GESELLSCHAFT
FÜR INFORMATIK

Berlin, 31. Juli 2024

Positionspapier

der Fachgruppe Informatik und Ethik der Gesellschaft für Informatik e.V. (GI)

zum EU Artificial Intelligence Act

1. Vorwort

Dieses Positionspapier richtet sich in erster Linie an die Mitglieder der Gesellschaft für Informatik und darüber hinaus an Informatikfachleute, die IT-Systeme entwerfen, herstellen, betreiben oder verwenden. Darunter fallen auch Fachleute, die interdisziplinär an der Schnittstelle zur Informatik arbeiten. Es soll eine Orientierung in der gegenwärtigen Debatte rund um die Ethik und Regulierung von Künstlicher Intelligenz anbieten.
Als Fachgruppe Informatik und Ethik im Fachbereich Informatik und Gesellschaft der Gesellschaft für Informatik e.V. beobachten wir sehr aufmerksam und genau die Entwicklungen im Umfeld der Informatik hinsichtlich Regulation und möglicher Auswirkungen von Informatik-Systemen auf Gesellschaft, Wissenschaft und Wirtschaft. Der EU Artificial Intelligence Act (deutsch: EU-Verordnung über künstliche Intelligenz) oder im Folgenden abgekürzt mit AIA stellt dabei ein ganz entscheidendes Regelwerk dar, das die vermeintlich unregulierte Entwicklung der Schlüsseltechnologie „Künstliche Intelligenz“ (KI) kritisch und rechtssicher begleiten soll. Neben wirtschaftlichen und europarechtlichen Erwägungen sind es insbesondere die moralisch begründeten europäischen Werte, die eine gemeinwohlverträgliche Entwicklung im Bereich KI fördern soll in Abgrenzung zu Entwicklungen in konkurrierenden Wirtschaftsräumen z. B. in Nordamerika und Asien.
Dieses Positionspapier beleuchtet die moralische Dimension des AIA sowie den Entstehungskontext und das regulatorische Umfeld. Es soll eine kontextuelle Einordnung bieten, um letztendlich zur Beteiligung am Diskurs anzuregen, ganz im Einklang mit den Ethischen Leitlinien der GI [1].
Wir beziehen uns in unseren Ausführungen auf die im Amtsblatt der Europäischen Union vom 12. Juli 2024 veröffentlicht rechtsverbindliche Dokumentenversion vom 13. Juni 2024 mit dem Verordnungskennzeichen 2024/1689 (Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828) [2]. Darüber hinaus nutzten wir auch zur leichteren Navigation durch den Gesetzestext den „AI Act Explorer“ (Basis war jeweils die aktuellste verfügbare Dokumentenversion, deren Aktualisierung mit der Berichtigung/Corrigendum vom 17.04.2024 endet, führend in unseren Ausführungen war jedoch stets die rechtsverbindliche Dokumentenversion vom 13. Juni 2024) [3].

2. Entwicklung des AIA

In den ersten Verordnungsvorschlag der EU-Kommission vom 21. April 2021 sind einige Vorarbeiten auf EU-Ebene eingeflossen. Dazu zählen insbesondere die Ethik-Leitlinien für eine vertrauenswürdige KI, erstellt durch die Hochrangige Expertengruppe für KI [4], das EU-Weißbuch zur KI [5] sowie die EU-Digitalstrategie [6].

Wesentliche Grundlagenarbeit wurde auch auf nationaler Ebene erbracht, z. B. in Deutschland durch das Gutachten der Datenethikkommission [7], dem Abschlussbericht der Enquete-Kommission KI [8] oder der 2. Ausgabe der Normungsroadmap KI, kurz NRM KI [9].
Nach den erfolgreich abgeschlossenen Trilog-Verhandlungen am 6. und 8. Dezember 2023, der Einigung des Ausschusses der Ständigen Vertreter der Mitgliedstaaten über Ausformulierungen einiger Detailpassagen insbesondere in den Bereichen Generativer KI und Biometrischer Überwachung am 2. Februar 2024, der Billigung durch die Ausschüsse für Binnenmarkt (IMCO) und für bürgerliche Freiheiten, Justiz und Inneres (LIBE) am 13. Februar 2024, der Verabschiedung mit 523 Ja-Stimmen zu 46 Nein-Stimmen bei 49 Enthaltungen durch das EU-Parlament am 13. März 2024, einer abschließenden Prüfung von Rechts- und Sprachsachverständigen, der Verabschiedung durch den Rat der Europäischen Union am 21. Mai 2024 und der Veröffentlichung im Amtsblatt der Europäischen Union (ABI) am 12. Juli 2024, ist die Verordnung über Künstliche Intelligenz nach einer Einspruchsfrist von 20 Tage im Anschluss ihrer Veröffentlichung im ABI am 1. August 2024 in Kraft getreten.
Der Weg über die letzten Jahre hin zu einer Regulierung innerhalb der EU und die Entstehung eines weltweit ersten umfassenden Rechtsrahmen für die Nutzung und Anwendung von Künstlicher Intelligenz (KI) wurde konsequent, demokratisch und letztendlich erfolgreich beschritten.

Der AIA steht nicht für sich allein

Der AIA hat mehr oder weniger starke Bezüge zu weiteren bereits existierenden oder aber auch in Arbeit befindlichen Gesetzen der EU. Grundsätzlich unterscheidet die europäische Gesetzgebung zwischen Verordnungen und Richtlinien. Im Gegensatz zu Verordnungen, die unmittelbar nach Inkrafttreten in den Mitgliedstaaten gelten, sind Richtlinien nicht unmittelbar in den Mitgliedstaaten gültig. Sie müssen zunächst von einzelnen Mitgliedstaaten in nationales Recht umgesetzt werden, um Gültigkeit zu erlangen.
Starke Bezüge zum AIA haben beispielsweise die EU-Grundrechtecharta, der New Legislative Framework (NLF, Marktüberwachung und Rahmen für die Akkreditierung von Produktkonformitätsprüforganisationen), die EU-Datenschutzgrundverordnung oder der Digital Services Act (DSA) zusammen mit dem Digital Markets Act (DMA), zur Regulierung der Aktivitäten von Anbietern digitaler Dienste und Beschränkung marktbeherrschender Digitalkonzerne, sogenannte Gatekeeper, und Ihrer zentralen Online-Plattformen. Ebenso relevante Bezüge können zu sektorspezifischen Regularien bestehen, die vereinzelt nicht immer redundanz- oder gar widerspruchsfrei sein müssen, wie z. B. teilweise Überschneidungen mit der Medizinprodukteverordnung (MDR). Es wirkt grundsätzlich bei EU-Gesetzten im Zweifels- und Einzelfall immer der höhere Schutzbedarf bzw. die stärkere Anforderung, um Doppelaufwände und Überregulierung durch die Anwendung mehrerer konkurrierender Verordnungen oder Richtlinien zu ein und demselben Aspekt zu vermeiden.

3. Kurzüberblick: Status Quo

Der grundsätzliche Aufbau des AIA ist analog der üblichen Gliederung anderer umfassender EU-Rechtsakte strukturiert und zum einen römisch durchnummeriert nach Kapitel I bis XIII und zum anderen über das gesamte Dokument und alle Kapitel hinweg durchgehend nach Artikel von 1 bis 113 arabisch durchnummeriert. Die Anhänge sind wieder in römischen Ziffern von I bis XIII durchnummeriert.
Im Folgenden ein kleiner Überblick zu aus unserer Sicht besonders erwähnenswerten Kapiteln, Artikeln und deren Inhalte, sowie referenzierenden Anhängen, die teilweise auch für unsere weitere Diskussion relevant sind:

Kapitel I, Artikel 1 nennt die Zwecke des AIA

Die KI-Verordnung beruht auf den Werten und Grundrechten der EU, ist menschenzentriert, risikobasiert und kritikalitätsbewertend und verfolgt unter anderem folgende Ziele:

- Verankerung europäischer Werte in KI-Systemen
- Gewährleistung der EU-Grundrechte
- Vertrauen in KI-gestützte Lösungen in der gesamten Gesellschaft erhöhen und gleichzeitig Unternehmen Anreize geben, diese zu entwickeln
- Stärkung von Innovation im Bereich Künstliche Intelligenz
- Rechtssicherheit zur Förderung von Innovation und Investition in KI
- Definition von Künstlicher Intelligenz und KI-Systemen, angelehnt an die der OECD [10]

Kapitel I, Artikel 2 legt Anwendungsbereiche und betroffene Akteure fest

- #1 Anbieter weltweit, die KI-Systeme in der EU in Verkehr oder in Betrieb bringen, … unabhängig davon, ob diese Anbieter in der Union oder in einem Drittland niedergelassen oder ansässig sind;
- #2 Betreiber von KI-Systemen, die ihren Sitz in der Union haben oder dort ansässig sind sowie
- #3 Anbieter und Betreiber weltweit, wenn das vom KI-System hervorgebrachte Ergebnis in der EU verwendet wird bzw. Wirkung entfaltet

Weitere in den Anwendungsbereich eingeschlossene Akteure sind

- Importeure und Händler von KI-Systemen,
- Produkthersteller, die ein KI-System zusammen mit ihrem Produkt und unter ihrem eigenen Namen oder ihrer eigenen Marke in Verkehr bringen oder in Betrieb nehmen,
- Bevollmächtigte von Dienstleistungserbringern, die nicht in der Union niedergelassen sind;
- Betroffene, die in der Union ansässig sind.

Kapitel I, Artikel 3 definiert zentrale Begriffe

Die Definition von Künstlicher Intelligenz und KI-Systemen wird im AIA angelehnt an die der OECD [10].
Ein "KI-System" ist demnach " ein maschinengestütztes System, das für einen in wechselndem Maße autonomen Betrieb ausgelegt ist, das nach seiner Einführung anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ergebnisse wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen hervorgebracht werden, die physische oder virtuelle Umgebungen beeinflussen können. Englischer Originaltext: 'AI system' means a machine-based system designed to operate with varying levels of autonomy, that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. (Kapitel 1 Artikel 3)

Kapitel II, Artikel 5 behandelt verbotene Praktiken

Verboten sind beispielweise soziale Bewertungssysteme, die Manipulation von Verhalten, Emotionserkennungssysteme am Arbeitsplatz und in der Bildung, KI-Systeme, die Menschen auf der Grundlage sensibler Daten (wie ihrer politischen Gesinnung oder sexuellen Orientierung) kategorisieren, und es gilt ein teilweises Verbot für Predictive Policing Systeme. Die endgültige Liste ist länger als im ursprünglichen Entwurf vorgesehen.

Kapitel III beschreibt hochriskante KI-Systeme.

Hochrisiko-KI-Systeme sind KI-Systeme aus einem der folgenden Einsatzbereiche: Biometrische Daten, kritische Infrastruktur, allgemeine und berufliche Bildung, Bewerberauswahl und Bewerbermanagement, Rechtsdurchsetzung und Asylverfahren." (aus Anhang III)
Hochrisiko-KI-Systeme müssen eine Menge an Anforderungen und Verpflichtungen erfüllen, z. B. müssen Konformitätsbewertungen durchgeführt werden sowie ein Qualitäts- und Risikomanagement aufgebaut werden. Technische Dokumentation muss vor Inbetriebnahme erstellt werden und auf dem neuesten Stand gehalten werden. (Kapitel III, Artikel 11, Absatz 1)
In Kapitel VIII Artikel 71 wird die EU-Datenbank für die Erfassung von Hochrisiko-KI-Systeme erklärt, die im Anhang III gelistet sind.

Kapitel V nennt besondere Pflichten für KI-Modelle mit allgemeinem Verwendungszweck (general purpose AI-model)

Als Reaktion auf generative KI-Modelle wurden weitere Regeln in den AIA aufgenommen. Diese betreffen Transparenzpflichten und Dokumentationspflichten wie z. B. Auskunft über verwendete Trainingsdaten sowie Trainings- und Testverfahren. (Anhang IV, Absatz 2 d) Noch härteren Vorgaben unterliegen KI-Modelle mit allgemeinem Verwendungszweck mit

systemischen Risiken. Darunter versteht man ein Modell, das „über Fähigkeiten mit hohem Wirkungsgrad [verfügt], die mithilfe geeigneter technischer Instrumente und Methoden, einschließlich Indikatoren und Benchmarks, bewertet werden“ (Kapitel V, Artikel 51, Absatz 1a) und wenn die „ kummulierte Menge des für sein Training verwendeten Berechnungen , gemessen in FLOPS, mehr als 10^25 Flops beträgt.“(Kapitel V, Artikel 51, Absatz 2)
Die Anbieter müssen schwerwiegende Vorfälle sowie die Energieeffizienz der Systeme melden. Außerdem wird auf EU-Ebene ein sogenanntes KI-Office eingerichtet, dessen Aufgabe es ist, derartige Modelle zu überwachen und die gemeinsamen Vorschriften in allen Mitgliedstaaten durchzusetzen. Es wird also deutlich zwischen kleineren und sehr großen Sprachmodellen unterschieden.

Kapitel XII, Artikel 99 legt Strafzahlungen fest

- bei Nichtkonformität (Absatz (3), Missachtung Artikel 5 „Verbotene KI-Praktiken“) bis zu 7 % des Umsatzes oder 35 Mio. €,
- bei Verstoß gegen Bestimmungen für Akteure oder notifizierte Stellen (Absatz (4)) bis zu 3 % oder 15 Mio. € und
- bei falschen, unvollständigen oder irreführenden Angaben (Absatz (5)) bis zu 1% oder 7,5 Mio. €.

Das Schädigungspotential definiert die Anforderungen

Zentrales Regulierungsinstrument ist ein risikobasierter Ansatz für Algorithmische Systeme und unterscheidet zwischen unannehmbarem Risiko und einem damit verbundenen Verbot, Hochrisiko mit entsprechenden Konformitätspflichten, beschränktes Risiko mit Transparenzpflichten, sowie kein bzw. geringes Risiko ohne Auflagen. Das Prinzip ist dabei generell: je höher das Schädigungspotential, das Risiko bzw. die Kritikalität, desto höher sind die Anforderungen an spezifische Kriterien und damit an die Regulierung.

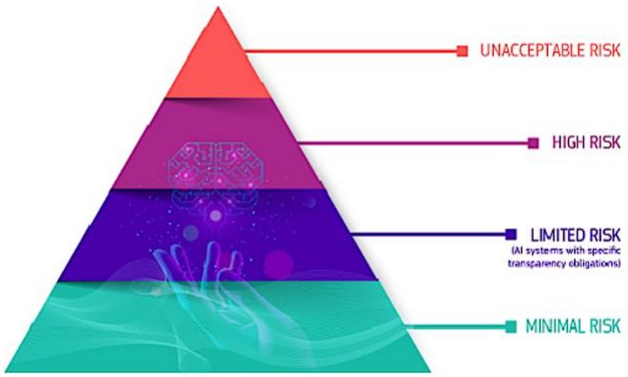

Abb. 1: Europäische Kommission,
https://digital-strategy.ec.europa.eu/de/policies/regulatory-framework-ai

Wie man solch spezifische Kriterien anwenden könnte, findet man beispielsweise in der VDE SPEC 90012, 2022 (VCIO), VCIO based description of systems for AI trustworthiness Characterisation“ (VDE SPEC 90012, 2022) mit den dort enthaltenen Werten:

1.) Transparenz,
2.) Rechenschaftspflicht,
3.) Datenschutz,
4.) Fairness,
5.) Verlässlichkeit,

oder entsprechend in der NRM KI (Deutsche Normungsroadmap Künstliche Intelligenz Ausgabe 2, 2022), angelehnt sowohl an die „Ethik-Leitlinien für eine vertrauenswürdige KI“, erstellt durch die Hochrangige Expertengruppe für KI (HLEG, 8. April 2019), als auch an den Abschlussbericht der Enquete-Kommission KI (28. Oktober 2020), mit dem dort in der NRM KI beschriebenen Werte-Setup:
1.) Vorrang menschlicher Aufsicht von KI-Systemen sowie die Einhaltung und Sicherstellung von Grundrechten,
2.) Technische Robustheit und Sicherheit,
3.) Schutz der Privatsphäre und Datenqualitätsmanagement,
4.) Transparenz, Nachvollziehbarkeit und Erklärbarkeit.
5.) Fairness, Nichtdiskriminierung und Vielfalt,
6.) Gesellschaftliches und ökologisches Wohlergehen,
7.) Rechenschaftspflicht

4. Einordnung des AIA

Befürworter des AIA halten es für richtig, KI frühzeitig zu regulieren und unerwünschte Anwendungen schnellstmöglich zu verbieten und sehen dies im AIA geregelt. Die Kritik kommt demgegenüber von zwei Seiten; so gibt es die wirtschaftsliberale Position, welche die Meinung vertritt, dass eine Regulierung zum jetzigen Zeitpunkt zu früh komme und eine zu starke Regulierung oder gar eine Überregulierung innovationshemmend und gerade für die europäische KI-Industrie schädlich sein könnte.
Auf der anderen Seite gehen Bürgerrechtlern die Ausnahmeregelungen und Befugnisse bezüglich automatischer Gesichtserkennung für Strafvollzugsbehörden und andere staatliche Stellen zu weit. Das KI-Gesetz, ursprünglich politisch verabschiedet Anfang Dezember 2023 im Trilog, wurde dahingehend noch in letzter Minute geändert. Die Verwendung von biometrischen Fernidentifizierungssystemen in Echtzeit in öffentlich zugänglichen Räumen ist zwar grundsätzlich verboten, es gibt aber eine Liste von Ausnahmen und damit einige Schwachstellen. Auch der Richtervorbehalt wurde aufgeweicht [11, 12].

Eine grundrechtliche Folgenabschätzung und öffentliche Transparenz beim Einsatz von Hochrisiko-KI durch Behörden sieht das Gesetz erst dank einer intensiven Lobbyarbeit zivilgesellschaftlicher Organisationen vor [13].
Aus unserer Sicht birgt in diesem Kontext die europaweit beliebte Ausrede der „Nationalen Sicherheit" das größte Risiko, den AIA komplett auszuhebeln, aber auch weitere kritische Stellen im AIA sind öffentlich zu diskutieren.

Kritische Stelle „Militär-/Verteidigungszwecke" oder „Nationalen Sicherheit"

KI-Systeme im Kontext militärische Zwecke, Verteidigungszwecke oder Zwecke der nationalen Sicherheit sind grundsätzlich vom Anwendungsbereich dieser Verordnung ausgenommen, unabhängig davon, ob es sich um eine öffentliche oder private Einrichtung handelt, die in diesem Kontext tätig ist.
Was den Zweck der Nationalen Sicherheit und damit den Ausschluss des Anwendungsbereiches des AIA betrifft, liegt nach Artikel 4 Absatz 2 des Vertrages über die Europäische Union (EUV, Zitat daraus: „Insbesondere die nationale Sicherheit fällt weiterhin in die alleinige Verantwortung der einzelnen Mitgliedstaaten") die Feststellung einer Gefährdung der Nationalen Sicherheit einzig und allein in der Zuständigkeit der einzelnen Mitgliedstaaten.
Damit ist es erst einmal jedem Mitgliedstaat möglich, jederzeit zu jedem Thema die Nationale Sicherheit festzustellen, jedes beliebige KI-System dem Anwendungsbereich des AIA zu entziehen, sei es verboten oder als Hochrisiko bezeichnet. Die Geschichte lehrt uns, dass davon auch Gebrauch gemacht werden wird, unbenommen der Möglichkeit, dass natürlich auch jederzeit und zu jedem Thema Klage beim Europäischen Gerichtshof eingereicht werden kann.

Kritische Ausnahmen bezüglich biometrischer Echtzeit-Fernidentifizierung

Trotz klarem Verbot biometrischer Echtzeit-Fernidentifizierung im öffentlichen Raum für die Strafverfolgung und auch ohne das Feststellen der Nationalen Sicherheit sind Ausnahmen möglich. Die ursprünglich vom Parlament geforderten Verbote von Echtzeitüberwachung wurden im Trilog erheblich aufgeweicht. Spezifisch notwendige Zwecke, wie die Suche nach Opfern oder vermissten Personen, der Identifizierung von Verdächtigen bei schweren Straftaten, sind explizit vom Verbot ausgenommen und somit erlaubt. Dies ermöglicht sozusagen eine Massenüberwachung durch die Hintertür.
Angesichts der jüngsten Festnahme einer RAF-Terroristin, die teilweise einer von Journalisten angewendeten Gesichtserkennungssoftware zu verdanken ist, beklagt die Polizei, dass biometrische Gesichtserkennung bisher nicht in der Polizeiarbeit zur Gefahrenabwehr und Strafverfolgung erlaubt wurde. Dies zeigt deutlich, wie unter dem Deckmantel der Nationalen Sicherheit, biometrische Gesichtserkennung als unbedingt notwendige Technologie gefordert wird [14].

Fragwürdige Emotionserkennung

Ebenso gibt es eine Schwachstelle bei KI-Systeme zur Emotionserkennung, die generell am Arbeitsplatz oder in Bildungseinrichtungen verboten sind, jedoch aus medizinischen oder Sicherheitsgründen erlaubt sind (Artikel 5, Absatz (1), f). Zudem sind KI-Systeme erlaubt, um die Emotionen von Asylbewerbenden zu „erkennen“ [12].
Dies würde die Praxis von Behörden ergänzen, die beispielsweise KI-Systeme zur Dialekterkennung und somit zur Einschätzung der Wahrhaftigkeit der Fluchtgeschichte nutzen, wie in einem Pilotprojekt des Bundesamts für Migration und Flüchtlinge erprobt wurde [15].

Einstufung als Hochrisiko moralisch problematisch

Weitere Schwachstellen öffnen sich z. B. auch durch das Mitspracherecht von KI-Entwicklern bei der Frage, ob ihre Systeme als Hochrisiko-KI gelten.
Außerdem gelten in den weiteren Bereichen Strafverfolgung und Migration ebenso Ausnahmen für Hochrisiko-KI-Systeme, durch die Behörden wesentliche Kernbestimmungen des Gesetzes umgehen können.
Der Kompromiss zur KI-Verordnung offenbart einen systemischen Fehler bei der EU-Gesetzgebung:
Nationale Regierungen und lobbyierende Hersteller von Techniken für die Strafverfolgung haben einen unverhältnismäßig großen Einfluss, so dass staatliche und wirtschaftliche Interessen gegenüber dem öffentlichen Interesse und dem Schutz der Menschenrechte überwiegen.
Durch die Risiko-Einstufung können bestimmte aus moralischer Sicht zu ächtende KI-Systeme zur Hochrisiko-Technik „geadelt“ werden. Damit werden indiskutable Vorschläge Gegenstand einer Diskussion wie beispielsweise der militärische Einsatz und die menschenrechtsverletzenden Praktiken im Umgang mit geflüchteten Personen.

Noch erheblicher Forschungsbedarf zu Transparenzpflichten bestimmter KI-Systeme

„Anbieter von KI-Systemen, einschließlich KI-Systeme mit allgemeinem Verwendungszweck, die synthetische Audio-, Bild-, Video- oder Textinhalte erzeugen, stellen sicher, dass die - Ergebnisse des KI-Systems in einem maschinenlesbaren Format gekennzeichnet sind und als künstlich erzeugt oder manipuliert erkennbar sind.“
„Die Anbieter - sorgen dafür, dass – soweit technisch möglich - ihre technischen Lösungen wirksam, interoperabel, belastbar und zuverlässig sind und berücksichtigen dabei die Besonderheiten und Beschränkungen der verschiedenen Arten von Inhalten, die Umsetzungskosten und den allgemein anerkannten Stand der Technik, was sich in den einschlägigen technischen Normen niederschlagen kann“ (Kapitel IV, Artikel 50, Absatz (2).
Hier ist wohl noch erheblicher Forschungsbedarf notwendig. Es gibt Ansätze KI-generierte Inhalte mit Wasserzeichen zu kennzeichnen oder Metadaten einzuführen. Derzeit genutzte Wasserzeichen werden als nicht sicher

eingeschätzt. Es ist technisch anspruchsvoll, dass Wasserzeichen erhalten bleiben, wenn Bilder skaliert oder komprimiert werden [16].

Keine Berücksichtigung immaterieller Nachhaltigkeit

Eine explizite Einbeziehung von Nachhaltigkeit, die sich neben materiellen Gütern auch auf immaterielle / soziale Güter bezieht (zum Beispiel sogenannte Wissensgüter und kulturelle Artefakte digital repräsentiert als Text, Bild, Audio, Video, Software oder algorithmische Systeme), haben wir leider vermisst. Zumindest ein klares Bekenntnis zu einem bewussten Umgang mit materiellen sowie immateriellen Ressourcen in der Weise, dass deren heutige langfristig orientierte Erstellung, freie Verwendung und kollaborative Weiterentwicklung die Bedürfnisse kommender Generationen nicht beeinträchtigt und ein gutes Leben heute und in Zukunft erlaubt, sollte dringend auch in den AIA Einzug finden.
Diverse Nachhaltigkeitsmodelle mit dem klassischen Drei- oder Vier-Säulen Modell (Ökologie/Ökonomie/Soziales/Option 4 Kultur), den Sustainable Development Goals (SDGs) der UN bzw. deren Überführung in die Europäische Nachhaltigkeitsstrategie inkl. Green Deal oder Anleihen des Club of Rome, wären denkbar gewesen, auch im AIA sinnvoll im Kontext KI zu verankern. Dies ist schon bei der Digitalisierung versäumt worden, wurde aber zumindest auf bundesdeutscher Ebene durch den „Wissenschaftlichen Beirat der Bundesregierung Globale Umweltveränderungen“ (WBGU) in seinem Hauptgutachten „Unsere gemeinsame digitale Zukunft“ von 2019 umfassend nachgeholt [17].
Nachhaltigkeits-Potentiale von KI sind in allen Bereichen durchaus gegeben. Ökologisch bedeutet den Schutz der natürlichen Umwelt und Ressourcen, ökonomisch bedeutet eine langfristig tragfähige und faire wirtschaftliche Entwicklung und sozial bedeutet die Förderung von Gerechtigkeit, Gleichheit und der Lebensqualität aller Menschen.
Soziale Nachhaltigkeit im Kontext der Künstlichen Intelligenz zielt darauf ab, eine zukunftsfähige, global gerechte und lebenswerte Gesellschaft für alle sozialen Gruppen zu schaffen. Dies beinhaltet die Erfüllung grundlegender menschlicher Bedürfnisse wie Nahrung und Wohnen, die Gewährleistung guter Lebensbedingungen durch Bildung und die Förderung von Chancengleichheit und sozialer Teilhabe.
"Aus sozialer Perspektive sollte eine nachhaltige KI beispielsweise die Privatsphäre von Menschen und ihre Selbstbestimmung mit Blick auf personenbezogene Daten schützen, ihre Autonomie und Handlungsfreiheit gegenüber KI-Systemen stärken und im Sinne des Allgemeinwohls Diskriminierung und Bias vermeiden. Mit Blick auf eine sozial nachhaltige KI setzt dies beispielsweise voraus, dass ein ausreichendes Maß an Transparenz, Nicht-Diskriminierung und Fairness, menschliche Aufsicht sowie Selbstbestimmung und Datenschutz gewährleistet ist“ [18].
Dem Potential bei der ökologischen Nachhaltigkeit, beispielsweise in den Bereichen der Kreislaufwirtschaft und der Energiewende, steht der ungeheure

Ressourcenverbrauch von KI-Systemen entgegen, sowohl bei der Herstellung der Hardware, dem Betrieb der riesigen Serverfarmen besonders beim Training und der Entsorgung der am Ende des Lebenszyklus obsolet gewordenen Technik.
Bezüglich ökonomischer Nachhaltigkeit gibt es Diskussionen hinsichtlich der Marktmacht und Machtkonzentration einzelner Tech-Giganten. Letztere bestimmen häufig, für welchen Zweck KI-Systeme und Technologien entwickelt werden. Wettbewerbsverzerrungen, Wettbewerbsvorteile z. B. durch nicht frei zugängliche Datenpools und einen damit gegebenen Innovationsvorsprung schränken aber die Auswahlmöglichkeiten von KI-Systemen für Nutzer*innen weiter ein. Bezüglich ökonomischer Nachhaltigkeit wären eine Marktvielfalt sowie eine Ausschöpfung des Innovationspotentials wünschenswert. Dazu sind faire Zugangsmöglichkeiten zu den erforderlichen Datenquellen und Tools für möglichst viele Akteure aus unterschiedlichsten Wirtschaftsbereichen zwingend notwendig. Die Entwicklung von KI-Systemen sollte global besser verteilt sein. Es müssen z. B. Rahmenbedingungen geschaffen werden, dass insbesondere KMUs vergünstigten Zugang zu KI-Technologien erhalten.
KI-basierte Technologien berühren also mehrere Nachhaltigkeitsdimensionen. Am Beispiel der Automatisierung von Prozessen der Digitalwirtschaft zeigt sich, dass der Einsatz von KI-basierten Technologien Auswirkungen auf die Art der Arbeitsplätze und die Arbeitsbedingungen der Arbeitnehmer*innen sowie die Umwelt haben. Bestimmte Jobprofile können und werden ganz oder teilweise wegfallen, auch wenn durchaus neue Arbeitsplätze geschaffen werden. Dabei werden einerseits neue Jobprofile entstehen, die in der Regel eine höhere Qualifikation erfordern, also Möglichkeiten zur Weiterentwicklung bzw. persönlichen Entfaltung und oftmals besseres Gehalt bieten. Andererseits erhalten Arbeitnehmer*innen besonders im Niedriglohnsegment zumeist im Globalen Süden, wie beispielsweise Clickworker, die bei der Datenaufbereitung arbeiten, häufig eine sehr geringe Entlohnung und haben kaum Weiterbildungsmöglichkeiten. Für eine ökonomisch und sozial nachhaltige Perspektive müssen faire Löhne entlang der gesamten Wertschöpfungskette der KI-Entwicklung sichergestellt sein. Nehmen wir zusätzlich zur Software auch die KI-Hardware in den Blick, treten die Umweltauswirkungen deutlich zu Tage, von der Förderung der benötigten Materialien über den immensen Energieverbrauch von KI-Systemen bis zur Entsorgung der in immer kürzerer Zeit obsolet gewordenen Hardware. Unternehmen, aber auch der Staat sollten vielfältige Fort- und Weiterbildungsmöglichkeiten für KI-gestützte Tätigkeiten anbieten, sowie auf menschenwürdige Arbeitsplätze, umweltschützende Produktion und soziale Gerechtigkeit entlang der gesamten KI-Entwicklung achten [19].
Gleichwohl schlägt der AIA einen pragmatischen Weg mit positiver Wirkung für Wirtschaft, Wissenschaft und Zivilgesellschaft ein. Zentrale Faktoren hierfür werden im Folgenden vorgestellt.

Vereinfachte Betrachtung als Algorithmisches System

Die Einschätzung, ob ein Algorithmisches System oder eine Gesamtanwendung an irgendeiner Stelle KI enthält, ist alles andere als trivial. Unter Umständen auch tief verborgen, in einem von vielen Modulen oder Komponenten, vielleicht zugekauft oder von einem Drittdienstleister nach Auftrag gefertigt.
Auch das Verständnis, was ist überhaupt eine KI-Methode (z. B. Problemlösen, Optimieren, Planen, Entscheiden, …) und dessen Zusammenhang mit KI-Fähigkeiten (z. B. Wahrnehmen, Verarbeiten, Handeln, …) erfordert eine fundierte Ausbildung und vertieftes Wissen.
Die Definition eines KI-Systems nach OECD mag zwar vernünftig sein und auch einen guten Kompromiss darstellen, um nicht einen weiteren redundanten Versuch der Definition von KI zu generieren. Jedoch damit allein die Eingangsfrage der Relevanz eines Objektes in Bezug auf den AIA hinreichend zu beantworten, ist unrealistisch.
Die Datenethikkommission empfiehlt in ihrem Gutachten, auf die hinsichtlich der Regulierungsziele unnötige Unterscheidung zwischen „IT-mit-KI-System“ und „IT-ohne-KI-System“ zu verzichten, da es im Kern um den Output und nicht primär um die Funktionsweise Algorithmischer Systeme gehe. Dies ist unserer Ansicht nach zu empfehlen, da die mitunter sehr komplexe und zeitaufwendige Einordnung in Systeme mit oder ohne KI-Anteil entfällt. Es werden alle Algorithmischen Systeme betrachtet, denen ein relevantes Schädigungspotential innewohnt, unabhängig von einer Beurteilung, ob KI enthalten ist, oder nicht. Ein objektiv existierendes Schädigungspotential, z. B. durch eine spezifische Fehlfunktion eines autonom gelenkten Fahrzeuges, wird nicht weniger gefährlich oder aber gefährlicher durch die Existenz oder Abstinenz eines KI-basierten Codeanteils.

Ein angemessen breiter Anwendungsbereich

Durch den Kompromiss der Definition von KI-Systemen (Kapitel 1, Artikel 3), angelehnt an die bereits etablierte Begriffserklärung der OECD, wird nicht nur eine breite Akzeptanz zum gemeinsamen Verständnis von KI-Systemen geschaffen, auch erlaubt diese „Deutung“ aus unserer Sicht, entgegen auch anderer Meinungen, eine hinreichend breite Anwendung des AIA auf entsprechende Algorithmische Systeme (Artikel 2).
Eine ganz besonders gelungene Formulierung verhindert eine Umgehung des AIA, auch wenn Anbieter und Betreiber von KI-Systemen ihren Sitz in ein Drittland verlegen oder schon dort ansässig sind. Jeglicher auch in einem Drittland von einem KI-System erzeugter Output, der in der Europäischen Union verwendet wird oder Wirkung entfaltet, fällt unter den Anwendungsbereich des AIA.

Lebensnahe Untergliederung des relevanten Risikobereiches

Die Risiko- bzw. Kritikalitätsstufen gliedern sich im AIA nach unvertretbarem (verboten), hohem (Konformitätspflichten), beschränktem

(Transparenzpflichten) und minimalem Risiko (keinerlei Auflagen). Die beiden Stufen mit aktiven Pflichten (hohes und beschränktes Risiko) könnte man auch analog dem Gutachten der Datenethikkommission in drei Stufen differenzieren. Die Dreiteilung und Unterscheidung der die entsprechenden Schutzbedarfe mitigierenden Kontrollverfahren entspräche einer bereits in der Fehler-Ursachen-Wirkungs-Forschung etablierten intuitiven Logik: Ex-post-Kontrollen bei gewissen Schädigungspotential (beschränktes Risiko nach AIA), Ex-ante-Kontrollen bei regelmäßigem oder deutlichem Schädigungspotential (entsprechender Ex-ante-Anteil des hohen Risikos nach AIA) und kontinuierliche Live-Kontrollen bei erheblichem Schädigungspotential (entsprechender Live-Anteil des hohen Risikos nach AIA). Wichtig ist, dass sich wirksame Kontrollverfahren nun etablieren und intensiv weiter erforscht werden [20].

Orientierung bei der Einordnung als Hochrisiko-KI-Systeme

Grundsätzlich ist es eine gute Festlegung, explizit Hochrisiko-KI-Systeme sowohl in einem Anhang zur Verordnung (Anhang III), als auch parallel in einer eigens dafür bereitgestellten Hochrisiko-Datenbank aufzuführen und auch Verfahren zu etablieren, wie diese Inhalte möglichst aktuell zu halten sind (Artikel 7).
Bei exponentiellen Technologien ist nicht nur meist KI im Spiel, sondern oft ist es auch die Komponente, die eine Technologie erst exponentiell wachsen bzw. sich entwickeln lässt. Vor diesem Hintergrund werden viele neue KI-Entwicklungen und auch viele neue Hochrisiko-KI-Systeme viel zu schnell entstehen, um diese in einer angemessenen Zeit erst einmal als solche zu identifizieren und anschließend in einem Anhang zum AIA und in die Hochrisiko-Datenbank aufzunehmen.
Diese Lücke könnte mit einem generischen Ansatz zur objektiven, reproduzierbaren und nachvollziehbaren möglichst simplen arithmetischen Berechnung des Grades eines Risikos für neue KI-Systeme geschlossen werden. Dabei würde sich das Schädigungspotential in einem spezifischen Anwendungskontext im Wesentlichen klassisch aus der Eintrittswahrscheinlichkeit und der Schadenshöhe ableiten. Die Schadenshöhe berücksichtigt dabei umfassend KI-Spezifika wie z. B. die Schadensart (z. B. Individuen / Gruppen / Gesamtgesellschaft und Organisation / Branche / Gesamtwirtschaft), die Schadenstiefe bzw. -Umfang, Reversibilität, mögliche begünstigende und erschwerende Faktoren, beschlossene bzw. mögliche Gegenmaßnahmen, … Eine entsprechende Arithmetik berechnet noch in der Designphase eines neuen KI-Systems aus den Parametern und optionalen Gewichten das Schadenspotential, zum Beispiel normiert auf eine Skala von 0-100. Die Risiko- bzw. Kritikalitätsstufen könnten dann Intervallen auf der Skala zugeordnet werden, beispielsweise unvertretbares Risiko (verboten) dem Score 81-100, hohes Risiko (Konformitätspflichten) 41-80, beschränktes Risiko (Transparenzpflichten) 21-40 und minimales Risiko (keinerlei Auflagen) 0-20. Würde man die beiden Stufen mit aktiven Pflichten (hohes und beschränktes

Risiko) analog AIA wieder entsprechend dem Gutachten der Datenethikkommission in drei Stufen differenzieren bzw. hohes Risiko unterteilen in regelmäßiges / deutliches Schädigungspotential (Ex-ante-Kontrollen) und erhebliches Schädigungspotential (kontinuierliche Live-Kontrollen), wäre eine Zuordnung in die Intervalle 41-60 (Ex-ante-Kontrollen) und 61-80 (kontinuierliche Live-Kontrollen) denkbar.
Dieser generische Ansatz samt Arithmetik könnte Organisationen recht einfach bei der objektiven, reproduzierbaren und nachvollziehbaren Einordnung Ihrer neuen oder noch im Design oder der Entwicklung befindlichen KI-Systeme helfen, aber auch dem zentralen EU-KI-Office bei deren hoheitlichen Aufgabe der transparenten und unwillkürlichen Identifikation und Erfassung von Hochrisiko-KI-Systeme in den Anhang III wie auch in die Datenbank für Hochrisiko-KI-Systeme.
Die Aufgabe der Entwicklung eines solchen generischen Ansatz zur objektiven, reproduzierbaren und nachvollziehbaren möglichst simplen arithmetischen Berechnung des Grades eines Risikos für ein KI-System könnte man passenderweise der Normung und dem gemeinsamen technischen Ausschusses 21 zu Künstlicher Intelligenz (CEN-CENELEC JTC 21 "AI") übertragen.

Frühzeitige Beauftragung der Normungsorganisationen

Normung bzw. Standardisierung erfolgt zu vergleichbaren Themen in der Regel abgestimmt auf nationaler (in Deutschland DIN und DKE), europäischer (CEN/CENELC) und internationaler (z. B. ISO, IEC) Ebene und soll grundsätzlich Orientierung, Sicherheit und Klarheit geben. Sie soll Nutzen stiften für Verbraucher und Anwender ebenso wie für Gesellschaft, Wissenschaft und Wirtschaft, sowie innovative Technologien und erhöhte Sicherheit in allen Bereichen fördern [21].
Normen und Standards stellen Wissen in Form von Regeln und Leitlinien dar, ganz besonders im Umfeld Elektrotechnik, Elektronik und der Informations- und Kommunikationstechnologie.
Bereits in einer relativ frühen Entwicklungsphase des AIA, noch weit vor Inkrafttreten und Wirkung entfalten der Verordnung, hat die Europäische Kommission Ende 2022 in Form eines sogenannten „draft standardisation request“ die europäischen Standardisierungsorganisationen aufgerufen bzw. formal beauftragt, die Arbeiten aufzunehmen zur Unterstützung einer sicheren und vertrauenswürdigen künstlichen Intelligenz.
Die Endfassung des Normungsauftrages der Europäischen Kommission zur Unterstützung der EU-Politik im Bereich der künstlichen Intelligenz wurde im Mai 2023 veröffentlicht.
Vor diesem Hintergrund wurde innerhalb des bereits auf europäischer Ebene existierenden gemeinsamen technischen Ausschusses 21 zu Künstlicher Intelligenz (CEN-CENELEC JTC 21 "AI") eine spezielle Arbeitsgruppe (TG) für die Unterstützung der Umsetzung des AIA und des „standardisation request“ durch entsprechende Normen und Standards gebildet.

Ein Beispiel für die insgesamt 10 im „standardisation request“ zum AIA mit April 2025 als Fertigstellungstermin beauftragten Themenfelder ist an erster Stelle aufgeführt die Unterstützung bei der Entwicklung eines „Risikomanagementsystem für KI-Systeme“ durch entsprechende Normen bzw. Normungsunterlagen und Standards. (Durchführungsbeschluss der Kommission über einen Normungsauftrag an das Europäische Komitee für Normung und das Europäische Komitee für elektrotechnische Normung zur Unterstützung der Unionspolitik im Bereich der künstlichen Intelligenz [22].

Förderung der Wettbewerbsfähigkeit besonders für KMUs und Start-ups

Der AIA bietet einen einheitlichen Rechtsrahmen, er fördert Transparenz und Vertrauen in KI-Anwendungen und damit auch Innovation und Wettbewerbsfähigkeit, Voraussetzungen die wesentlich für den Erfolg besonders von KMUs und Start-ups sind. Der AIA umfasst Bestimmungen zur Förderung eines einheitlichen Marktes, was eine Stärkung der Innovationskraft für Unternehmen bedeutet. Kapitel VI listet Maßnahmen zur Unterstützung der Innovation auf. Besonders erwähnenswert sind die KI-Reallabore, die in Artikel 57 eingeführt werden:
„Die nach Absatz 1 eingerichteten KI-Reallabore bieten eine kontrollierte Umgebung, um Innovation zu fördern und die Entwicklung, das Training, das Testen und die Validierung innovativer KI-Systeme für einen begrenzten Zeitraum vor ihrem Inverkehrbringen oder ihrer Inbetriebnahme nach einem bestimmten zwischen den zukünftigen Anbietern und der zuständigen Behörde vereinbarten Reallabor-Plan zu erleichtern. In diesen Reallaboren können auch im Reallabor beaufsichtigte Tests unter realen Bedingungen durchgeführt werden“ (Absatz (5)).
Artikel 62 regelt, dass die Mitgliedstaaten „KMUs, einschließlich Start-up-Unternehmen, die - ihren Sitz oder eine Zweigniederlassung in der Union haben, soweit sie die Voraussetzungen und Auswahlkriterien erfüllen, vorrangigen Zugang zu den KI-Reallaboren [gewähren].“
Auch bei der Festsetzung der Gebühren für die Konformitätsbewertung werden die Interessen und Bedürfnisse der KMUs und Start-ups berücksichtigt und die Gebühren proportional zur Größe der Unternehmen, der Marktgröße und anderer einschlägiger Kennzahlen gesenkt (Artikel 62, Absatz 2).
Die EU-Kommission erarbeitet Leitlinien für das für Hochrisiko-KI-Systeme geforderte Qualitätsmanagement. Gemäß diesen Leitlinien können Kleinstunternehmen dieses Qualitätsmanagement in vereinfachter Form erfüllen, ohne dass das erforderliche Schutzniveau für Hochrisiko-KI-Systeme beeinträchtigt wird (Artikel 63, Absatz 1).

Von der Regulierung ausgenommen KI-Systeme

Open Source KI-System sind grundsätzlich vom AIA ausgenommen. (Artikel 53, Absatz 2) Anbieter von Open Source KI-Tools, -Diensten, -Komponenten oder Prozessen sind von der Regulierung nur betroffen, wenn es sich dabei um verbotene Anwendungen oder KI-Modelle mit allgemeinem Verwendungszweck

mit systemischen Risiken (Kapitel II und IV Artikel 51) oder Hochrisiko-Systeme handelt [23].
Ebenso sind KI-Systeme und Modelle, die rein zum Zweck der Wissenschaftlichen Forschung und Entwicklung erstellt wurden, nicht von der Regulierung betroffen (Artikel 2, Absatz 6).
„This Regulation should support innovation, respect freedom of science, and should not undermine research and development activity. It is therefore necessary to exclude from its scope AI systems and models specifically developed and put into service for the sole purpose of scientific research and development“ [24].

Positive Wirkung des AIA für Firmen in Deutschland und Europa

„Während bindende Regularien wie der AI Act und zugehörige Umsetzungsverfahren noch nicht etabliert sind, suchen Unternehmen individuell nach Lösungen, um KI-Systeme verantwortungsbewusst und wertebasiert zu gestalten. Einige Unternehmen haben beispielsweise ethische Leitlinien veröffentlicht, in denen sie Werte wie Fairness, Robustheit und menschliche Autonomie artikulieren. Jedoch reicht das allein nicht aus, damit KI tatsächlich ethisch unbedenklich gestaltet wird“ [25].
Im Gegenteil, es drängt sich der Eindruck auf, dass mit Hilfe des Feigenblatts „Ethik“ einer rechtlich verbindlichen Regulierung der Wind aus den Segeln genommen werden soll („Ethics Washing“).
Der EU AI ACT liefert den Firmen einen sicheren Rechtsrahmen. Dieser ist notwendig, da in den letzten Jahren ethische Konflikte drastisch zugenommen haben.
„Zwischen 2012 und 2021 stieg die Anzahl gemeldeter Zwischenfälle wie unzulässige Überwachung und Diskriminierung, Verstöße gegen Daten- und Urheberrechtsschutz sowie Manipulation durch Deepfakes im Zusammenhang mit KI um das 26-fache“ [26].
Die bisherigen Leitlinien von Firmen oder NGOs enthielten keine verpflichtenden Praktiken oder definierten Verantwortlichkeiten. Es gab z. B. keine Sanktionen bei Verstößen [27].
Der AIA bietet auf Basis unserer gemeinsamen europäischen Werte und der EU-Grundrechte aus unserer Sicht eine sehr gelungene Voraussetzung und hinreichend Anreize, vertrauenswürdige, menschenzentrierte und rechtssichere KI-Systeme zu entwickeln und Innovation, sowie die dafür notwendigen Investitionen im Bereich Künstliche Intelligenz zu fördern (Kapitel I, Artikel 1).

Auswirkungen auf die Wissenschaft

Die Freiheit der Wissenschaft ist ein hohes Gut, das in Deutschland sogar recht weit vorn im Grundgesetz geschützt wird. Diese Freiheit entbindet allerdings nicht von der Verfassungstreue. Dieser Allgemeinplatz soll vorangestellt sein, wenn wir über Grundlagenforschung im Bereich der KI-Entwicklung sprechen.
Im Bereich der öffentlichen Forschung schränkt der AIA die Forschenden nicht mehr ein als bereits bestehende Gesetze und Ethische Leitlinien. Bei der

Erstellung von Anwendungen, aber eben auch Basismodellen, soll die eigene Forschung auf Diskriminierungspotential und Beeinträchtigung des Jugendschutzes überprüft werden. Auch bei der Akquise der Daten sollen rechtliche und moralische Normen beachtet werden, auch und besonders im Falle riesiger Datenmengen. Open-Source-Modelle werden anders behandelt, was der Arbeitsweise der öffentlichen Wissenschaft entgegenkommt, ist sie doch offen und frei zugänglich. Über Auftrags- und Militärforschung haben wir an anderer Stelle oben bereits ein paar Worte verloren. Insgesamt sehen wir keine Einschränkungen im Bereich der öffentlich geförderten KI-Forschung.

Auswirkungen auf die Zivilgesellschaft

Die Zivilgesellschaft kommt im AIA meist in der Rolle der Bereitstellung von Sachverständigen und als Kontrollgremium vor, das konsultiert werden kann, aber nicht muss. Auch hier zeigt sich nur die Formulierung gängiger Praxis in allen Bereichen der gegenwärtigen Technikentwicklung: Die Zivilgesellschaft wird zwar hoch geschätzt, aber nur nach Ermessensspielraum berücksichtigt („sollte", „könnte", „gegebenenfalls"). Die GI ist eine starke Stimme der aktiven Zivilgesellschaft im technischen Bereich, es ist zu erwarten, dass entsprechende Anfragen für Gremienbeteiligung und Expertisen gestellt werden.
Interessant ist der Whistleblower-Schutz, der angesprochen wird. Im AIA wird der Wunsch formuliert, das die Whistleblower-Richtlinie (EU) 2019/1937 zur Anwendung kommen soll, wenn auf Missstände hingewiesen bzw. Verstöße gegen den AIA öffentlich gemacht werden.
Über diesen beiden eher politischen Interessensvertretungen hinaus wird die Zivilgesellschaft nicht weiter adressiert. Insbesondere die Maker-Szene und freie Entwickler*innen kommen so nicht im Text vor. Sie sind dennoch betroffen, denn auch wenn es Ausnahmen für Open Source gibt, so wird diese Ausnahme wieder eingeschränkt, wenn es um Fragen des Jugendschutzes geht. Nicht publizierte KI-Modelle und Herstellung von KI-Systemen für den Privatgebrauch sind von der Regulierung durch den AIA ausgenommen.
An dieser Stelle hätten wir uns mehr Rechtssicherheit für die KI-Community gewünscht, insbesondere was den Umgang mit großen Plattformen wie Hugging Face betrifft.

5. Fazit und Ausblick

EU-Gesetzgebungsverfahren, insbesondere Verordnungen, dauern von der Idee, über die Einigung bis hin zur Umsetzung und Inkrafttreten, speziell bei der Regulierung exponentieller Technologien, eine gefühlte Ewigkeit. Nach Inkrafttreten des AIA muss der überwiegende Pflichtenkatalog erst 24 Monate später umgesetzt sein. Bestimmte Pflichten gelten z. B. bereits nach 6, 12 oder 36 Monaten (Artikel 113). Auch wenn im AIA nicht alles perfekt ist, so ist es doch das weltweit erste umfassende und in vieler Hinsicht zukunftsweisende Regelwerk seiner Art (Wir halten die KI-Regulierungsansätze von den USA und

China nicht für umfassend und verbindlich). Beim AIA handelt es sich um eine verbindliche Verordnung und ein umfassendes Regelwerk. Es ist eine Einigung zwischen immerhin 27 zum Teil recht unterschiedlichen Staaten mit diversen Interessenslagen. Es ist eine Einigung nicht vordergründig auf Basis gemeinsamer Wirtschaftsinteressen, sondern allen voran ein klares Bekenntnis zu gemeinsamen Werten innerhalb der EU und zwischen den einzelnen Mitgliedsstaaten, klar kommuniziert per Gesetz gegenüber dem Rest der Welt. Mit einer Präambel vorangestellt mit dem 1. Artikel, die sich sehen lassen kann – im Folgenden frei widergegeben: „...Einführung menschenzentrierter und vertrauenswürdiger künstlicher Intelligenz zu fördern und gleichzeitig ein hohes Maß an Schutz der Gesundheit, der Sicherheit, der in der Charta verankerten Grundrechte, einschließlich der Demokratie, der Rechtsstaatlichkeit und des Umweltschutzes, vor schädlichen Auswirkungen von Systemen der künstlichen Intelligenz in der Union zu gewährleisten...“ „... das Funktionieren des Binnenmarktes zu verbessern ... und Innovationen zu unterstützen“.
Und dennoch trotz AIA: „Ethik bleibe ein unabdingbares Instrument, um den Sinn von Regulierungsmaßnahmen überhaupt zu verstehen und beständig zu überprüfen, ob diese uns auch effektiv schützen. Auch helfe Ethik dabei, den Einfluss von KI als soziotechnisches System auf die Gesellschaft umfassend wahrzunehmen – und gleichzeitig den Einfluss, den Menschen selbst auf KI-Systeme haben“ [25].

6. Quellen

[1] Gesellschaft für Informatik e.V., „Ethische Leitlinien“, 2018. Verfügbar unter: https://gi.de/ueber-uns/organisation/unsere-ethischen-leitlinien/. [Abgerufen am 8.7.2024].

[2] Europäisches Parlament, „Gesetz über künstliche Intelligenz“, 2024. Verfügbar unter: https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_DE.html. [Abgerufen am 8.7.2024].

[3] Future of Life Institute, „Der AI Act Explorer“. Verfügbar unter: https://artificialintelligenceact.eu/de/ai-act-explorer/. [Abgerufen am 8.7.2024].

[4] HEG-KI, „Ethik-Leitlinien für eine vertrauenswürdige KI“, 2019. Verfügbar unter: https://op.europa.eu/de/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1. [Abgerufen am 8.7.2024].

[5] Europäische Kommission, „Weißbuch zur Künstlichen Intelligenz – Ein europäisches Konzept für Exzellenz und Vertrauen“, 2020. Verfügbar unter: https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:52020DC0065. [Abgerufen am 8.7.2024].

[6] Europäische Kommission, „Digitaler Kompass: Der Weg in die die digitale Zukunft“, 2021. https://www.consilium.europa.eu/de/policies/a-digital-future-for-europe. [Abgerufen am 16.7.2024].

[7] Datenethikkommission der Bundesregierung, „Gutachten der Datenethikkommission der Bundesregierung“, 2019. Verfügbar unter: https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.html. [Abgerufen am 16.7.2024].

[8] Enquete-Kommission Künstliche Intelligenz, “Bericht der Enquete-Kommission Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale”, 2020. Verfügbar unter: https://dserver.bundestag.de/btd/19/237/1923700.pdf. [Abgerufen am 16.7.2024].

[9] DIN, DKE, „Deutsche Normungsroadmap Künstliche Intelligenz (Ausgabe 2), 2022. Ausgabe der Normungsroadmap KI, kurz NRM KI [9]. Verfügbar unter: https://www.din.de/de/forschung-und-innovation/themen/kuenstliche-intelligenz/fahrplan-festlegen. [Abgerufen am 16.7.2024].

[10] Bertuzzi, L., „OECD updates definition of Artificial Intelligence ‘to inform EU’s AI Act’”, 2023. Verfügbar unter: https://www.euractiv.com/section/artificial-intelligence/news/oecd-updates-definition-of-artificial-intelligence-to-inform-eus-ai-act/. [Abgerufen am 16.7.2024].

[11] Industry of Things, „Kritik am AI Act von deutschen Verbänden“, 2024.Verfügbar unter: https://www.industry-of-things.de/kritik-am-ai-act-von-deutschen-verbaenden-a-92e8d32b2c8dd1fabb1eb379b7301621/. [Abgerufen am 16.7.2024].

[12] Müller, A., Spielkamp, M., „Einigung zum AI Act: Wichtiger Schutz und gefährliche Schlupflöcher Key safeguards and dangerous loopholes”, 2023. Verfügbar unter: https://algorithmwatch.org/de/einigung-zum-ai-act-wichtiger-schutz-und-gefahrliche-schlupflocher/. [Abgerufen am 16.7.2024].

[13] Hahn, S., „AI Act. Es droht ein Frontalangriff auf unsere Bürgerrechte“, 2024. Verfügbar unter: https://www.svenja-hahn.eu/post/pm-ai-act-es-droht-frontalangriff-auf-unsere-burgerrechte-svenja-hahn-zur-biometrischen-uberwachung. [Abgerufen am 16.7.2024].

[14] Gewerkschaft der Polizei, „GdP fordert deutschlandweite Gesichterkennungssoftware“, 2024. Verfügbar unter; https://www.gdp.de/bund/de/stories/2024/03/2024-03-01_kopelke-moderne-software-und-ki-koennen-bei-der-verbrechensbekaempfung-extrem-schnell-helfen. [Abgerufen am 16.7.2024].

[15] DIP, „Einsatz von Dialekterkennungssoftware im Bundesamt für Migration und Flüchtlinge“, 2022. Verfügbar unter: https://dip.bundestag.de/vorgang/einsatz-von-dialekterkennungssoftware-im-bundesamt-f%C3%BCr-migration-und-fl%C3%BCchtlinge/290504. [Abgerufen am 16.7.2024].

[16] Leisegang, L., „Mit Wasserzeichen gegen die babylonische Verwirrung“, 2023. Verfügbar unter: https://netzpolitik.org/2023/kuenstliche-intelligenz-mit-wasserzeichen-gegen-die-babylonische-verwirrung/. [Abgerufen am 16.7.2024].

[17] WBGU, „Unsere gemeinsame digitale Zukunft“, 2019. Verfügbar unter: https://www.wbgu.de/de/publikationen/publikation/unsere-gemeinsame-digitale-zukunft. [Abgerufen am 16.7.2024].

[18] Mollen, A., „Nachhaltige KI und digitale Selbstbestimmung“, 2022. Verfügbar unter: https://codina-transformation.de/wp-content/uploads/CODINA_Nachhaltige_KI_u_Digitale_Selbstbestimmung.pdf. [Abgerufen am 16.7.2024].

[19] Rohde, F., Wagner, J., Reinhardt, P., Petschow, U., Meyer, A., Voß, M., Mollen, A., „Nachhaltigkeitskriterien für künstliche Intelligenz“, 2021. Verfügbar unter: https://www.ioew.de/fileadmin/user_upload/BILDER_und_Downloaddateien/Publikationen/2021/IOEW_SR_220_Nachhaltigkeitskriterien_fuer_Kuenstliche_Intelligenz.pdf. [Abgerufen am 16.7.2024].

[20] Gesellschaft für Informatik e.V. (Hrsg.), „Abschlussbericht ExamAI – KI-Testing und Auditing“, 2021. Verfügbar unter: https://gi.de/fileadmin/PR/Testing-AI/Abschlussbericht_ExamAI_-_KI_Testing_und_Auditing.pdf. [Abgerufen am 16.7.2024].

[21] DKE, „Die Bedeutung der Normung – Nutzen und Vorteile“, 2023. Verfügbar unter: https://www.dke.de/de/normen-standards/bedeutung-der-normung. [Abgerufen am 16.7.2024].

[22] Europäische Kommission, „Durchführungsbeschluss der Kommission über einen Normungsauftrag an das Europäische Komitee für Normung und das Europäische Komitee für elektrotechnische Normung zur Unterstützung der Unionspolitik im Bereich der künstlichen Intelligenz”, 2023. Verfügbar unter: https://ec.europa.eu/transparency/documents-register/detail?ref=C(2023)3215&lang=de. [Abgerufen am 16.7.2024].

[23] Keller, P., „A Frankenstein-Like Approach: Open Source in the AI Act”, 2023. Verfügbar unter: (https://openfuture.eu/blog/a-frankenstein-like-approach-open-source-in-the-ai-act/. [Abgerufen am 16.7.2024].

[24] Europäischer Rat, „Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts – Analysis of the final compromise text with a view to agreement“, 2024. Verfügbar unter: https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf (12c). [Abgerufen am 16.7.2024].

[25] Gesellschaft für Informatik e.V. (Hrsg.), „KI-Ethik meets Design Thinking – Ergebnisse aus der Roundtable-Reihe ethische KI-Entwicklung“, 2024. Verfügbar unter: https://gi.de/fileadmin/PR/Roundtable-KI/RTeKI_Abschlusspublikation_digital.pdf. [Abgerufen am 16.7.2024].

[26] Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Parli, V., Shoham, Y., Wald, R., Clark, J., Perrault, R., „The AI Index 2023 Annual Report", 2023. Verfügbar unter: https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf. [Abgerufen am 16.7.2024].

[27] Henriksen, A., Enni, S., Bechmann, A., „Situated Accountability: Ethical Principles, Certification Standards, and Explanation Methods in Applied AI", Situated Accountability: Ethical Principles, Certification Standards, and Explanation Methods in Applied AI, 2021. Verfügbar unter: https://dl.acm.org/doi/10.1145/3461702.3462564. [Abgerufen am 16.7.2024].

Kontakt

Fachgruppe Informatik und Ethik
Dr. Stefan Ullrich (stellv. Sprecher)
Gesellschaft für Informatik e.V. (GI)
Geschäftsstelle Berlin im Spreepalais am Dom
Anna-Louisa-Karsch-Str. 2, 10178 Berlin
E-Mail: stefan.ullrich[at]gi.de

Über die Fachgruppe Informatik und Ethik

Die Fachgruppe Informatik und Ethik der GI hat es sich zur Aufgabe gemacht, Diskurse zu ethischen Problemen der Informatik zu initiieren und zu fördern. Neben Vorträgen, Workshops, Beratung und Aufsätzen stellen die Mitglieder der Fachgruppe „Informatik und Ethik" regelmäßig hypothetische, aber realistische Fallbeispiele vor, die zur Diskussion anregen sollen. Die Fälle können jeweils von Interessierten im Blog der Fachgruppe auf der GI-Website https://gewissensbits.gi.de kommentiert und diskutiert werden.

Über die Gesellschaft für Informatik e.V. (GI)

Die Gesellschaft für Informatik e.V. (GI) ist die größte Fachgesellschaft für Informatik im deutschsprachigen Raum. Seit 1969 vertritt sie die Interessen der Informatikerinnen und Informatiker in Wissenschaft, Gesellschaft und Politik und setzt sich für eine gemeinwohlorientierte Digitalisierung ein. Mit 14 Fachbereichen, über 30 aktiven Regionalgruppen und unzähligen Fachgruppen ist die GI Plattform und Sprachrohr für alle Disziplinen in der Informatik. Weitere Informationen finden Sie unter www.gi.de.