auf Ermessensentscheidungen und Normen mit Beurteilungsspielraum nicht gesehen.

Sobald jedoch Formen von KI oder maschinellem Lernen – gleich in welcher Form – in Sachgebieten mit gesetzlichen Entscheidungsprogrammen zur Anwendung kommen, die durch Beurteilungsspielräume auf der Tatbestandsseite oder durch Ermessensspielräume auf der Rechtsfolgenseite gekennzeichnet sind, oder in denen die Sachverhaltsermittlung in diesem Sinne KI-basiert erfolgt und damit auch bei gebundenen Entscheidungen diese präjudizieren kann, bedarf es ergänzender Mechanismen, die die demokratische Verwaltungslegitimation absichern. Hier bietet sich ein Rahmengesetz an:

Als Elemente eines solchen Rahmengesetzes könnten grundsätzlich in Orientierung an den Grundsätzen der *Responsible AI* und ihrer allgemein anerkannten Konkretisierung in Betracht kommen (nachfolgend nicht im Sinne einer Wertungsrangfolge):

- Definition zulässiger bzw. unzulässiger Einsatzgebiete
- Begriffsbestimmungen
- Vorkehrungen zur Gewährleistung der Diskriminierungsfreiheit bei der Zusammenstellung von Trainingsdaten für Maschinelles Lernen

- Dokumentations- und Archivierungspflichten
- Begutachtungspflichten
- Technikfolgenabschätzung
- Ergänzende Rechtsmittel- und Rechtsschutzmechanismen
- Evaluationspflichten
- Technische und fachliche Anforderungen an Anwender:innen, die die Systeme erstellen bzw. anwenden
- Technische Sicherheit
- Datenschutz

Bei der Ausgestaltung dieses Rahmengesetzes ist ein bestmöglicher Ausgleich zwischen Gewährleistung der normativen Anforderungen der demokratischen Verwaltungslegitimation und der Funktionsfähigkeit des KI-Einsatzes herzustellen; dies ist eine sehr komplexe Aufgabe. Das Rahmengesetz sollte den Einsatz von KI in der Verwaltung strukturieren und demokratisch absichern, aber den KI/ML[57]-Einsatz nicht unmöglich machen. Im Besonderen ist durch prozedurale Vorgaben das Interaktionsverhältnis zwischen den KI-Systemen und den menschlichen Entscheidungsträger:innen auszugestalten. Ferner wird insbesondere der Prozess des ML-Trainings durch normative Vorgaben zu strukturieren sein.

Bei der Beschaffung und dem Einsatz von KI auf den o.g. Handlungsebenen ist die vorausschauende Beachtung des KI-Acts und anderer Regelungen geboten. Sollte die deutschen Gesetzgeber (Bundestag und Landtage) den Ansatz jeweiliger Rahmengesetze wählen, wird insbesondere die Kompatibilität zum KI-Act eine rechtsetzungstechnische Gestaltungsaufgabe sein. Es ist aber denkbar, dass sich beide Normensysteme in Einklang bringen lassen, da sie dieselben Regelungsziele haben werden.

Bei einer Entwicklung von Rahmengesetzen auf Bundes- und Landesebene sollten die Fachbehörden aufgefordert werden zu prüfen, ob ein System in einem spezifischen Anwendungsbereich genutzt werden kann (inkl. Rechtsfolgenabschätzung und Kosten/Nutzen Analyse). Auch könnte es sich im Hinblick auf mögliche Veränderungen im KI-Bereich als auch zwischenzeitlich gewonnene positive wie negative Erkenntnisse aus der Anwendung eines solchen Gesetzes anbieten, eine Evaluationsklausel aufzunehmen: z.B. alle 5 Jahre muss eine Kontrolle bzw. Evaluation stattfinden.

Alternativ zu einem „KI-Rechtsrahmengesetz“ könnte auch der Weg einer Aufnahme entsprechender Regelungen in einen eigenen Abschnitt in die Verwaltungsverfahrensgesetze des Bundes und der Länder gewählt werden. Soweit sich in bestimmten Regelungsbereichen wie Planungsrecht die Notwendigkeit des Einsatzes besonderer Anwendungen ergeben, könnte es sich anbieten, diese in den Fachgesetzen mit aufzunehmen bzw. soweit schon eigene Regelungen wie im SGB X vorhanden entsprechend zu ergänzen.

Insgesamt bietet es sich an, die vorgenannten Fragen durch rechtliche Vorschriften zu gestalten, um so eine (bessere) Gerichtsfestigkeit des Einsatzes von KI in diesen Fällen zu erreichen.

[57] Maschinelles Lernen (ML)

7.12 Richtlinien, Verwaltungsvorschriften zum verantwortungsvollen Einsatz von Künstlicher Intelligenz (RAI-RL) in der Verwaltung der Freien und Hansestadt Hamburg – Eckpunkte

Es wird empfohlen für den Einsatz von Systemen, die Künstliche Intelligenz nutzen, in der Verwaltung der Freien und Hansestadt Hamburg geeignete verwaltungsinterne Richtlinien einzusetzen. Deren Sinn und Zweck ist es, Regeln für den verantwortungsvollen Einsatz von Künstlicher Intelligenz *(Responsible AI)* aufzustellen, zum Beispiel und ohne Anspruch auf Vollständigkeit und die richtige Wertungsreihenfolge im Zusammenhang mit

- der Beschaffung
- der dauerhaften Betreuung einschließlich der menschlichen Aufsicht
- den Anforderungen an Eingangsdatensätze
- der Umsetzung von Dokumentationspflichten
- der verwaltungspraktischen Anwendung

von Systemen, die KI verwenden.

Zunächst ist eine Evaluierung der bestehenden Regelungswerke vorzunehmen, inwieweit sie den Grundsätzen der *Responsible AI* und des KI-Acts entsprechen. Es ist zu entscheiden, ob eine Überarbeitung bzw. Neufassung erforderlich ist oder eine Neufassung in einer oder mehrerer sich ergänzender Richtlinien vorzunehmen ist. Es könnte sich anbieten, daneben eine „Rahmenrichtlinie“ zu erlassen, die sich an Anwender:innen sowie zentrale und dezentrale fachliche Stellen wendet, und vorhandene bzw. ggfs. überarbeitete Vorschriften erweitert bzw. diese einordnet und einen einheitlichen Handlungsrahmen setzt. Damit wird das Konzept verantwortungsvoller Künstlicher Intelligenz (RAI) in Verwaltungshandeln übertragen. Der Einsatz von Systemen, die Künstliche Intelligenz nutzen, wird damit befördert, unter Beachtung ethischer, rechtsstaatlicher und ökonomischer Grenzen.

Es empfiehlt sich, dieses frühzeitig auf den Weg zu bringen. Die Verwaltung der Hansestadt Hamburg verbessert damit qualitativ schon vor dem Inkrafttreten des KI-Acts den verantwortungsvollen Einsatz von Künstlicher Intelligenz und setzt das RAI-Konzept um.

Eine Rahmenrichtlinie könnte einen einheitlichen Handlungsrahmen für die betroffenen Stellen schaffen, der dem RAI-Konzept entspricht. Dieser Rahmen kann auch nach dem Inkrafttreten des KI-Acts für die KI-Anwendungen fortgelten, die vom KI-Act nicht erfasst werden, für die aber entsprechend der obigen Ausführungen die Anwendung der Grundsätze des RAI Konzeptes sowohl zur Vermeidung von Schäden wie auch der Förderung gesellschaftlicher Akzeptanz geboten ist.

Für den Fall des Erlasses eines KI-Rahmengesetzes (vgl. Empfehlung 7.11) können in einer Rahmenrichtlinie Konkretisierungen der Eckpunkte des Gesetzes vorgenommen werden. Sollte von dem Erlass eines Rahmengesetzes Abstand genommen werden und notwendige gesetzliche Regelungen etwa durch Ergänzungen bestehender Regelungen z.B. im Verwaltungsverfahrensgesetz vorgenommen werden, könnte eine „Rahmenrichtlinie“ die wesentlichen Anforderungen an den Einsatz von KI festlegen.

Nach Verabschiedung des KI-Acts sind die Richtlinien bzw. Rahmenrichtlinie gegebenenfalls anzupassen. Das gleiche gilt für den Fall, dass die Freie und Hansestadt Hamburg ein KI-Gesetz entsprechend Ziffer 11

dieser Studie auf den Weg bringt. Im Hinblick auf die technische Entwicklung sollten die Richtlinien regelmäßig evaluiert und ggfs. angepasst werden.

Es empfiehlt sich den Einsatz von KI-Systemen zu Zwecken der Strafverfolgung nicht zu erfassen, da sich hierfür im KI-Act wesentliche Sonderregelungen abzeichnen und die Materie vom Einsatz in allgemeinen und besonderen Verwaltungsangelegenheiten abweicht.

Mögliche Eckpunkte für die Überarbeitung bestehender Richtlinien bzw. die Erarbeitung einer Rahmenrichtlinie zum verantwortungsvollen Einsatz Künstlicher Intelligenz (RAI-RL) in der Verwaltung der Freien und Hansestadt Hamburg werden in der Anlage 3 ohne Anspruch auf Vollständigkeit und nicht im Sinne einer Wertungsreihenfolge aufgeführt.

7.13 Kommunikation nach Außen – Botschaften

Im Interesse der Vertrauensförderung bzw. Vertrauensbildung bei den Bürger:innen ist die Leitlinie einer gesellschaftlich verantwortlichen Digitalisierung der Hamburger Verwaltung nach den Grundsätzen einer *Responisble AI* und dem KI-Act in der Öffentlichkeit nachhaltig zu vertreten. Hierfür bieten sich entsprechende, noch zu entwickelnde Formate und Veranstaltungen an. Wichtig ist es, Räume zu schaffen, um diejenigen, die den Wandel herbeiführen, die unsere Gesellschaft, unsere Lebens- und Arbeitswelt verändern und die unsere Einstellung gegenüber der Technologie prägen, mit den Entscheidungsträger:innen der Digitalisierung der Verwaltung zusammenzubringen.

Zugleich kann Hamburg mit diesem Ansatz auch im Bundesvergleich eine Standortvorteil geltend machen. Hinzu kommt, dass die Hamburger Verwaltung so auch bei der Gewinnung von einschlägigem Nachwuchs durch die Botschaft, auch im Bereich des Einsatzes von KI gesellschaftlich verantwortlich zu handeln, von zusätzlicher Attraktivität ist. Im Arbeitsmarkt zeigt sich, dass für die Arbeitssuche das Thema gesellschaftlich verantwortlichen Handelns zunehmend von Bedeutung ist.

Mögliche Botschaften für die gesellschaftlich verantwortliche, ethische Digitalisierung und den Einsatz von KI können im Einklang mit der Digitalstrategie für Hamburg sein:

- Hamburg sieht in der Digitalisierung der Verwaltung einen entscheidenden Schritt zur Verbesserung der Dienstleistungen für alle Bürger:innen.
- Hamburg versteht die Digitalisierung der Gesellschaft und ihre Auswirkungen als technisch geprägtes, rechtlich gestaltbares und gesellschaftlich verantwortliches Thema.
- Hamburg denkt Digitalisierung der Verwaltung interdisziplinär.
- Hamburg ist bundes- und europaweit führend bei der gesellschaftlich verantwortlichen Digitalisierung der Verwaltung.
- Hamburg treibt die Überlegungen zu Transparenz, Nachvollziehbarkeit, für demokratische Legitimation, für fairen Umgang und die Gewährleistung effektiven Rechtsschutzes und Akzeptanz bei Bürger:innen voran.
- Hamburg stellt sich der Verantwortung, die die Digitalisierung der Gesellschaft mit sich bringt, weitsichtig, mit Kreativität und Mut zur Veränderung.
- Hamburg wird seinen Beitrag leisten, die Überlegungen zum verantwortungsvollen Umgang bei dem Einsatz von KI-Systemen in der Verwaltung auch auf andere Bereiche wie Internetplattformen, Sozialen Medien etc. zu übertragen.

- Hamburg wird bei der Entwicklung von technischen Lösungen zur Gewährleistung der Anforderungen an einen verantwortungsvollen Umgang die Hamburger Digitalwirtschaft einbinden und stärken.

Anhang 1 Checkliste

Vorbemerkung

Der Checkliste liegt der Entwurf des KI-Acts vom 21.4.2021 zu Grunde. Es ist davon auszugehen, dass es im weiteren Normgebungsverfahren zu Änderungen kommen wird, sodass die Checklist entsprechend anzupassen wäre.

1. Ist es KI gemäß KI-Act-Definition?

[Art. 3 Abs. 1 KI-Act Entwurf & Anhang I]

Zutreffendes bitte ankreuzen:

1. Handelt es sich um eine Software, die mit einer oder mehreren der folgenden Techniken und Konzepten entwickelt worden ist? Zutreffendes bitte ankreuzen:	
A) Handelt es sich um maschinelles Lernen • mit beaufsichtigtem, unbeaufsichtigtem und bestärkendem Lernen • unter Verwendung einer breiten Palette von Methoden, einschließlich des tiefen Lernens (Deep Learning)?	
B) Handelt es sich um Logik- und wissensgestützte Konzepte, einschließlich • Wissensrepräsentation, • induktiver (logischer) Programmierung, • Wissensgrundlagen, • Inferenz- und Deduktionsmaschinen, • (symbolischer) Schlussfolgerungs- und Expertensysteme?	
C) Handelt es sich um • statistische Ansätze, • Bayessche Schätz-, Such- und Optimierungsmethoden?	
2. Kann die Software im Hinblick auf eine Reihe von Zielen, die vom Menschen festgelegt werden, Ergebnisse (z.B. Inhalte, Vorhersagen, Empfehlungen oder Entscheidungen) hervorbringen, die das Umfeld beeinflussen, mit dem sie interagieren?	

Wurde bei 1. & 2. ein Kreuz gesetzt, handelt es sich um KI im Sinne des KI-Act-Entwurfs und Sie müssen die darin festgehaltenen Anforderungen erfüllen. Weiter zu Punkt 2. der Checkliste.

Wurde in keinem oder nur einem der Felder ein Kreuz gesetzt, fällt ihr System nicht unter die KI-Definition des KI-Act Entwurfs und Sie können die Checkliste beenden.

2. Welche Risikokategorie?

Der KI-Act Entwurf verfolgt einen risikobasierten Ansatz, bei dem Anwendungen, die einen höheren potentiellen Schaden oder Nachteil für Menschen nach sich ziehen können, höheren rechtlichen Anforderungen unterliegen. Es wird zwischen den folgenden vier Risikokategorien unterschieden:

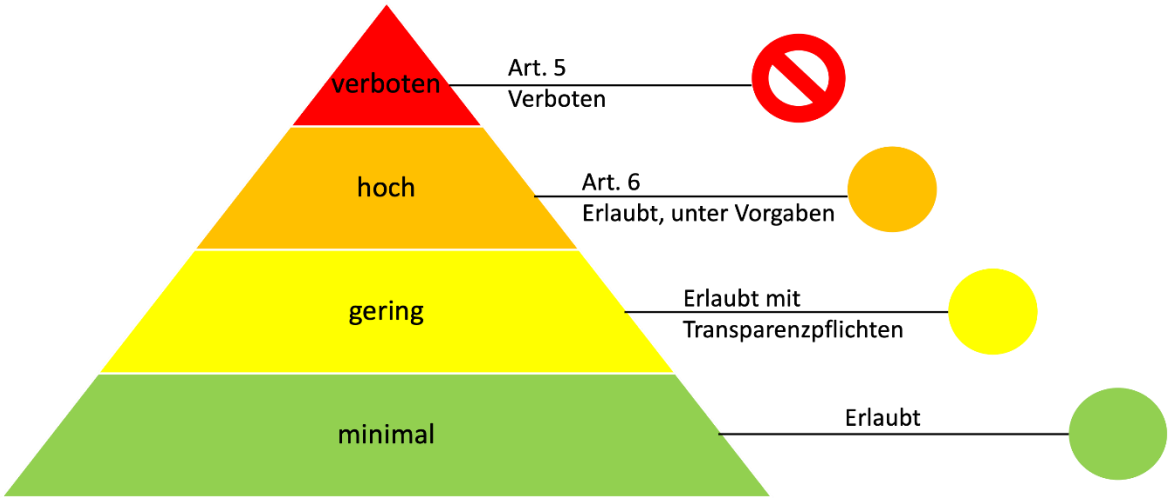

Abb.1.: Eigene Darstellung: Der risikobasierte Ansatz nach KI-Act

<u>Verbotene KI-Systeme</u> (Titel II, Art. 5 KI-Act Entwurf)

KI-Systeme, die zu folgenden Zwecken eingesetzt werden können, sind verboten:

Die Unterschwellige Beeinflussung von Personen zu deren (potentiellen) Schaden Beeinflusst das Produkt unterschwellig das Verhalten einer Person, außerhalb deren Bewusstseins in einer Weise so wesentlich, dass es dieser oder eine andere Person physischen oder psychischen Schaden zufügt oder zufügen könnte?	
Das Ausnutzen der Schutzbedürftigkeit bestimmter Personengruppen Beeinflusst das Produkt das Verhalten einer Person, die einer schutzbedürftigen Personengruppe angehört in einer Weise, die die Schwäche oder Schutzbedürftigkeit (bspw. bei körperlicher oder geistiger Behinderung) ausnutzt und dieser oder einer anderen Person psychischen oder physischen Schaden zufügt oder zufügen könnte?	
Behördliche Bewertung oder Klassifizierung natürlicher Personen Handelt es sich um ein Produkt, welches zur behördlichen Bewertung oder Klassifizierung von der Vertrauenswürdigkeit natürlicher Personen, aufgrund von Verhaltensweisen oder Persönlichkeitsmerkmalen? Resultiert daraus eine Schlechterstellung oder Benachteiligung bestimmter natürlicher Personen oder ganzer Personengruppen in sozialen Zusammenhängen, die sich von den Umständen unterscheiden, unter denen die Daten ursprünglich erzeugt oder erfasst wurden? Resultiert daraus eine ungerechtfertigte oder unverhältnismäßige Schlechterstellung oder Benachteiligung bestimmter natürlicher Personen oder ganzer Personengruppen in Bezug auf deren soziales Verhalten?	
Strafverfolgung in öffentlichen Räumen mit biometrischen Echtzeit-Identifizierungssystemen Handelt es sich bei dem Produkt um ein biometrisches Echtzeit-Fernidentifizierungssystem, welches in öffentlichen Räumen zu Strafverfolgungszwecken, ausgenommen der folgend aufgelisteten, eingesetzt wird?	

Mit Ausnahme von folgenden Zwecken: - Gezielte Suche nach bestimmten potentiellen Opfern von Straftaten oder nach vermissten Kindern - Abwenden einer konkreten, erheblichen und unmittelbaren Gefahr für das Leben oder die körperliche Unversehrtheit natürlicher Personen oder eines Terroranschlags - Erkennen, Aufspüren, Identifizieren oder Verfolgen eines Täters oder Verdächtigen einer Straftat im Sinne des Art. 2 Abs. 2 des Rahmenbeschlusses 2002/584/JI des Rates, der in dem betreffenden Mitgliedstaat nach dessen Recht mit einer Freiheitsstrafe oder einer freiheitsentziehenden Maßregel der Sicherung im Höchstmaß von mindestens drei Jahren bedroht ist.	
Biometrische Echtzeit-Fernidentifizierungssysteme Für die Verwendung biometrischer Echtzeit-Fernidentifizierungssysteme gelten besondere Ausnahmen. Diese werden an dieser Stelle nicht aufgeführt und können nachgelesen werden (Titel II, Art. 5 Abs. 1 d) & Abs. 2 KI-Act Entwurf	

Wurde bei einem dieser Felder ein Kreuz gesetzt, handelt es sich um ein verbotenes KI-System. Das Inverkehrbringen, die Inbetriebnahme oder die Verwendung ist verboten (Titel II, Art. 5 KI-Act Entwurf). Die Checkliste ist an dieser Stelle zu beenden.

Wurde kein Kreuz gesetzt, gilt es zu prüfen, welche der übrigen Risikokategorien zutreffen.

Hochrisiko-Anwendungen

Handelt es sich bei dem KI-System um eine Hochrisiko-Anwendung gemäß KI-Act Entwurf? (Anhang III, gemäß Art. 6 Abs. 2 KI-Act Entwurf)

Im Wesentlichen gilt es zu prüfen, ob ein KI-System in die Hochrisikokategorie fällt oder nicht. Denn tut sie das nicht, ist sie entweder verboten oder die Anforderungen sind sehr gering. Zur Abschätzung des Aufwandes ist also die entscheidende Trennlinie Hochrisiko oder nicht (Anhang III, gemäß Art. 6 Abs. 2 KI-Act Entwurf)

Zutreffendes bitte ankreuzen:

Biometrische Identifizierung und Kategorisierung natürlicher Personen Soll das KI-System bestimmungsgemäß für die biometrische Echtzeit-Fernidentifizierung und nachträgliche biometrische Fernidentifizierung natürlicher Personen verwendet werden?	
Verwaltung und Betrieb kritischer Infrastrukturen Soll das KI-System bestimmungsgemäß als Sicherheitskomponente in der Verwaltung und im Betrieb des Straßenverkehrs sowie in der Wasser-, Gas-, Wärme- und Stromversorgung verwendet werden? *[Def. Sicherheitskomponente Titel I, Art. 3, 14 K-Act Entwurf.]*	
Allgemeine und berufliche Bildung Soll das KI-System bestimmungsgemäß für Entscheidungen über den Zugang oder die Zuweisung natürlicher Personen zu Einrichtungen der allgemeinen und beruflichen Bildung verwendet werden?	

Soll das KI-System bestimmungsgemäß für die Bewertung von Schülern in Einrichtungen der allgemeinen und beruflichen Bildung und für die Bewertung der Teilnehmer an üblicherweise für die Zulassung zu Bildungseinrichtungen erforderlichen Tests verwendet werden?	
Beschäftigung, Personalmanagement und Zugang zur Selbstständigkeit: Soll das KI-System bestimmungsgemäß für die Einstellung oder Auswahl natürlicher Personen verwendet werden (insbesondere für die Bekanntmachung freier Stellen, das Sichten oder Filtern von Bewerbungen und das Bewerten von Bewerbern in Vorstellungsgesprächen oder Tests)?	
Soll das KI-System bestimmungsgemäß für Entscheidungen über Beförderungen und über Kündigungen von Arbeitsvertragsverhältnissen, für die Aufgabenzuweisung sowie für die Überwachung und Bewertung der Leistung und des Verhaltens von Personen in solchen Beschäftigungsverhältnissen verwendet werden?	
Zugänglichkeit und Inanspruchnahme grundlegender privater und öffentlicher Dienste und Leistungen: Soll das KI-System bestimmungsgemäß von Behörden oder im Namen von Behörden verwendet werden, um zu beurteilen, ob natürliche Personen Anspruch auf öffentliche Unterstützungsleistungen und -dienste haben und ob solche Leistungen und Dienste zu gewähren, einzuschränken, zu widerrufen oder zurückzufordern sind?	
Soll das KI-System bestimmungsgemäß für die Kreditwürdigkeitsprüfung und Kreditpunktebewertung natürlicher Personen verwendet werden (Ausnahme: KI-Systeme, die von Kleinanbietern für den Eigengebrauch in Betrieb genommen werden)?	
Soll das KI-Systeme bestimmungsgemäß für die Entsendung oder Priorisierung des Einsatzes von Not- und Rettungsdiensten, einschließlich Feuerwehr und medizinischer Nothilfe, verwendet werden?	
Strafverfolgung: Soll das KI-System bestimmungsgemäß von Strafverfolgungsbehörden für individuelle Risikobewertungen natürlicher Personen verwendet werden, um das Risiko abzuschätzen, dass eine natürliche Person Straftaten begeht oder erneut begeht oder dass eine Person zum Opfer möglicher Straftaten wird?	
Soll das KI-System bestimmungsgemäß von Strafverfolgungsbehörden als Lügendetektoren und ähnliche Instrumente oder zur Ermittlung des emotionalen Zustands einer natürlichen Person verwendet werden?	
Soll das KI-Systeme bestimmungsgemäß von Strafverfolgungsbehörden zur Aufdeckung von Deepfakes gemäß Art. 52 Abs. 3 KI-Act Entwurf verwendet werden?	
Soll das KI-Systeme bestimmungsgemäß von Strafverfolgungsbehörden zur Bewertung der Verlässlichkeit von Beweismitteln im Zuge der Ermittlung oder Verfolgung von Straftaten verwendet werden?	
Soll das KI-Systeme bestimmungsgemäß von Strafverfolgungsbehörden zur Vorhersage des Auftretens oder erneuten Auftretens einer tatsächlichen oder potentiellen Straftat auf der Grundlage des Profils natürlicher Personen gemäß Art. 3 Abs. 4 der Richtlinie (EU) 2016/680 oder zur Bewertung von Persönlichkeitsmerkmalen und Eigenschaften oder vergangenen kriminellen Verhaltens natürlicher Personen oder von Gruppen verwendet werden?	
Soll das KI-Systeme bestimmungsgemäß von Strafverfolgungsbehörden zur Erstellung von Profilen natürlicher Personen gemäß Art. 3 Abs. 4 der Richtlinie (EU) 2016/680 im Zuge der Aufdeckung, Ermittlung oder Verfolgung von Straftaten verwendet werden?	

Soll das KI-Systeme bestimmungsgemäß zur Kriminalanalyse natürlicher Personen eingesetzt werden und es den Strafverfolgungsbehörden ermöglichen, große komplexe verknüpfte und nicht verknüpfte Datensätze aus verschiedenen Datenquellen oder in verschiedenen Datenformaten zu durchsuchen, um unbekannte Muster zu erkennen oder verdeckte Beziehungen in den Daten aufzudecken?	
Migration, Asyl und Grenzkontrolle: Soll das KI-System bestimmungsgemäß von zuständigen Behörden als Lügendetektoren und ähnliche Instrumente oder zur Ermittlung des emotionalen Zustands einer natürlichen Person verwendet werden?	
Soll das KI-System bestimmungsgemäß von zuständigen Behörden zur Bewertung eines Risikos verwendet werden (einschließlich eines Sicherheitsrisikos, eines Risikos der irregulären Einwanderung oder eines Gesundheitsrisikos, das von einer natürlichen Person ausgeht, die in das Hoheitsgebiet eines Mitgliedstaats einzureisen beabsichtigt oder eingereist ist)?	
Soll das KI-System bestimmungsgemäß von zuständigen Behörden zur Überprüfung der Echtheit von Reisedokumenten und Nachweisunterlagen natürlicher Personen und zur Erkennung unechter Dokumente durch Prüfung ihrer Sicherheitsmerkmale verwendet werden?	
Soll das KI-Systeme bestimmungsgemäß zuständige Behörden bei der Prüfung von Asyl- und Visumanträgen sowie Aufenthaltstiteln und damit verbundenen Beschwerden im Hinblick auf die Feststellung der Berechtigung der den Antrag stellenden natürlichen Personen unterstützen?	
Rechtspflege und demokratische Prozesse: Soll das KI-System bestimmungsgemäß Justizbehörden bei der Ermittlung und Auslegung von Sachverhalten und Rechtsvorschriften und bei der Anwendung des Rechts auf konkrete Sachverhalte unterstützen?	

Wurde in einem der Felder ein Kreuz gesetzt, kann weiter zu Punkt 3. der Checkliste gegangen werden.

Wurde kein Kreuz gesetzt, gilt es zu prüfen, ob das KI-System eine Sicherheitskomponente eines Produkts oder in einem Teil eines Produktes darstellt, oder welche der übrigen Risikokategorien zutreffen.

Handelt es sich bei dem KI-System um eine Sicherheitskomponente in einem Produkt oder um einen Teil eines Produktes, das den Harmonisierungsvorschriften der Union unterliegt (Anhang II)? *[Def. Sicherheitskomponente in Titel I, Art. 3, 14 KI-Act Entwurf]*

Maschinen Handelt es sich bei dem Produkt um Maschinen gemäß Richtlinie 2006/42/EG? z.B. **Sensorik Anwendungen** in Mikrocontrollern, Embedded GPUs, Smartphones und Co?	
Spielzeug Handelt es sich bei dem Produkt um Spielzeug gemäß Richtlinie 2009/48/EG?	
Sportboote und Wassermotorräder Handelt es sich bei dem Produkt um Sportboote und Wassermotorräder gemäß Richtlinie 2013/53/EU?	
Aufzüge Handelt es sich bei dem Produkt um Aufzüge gemäß Richtlinie 2014/33/EU?	

Geräte und Schutzsysteme zur bestimmungsgemäßen Verwendung in explosionsgefährdeten Bereichen Handelt es sich bei dem Produkt um Geräte und Schutzsysteme zur bestimmungsgemäßen Verwendung in explosionsgefährdeten Bereichen gemäß Richtlinie 2014/34/EU?	
Funkanlagen Handelt es sich bei dem Produkt um Funkanlagen gemäß Richtlinie 2014/53/EU?	
Druckgeräte Handelt es sich bei dem Produkt um Druckgeräte gemäß Richtlinie 2014/68/EU?	
Personenschutzausrüstung Handelt es sich bei dem Produkt um Personenschutzausrüstung/ persönliche Schutzausrüstung gemäß der Richtlinie 2016/425/EU?	
Medizinprodukte Handelt es sich bei dem Produkt um ein Medizinprodukt gemäß der Richtlinie 2017/745/EU?	
In-vitro-Diagnostika Handelt es sich bei dem Produkt um ein In-vitro-Diagnostika gemäß der Richtlinie 2017/746/EU?	

Handelt es sich bei dem KI-System um eine Sicherheitskomponente in einem Produkt oder selbst um ein Produkt, das folgenden Harmonisierungsvorschriften der Union unterliegt (Titel I, Art. 2 Abs. 2 KI-Act Entwurf sowie Anhang II B)? *[Def. Sicherheitskomponente in Titel I, Art. 3, 14 KI-Act Entwurf.]*

Zivilluftfahrt Verordnung (EG) Nr. 300/2008: gemeinsamen Vorschriften für die Sicherheit in der Zivilluftfahrt	
Genehmigung und Marktüberwachung von Fahrzeugen und Kraftfahrzeugen Verordnung (EU) Nr. 168/2013 des Europäischen Parlaments und des Rates vom 15. Januar 2013 über die Genehmigung und Marktüberwachung von zwei- oder dreirädrigen und vierrädrigen Fahrzeugen	
Genehmigung und Marktüberwachung von land- und forstwirtschaftlichen Fahrzeugen Verordnung (EU) Nr. 167/2013 des Europäischen Parlaments und des Rates vom 5. Februar 2013 über die Genehmigung und Marktüberwachung von land- und forstwirtschaftlichen Fahrzeugen	
Schiffsausrüstung Richtlinie 2014/90/EU des Europäischen Parlaments und des Rates vom 23. Juli 2014 über Schiffsausrüstung	
Interoperabilität des europäischen Eisenbahnsystems Richtlinie (EU) 2016/797 des Europäischen Parlaments und des Rates vom 11. Mai 2016 über die Interoperabilität des Eisenbahnsystems in der Europäischen Union	
Genehmigung und Marktüberwachung von Kraftfahrzeugen, Kraftfahrzeuganhänger und anderen Systemteilen technischer Einheiten	

Verordnung (EU) 2018/858 des Europäischen Parlaments und des Rates vom 30. Mai 2018 über die Genehmigung und die Marktüberwachung von Kraftfahrzeugen und Kraftfahrzeuganhängern sowie von Systemen, Bauteilen und selbstständigen technischen Einheiten für diese Fahrzeuge Verordnung (EU) 2019/2144 des Europäischen Parlaments und des Rates vom 27. November 2019 über die Typgenehmigung von Kraftfahrzeugen und Kraftfahrzeuganhängern sowie von Systemen, Bauteilen und selbstständigen technischen Einheiten für diese Fahrzeuge im Hinblick auf ihre allgemeine Sicherheit und den Schutz der Fahrzeuginsassen und von ungeschützten Verkehrsteilnehmern	
Zivilluftfahrt und Flugsicherheit Verordnung (EU) 2018/1139 des Europäischen Parlaments und des Rates vom 4. Juli 2018 zur Festlegung gemeinsamer Vorschriften für die Zivilluftfahrt und zur Errichtung einer Agentur der Europäischen Union für Flugsicherheit	

Wurde in einem der Felder ein Kreuz gesetzt, gilt Titel IV, Art. 84 KI-Act Entwurf. An dieser Stelle kann die Checkliste beendet werden.

Wurde in der bisherigen Checkliste noch kein Kreuz gesetzt, gilt es zu prüfen, welche der übrigen Risikokategorien zutreffen:

<u>Geringes Risiko</u>

Beabsichtigen Sie die Entwicklung, das Inverkehrbringen oder die Inbetriebnahme eines KI-Systems mit folgenden Eigenschaften: Interaktion mit Menschen Das KI-System ist zur Interaktion mit natürlichen Personen bestimmt Handelt es sich um ein KI-System, welches für die Interaktion mit natürlichen Personen bestimmt ist?	
Emotionserkennung und biometrische Kategorisierung Das KI-System wird zur Erkennung von Emotionen oder zur Assoziierung von (gesellschaftlicher) Kategorien anhand biometrischer Daten eingesetzt	
Deepfakes Das KI-System erzeugt oder manipuliert Bild-, Ton- oder Videoinhalte, die wirklichen Personen, Gegenständen, Orten oder anderen Einrichtungen oder Ereignissen merklich ähneln und einer Person fälschlicherweise als echt oder wahrhaftig erscheint („Deepfake")?	

Wurde in dieser Checkliste ein Kreuz gesetzt, gilt die Pflicht zur Offenlegung der Tatsache (Titel IV, Art. 52 KI-Act Entwurf). Die KI-Systeme müssen so konzipiert und entwickelt werden, dass natürlichen Personen mitgeteilt wird, dass es sich um ein KI-System handelt.

Wenn bis hierhin kein Kreuz gesetzt wurde, fällt man nach dem Ausschlussverfahren in die minimale Risikokategorie und muss keine der Anforderungen nach dem KI-Act erfüllen (kann aber freiwillig).

Die Checkliste kann an dieser Stelle beendet werden.

3. Welche Rolle?

Rechte und Pflichten in Bezug auf Hochrisiko-KI-Systeme unterscheiden sich nach Rollen [Titel III, Kapitel 3 KI-Act]. Ein Akteur kann je nach Tätigkeit mehrere Rollen einnehmen, bzw. die Rollen können bei Änderung der Tätigkeiten wechseln.

Zutreffendes ankreuzen:

Anbieter [Titel III, Kapitel 3, Art. 16-23 KI-Act] Unter Anbieter fallen folgende Akteure: Natürliche oder juristische Personen, Behörden, Einrichtungen oder sonstige Stellen, die ein KI System entwickeln oder entwickeln lassen, um es unter ihrem eigenen Namen oder ihrer eigenen Marke – entgeltlich oder unentgeltlich – in Verkehr zu bringen. Natürliche oder juristische Personen, Behörden oder sonstige Stellen, die wesentliche Änderungen an einem KI System vorgenommen haben. Natürliche oder juristische Personen, Behörden oder sonstige Stellen, die die Zweckbestimmung eines bereits im Verkehr befindlichen KI Systems verändert haben oder verändern werden.	
Produkthersteller [Art. 24 KI-Act] Zu Produktherstellern zählen natürliche oder juristische Personen, die Produkte herstellen oder entwickeln und herstellen lassen und diese unter ihrem eigenen Namen oder ihrer eigenen (Handels-)Marke vermarkten oder für ihre eigenen Zwecke verwenden. Darunter fallen Produkte folgender Rechtsakte: • Maschine (Richtlinie 2006/42/EG) • Spielzeug (Richtlinie 2009/48/EG) • Sportboote & Wassermotorräder (Richtlinie 2013/53/EU) • Aufzüge (Richtlinie 2014/33/EU) • Geräte und Schutzsysteme zur bestimmungsgemäßen Verwendung in explosionsgefährdeten Bereichen (Richtlinie 2014/34/EU) • Funkanlagen (Richtlinie 2014/53/EU) • Druckgeräte (Richtlinie 2014/68/EU) • Seilbahnen (Verordnung (EU) 2016/424) • persönliche Schutzausrüstungen (Verordnung (EU) 2016/425) • Geräte zur Verbrennung gasförmiger Brennstoffe (Verordnung (EU) 2016/426) • Medizinprodukt (Verordnung (EU) 2017/745) • In-vitro-Diagnostika (Verordnung (EU) 2017/746)	
Bevollmächtigte [Art. 25 KI-Act] Bin ich eine in der Union ansässige oder niedergelassene natürliche oder juristische Person, die vom Anbieter eines KI Systems schriftlich dazu bevollmächtigt wurde, in seinem	

Namen die in der KI Verordnung festgelegten Pflichten zu erfüllen bzw. Verfahren durchzuführen?	
Einführer [Art. 26, 28 KI-Act] Bin ich eine in der Union ansässige oder niedergelassene natürliche oder juristische Person, die ein KI System, das den Namen oder die Marke einer außerhalb der Union ansässigen oder niedergelassenen natürlichen oder juristischen Person trägt, in der Union in Verkehr bringt oder in Betrieb nimmt?	
Händler [Art. 27, 28 KI-Act] Bin ich eine natürliche oder juristische Person in der Lieferkette, die ein KI-System ohne Änderungen seiner Merkmale auf dem Unionsmarkt bereitstellt, mit Ausnahme des Herstellers oder des Einführers?	
Nutzer [Art. 28, 29 KI-Act] Bin ich eine natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, die ein KI System in eigener Verantwortung oder verwendet, es sei denn, das KI System wird im Rahmen einer persönlichen und nicht beruflichen Tätigkeit verwendet?	
Sonstige Dritte [Art. 28 KI-Act] Bringe ich ein Hochrisiko-KI-System unter meinem Namen oder meiner Marke in Verkehr oder nehme es in Betrieb? Bin ich eine natürliche oder juristische Person, die die Zweckbestimmung eines sich bereits im Verkehr befindenden oder in Betrieb genommenen Hochrisiko-KI-Systems verändert? Bin ich eine natürliche oder juristische Person, die eine wesentliche Änderung an einem Hochrisiko-KI-System vornimmt?	
Kleinanbieter Bin ich eine natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, bei der es sich um ein Kleinst- oder Kleinunternehmen im Sinne der Empfehlung 2003/361/EG und ein KI-System entwickelt oder entwickeln lässt, um es unter meinem Namen oder den meiner Marke in Verkehr zu bringen oder in Betrieb zu nehmen?	

Wenn Sie ein oder mehrere Kreuze gesetzt haben, gehen Sie weiter zu 4.

Wenn Sie bei „Sonstige Dritte“ ein Kreuz gesetzt haben, zählen Sie als neuer Anbieter des Produkts und unterliegen den Anbieterpflichten gemäß Art. 16 KI-Act.

4.Was muss ich tun? (Output)

An dieser Stelle würden nun die Anforderungen stehen, die sich aus dem Durchlaufen der Checkliste ergeben haben. In analoger Form ist dies zu unübersichtlich, weshalb an dieser Stelle darauf verzichtet wird und in digitaler Form ausgearbeitet werden sollte.

Anhang 2 Identifikation technischer Möglichkeiten zur Gewährleistung des gesellschaftlich verantwortlichen Einsatzes von KI

Um den menschlichen Anwender:innen eine Möglichkeit zu bieten, den Output von KI-Systemen zu verstehen und ihnen zu vertrauen, bedarf es, über die rechtlich Ebene hinaus, auch einer Auseinandersetzung mit dem Konzept der *Responsible AI* auf technischer Ebene. Der Bereich „*Responsible AI*" umfasst die in der Studie bereits genannten Punkte. Jeder Punkt muss gesondert betrachtet werden, da diese technisch und logisch unterschiedlich behandelt werden müssen. Die **Eine** technische Lösung wird es hierfür nicht geben. Im Folgenden werden hierzu einige Vorschläge gemacht und Hinweise gegeben. Da die Identifikation passender technischer Lösungen auch von dem komplexen Gefüge des Anwendungskontextes und des Systems abhängt, erhebt die Aufstellung keinen Anspruch auf Vollständigkeit.

Ethik

Zur Vermeidung eines sog. Bias (Voreingenommenheit) muss u.a. die Verteilung der Cluster von einem Menschen vorgenommen werden. Könnte diese Aufgabe ein System übernehmen, bräuchte man dieses System nicht noch einmal entwickeln. Für eine Beurteilung, ob dem System ein Bias antrainiert wurde, braucht es fachliche Unterstützung durch Personen, die die Funktionsweise der unterliegenden Systeme (Algorithmen) kennen. Da diese Fachleute meist kein tieferes Domain-Verständnis der zu beurteilenden Aspekte kennen, kann dies nur durch eine technisch-fachliche Unterstützung erfolgen. In diesem Bereich muss zwingend interdisziplinär, fachübergreifend und gleichberechtigt zusammengearbeitet werden.

Zur Vermeidung eines Bias (z.B. bei Personen) müssen alle Personengruppen in gleicher Anzahl im Trainingsdatensatz repräsentiert werden. Dies ist besonders dann zu beachten, wenn große, nicht überschaubare Datenmengen zum Training eines Systems verwendet werden. In diesem Fall muss bei der Qualitätssicherung ein erhöhter Aufwand investiert werden, das Verhalten des Systems mit möglicherweise unterrepräsentierten Personengruppen getestet werden, um in der Folge die Gleichbehandlung sicherzustellen.

Erklärbarkeit / Transparenz

Um die vollständige Transparenz eines Systems herzustellen, muss das Verhalten / die Entscheidung des Systems

a. erklärbar,
b. nachvollziehbar und
c. wiederholbar

sein.

Um diese Anforderungen zu erfüllen, verwendet man den wissenschaftlichen Forschungszweig der Informatik – Explainable AI. Fachlich gehört dieser Prozess zum Testing (ISO-Reihe 250xx). Bei vielen KI-Verfahren/Algorithmen kann die Erklärbarkeit vollständig hergestellt werden (beispielweise bei KI-Systemen, die in Form von Entscheidungsbäumen zu ihrem Ergebnis kommen). Hierfür ist das übliche Testvorgehen ausreichend. Erklärbarkeit und Transparenz sind daher auch eng mit der Robustheit und

dem Testing verknüpft. Bei Systemen mit hohem Risiko im Anwendungsfall, wird meistens ein mathematischer Beweis (z.B. durch die Herleitung über Gleichungssysteme) gefordert. Diese KI-Verfahren/Algorithmen werden als White Box Verfahren bezeichnet, da ihr Verhalten vorhersehbar, belegbar ist und sie damit als transparent bezeichnet werden können. Im Gegensatz zu White Box Verfahren gelten sog. Black Box KI-Verfahren/Algorithmen, als mathematisch nicht beweisbar. Hierzu zählen überwiegend Verfahren, die künstliche neuronale Netze (Deep Learning, Reinforcement Learning etc.) enthalten. Hier kann keine vollständige Transparenz und Erklärbarkeit in der Entscheidungsfindung des Systems sichergestellt werden. Bei diesen Verfahren kann die Funktionsfähigkeit nur durch Annäherung mit erhöhtem Testaufwand erfolgen, es werden hierzu die Methoden des sog. Black Box-Testing herangezogen. Die meisten Testverfahren beruhen auf dem Grundsatz des „Ursache- Wirkungsgrads“, d.h. es besteht die Frage, ob das Ergebnis nach einer Eingabe logisch und erklärbar ist, was nicht gleich bedeutet, dass die Entscheidungsfindung des KI-Systems erklärbar wäre. Hierfür muss eine Teststrategie und ein Testdatensatz bereitgehalten werden, den das KI-System nicht kennt und welches nicht in das Training eingeflossen ist. Diese Testdaten müssen im Voraus von Menschen bewertet und klassifiziert werden, damit eine Aussage über die Funktionsfähigkeit möglich ist. Automatisierte Tests können hier nur Anhaltspunkte liefern und zeigen nur einen Trend der Funktion in der Entscheidungsfindung des KI-Systems auf. Hierzu wird oft das sog. Fuzzing benutzt, wobei das System über lange Zeit mit Zufallsdaten oder synthetischen Daten getestet wird. Dies lässt aber keine qualitative Aussage der Entscheidungsfindung zu. Wie bei allen Softwaretests sollten auch die Grenzfälle und Negativtests mit einbezogen werden. Bei Hochrisikosystemen werden oft hybride KI-Verfahren verwendet. Bei dieser Vorgehensweise wird das KI-System in den Grenzfällen durch mathematisch beweisbare Algorithmen in seiner Entscheidungsfreiheit begrenzt, um fatale Fehleinschätzungen des KI-Systems zu verhindern. Black Box KI-Verfahren/Algorithmen sind zurzeit noch der Forschungsgegenstand vieler Organisationen. Als Leitlinie können nur sog. Frameworks herangezogen werden, wie beispielweise „Testing Framework for Black Box AI Models“ (IEEE, Print ISBN:978-1-6654-1219-3/DOI: 10.1109/ICSE-Companion52605.2021.00041). Bei KI-Systemen, die auf Black Box-Verfahren beruhen, wird dringend empfohlen, diese Systeme von Dritten und unabhängigen Expert:innen prüfen zu lassen. Um die Funktionsweise von Black Box-Verfahren fundiert beurteilen zu können, bedarf es dringend weiterer Erfahrungswerte aus diesem Bereich.

Robustheit / Zuverlässigkeit

Im Hinblick auf die Gewährleistung von Robustheit und Zuverlässigkeit eines Systems gilt grundsätzlich:

Ein KI-System nicht nach den Regeln und Normen des Software Engineerings zu entwickeln ist grob fahrlässig.

Bei der Einführung eines neuen Systems muss zwischen fertigen Software-Produkten und einer Neuentwicklung unterschieden werden. Basiert das KI-System auf einem fertigen Produkt, muss darauf geachtet werden, dass dieses Produkt nach den Vorgaben des Software Engineering entwickelt wurde (Kenntlich durch ISO oder ggf. IEEE-Zertifizierungen). Bei der Vergabe von Aufträgen sollte überprüft werden, ob der Auftragnehmer die nachfolgenden Prozesse anwendet. Wird eine Neuentwicklung angestrebt, müssen die Vorgehensmodelle erst geschaffen werden, um zu garantieren, dass diese eingehalten werden. Dieser Punkt wird aus Unwissenheit oder Budgetmangel oft vernachlässigt, wodurch die Robustheit

und Zuverlässigkeit eines Systems nicht garantiert werden kann. Ein KI-System wird in der Entwicklung wie ein Software-System behandelt und hat zusätzlich erhöhte Anforderungen.

Prozesse und Normen des Software Engineering

Zur Hilfestellung im Hinblick auf die in der Entwicklungsphase relevanten Anforderungen im Sinne der RAI, können bereits in den verschiedenen Phasen der Software Architektur folgende Prozesse und Normen[58]eine Rolle spielen:

- Planung
    - Anforderungserhebung (definiert durch IEEE)
    - Lastenheft (Anforderungsdefinition) (IEEE 830-1998)
    - Pflichtenheft (Mit technischen Ansätzen verfeinertes Lastenheft VDI-Richtlinie 3694)
    - Aufwandsschätzung (z. B. mittels Function-Point-Verfahren oder COCOMO)
    - Vorgehensmodell (ISO/IEC 12207)
- Analyse
    - Mock-up
    - Prozessanalyse/Prozessmodell (ISO/IEC 12207)
    - Systemanalyse
    - Strukturierte Analyse (SA)
- Entwurf
    - Softwarearchitektur (IEEE 1471:2000) oder International Software Architect Qualification
    - Board (iSAQB)
    - Strukturiertes Design (SD)
    - Fundamental Modeling Concepts (FMC)
- Programmierung
    - Normierte Programmierung (DIN 66260)
      oder
    - Objektorientierte Programmierung (OOP IEEE 830)
- Validierung und Verfikation
    - Modultests (Low-Level-Test)
    - Integrationstests (Low-Level-Test)
    - Systemtests (High-Level-Test)
    - Akzeptanztests (High-Level-Test)
- Anforderungsmanagement
    - Minimum: ISO/IEC 15504 (SPICE)
- Projektmanagement
    - Risikomanagement
    - Projektplanung

[58] Vgl. auch JRC Technical Report der EU Commission, AI Watch: Artificial Intelligence Standardisation Landscape Update - Analysis of IEEE standards in the context of the European AI Regulation, 2023, https://publications.jrc.ec.europa.eu/repository/handle/JRC131155 (zuletzt aufgerufen am 04.03.2023)

        - Projektverfolgung und -steuerung
- Qualitätsmanagement
        - SPICE (Software Process Improvement and Capability Determination ISO/IEC 15504-5)
        - Incident Management (ITIL 4)
        - Problem-Management (ISO/IEC 15504)
        - Softwaremetrik (Messung von Softwareeigenschaften, min. Restfehler, MTBF, Tests)
        - statische Analyse (Berechnung von Schwachstellen)
        - Software-Ergonomie (Arbeitsstättenverordnung (ArbStättV) sowie in der Norm EN ISO 9241)
- Konfigurationsmanagement
        - Versionsverwaltung
        - Änderungsmanagement/Veränderungsmanagement
        - Releasemanagement
        - Application-Management (ITIL)
- Softwareeinführung
        - Iterative Einführung
- Dokumentation
        - Technische Dokumentation (VDI 4500)
        - Softwaredokumentation (veraltet, gute Leitlinie: DIN 66230, DIN 66231, DIN 66232)
        - ggf. Systemdokumentation (Weiterentwicklung und Fehlerbehebung)
        - Betriebsdokumentation (Betreiber/Service, gesetzlich vorgeschrieben, je nach Geschäftszweig)
        - Bedienungsanleitung (Anwender EN 82079-1)
        - Verfahrensdokumentation (Beschreibung rechtlich relevanter Softwareprozesse EN 82079, GoBD, HGB, AO)

Qualitätssicherung / Validierung und Verifikation

Das sog. Testen/Testing ist ein Vorgehen im Qualitätsmanagement von Software und wird über die ISO/IEC 25000 bis ISO/IEC 25064 definiert. Die meisten der oben aufgezeigten Problemstellungen fallen fachlich in den Bereich der Qualitätssicherung/Softwaretestings. Aufgrund der Komplexität moderner KI-Systeme ist eine Software nie fehlerfrei, da die Komplexität ins Unendliche steigen kann. Lediglich die Fehlerzahl kann reduziert werden.

Pol, Koomen, Spillner[59] erläutern 'Testen' wie folgt:

[59] Martin Pol, Tim Koomen, Andreas Spillner, Management und Optimierung des Testprozesses. Ein praktischer Leitfaden für erfolgreiches Testen von Software mit TPI und TMap, 2. aktualisierte Auflage, dpunkt.Verlag, Heidelberg 2002

„Tests sind nicht die einzige Maßnahme im Qualitätsmanagement der Softwareentwicklung, aber oft die letztmögliche. Je später Fehler entdeckt werden, desto aufwändiger ist ihre Behebung, woraus sich der Umkehrschluss ableitet: Qualität muss (im ganzen Projektverlauf) implementiert und kann nicht 'eingetestet' werden.“

„Beim Testen in der Softwareentwicklung wird i. d. R. eine mehr oder minder große Fehleranzahl als 'normal' unterstellt oder akzeptiert. Hier herrscht ein erheblicher Unterschied zur Industrie: Dort werden im Prozessabschnitt 'Qualitätskontrolle' oft nur noch in Extremsituationen Fehler erwartet.“

Gerade bei KI-Systemen ist deshalb das Testen umso wichtiger, damit fatales Fehlverhalten verhindert, oder zumindest auf ein Minimum reduziert wird. Als Schätzwert im Sinne des Resourcenmanagements lässt sich festhalten: Es ist ratsam dieselben zeitlichen Ressourcen zum Testen des Systems einzuplanen, die für die Entwicklungsphase benötigt wurden. Software und KI-Systeme können durch die Einhaltung der ISO-Norm gut beurteilt werden. Weiterhin gibt auch das Vorgehen der Entwickler:innen gute Hinweise auf ein ausreichendes Testen. Wird eine Software im sog. „Pair programming“ entwickelt, gibt es eine definierte Vorgehensweise nach dem „Vier Augen Prinzip“ (ISBN 978-0-321-27865-4, Kap. 10, S. 58). Bei Hochrisikoanwendungen wird oft auch das Tripple Programming (Extreme Programming - XP) verwendet, wobei das Testen in die tägliche Software-Entwicklung integriert ist.

Anhang 3 Mögliche Eckpunkte für die Überarbeitung bestehender Richtlinien und Verwaltungsvorschriften bzw. die Erarbeitung einer Rahmenrichtlinie

Bei den folgenden Eckpunkten handelt es sich um Empfehlungen, die sich an den Grundsätzen von *Responsible AI*, dem KI-Act Entwurf sowie den vorherigen Ausführungen der Studie orientieren. Sie haben weder den Anspruch auf Vollständigkeit noch auf die zwingend richtige Systematik ihrer Auflistung. Zu Berücksichtigen dürften ferner sein die in der vorherigen Anlage 2 genannten technischen Möglichkeiten zur Umsetzung und Gewährleistung wesentlicher Anforderungen an den gesellschaftlich verantwortlichen Einsatz von KI in der Verwaltung.

1. Sinn und Zweck

Sinn und Zweck einer Richtlinie ist die Gewährleistung eines verantwortungsvollen Einsatzes von Systemen, die KI verwenden (RAI). Es wird empfohlen, die Regelungen für den Einsatz von Software in der Verwaltung der Freien und Hansestadt Hamburg (FHH) um Regeln für den Einsatz von Systemen Künstlicher Intelligenz zu erweitern. Es geht darum, Vertrauen in staatliches Handeln zu stärken.

An welcher Stelle das Thema zweckmäßig geregelt wird, ist nicht Gegenstand dieser Studie. Es muss jedoch beachtet werden, dass verschiedene bereits vorhandene Verwaltungsvorschriften die Thematik berühren. Bestehende Regelungen wie z.B. die Freigaberichtlinie vom 4. April 2005 in der Fassung vom 18. November 2010 und Verwaltungsvorschriften nach § 74 LHO bleiben unberührt bzw. sind nach einer entsprechenden Evaluierung zu überarbeiten, ggfs. neu zu konzipieren und möglicherweise um eine Rahmenrichtlinie zu ergänzen.

2. Definition

Der Begriff „Systeme, die Künstliche Intelligenz nutzen“ (umgangssprachlich KI) muss definiert werden. Es bietet sich hierfür an auf Art. 3 Ziffer 1 des KI-Act Entwurfs in Verbindung mit Anhang I in der jeweiligen Fassung zurückzugreifen. Allerdings sind die Diskussionen hierzu auf europäischer Ebene noch nicht abgeschlossen.

Alternativ und prägnanter bietet sich folgende Definition an: Systeme, die Künstliche Intelligenz nutzen, sind Systeme, die sich selbst anpassen können.

3. Einsatzbereiche

Systeme Künstlicher Intelligenz können für die Vorbereitung von Verwaltungsentscheidungen im Rahmen dieser Richtlinien eingesetzt werden. Sie unterstützen die Verwaltungsmitarbeiter:innen bei der Erledigung ihrer Aufgaben.

Es wird empfohlen, dass die Richtlinie grundsätzlich dazu ermutigt, Innovationen zu wagen. Der Fokus liegt auf der Vorbereitung von Entscheidungen als Assistenzsystem. Gleichzeitig ist darauf zu achten, dass der Einsatz Sinn macht und sich rechnet. Die Grundsätze der Wirtschaftlichkeit und Sparsamkeit sind zu beachten.

Soweit die Sorge besteht, dass Personal durch KI-Systeme ersetzt wird, ist auf die aktuellen und sich abzeichnenden Schwierigkeiten bei der Personalrekrutierung und die notwendige technische und fachliche Betreuung zu verweisen.

Im Verwaltungsakt ist zu dokumentieren, wenn bei der Entscheidungsfindung Systeme Künstlicher Intelligenz genutzt wurden. Dies dient der Transparenz.

Prognoseentscheidungen sind gesondert zu begründen. Sie sollen nicht ausschließlich auf Erkenntnisse, die durch den Einsatz von Systemen Künstlicher Intelligenz ermittelt wurden, gestützt sein. Es ist zumindest eine Plausibilitätsprüfung notwendig.

4. Beschränkungen

Der vollständig automatisierte Erlass eines Verwaltungsakts steht unter Gesetzesvorbehalt, dabei muss dem Einsatz von Systemen Künstlicher Intelligenz vom Gesetzgeber ausdrücklich zugestimmt werden. Es darf weder ein Ermessen noch ein Beurteilungsspielraum bestehen, § 35a VwVfG.

Systeme Künstlicher Intelligenz dürfen weder zur unterschwelligen Beeinflussung noch zur Bewertung der Vertrauenswürdigkeit von natürlichen Personen aufgrund ihres Verhaltens oder von Persönlichkeitsmerkmalen genutzt werden. Dies wäre Grundrechtsrelevant.

Die Bewertung der Vertrauenswürdigkeit muss im Sinne der Gleichbehandlung insgesamt ausgeschlossen sein. Hierunter fällt auch das sogenannte „Social Engineering“. Hierfür spricht sowohl die Achtung der Menschenwürde, als auch die Gefahr der Manipulation und von Fehlinterpretationen. Art. 5 Abs. 1 a) und c) KI-Act Entwurf will die Schlechterstellung bzw. Benachteiligung ausschließen.

5. Responsible Artificial Intelligence

Der Einsatz von Systemen Künstlicher Intelligenz muss nachvollziehbar, unvoreingenommen, sicher und unter Beachtung des geltenden Rechts erfolgen. Die Kontrolle und Steuerung des Einsatzes durch Verwaltungseinheiten bzw. Mitarbeiter:innen ist zu gewährleisten. (Konzept RAI)

Zu den einzelnen Kriterien gibt es teilweise Zielkonflikte und (technische) Grenzen. Diese müssen bei der Bewertung durch Verwaltungsmitarbeiter:innen vor der Entscheidungsfindung einfließen. Es geht darum, das Vertrauen in staatliches Handeln zu stärken. Bei der Entscheidungsfindung im Verwaltungsverfahren, sind deshalb die aus dem Einsatz von Systemen Künstlicher Intelligenz gewonnen Erkenntnisse stets entsprechend zu bewerten.

6. Eingangsdatensätze

Eingangsdatensätze sind in Abhängigkeit von der Komplexität und der (zukünftigen) technischen Entwicklung zu wählen.

7. Dokumentation

Es ist in Abhängigkeit von der Komplexität und der technischen Möglichkeiten zu dokumentieren, welche Datensätze zu welchen Ergebnissen führen. Diese Dokumentationspflicht dient ausschließlich der internen Kontrolle und Steuerung sowie der Bewertung. Je komplexer das System, desto weniger ist es erklärbar. Im Zielkonflikt Nachvollziehbarkeit vs. Manipulationsgefahr entscheidet sich Empfehlung 7 für die Nachvollziehbarkeit, soweit dies möglich ist. Im Hinblick auf den KI-Act ist ohnehin von Dokumentationspflichten auszugehen, der Entwurf spricht in Art. 12 KI-Act Entwurf von „Protokollierung“.

Das Wissen sollte intern bleiben, um Manipulationen zu vermeiden. Die Offenlegung gegenüber Dritten sollte nur auf gesetzlicher Grundlage erfolgen.

8. „Class Balance“

Die Datensätze müssen ausgeglichen sein und jegliche Voreingenommenheit (Bias) ausschließen. In Zweifelsfällen ist der Einsatz von Systemen Künstlicher Intelligenz zurückzustellen.

9. Fortlaufende Pflichten

Ziffer 7 und 8 sind bei der Beschaffung und während der gesamten Einsatzzeit zu gewährleisten, um Fehlentwicklungen im Betrieb Systeme Künstlicher Intelligenz auszuschließen. Die Dokumentationspflichten sind fortlaufend. Das System muss fortlaufend und angemessen kontrolliert und gewartet werden.

10. Technische Betreuung

Mit der Beschaffung ist eine Betreuung und eingreifende Korrektur für die gesamte Einsatzzeit sicherzustellen, entweder durch den Anbieter oder behördliche Stellen. Es reicht nicht, ein einwandfreies System zu beschaffen, es muss auch fortlaufend betreut und gewartet werden.

An die Eignung des Anbieters werden erhöhte Anforderungen im Sinne des Vergaberechts an Fachkunde, Leistungsfähigkeit und Zuverlässigkeit gestellt. Er muss langfristig in der Lage sein, den vertraglichen Pflichten nachzukommen. Der Stand der Technik ist zu gewährleisten, (z.B. entsprechend dem Software Engineering Body of Knowledge der IEEE Computer Society). Die erhöhten vergaberechtlichen Anforderungen an die Eignung des Anbieters können jedoch für Start-ups im Vergabeverfahren ein Ausschlusskriterium sein, hier wäre mit behördlichen Mitteln gegenzusteuern.

Der Anbieter muss Gebrauchsanweisungen (vgl. Art. 13 KI-Act Entwurf) und Testmöglichkeiten zur Verfügung stellen.

11. Menschliche Aufsicht

Die anwendende behördliche Stelle muss in der Lage sein, die grundlegende Funktionsweise des Systems zu verstehen, um als menschliche Aufsicht zu fungieren und bei Fehlentwicklungen Gegenmaßnahmen zu ergreifen. Die Beschäftigten sind entsprechend zu qualifizieren. Hilfsweise ist zentraler behördlicher Sachverstand einzubeziehen.

Hierfür müssen Ressourcen zur Verfügung gestellt werden, die bei der Wirtschaftlichkeitsbetrachtung berücksichtigt werden müssen, siehe Ziffer 3, zur menschlichen Aufsicht vgl. Art. 14. EU KI-Act Entwurf.

Behördlicherseits ist mit Beginn des Beschaffungsvorgangs und im Betrieb ein Höchstmaß an interdisziplinärer Zusammenarbeit zwischen IT-Expert:innen und Anwender:innen geboten, was entsprechende organisatorische Vorkehrungen und Arbeitsweisen erfordert.

12. Sicherheit

An die Sicherheit werden beim Einsatz von Systemen Künstlicher Intelligenz erhöhte Anforderungen gestellt. Das System ist grundsätzlich angreifbar. Gegenmaßnahmen und Sicherheitsvorkehrungen einschließlich des Risikomanagements müssen den einschlägigen technischen Vorschriften entsprechen; es wird insbesondere auf ISO 27001 verwiesen.

Der Verweis auf ISO 27001 oder Teile davon setzt Sicherheits-Standards. Die Norm spezifiziert die Anforderungen für Einrichtung, Umsetzung, Aufrechterhaltung und fortlaufende Verbesserung eines dokumentierten Informationssicherheits-Managementsystems unter Berücksichtigung des Kontextes einer Organisation.

Die EU-Kommission arbeitet derzeit an technischen Vorschriften zur Untersetzung des KI-Acts.

13. Evaluation

Die Richtlinie ist alle zwei Jahre zu evaluieren. Eine Evaluierung empfiehlt sich im Hinblick auf die sich abzeichnende dynamische technische Entwicklung und die zunehmenden praktische Erfahrungen im behördlichen Einsatz von Systemen, die Künstliche Intelligenz nutzen.

Anhang 4 Glossar

Begriff	Erklärung
KI	Unter KI versteht man im Allgemeinen einen Teilbereich der Informatik, wobei Softwareprogramme in der Lage sind, bestimmte Aufgaben zu erledigen, für die es ansonsten menschliches intelligentes Handeln braucht. Zur KI werden sowohl Verfahren des maschinellen Lernens als auch eine Reihe wissensbasierter Verfahren („Expertensysteme") verstanden. Definition im KI-Act ist im Endeffekt richtungsgebend.
Maschinelles Lernen	Teilbereich von KI, bei dem das intelligente Verhalten nicht etwa vom Menschen einprogrammiert wird, sondern über viele Beispieldaten erlernt wird. Bsp. Bilderkennung
Algorithmen	Algorithmus ist ein allgemeinerer, nicht IT-spezifischer Begriff, unter dem eine Verarbeitungsvorschrift (Abfolge von Anweisungen) verstanden wird. So sind beispielsweise auch Kochrezepte oder Notenblätter Algorithmen, da sie dazu anleiten, einzelne Zutaten oder Noten so zu verarbeiten und zu kombinieren, dass danach das gewünschte Gericht oder Musikstück erstellt werden kann. In der Informatik stellen Algorithmen eine Verarbeitungsvorschrift für Maschinen dar, die anhand dieser aus Eingaben die gewünschten Ausgaben berechnen. Ein Algorithmus folgt logischen und mathematischen Gesetzen.
Deepfake	Anwendungen von KI, bei denen Medieninhalte (v.a. Audio & Video) manipuliert oder künstlich erstellt werden.
Black Box	KI-Verfahren, die mathematisch nicht beweisbar sind. In diesen Verfahren kann keine vollständige Transparenz und Erklärbarkeit der Entscheidungsfindung des Systems gewährleistet werden. Darunter fallen bspw. künstliche neuronale Netze.
White Box	KI-Verfahren, die mathematisch beweisbar sind. In diesen Verfahren kann eine Transparenz und Erklärbarkeit gewährleistet werden. Darunter zählen bspw. Entscheidungsbäume.
Bias	Eine systematische Verzerrung, die die Entscheidungsfindung beeinflusst. Im Zusammenhang mit KI-Anwendungen häufig aufgrund unausgeglichener Trainingsdatensätze, wodurch die Entscheidungsfindung in eine bestimmte Richtung verzerrt wird. Beispielsweise die Benachteiligung bestimmter Personengruppen, die aus soziokulturellen Gründen bereits unfaire Diskriminierung erfahren, wie bspw. Frauen oder Personen ethnischer Minderheiten.
Reinforcement Learning	Lernmethode, bei der gewünschtes Verhalten über häufiges Wiederholen mit Feedback (Belohnung und Bestrafung) antrainiert wird