

# UNESCO- Empfehlung zur Ethik Künstlicher Intelligenz

Bedingungen zur Implementierung  
in Deutschland



**unesco**

Deutsche  
UNESCO-Kommission

# UNESCO-Empfehlung zur Ethik Künstlicher Intelligenz. Bedingungen zur Implemen- tierung in Deutschland

Prof. Dr. Matthias C. Kettemann

# Inhalt

---

Abkürzungs- verzeichnis	6
----------------------------	---

---

Vorwort	8
---------	---

---

Synthese und Handlungsansätze für Politik und Verwaltung	11
-------------------------------------------------------------------	----

---

---

I. Einleitung: Die neue UNESCO- Empfehlung zur Ethik Künstlicher Intelligenz – Globaler Referenzrahmen für KI-Ethik	14
------------------------------------------------------------------------------------------------------------------------------------	----

---

II. Mehrwert der UNESCO KI-Ethik- Empfehlung im Vergleich zu anderen inter- nationalen Initiativen	20
----------------------------------------------------------------------------------------------------------------	----

---

II.1. OECD-Empfehlung zu Künstlicher Intelligenz	22
-----------------------------------------------------	----

---

II.2. Europarat: CAHAI	23
------------------------	----

---

II.3. Gesetzgebungsvorhaben der Europäischen Union	25
-------------------------------------------------------	----

---

II.4. UNESCO KI-Ethik-Empfehlung: Ansatz und Mehrwert	26
----------------------------------------------------------	----

---

II.5. Zusammenfassung	27
-----------------------	----

---

---

III. Bedingungen in Deutschland für die Implementierung der UNESCO KI-Ethik- Empfehlung	28
-----------------------------------------------------------------------------------------------------	----

---

III.1. Einführung	30
-------------------	----

---

III.2. Ethical Impact Assessment (Policy Area 1)	30
-----------------------------------------------------	----

---

III.3. Ethical Governance and Stewardship (Policy Area 2)	34
--------------------------------------------------------------	----

---

III.4. Data Policy (Policy Area 3)	36
------------------------------------	----

---

III.5. Development and International Cooperation (Policy Area 4)	38
------------------------------------------------------------------------	----

---

III.6. Gender (Policy Area 6)	40
-------------------------------	----

---

---

IV. Fazit	44
-----------	----

---

Impressum	48
-----------	----

---

# Abkürzungsverzeichnis

4IR	Expert National Task Force on the Fourth Industrial Revolution	EU-AI	Vorschlag für eine KI-Verordnung der EU
AI	Artificial Intelligence	GIZ	Gesellschaft für Internationale Zusammenarbeit
AIGO	OECD Expert Group on Artificial Intelligence	HEG-KI	Hochrangige Expertengruppe für Künstliche Intelligenz der Europäische Kommission
BMFSFJ	Bundesministerium für Familie, Senioren, Frauen und Jugend	KI	Künstliche Intelligenz
BMWi	Bundesministerium für Wirtschaft und Energie	LDCs	Least Developed Countries
CAHAI	Ad hoc Committee on Artificial Intelligence of the Council of Europe	LLDCs	Landlocked Developing Countries
CAHAI-LFG	CAHAI Legal Frameworks Group	LMICs	Low- and Middle-Income Countries
CDEP	Committee on Digital Economy Policy	NKIS	Nationale KI-Strategie der Bundesregierung
DEK	Datenethikkommission der Bundesregierung	OECD	Organisation for Economic Cooperation and Development
DIN	Deutsches Institut für Normung	SDGs	Sustainable Development Goals – Ziele für nachhaltige Entwicklung
DKE	Deutsche Kommission Elektrotechnik Elektronik Informationstechnik	SIDS	Small Island Developing States
DUK	Deutsche UNESCO-Kommission	UNESCO	United Nations Educational, Scientific and Cultural Organization
EIA	Ethical Impact Assessment		
EKKI	Enquete-Kommission Künstliche Intelligenz des Bundestages		
EMRK	Europäische Menschenrechtskonvention		
EU	Europäische Union		

„Auch wenn vieles noch offen ist, eins ist klar: Noch haben wir als Menschheit die Möglichkeit, die künftige Entwicklung von KI zu lenken und in einer humanistischen, menschenrechtlichen Perspektive ethisch einzuhegen. Die UNESCO hat sich im Rahmen ihres Mandats diesem Thema angenommen.“  
— Prof. Dr. Maria Böhmer, Präsidentin der Deutschen UNESCO-Kommission



Prof. Dr. Maria Böhmer,  
Präsidentin der Deutschen  
UNESCO-Kommission

Künstliche Intelligenz (KI) hat in den letzten Jahren Einzug in viele Bereiche unseres Lebens gehalten und sie grundlegend verändert. Sie sorgt für personalisierte Suchergebnisse im World Wide Web, für Gesichtserkennung auf dem Smartphone und teilautomatisiertes Fahren. Der Einsatz selbstlernender Algorithmen ist ein Megatrend unserer Zeit, gleichzeitig sind die langfristigen Folgen des Einsatzes von KI kaum prognostizierbar. So spannt der aktuelle Diskurs zu KI und ihren Auswirkungen einen weiten Bogen und reicht von Utopien einer besseren und gerechteren Welt bis zu Dystopien einer unkontrollierbaren und den Menschen beherrschenden Superintelligenz. Auch wenn vieles noch offen ist, eins ist klar: Noch haben wir als Menschheit die Möglichkeit, die künftige Entwicklung von KI zu lenken und in einer humanistischen, menschenrechtlichen Perspektive ethisch einzuhegen.

Die UNESCO hat sich im Rahmen ihres Mandats diesem Thema angenommen. Nach einem langwierigen Aushandlungsprozess der 193 Mitgliedsstaaten, in denen auch vielfältige Stakeholder eingebunden waren, hat die UNESCO im November 2021 eine wegweisende Empfehlung zur Ethik Künstlicher Intelligenz verabschiedet. Die UNESCO-Empfehlung ist die erste global verhandelte Antwort auf eine der größten Herausforderungen unserer Zeit. Sie überzeugt durch ihre starke Menschenrechtssprache, ein klares ethisches Fundament und konkrete Gestaltungsaufgaben für die Politik und erfüllt alle Voraussetzungen, um zu einem globalen Kompass für eine verantwortliche KI-Entwicklung und -Nutzung zu werden.

Die Empfehlung definiert einerseits ethische Leitplanken und gibt konkrete Umsetzungsaufgaben für Politikfelder, in denen KI schon heute breit diskutiert wird. Zugleich hat der inklusive Aushandlungsprozess auch Felder in den Blick gerückt, die im KI-Diskurs bislang eher eine Nebenrolle gespielt haben, jedoch im Kern des Mandats der UNESCO stehen. So zeichnet sich die UNESCO-Empfehlung durch konkrete Bezüge zu elf verschiedenen Politikfeldern aus, darunter Bildung und Wissenschaft, Kommunikation, Gesundheit und Umwelt. Bemerkenswert ist auch der Fokus der Empfehlung auf bisherige „blind spots“ des KI-Rechts, wie etwa KI und Gender, KI und Nachhaltigkeit, KI und der Globale Süden sowie KI und Bildung. Außerdem benennt sie klare Verfahrens- und Transparenzpflichten, wie zum Beispiel eine multistakeholderbasierte, offene und responsive Ausgestaltung der entsprechenden Prozesse.

Die Deutsche UNESCO-Kommission, als Bindeglied zwischen der UNESCO, der nationalen Politik und der Zivilgesellschaft, unterstützt die Umsetzung dieses wichtigen Völkerrechtstextes in Deutschland. Als inhaltliches Fundament für diese Aufgabe dient uns künftig die vorliegende Studie. Sie nimmt den Status quo in den Blick und legt dar, welche Prozesse, Instrumente und Gesetze es in Deutschland bereits gibt, die eine KI-Entwicklung gemäß der Empfehlung unterstützen. Gleichzeitig analysiert sie auch, wo noch

Regulierungs- oder Handlungsbedarf besteht, und macht konkrete Vorschläge für die Implementierung der UNESCO-Empfehlung in Deutschland. Zudem arbeitet sie den Mehrwert der UNESCO-Empfehlung im Vergleich zu anderen Initiativen heraus, wie beispielsweise der OECD-Empfehlung oder dem Richtlinienentwurf der Europäischen Kommission. So weist die UNESCO-Empfehlung etwa als einzige auf kulturelle Aspekte hin, die von dem Einsatz von KI-Systemen berührt sein können, und setzt einen Akzent auf die Bedeutung und Schutzwürdigkeit marginalisierter Gruppen.

Unser großer Dank gilt dem Verfasser der Studie, Professor Matthias C. Kettmann. Dafür, dass die Entwicklung von KI in ethisch-menschenrechtlichen Bahnen verläuft, brauchen wir jegliche Anstrengung von Politik und Verwaltung. KI-Dystopien müssen Science-Fiction bleiben!



# Synthese und Handlungsansätze für Politik und Verwaltung

## Synthese: Mehrwerte der UNESCO KI-Ethik-Empfehlung

→ Die KI-Ethik-Empfehlung der UNESCO ist der erste, in einem offenen und einschließenden Multi-Stakeholder-Prozess und zugleich von Staaten ausverhandelte, globale Völkerrechtstext zu KI-Ethik. Sie umfasst weltweit akzeptierte ethische Standards für KI-Technologien unter voller Beachtung des Völkerrechts, insbesondere der Menschenrechte, die dann in der Folge eine zentrale Rolle bei der Entwicklung von KI-bezogenen Normen auf der ganzen Welt spielen können. 193 Mitgliedstaaten haben sich zu ihr bekannt. Der zweijährige, offene Prozess, bei dem auch nichtstaatliche Akteur\*innen ihre Stimmen einbringen konnten, stellt einen bedeutenden Mehrwert an Legitimation dar.

→ Inhaltlich bemerkenswert ist der ganzheitliche Ansatz – KI-Ethik als Gesellschaftsaufgabe – und die differenzierte Ausformulierung von Regeln für KI-Entwicklung und -Nutzung in elf verschiedenen Politikfeldern, darunter Bildung und Wissenschaft, Kommunikation, Gesundheit sowie Umwelt.

→ Der Fokus auf ethische Regeln, die aber stets eng mit menschenrechtlichen Verpflichtungen verknüpft und begründet werden, ist bewusst gewählt. Klar konturierte ethische Werte und Grundsätze beziehen sich in verschiedener Weise auf das Recht; sie können bei der Entwicklung und Umsetzung von politischen Maßnahmen und bei der Interpretation von Rechtsnormen helfen, indem sie die technologische Entwicklung orientieren.

→ Bemerkenswert ist weiter der Fokus auf „blind spots“, wie etwa KI und Gender, KI und Nachhaltigkeit, KI und der Globale Süden (Entwicklung), KI und Bildung sowie auf klare Verfahrens- und Transparenzpflichten: Konkret werden ethische Impact Assessments gefordert und eine multistakeholderbasierte, offene und responsive Ausgestaltung der entsprechenden Prozesse.



# Handlungsansätze für KI-Akteur\*innen in Politik und Verwaltung

## 1 KI muss Menschenrechte achten, sichern und fördern

Weltweit und gerade auch in Deutschland muss jede Entscheidung zu KI – als kontingente Technologie – im Prozess der Planung und des Einsatzes an ethischen Werten, Menschenrechten und Völkerrecht ausgerichtet werden.

## 2 Ethische (und rechtliche) Leitplanken müssen im gesamten Lebenszyklus von KI-Systemen eingehalten werden

KI ist kein Deus ex Machina: Alle KI-Akteur\*innen müssen daher den gesamten Lebenszyklus von KI-Systemen in den Blick nehmen: ein Prozess, der mit der Ausformulierung der Anforderungen an das KI-System und der Zusammenstellung des Teams, das über Trainingsdaten und die Bedingungen maschinellen Lernens entscheidet, beginnt.

## 3 KI-Daten müssen möglichst öffentlich verfügbar sein

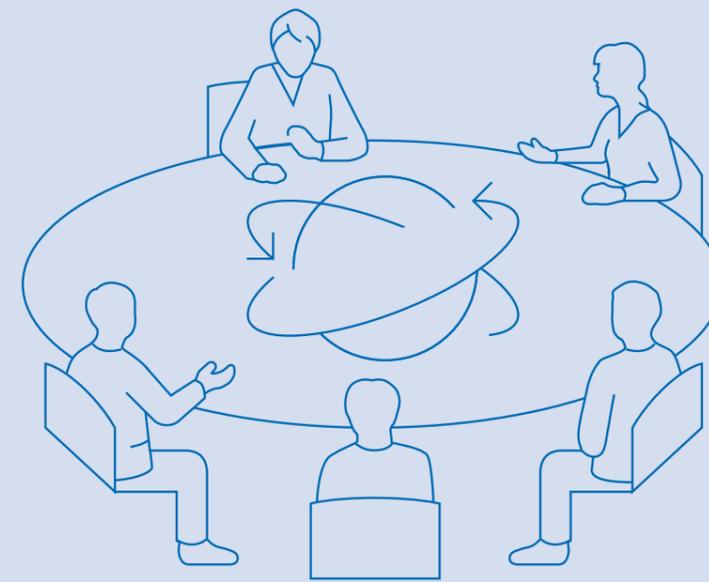
Die Bundesregierung muss sicherstellen, dass die Wissenschaft datenschutzwahrend Zugang zu KI-Forschungsdaten bekommt und diese nicht von großen Unternehmen monopolisiert werden. Dieser Zugang muss neben der Wissenschaft auch ausgewählten zivilgesellschaftlichen Akteur\*innen ermöglicht werden, die anderen Logiken und Interessen folgen als die Wissenschaft und eine Public-Watchdog-Funktion innehaben können.

## 4 KI muss zu gesellschaftlicher Vielfalt beitragen

Gender, Inclusiveness und Diversity Mainstreaming in allen Prozessen während des gesamten Lebenszyklus von KI-Anwendungen helfen entscheidend, Gefahren der Diskriminierung vorzubeugen. Damit die Inhalte von Trainingsdaten frei von Stereotypen und Verzerrungen sind, müssen entsprechende Dokumentationspflichten für KI-Akteur\*innen in Deutschland geschaffen werden. Entwickler\*innen und Forscher\*innenteams sind zudem diverser aufzustellen. Aber Nichtdiskriminierung und Gleichbehandlung sind systemische Fragen: Politik und Verwaltung in Deutschland sollten daher in allen Politikfeldern Maßnahmen ergreifen, um die Gleichberechtigung und Gleichstellung von Frauen und diversen Personen aktiv zu fördern.

## 5 KI muss öffentlich überwacht werden

Der deutsche Gesetzgeber muss klarere Rahmenbedingungen für die Durchführung von ethischen Impact Assessments und Überwachungsmechanismen erarbeiten und umsetzen. Diese müssen einen Anspruch auf Herausgabe von Transparenzprotokollen für Nutzer\*innen enthalten. Nur interne Überprüfungen sind nicht ausreichend. Die unternehmerischen KI-Akteur\*innen müssen sicherstellen, dass Ergebnisse von KI-Entscheidungsprozessen verständlich, erklärbar und rechtfertigbar sind. Menschen dürfen sich der KI nicht ausgeliefert fühlen. Im Gegenteil: KI ist eine Gesellschaftsaufgabe und alle KI-Akteur\*innen sind verpflichtet, Bürger\*innen durch moderne Wissensformate verstärkt an die Rolle und das Potenzial der KI heranzuführen.



## 6 KI muss global zugänglich werden

KI-Akteur\*innen, hier besonders die Bundesregierung, sollten im Rahmen der technischen Entwicklungszusammenarbeit KI-Kapazitäten in Ländern mit geringem und mittlerem Einkommen (Low- and Middle-Income Countries, LMICs) und in Least Developed Countries (LDCs) aufbauen helfen, um kontextbezogene Verzerrungen ‚westlicher KI‘ zu korrigieren. Gerade durch den fehlenden Zugang zu großen Datenmengen, die bei der KI-Entwicklung entscheidend sind, werden bestehende Gräben noch vertieft, wenn nicht bewusst gegengesteuert wird. Neben der staatlichen Entwicklungszusammenarbeit ist hier besonders die Rolle der Privatwirtschaft maßgeblich, die als Partnerin auftreten und von der Bundesregierung zunehmend in die Pflicht genommen werden sollte.

## 7 KI muss zur Realisierung der SDGs beitragen

Digitalisierung muss nachhaltig und Nachhaltigkeit digital ausgestaltet werden. KI muss stets mit Blick auf eine Förderung der Realisierung der Ziele für nachhaltige Entwicklung (SDGs) geprüft werden. Generell muss die KI-Nutzung insbesondere (auch) auf ökologische Nachhaltigkeit eingeordnet werden.

## 8 KI-Standards müssen in inklusiven Prozessen ausverhandelt werden

Die Bundesregierung und alle involvierten KI-Unternehmen und Standardisierungskörper müssen stärkeres zivilgesellschaftliches Engagement bei Standardisierungsprozessen ermöglichen, da bei diesen traditionell Industrieinteressen stark vertreten sind. Möglichst alle KI-Akteur\*innen sollen involviert sein, insbesondere Vertreter\*innen der Zivilgesellschaft, Datenschutzbehörden und Verbraucher\*innenschutzorganisationen sowie Vertreter\*innen vulnerabler Gruppen.

Einleitung: Die neue UNESCO-Empfehlung  
zur Ethik Künstlicher Intelligenz –  
Globaler Referenzrahmen für KI-Ethik



Die UNESCO-Empfehlung zur Ethik Künstlicher Intelligenz ist ein Beitrag zu global gemeinsamer, digitaler Menschenrechtspolitik, die leider noch in zu vielen Staaten nur national gedacht wird. Der 14. Bericht der Bundesregierung über ihre Menschenrechtspolitik<sup>1</sup> enthält in dem „Aktionsplan Menschenrechte“ für die Jahre 2021–2022 als Priorität die Stärkung des Menschenrechtsschutzes „angesichts des digitalen Wandels“. Darunter fällt auch der Schutz der Menschenrechte im Rahmen der Entwicklung der Künstlichen Intelligenz (KI).

Angesichts der Vielfalt der wirtschaftlichen, sozialen und kulturellen, bürgerlichen und politischen Rechte, deren Achtung, Schutz und Durchsetzung durch die zunehmende Nutzung von KI berührt werden, besteht ein Bedarf an Prinzipien und Regeln, um die Potenziale von KI zu realisieren, dabei aber die individuellen Freiheiten zu schützen und gesellschaftlichen Zusammenhalt zu garantieren. Vor diesem Hintergrund hat die UNESCO ihre KI-Ethik-Empfehlung ausgearbeitet (im Folgenden: „Empfehlung“).<sup>2</sup>

Im Koalitionsvertrag 2021–2025 werden Menschenrechte nicht im Kontext der KI erwähnt, während auf die KI als „Zukunftsfeld“ verwiesen wird. Gleichzeitig wird die Nutzung der Potenziale von KI festgeschrieben sowie die Bedeutung eines intensiven transatlantischen Dialogs zu Datensouveränität, Netzfreiheit und KI hervorgehoben.<sup>3</sup> Dies zeigt, dass eine Debatte über die menschenrechtlichen und ethischen Aspekte der Entwicklung und Nutzung von KI sinnvoll und, darauf aufbauend, die Festlegung klarer ethischer Prinzipien und operationalisierbarer Aufgaben für einschlägige Akteursgruppen nötig ist.<sup>4</sup> Vor diesem Hintergrund ist die UNESCO-Empfehlung zur Ethik Künstlicher Intelligenz ein Beitrag zu einer global gedachten, gemeinsamen digitalen Menschenrechtspolitik.

1 Auswärtiges Amt, 14. Bericht der Bundesregierung über ihre Menschenrechtspolitik (Berichtszeitraum 1. Oktober 2018 bis 30. September 2020), <https://www.auswaertiges-amt.de/blob/2422192/f01891c5efa5d6d89df7a5693eab5c9a/201202-mrb-14-download-data.pdf>. (Alle Onlinequellen wurden zuletzt am 10.12.2021 eingesehen).

2 UNESCO, *Draft text of the Recommendation on the Ethics of Artificial Intelligence*, SHS/IGM-AIETHICS/2021/JUN/3 Rev.2., 2021 („Die Empfehlung“).

3 Koalitionsvertrag 2021–2025 zwischen SPD/Bündnis 90/Die Grünen/FPD, [https://www.spd.de/fileadmin/Dokumente/Koalitionsvertrag/Koalitionsvertrag\\_2021-2025.pdf](https://www.spd.de/fileadmin/Dokumente/Koalitionsvertrag/Koalitionsvertrag_2021-2025.pdf), S.19, 21, 154.

4 Der stark genutzte Begriff „Ethik“ in der aktuellen Technologie-Debatte bedarf ebenso der Reflexion. Wichtig ist wohl die selbstständige Betrachtung nicht-rechtlicher Normsysteme, der Mehrwert, der in ihrer formalen Unverbindlichkeit liegt und in ihrer Bedeutung für den flexiblen Umgang mit sich wandelnder Technik, wobei auf die unbestreitbaren Vorteile (notfalls mit Zwangsgewalt durchsetzbarer) rechtlicher Ordnungssysteme hinzuweisen ist.

Die am 23.11.2021 angenommene Empfehlung stellt einen „ethischen Leitkompass und ein globales normatives Fundament“ dar, das es ermöglichen wird, „eine starke Achtung der Rechtsstaatlichkeit in der digitalen Welt aufzubauen.“<sup>5</sup> Was dies für Deutschland bedeuten kann, ist zentraler Gehalt dieser Studie.

Zunächst widmet sich die Studie dem Mehrwert der UNESCO-Empfehlung im Vergleich zu anderen internationalen Initiativen. Dies ist nötig vor dem Hintergrund der beachtlichen Zahl von 173 Regelwerken zur KI-Ethik, die das bis Mitte 2020 aktualisierte „AI Ethics Guidelines Global Inventory“ von AlgorithmWatch enthält.<sup>6</sup>

Prinzipien müssen in der Praxis relevant werden. Daher widmet sich die Studie – mit Blick auf bestimmte Politikfelder – den Bedingungen in Deutschland für die Implementierung der Empfehlung. Ein wichtiger Output der Studie ist eine Synthese sowie Handlungsansätze für Politik, Verwaltung und Multiplikator\*innen.

**„Die UNESCO-Empfehlung zur Ethik Künstlicher Intelligenz ist ein Beitrag zu global gemeinsamer, digitaler Menschenrechtspolitik, die leider noch in zu vielen Staaten nur national gedacht wird.“**

5 UNESCO, *Recommendation on the ethics of artificial intelligence*, 2021 (Übersetzung durch den Verf.).

6 Vgl. Algorithm Watch, *AI Ethics Guidelines Global Inventory*, <https://inventory.algorithmwatch.org>.

## Charakteristik

Bei der Empfehlung handelt es sich um den ersten global verhandelten, völkerrechtlich relevanten Text im Bereich der KI-Ethik. Sie ist nicht nur global, sondern auch inhaltlich holistisch konzipiert. Die Empfehlung bietet den 193 Mitgliedstaaten der UNESCO einen Handlungsrahmen in diesem wichtigen Zukunftsfeld (zu diskutierende KI-Regeln des Europarats hingegen gelten nur für 47 Staaten; 46 Staaten waren an KI-Normen der OECD beteiligt). Zu den UNESCO-Mitgliedstaaten gehören neben rechtsstaatlich soliden Staaten wie Deutschland, in denen – über die europäische Ebene – KI-Normen in Arbeit sind, auch Staaten wie China, die den Löwenanteil an der KI-Produktion haben und KI bereits in menschenrechtlich herausfordernder Weise nutzen. Daneben gibt es auch solche Mitgliedstaaten, die weder über einen grundrechtlichen Schutz vor KI noch über nationale KI-Kapazitäten verfügen. Dass China, das sich im Laufe der Verhandlungen mehrfach kritisch zu bestimmten menschenrechtlichen Festlegungen der Empfehlung positionierte, schlussendlich den Konsens mittrug, wurde von Beobachter\*innen als besonders bemerkenswert bezeichnet.<sup>7</sup> Der größte Nutzer von KI (und das Sitzland der meisten KI-basierten Start-ups), die USA, ist derzeit nicht UNESCO-Mitglied, es ist aber bekannt, dass die USA im Rahmen der Möglichkeiten, die ihnen der Beobachterstatus einräumt, den Prozess zur Ausarbeitung der Empfehlung besonders eng verfolgt haben.

## Prozess

Die Empfehlung der UNESCO wurde in einem zweijährigen, intensiven und teils kontroversen zwischenstaatlichen Verhandlungsprozess erarbeitet. Dieser Multi-Stakeholder-Ansatz, der unter anderem mit mehreren regionalen Stakeholder-Anhörungen bereichert wurde, bietet bereits einen bedeutenden legitimatorischen Mehrwert. Durch Einbeziehung verschiedener Stakeholder-Gruppen folgt die Empfehlung somit einem „Best Practice“-Modell für internationale Normsetzung.

<sup>7</sup> Lina Rusch, „Künstliche Intelligenz: Soft Law statt Regulierung?“, Tagesspiegel Background Digitalisierung, 16. November 2021, <https://background.tagesspiegel.de/digitalisierung/kuenstliche-intelligenz-soft-law-statt-regulierung>.

## Inhalt

Mit konkretem Bezug auf elf Politikfelder, darunter Bildung und Wissenschaft, Kommunikation, Gesundheit sowie Umwelt, übersetzt die Empfehlung Prinzipien für eine würdesensible KI-Nutzung in politische Gestaltungsaufgaben.

Das Ziel der Empfehlung ist, die KI an den Menschenrechten auszurichten. Sie gibt der KI eine ethische Grundlage, die Menschenrechte und Menschenwürde nicht nur schützt, sondern alle drei Dimensionen des Menschenrechtsschutzes anspricht: Achtung, Schutz und Förderung/Umsetzung.

## Ethik und Rechte

Der Fokus auf ethische Regeln und nicht primär auf menschenrechtliche Verpflichtungen ist bewusst gewählt. Einerseits hat die UNESCO eine besondere Verantwortung hinsichtlich der ethischen Gestaltung wichtiger gesellschaftlicher Agenden. Andererseits beziehen sich klar konturierte ethische Werte und Grundsätze in verschiedener Weise auf das Recht; sie können bei der Entwicklung und Umsetzung von politischen Maßnahmen und bei der Interpretation von Rechtsnormen helfen, indem sie, wie die Empfehlung es formuliert, „im Hinblick auf die rasante technologische Entwicklung Orientierung bieten“. Ethik ist nicht ein „Weniger“ als Menschenrechte, sondern eine (wenn auch bezüglich der Schutzintention in vielen Bereichen deckungsgleiche) Alternative. Ethische Regeln sind anders konzipiert und aufgebaut und werden unterschiedlich, nicht zentral gesteuert, und regelmäßig nicht mit Zwangsgewalt durchgesetzt. Staaten, die menschenrechtliche Verpflichtungen eingehen, sind völkerrechtlich an diese gebunden. Staaten, die sich zu ethischen Verpflichtungen bekennen, können lediglich – aber immerhin – durch internationalen Druck zu verpflichtungskonformem Verhalten veranlasst werden. Während sich die Pflicht aus eingegangenen Menschenrechtsverträgen bei manchen Staaten in der Unterschrift erschöpft, ist es bei ethischen Commitments einleuchtend, dass diese durch ‚strebendes Bemühen‘ von Staaten erfüllt werden müssen. Besonders in einem recht neuen Bereich wie der KI sind rechtliche Regeln auch noch nicht – zumal international – konsensfähig geworden. Hier haben ethische Regeln eine wichtige Funktion.



## Ziel

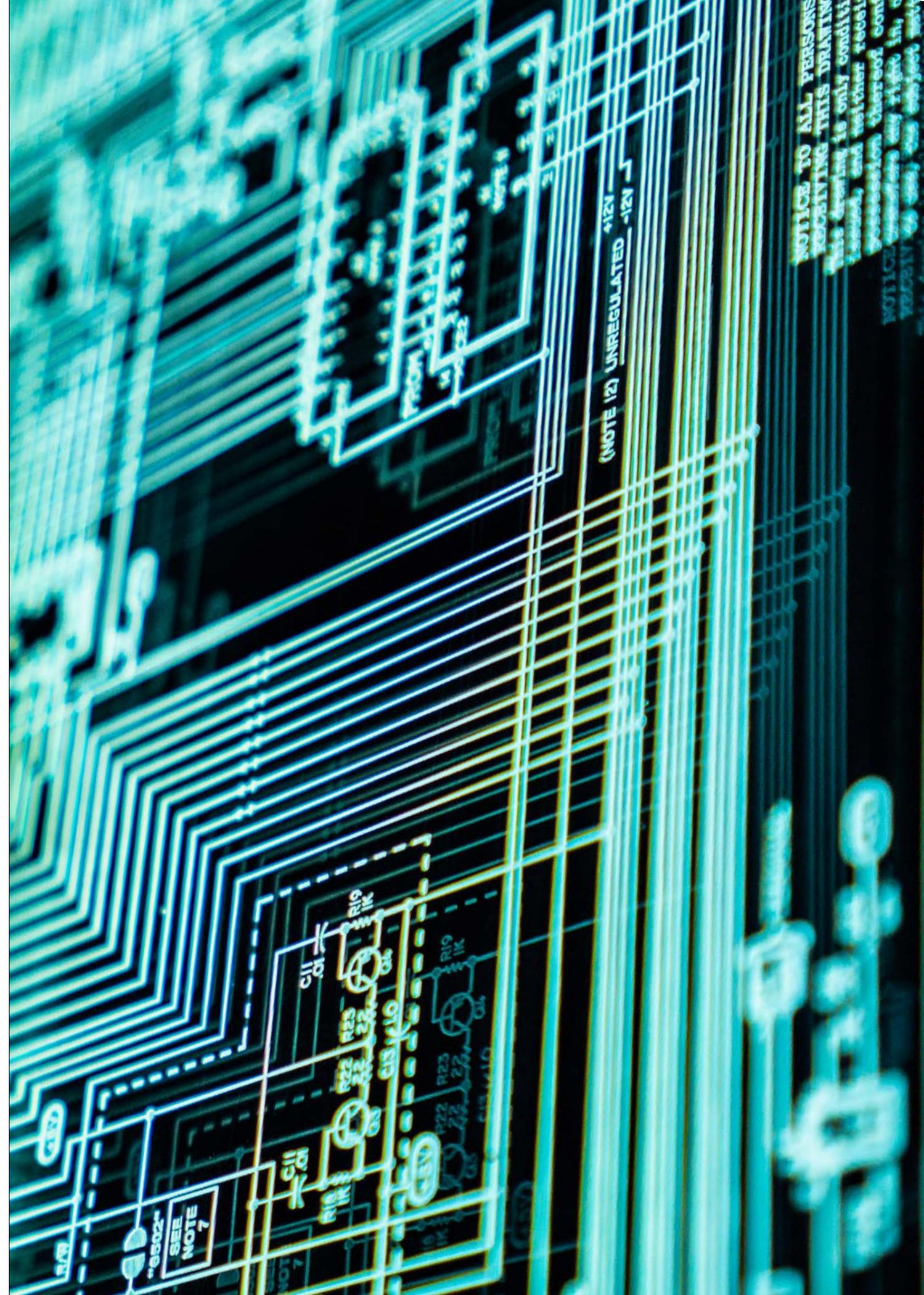
Die Empfehlung setzt sich zum Ziel, weltweit akzeptierte ethische Standards für KI-Technologien unter voller Beachtung des Völkerrechts, insbesondere der Menschenrechte, auszuformulieren, die dann in der Folge eine zentrale Rolle bei der Entwicklung von KI-bezogenen Normen auf der ganzen Welt spielen können.

Bemerkenswert sind der ganzheitliche Fokus auf die verschiedenen Politikfelder, das Bewusstsein, dass diese je unterschiedliche Regulierungsansätze verlangen, sowie der Fokus auf „blind spots“ bisheriger KI-Regeln. Dazu gehören der Umweltschutz, die nachhaltige und ressourcenschonende Verwendung von KI, und die Nutzung von KI im Bildungsbereich unter voller Anerkennung des Rechts auf Bildung für alle.

## Normative Technik

Je nach Handlungsfeld wohnt den empfohlenen Maßnahmen ein unterschiedlicher Verpflichtungsgrad inne. Zwar stellt die Empfehlung in ihrer Gesamtheit einheitlich „soft law“ da, also nicht formal bindendes Recht, doch wurde die Empfehlung in einem mehrjährigen Ausverhandlungsprozess so detailliert verhandelt wie ein völkerrechtlicher Vertrag, sodass sie sich qualitativ von einfachen Resolutionen oder Erklärungen unterscheidet. Die Empfehlung kann nicht vor Gerichten durchgesetzt werden, wird jedoch gegenüber Staaten wirkmächtig. Sie fordert unter anderem konkrete Governance-Maßnahmen (wie das Ethical Impact Assessment (EIA) für KI-Systeme oder den Aufbau eines Netzwerks unabhängiger AI Ethics Officer zur Kontrolle des EIA), internationale Zusammenarbeit und Forschung im KI-Bereich sowie die mittel-indifferente Auswahl nötiger Maßnahmen durch Nationalstaaten.

Mehrwert der  
UNESCO KI-Ethik-Empfehlung  
im Vergleich zu anderen  
internationalen Initiativen



# II. Mehrwert der UNESCO KI-Ethik-Empfehlung im Vergleich zu anderen internationalen Initiativen

Um den Mehrwert der UNESCO-Empfehlung fundiert einschätzen zu können, ist es erforderlich, zunächst auch andere internationale ethisch fundierte KI-Initiativen zu betrachten. So wird deutlich, in welchen Policy- und Governance-Feldern die UNESCO-Empfehlung neue, weiterführende Vorschläge macht, sie neu strukturiert oder neu konzipierte Themen, Instrumente oder Governance-Ansätze einführt, und ob neben derartigen inhaltlichen Aspekten die UNESCO-Empfehlung zusätzlich einen strukturellen Mehrwert beinhaltet.

Neben der Empfehlung der UNESCO sind für Deutschland besonders relevant der Europarat mit seinem dafür eingesetzten Ad-hoc-Ausschuss für Künstliche Intelligenz (CAHAI), die Europäische Union sowie die OECD, welche jeweils Leitlinien und Gesetze im Themenfeld der ethischen KI-Nutzung erlassen haben bzw. diese aktuell verhandeln.

## II.1. OECD-Empfehlung zu Künstlicher Intelligenz

Die Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (Organisation for Economic Co-operation and Development, OECD) hat vor dem Hintergrund der umgreifenden Entwicklung und Einführung von Systemen der KI das Erfordernis der Entwicklung internationaler Policy-Instrumente erkannt. Im Zuge dessen hat das OECD Committee on Digital Economy Policy (CDEP) unter Einbeziehung von Regierungen, Industrie, Zivilgesellschaft und Gewerkschaften mittels Bildung einer Expertengruppe (AIGO) Empfehlungen erarbeitet, was schlussendlich in

deren Verabschiedung als „Empfehlung des OECD-Rats zu Künstlicher Intelligenz“ am 22. Mai 2019 mündete.<sup>8</sup>

In der OECD-Empfehlung werden vorangestellt einige Termini definiert [I.]: Die Definition von KI-Systemen fällt dabei so breit aus, dass die Frage aufkommt, worin der Unterschied von KI-Systemen und sonstigen Softwaresystemen bestehen soll. Auch das Begriffsmerkmal „maschinenbasiertes System“<sup>9</sup> hilft nicht besonders weiter. Bemerkenswert ist hier, dass der Lebenszyklus eines KI-Systems samt seiner Phasen ausdrücklich definiert wird und insbesondere die Phase von Design, Datensammlung und Verarbeitung und Modellierung des Systems als „kontextabhängige Abfolge“ bezeichnet wird.<sup>10</sup> Als nach eigenem Dafürhalten erste zwischenstaatliche Standardisierungsinitiative zu Künstlicher Intelligenz<sup>11</sup> stellt die OECD-Empfehlung in einem ersten Abschnitt folgende Grundsätze für eine verantwortungsvolle Steuerung vertrauenswürdiger Künstlicher Intelligenz auf: inklusives Wachstum, nachhaltige Entwicklung und Lebensqualität als Ziel des Einsatzes von Systemen Künstlicher Intelligenz [IV.1.1.], menschenzentrierte Werte und Fairness [IV.1.2.], Transparenz und Erklärbarkeit [IV.1.3.], Robustheit und Sicherheit der Systeme [IV.1.4.] sowie Rechenschaftspflichten der „KI-Akteure“ [IV.1.5.].<sup>12</sup>

Die OECD-Empfehlung sieht in ihrem zweiten Abschnitt betreffend einzelstaatlicher Maßnahmen und internationaler Zusammenarbeit die Förderung eines „digitalen Ökosystems“ vor, welches die für einen angemessenen Austausch von „KI-Wissen“ erforderlichen digitalen Technologien, Infrastrukturen sowie Mechanismen wie ein Datentreuhandmodell umfassen soll [V.2.2.].<sup>13</sup> Hieran zeigt sich erneut, dass hier eine spezifisch wirtschaftliche Perspektive eingenommen wird. Verstärkt wird dieser Eindruck durch die Forderung nach einem „für KI günstigem Politikumfeld“ [V.2.3.].<sup>14</sup> Zum einen soll ein – zügiger – Übergang von einer Forschungs- und Entwicklungsphase hin zu Implementierung und Betrieb von vertrauenswürdigen KI-Systemen erreicht werden [V.2.3.a].<sup>15</sup> Zum anderen werden die

8 OECD, Empfehlung des Rats zu künstlicher Intelligenz, OECD/LEGAL/0449, 22. Mai 2019, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.

9 Ebenda, S. 7.

10 Ebenda.

11 So in den vorangestellten Hintergrundinformationen, siehe ebenda, S. 3 ff.

12 Ebenda, S. 7 ff.

13 Ebenda, S. 8.

14 Ebenda, S. 9.

15 Ebenda.

Länder angehalten, ihre politischen und regulatorischen Rahmenbedingungen einer Überprüfung und gegebenenfalls notwendigen Anpassung zuzuführen, um KI-bezogenen Innovationen und Wettbewerb Vorschub zu leisten [V.2.3.b)].<sup>16</sup> Eine wirtschaftsorientierte, grundsätzlich den Einsatz von Künstlicher Intelligenz befürwortende Position ist in der OECD-Empfehlung folglich nicht von der Hand zu weisen.

Weiterhin geht die OECD-Empfehlung davon aus, dass es notwendigerweise zu einer zunehmenden Implementation von KI-Systemen in der Wirtschaft kommen wird, womit ein Wandel von Arbeitswelt und Gesellschaft korrespondiere, sodass sich der Mensch der Fähigkeiten bemächtigen müsse, die für die effektive Nutzung von und Interaktion mit KI-Systemen erforderlich sind [V.2.4.].<sup>17</sup> Solche Einschätzungen werden allerdings zu Recht kritisch gesehen, da sie möglicherweise anderweitige Ursachen von gesellschaftlichen und volkswirtschaftlichen Problemen sowie hierauf bezogene Gegenmaßnahmen verdecken könnten.<sup>18</sup> Insgesamt ist die Empfehlung des OECD-Rats ein prägnantes Dokument, das aus wirtschaftsorientierter Perspektive erste Linien für den Einsatz von KI zieht. Die dort aufgeführten Werte sind einerseits von solch universeller und naheliegender Natur, dass sich wohl die meisten Stakeholder und Akteur\*innen einer Zustimmung nicht entziehen würden. Andererseits verbleiben diese auf einer abstrakten Ebene, sodass die Frage nach konkreten Ausformungen und Policy- wie Governance-Maßnahmen offenbleibt. Hier liegt ein bedeutender Unterschied zu der Empfehlung der UNESCO, die eben nicht abstrakt bleibt, sondern an verschiedenen Stellen normativ abgestufte, politikfeldspezifische Forderungen aufstellt, die unterschiedliche gesellschaftliche KI-Akteur\*innen konkret in die Pflicht nehmen.

## II.2. Europarat: CAHAI

Aktuell beschäftigt sich der Europarat in seinen Gremien intensiv mit Künstlicher Intelligenz und ihrer Regulierung. Am 11. September 2019 mandatierte das Ministerkomitee ein Ad-hoc-Komitee für Künstliche Intelligenz (Ad Hoc Committee on Artificial Intelligence, CAHAI) mit der Durchführung eine Studie, unter anderem auf Basis von Multi-Stakeholder-Konsultationen sowie der Menschenrechtsstandards des Europarates. Diese sollte die Machbarkeit eines rechtlichen Rahmenwerks und die möglichen Elemente eines solchen ermitteln.<sup>19</sup> Am

16 Ebenda.

17 Ebenda.

18 Vgl. wirtschaftshistorisch: Aaron Benanav, Automatisierung und die Zukunft der Arbeit (übers. von Felix Kurz, Suhrkamp 2021), S. 17 ff.

19 Lisa Peets/Marty Hansen/Sam Jungyun Choi/Marianna Drake, AI Update: The Council of Europe Publishes Feasibility Study on Developing a Legal Instrument for Ethical AI, Global Policy Watch, 14. Januar 2021, <https://www.globalpolicywatch.com/2021/01/ai-update-the-council-of-europe-publishes-feasibility-study-on-developing-a-legal-instrument-for-ethical-ai>.

17. Dezember 2020 wurde diese Studie schlussendlich veröffentlicht,<sup>20</sup> ergänzt um eine die bisher eingebrachten Perspektiven zusammenfassende Publikation<sup>21</sup>. Die so angestoßene Entwicklung schreitet weiter voran. Derzeit hat die Legal Frameworks Group of the CAHAI (CAHAI-LFG) nach Multi-Stakeholder-Konsultationen und Plenartreffen des CAHAI<sup>22</sup> einen ersten Entwurf der möglichen Elemente rechtlich bindender Instrumente im Zusammenhang mit Künstlicher Intelligenz, Menschenrechten, Demokratie und Rechtsstaatlichkeit konzipiert, der nun von dem CAHAI begutachtet und anschließend in finalisierter Form dem Ministerkomitee vorgelegt werden soll.<sup>23</sup>

Techniken der Künstlichen Intelligenz, hier als Dienste und Produkte erfasst, wird ein Potenzial zur Förderung der Wohlfahrt sowie des individuellen und gesellschaftlichen Wohlstands zugeschrieben, wobei auf der Kehrseite ebenso potenzielle Negativeffekte und Gefahren, insbesondere hinsichtlich der durch die Europäische Menschenrechtskonvention (EMRK) geschützten Rechte, anerkannt werden [Rn.2].<sup>24</sup> Beispiele für die in der Studie genannten Risiken für die Menschenrechte sind KI-Systeme, die das Recht auf Gleichheit und Nichtdiskriminierung unterminieren, indem sie Vorurteile und Stereotype (z.B. bei der Beschäftigung) perpetuieren, sowie KI-gesteuerte Überwachungs- und Tracking-Anwendungen, die das Recht des Einzelnen auf Versammlungsfreiheit und freie Meinungsäußerung gefährden.<sup>25</sup> Weiterhin erkennt CAHAI ausdrücklich das Fehlen einer uniformen, universell anerkannten Definition des Terminus der Künstlichen Intelligenz an und stuft diesen als Sammelbegriff ein, dem sich technologieneutral zu nähern sei [Rn. 8–10].<sup>26</sup>

In der angeschlossenen CAHAI-Machbarkeitsstudie wird ein breites Spektrum von Instrumenten zur Regulierung von Künstlicher Intelligenz betrachtet, darunter internationale Rechtsinstrumente wie die EMRK und die EU-Grund-

20 CAHAI, Feasibility Study, CAHAI(2020)23, 17. Dezember 2020, <https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da>.

21 Siehe Europarat, Towards Regulation of AI Systems, DGI(2020)16, Dezember 2020, <https://rm.coe.int/prems-107320-gbr-2018-compli-cahai-couv-texte-a4-bat-web/1680a0c17a>.

22 Siehe dazu CAHAI, Abridged Meeting Report and List of Decisions, CAHAI(2021)10, 7. Juli 2021, <https://rm.coe.int/cahai-2021-10-5th-plenary-abridged-report-2776-1003-8532-v-2/1680a31d48>.

23 Mitteilung des Europarates vom 03. November 2021, <https://www.coe.int/en/web/artificial-intelligence/-/the-cahai-lfg-holds-its-5th-and-final-meeting>.

24 CAHAI, Feasibility Study, CAHAI(2020)23, 17. Dezember 2020, <https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da>, S. 2 ff.

25 Lisa Peets/Marty Hansen/Sam Jungyun Choi/Marianna Drake, AI Update: The Council of Europe Publishes Feasibility Study on Developing a Legal Instrument for Ethical AI, Global Policy Watch, 14. Januar 2021, <https://www.globalpolicywatch.com/2021/01/ai-update-the-council-of-europe-publishes-feasibility-study-on-developing-a-legal-instrument-for-ethical-ai>.

26 CAHAI, Feasibility Study, CAHAI(2020)23, 17. Dezember 2020, <https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da>, S. 4.



Deutsche UNESCO-Kommission

rechtcharta [Rn. 68 ff.]<sup>27</sup>, Ethikrichtlinien für KI, einschließlich solcher, die von privaten Unternehmen und Organisationen des öffentlichen Sektors entwickelt wurden [Rn. 74 ff.]<sup>28</sup>, und nationale Instrumente und Strategien für KI [Rn. 78 ff.]<sup>29</sup>. Die Studie wägt die Vor- und Nachteile dieser Maßnahmen ab und stellt fest, dass es derzeit kein verbindliches internationales Rechtsinstrument gibt, das speziell auf die Herausforderungen durch KI-Systeme zugeschnitten ist [Rn. 82 ff.]<sup>30</sup>. Zudem führen die bereits in Vielzahl existierenden, sich überschneidenden und verstärkend wirkenden Instrumente zu einer Erschwerung konsistenter und umfassender Interpretation. Dies hat zur Folge, dass kein einheitliches Schutzniveau garantiert werden kann und die CAHAI-Studie zu dem Schluss kommt, dass der kontemporäre Menschenrechtsschutz in dieser Hinsicht gesamtheitlich Lücken aufweist [Rn. 83 ff.]<sup>31</sup>.

Neben anderen Faktoren wird besondere Betonung auf die Notwendigkeit ausreichender menschlicher Kontrolle und Aufsicht<sup>32</sup> der technischen Robustheit von KI-Anwendungen sowie

auf effektive Transparenz und Erklärbarkeit<sup>33</sup> gelegt.<sup>34</sup> Als Antwort auf die Herausforderungen stellt die CAHAI-Studie die Kernelemente eines Rechtsrahmens, in Form von Prinzipien, Rechten und Obligationen auf [Rn. 95 ff.]<sup>35</sup>. So und indem spezifiziert wird, was konkret von einzelnen Menschenrechten erfasst wird, wie sich Betroffene effektiv auf Rechte berufen können, und welchen Anforderungen Entwickler\*innen und Anwender\*innen von KI-Systemen genügen müssen, um zu einem hinreichenden Menschenrechtsschutz beizutragen, sollen die bestehenden Menschenrechte in den Kontext des Einsatzes von KI-Systemen integriert werden. Im Zuge dessen identifiziert das CAHAI sieben essenzielle Prinzipien, unter anderem die Menschenwürde [Rn. 98 ff.], das Gebot der Nichtdiskriminierung, Geschlechtergerechtigkeit, Fairness und Diversität [Rn. 105 ff.]<sup>36</sup>. Die Empfehlung der CAHAI-Studie an den Europarat ist schlussendlich ein rechtlich bindendes Instrument, das auf Grund-

27 Ebenda, S.18 ff.

28 Ebenda, S.20 ff.

29 Ebenda, S.21 ff.

30 Ebenda.

31 Ebenda.

32 Siehe dazu auch Riikka Koulu, „Proceduralizing Control and Discretion: Human Oversight in Artificial Intelligence Policy“, Maastricht Journal of European and Comparative Law 27.6 (2020).

33 Siehe dazu u.a. Thomas Wischmeyer, „Artificial Intelligence and Transparency: Opening the Black Box“, in: Thomas Wischmeyer/Timo Rademacher (Hg.), *Regulating Artificial Intelligence* (Springer International Publishing 2020); Margot E Kaminski, „Understanding Transparency in Algorithmic Accountability“, in: Woodrow Barfield (Hg.), *The Cambridge Handbook of the Law of Algorithms* (Cambridge University Press 2020).

34 Lisa Peets/Marty Hansen/Sam Jungyun Choi/Marianna Drake, *AI Update: The Council of Europe Publishes Feasibility Study on Developing a Legal Instrument for Ethical AI*, Global Policy Watch, 14. Januar 2021, <https://www.globalpolicywatch.com/2021/01/ai-update-the-council-of-europe-publishes-feasibility-study-on-developing-a-legal-instrument-for-ethical-ai>.

35 CAHAI, *Feasibility Study*, CAHAI(2020)23, 17. Dezember 2020, <https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da>, S.18 ff.

36 Ebenda, S.27 ff.

UNESCO-Empfehlung zur Ethik künstlicher Intelligenz

lage obiger Prinzipien als Basis nationaler Gesetzgebungsvorhaben dienen könnte und einen risikobasierten und -differenzierten Ansatz verfolgen sollte [Rn. 95 ff.]<sup>37</sup>. Daneben werden weitere Soft- und Hard-Law-Instrumente empfohlen, die präzise und detailliert auf die sektorspezifischen Besonderheiten des KI-Einsatzes zugeschnitten sein sollen.<sup>38</sup>

Während die CAHAI-Studie eine Expert\*innenstudie ist, stellt sich die Empfehlung der UNESCO als Ergebnis eines diplomatischen Ausverhandlungsprozesses zwischen UNESCO-Mitgliedstaaten dar. Die CAHAI-Studie musste daher nicht den Konsens suchen. Auch fällt auf, dass die Studie bewusst auf die nötige Ausarbeitung rechtlicher Regelungen verweist. Mit der Erarbeitung des Menschenrechtsacquis und der Herausstellung der Bezüge zu KI ist die Studie sehr wertvoll, aber kategorial anders einzuordnen als die politikfeldspezifische Empfehlung der UNESCO mit dem breiten ethisch grundierten Ansatz.

### II.3. Gesetzgebungsvorhaben der Europäischen Union

Als ambitioniert – und vom Bindungscharakter als europäisches Recht unmittelbare Anwendung und Vorrang genießend – lässt sich das auf eine Regulierung Künstlicher Intelligenz gerichtete Gesetzgebungsvorhaben der Europäischen Union bezeichnen. Am 21. April 2021 präsentierte die Europäische Kommission ihren Entwurf einer Verordnung zu Künstlicher Intelligenz<sup>39</sup> – eine erste, wesentliche Stufe auf einem möglicherweise langwierigen und komplexen Gesetzgebungsprozess. Dem ging bereits die Einsetzung einer High-Level Expert Group on AI, die einen umfassenden Bericht erarbeitete,<sup>40</sup> sowie die Vorstellung des „Weißbuch zur Künstlichen Intelligenz – Ein europäisches Konzept für Exzellenz und Vertrauen“<sup>41</sup> voraus; die Prozesse sind aber nicht bruchfrei miteinander verknüpft.

37 Ebenda.

38 Lisa Peets/Marty Hansen/Sam Jungyun Choi/Marianna Drake, *AI Update: The Council of Europe Publishes Feasibility Study on Developing a Legal Instrument for Ethical AI*, Global Policy Watch, 14. Januar 2021, <https://www.globalpolicywatch.com/2021/01/ai-update-the-council-of-europe-publishes-feasibility-study-on-developing-a-legal-instrument-for-ethical-ai>.

39 Europäische Kommission, *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*, COM(2021) 206 final, 21. April 2021, <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206>.

40 High-Level Expert Group on AI, *Policy and Investment Recommendations for Trustworthy AI*, 26. Juni 2019, S.48, <https://digital-strategy.ec.europa.eu/en/library/policy-and-investment-recommendations-trustworthy-artificial-intelligence>.

41 Europäische Kommission, *Weißbuch zur Künstlichen Intelligenz – Ein europäisches Konzept für Exzellenz und Vertrauen*, COM(2020) 65 final, 19. Februar 2020, [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_de.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_de.pdf).

Der Gesetzesentwurf der Europäischen Kommission lehnt sich in Struktur und Herangehensweise an etablierte Maßnahmen im Bereich der Produktsicherheit an.<sup>42</sup> Er verfolgt – so wie von der CAHAI-Studie empfohlen – einen nach eigener Aussage „gut-definierten“, risikobasierten und regulativen Ansatz.<sup>43</sup> In Art. 3 wird ein umfassender Katalog an Begriffsbestimmungen eingeführt, darunter unter Nr. 1 Systeme der Künstlichen Intelligenz als „eine Software, die mit einer oder mehreren der in Anhang I aufgeführten Techniken und Konzepte entwickelt worden ist und im Hinblick auf eine Reihe von Zielen, die vom Menschen festgelegt werden, Ergebnisse wie Inhalte, Vorhersagen, Empfehlungen oder Entscheidungen hervorbringen kann, die das Umfeld beeinflussen, mit dem sie interagieren“.<sup>44</sup> Durch diese Kreuzung von funktionaler und deskriptiver Definition fällt das regulative Spektrum des Kommissionsentwurfes weit aus, es werden zwar nicht jegliche Arten von Software erfasst, wohl aber einige Merkmale der meisten Systeme.<sup>45</sup>

Die Kommission differenziert zwischen vier verschiedenen Risiko-Leveln von Praktiken Künstlicher Intelligenz, namentlich unakzeptablen, hohen, beschränkten und minimalen Risiken.<sup>46</sup> So werden bestimmte Praktiken der Künstlichen Intelligenz in Art. 5 ausdrücklich untersagt, wenn auch unter gewissen Erlaubnisvorbehalten.<sup>47</sup> An Hochrisiko-KI-Systeme werden in Kapitel 2, Art. 8 ff. des Entwurfes bestimmte Anforderungen gestellt, darunter Transparenz- und Informationspflichten in Art. 13 und menschliche Aufsicht in Art. 14.<sup>48</sup> Kapitel 3, Art. 16 ff. statuiert Pflichten der Anbieter\*innen und Nutzer\*innen sowie anderer Beteiligter bei dem Einsatz von Hochrisiko-KI-Systemen.<sup>49</sup> In diesem Zusammenhang wurde bereits kritisiert, dass der Text des Entwurfes hinsichtlich der Definition von Hochrisiko-Systemen uneindeutig und daher verbesserungswürdig ist.<sup>50</sup> Art. 56 ff. sehen die Einrichtung eines „Europäischen Ausschusses für Künstliche Intelligenz“ zur Beratung und Unterstützung der Kommission vor.<sup>51</sup> Die Anwendung und Durchführung der Verordnung wird in Art. 59 nationalen Aufsichtsbehörden zugewiesen, die von den

42 Michael Veale/Frederik Zuiderveen Borgesius, „Demystifying the Draft EU Artificial Intelligence Act – Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach“, *Computer Law Review International* 22.4 (2021), S. 97 f.

43 Europäische Kommission, *Gesetz über Künstliche Intelligenz*, S.3.

44 Ebenda, S.39 ff.

45 Siehe dazu kritisch Michael Veale/Frederik Zuiderveen Borgesius, „Demystifying the Draft EU Artificial Intelligence Act – Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach“, *Computer Law Review International* 22.4 (2021), S.109.

46 Ebenda, S.98.

47 Europäische Kommission, *Gesetz über Künstliche Intelligenz*, S.43 ff.

48 Ebenda, S.46 ff.

49 Ebenda, S.52 ff.

50 Luciano Floridi, „The European Legislation on AI: A Brief Analysis of Its Philosophical Approach“, *Philosophy & Technology* 34 (2021), S.215, 219.

51 Europäische Kommission, *Gesetz über Künstliche Intelligenz*, S.72 ff.

Mitgliedstaaten benannt oder eingerichtet werden sollen.<sup>52</sup> Art. 60 sieht zudem die Einrichtung einer EU-Datenbank für eigenständige Hochrisiko-KI-Systeme durch die Kommission vor,<sup>53</sup> was als durchaus sinnvoller Schritt bewertet werden kann.<sup>54</sup>

Die gesamte regulative Struktur des Kommissionsentwurfes hier weiter umfassend zu beschreiben und kritisch zu erfassen, würde den Rahmen sprengen. Als Kritikpunkte gilt es hier gleichwohl – als Beispiel – zu erwähnen, dass der Kommissionsentwurf keine hinreichenden Beschwerde- und Rechtsschutzmechanismen für von KI-Systemen betroffene Individuen und Gemeinschaften beinhaltet, sowie, dass das Vollzugssystem ein gewisses Maß an Inkohärenz aufweist.<sup>55</sup> Dieser Kommissionsentwurf als weltweit erster Versuch einer horizontalen gesetzlichen Regulierung von Systemen der Künstlichen Intelligenz fügt sich gleichwohl in ein Bild von der Europäischen Union als digitalpolitischem Big Player ein, der in seiner kritischen Auseinandersetzung mit den normativen Herausforderungen der Digitalisierung konkurrenzlos ist – was freilich Risiken mit sich bringen könnte.

#### II.4. UNESCO KI-Ethik-Empfehlung: Ansatz und Mehrwert

Die UNESCO KI-Ethik-Empfehlung<sup>56</sup> verfolgt aus nachvollziehbaren Gründen keinen mit den Vorhaben der EU vergleichbar starken regulativen Anspruch. Als ergänzendes wie grundlegendes normatives Instrument [siehe Rn. 5 ff.] versucht die Empfehlung in einem umfassenden Dokument eigene Akzente zu setzen, indem es einerseits Mitgliedstaaten bei der Errichtung rechtlicher und regulativer Rahmenordnungen unterstützen und andererseits allen „KI-Akteuren“ als Grundlage für ethische Folgenabschätzung entlang des Lebenszyklus von KI-Systemen dienen will [Rn. 4]. Die so eingenommene Perspektive auf den gesamten Lebenszyklus eines KI-Systems ist zu begrüßen, da über die einzelne Entscheidung, Empfehlung oder Vorhersage hinausgehend vor- und nachgelagerte Maßnahmen und Handlungen sowie damit zusammenhängende Problematiken nicht aus einer notwendigerweise umfassenden Betrachtung herausfallen [Rn. 2]. In der Empfehlung wird auf solche Systeme abgestellt, die Kapazitäten zur Verarbeitung von Daten und Informationen besitzen in einer Art und Weise, die intelligentem Verhalten ähnele – freilich ohne wiederum den Begriff der Intelligenz zu definieren – und typischerweise Aspekte des logischen Denkens, des Lernens, der Wahrnehmung, der Vorhersage, der Planung oder der Kontrolle beinhaltet [Rn. 2].

Im weiteren Verlauf stellt die UNESCO-Empfehlung zunächst die maßgeblichen betroffenen Werte dar [III.1, Rn. 13 ff.] und konkretisiert diese wiederum in Prinzipien [III.2, Rn. 25 ff]. Diese werden ihrerseits in politischen Aktionsfeldern („areas of policy action“) operationalisiert [Rn. 48 ff]. Neben eher einschlägigen Bereichen, wie (ethische) Folgenabschätzungen [Rn. 50 ff.] und die Einrichtung starker Governance-Mechanismen [Rn. 54 ff.], werden dabei explizit und ausführlich bisher (wohl eher) vernachlässigte Bereiche, wie Schutz von Umwelt und Ökosystemen [Rn. 84 ff.], Gender [Rn. 87 ff.] und – als Kern-Mandatsfelder der UNESCO – Kultur [Rn. 94 ff.] sowie Bildung und Wissenschaft [Rn. 101 ff.] benannt. Hervorzuheben ist, dass sich hier etwaigen Vorstellungen in Politik, Wirtschaft und Gesellschaft entgegengestellt wird, der Einsatz von KI-Technologien sei per se ein Wundermittel, um das Wohlergehen von Mensch, Umwelt und Ökosystemen zu gewährleisten [Rn. 25]. Die Notwendigkeit menschlicher und öffentlicher Aufsicht wird ebenso expliziert wie, dass es im Ergebnis stets bei einer Verantwortlichkeit und Zurechenbarkeit von Menschen bleibe [Rn. 35 f.].

Im Übrigen verdienen verschiedene Aspekte der UNESCO-Empfehlung besondere Aufmerksamkeit, wie die in der Präambel eingebrachte Erkenntnis, dass der Einsatz von Systemen Künstlicher Intelligenz auch den menschlichen Verstand, die Denk-, Interaktions- und Entscheidungsweisen des Menschen und damit das Menschenbild beeinflusst. Die UNESCO-Empfehlung unterstreicht in ihrem globalen Anspruch eben auch die Besonderheiten und Herausforderungen von Ländern mit geringem und mittlerem Einkommen (Low- and Middle-Income Countries, LMICs) – unter denen sogenannte Least Developed Countries (LDCs), Landlocked Developing Countries (LLDCs) sowie Small Island Developing States (SIDS) hervorgehoben werden – und damit Anliegen insbesondere des Globalen Südens.<sup>57</sup> Die Kulturen, Werte und das Wissen der indigenen Bevölkerung, deren Anerkennung, Schutz und Förderung es zur Entwicklung nachhaltiger digitaler Ökonomien bedürfe, werden besonders herausgestellt und in den politischen Handlungsfeldern konkretisiert [vgl. bspw. Rn. 85].<sup>58</sup>

Damit weist die UNESCO-Empfehlung auf einzigartige Kontexte in anderen Staaten und Gesellschaften hin, die für die Entwicklung und Implementierung von Systemen Künstlicher Intelligenz überaus relevant sind. Bisher ist dieser Gesichtspunkt allenfalls am Rande ethischer Diskussionen, gerade mit Blick auf die beschriebenen Rahmenwerke, erwähnt worden. Dies mag nachvollziehbar sein; vor allem angesichts

des spezifisch europäischen Zuschnittes der CAHAI-Empfehlung und der Gesetzgebungsvorschläge der Europäischen Kommission. Für nationalstaatliche wie europäische digitalpolitische Erwägungen ist dies gleichwohl relevant angesichts des Umstandes, dass gerade das regulative Vorgehen der Europäischen Union international sehr genau beobachtet wird, mit erheblichen Auswirkungen auf Policy und Governance anderer Staaten; prominent ist zum Vergleich hier die Datenschutz-Grundverordnung zu nennen.

57 Ebenda, S.2.  
58 Ebenda, S.19.

Eine besondere Legitimation ergibt sich für die UNESCO-Empfehlung durch den bei der Erarbeitung genutzten Multi-Stakeholder-Ansatz: Bewusst wurde Input von Vertreter\*innen unterschiedlicher gesellschaftlicher Gruppen gesucht. Auch insofern generiert die UNESCO-Empfehlung Mehrwert, als sie als Beispiel für andere internationale Normbildungsprozesse dienen kann.<sup>59</sup>

Ein profundes Verständnis von Systemen der Künstlichen Intelligenz und den Folgen ihrer Implementation setzt ein fundiertes Wissen über die Gesellschaft und ihre Prozesse voraus. Damit im Zusammenhang ließe sich das Bedürfnis nach kontextsensitiven, langfristig orientierten Perspektiven auf die soziale Einbettung solcher Technologien formulieren. Viele der hierdurch verursachten Auswirkungen werden weniger durch die Technologie selbst als vielmehr durch die Bedingungen, unter denen sie eingesetzt wird, und die breitere wirtschaftliche Dynamik, mit der sie verbunden ist, bestimmt. Dies erfordert, dass die Regierungen eine umfassendere und stärker richtungsweisende Rolle übernehmen.<sup>61</sup> Die wertbezogene Empfehlung der UNESCO wird hierbei substantielle Stützfunktionen erfüllen.

#### II.5. Zusammenfassung

Viele ethische Leitlinien zu Künstlicher Intelligenz überschneiden sich in der Propagierung einiger herkömmlicher Prinzipien, divergieren aber im nächsten Schritt insofern hinsichtlich der praktischen Umsetzung, als dass konkrete Vorschläge sich in Details unterscheiden oder solche sogar nur fragmentarisch vorhanden sind.<sup>60</sup>

59 Ebenda, S.26 ff.

60 Siehe dazu schon CAHAI, Feasibility Study, CAHAI(2020)23, 17. Dezember 2020, <https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da>, Rn. 74 ff., S.20.

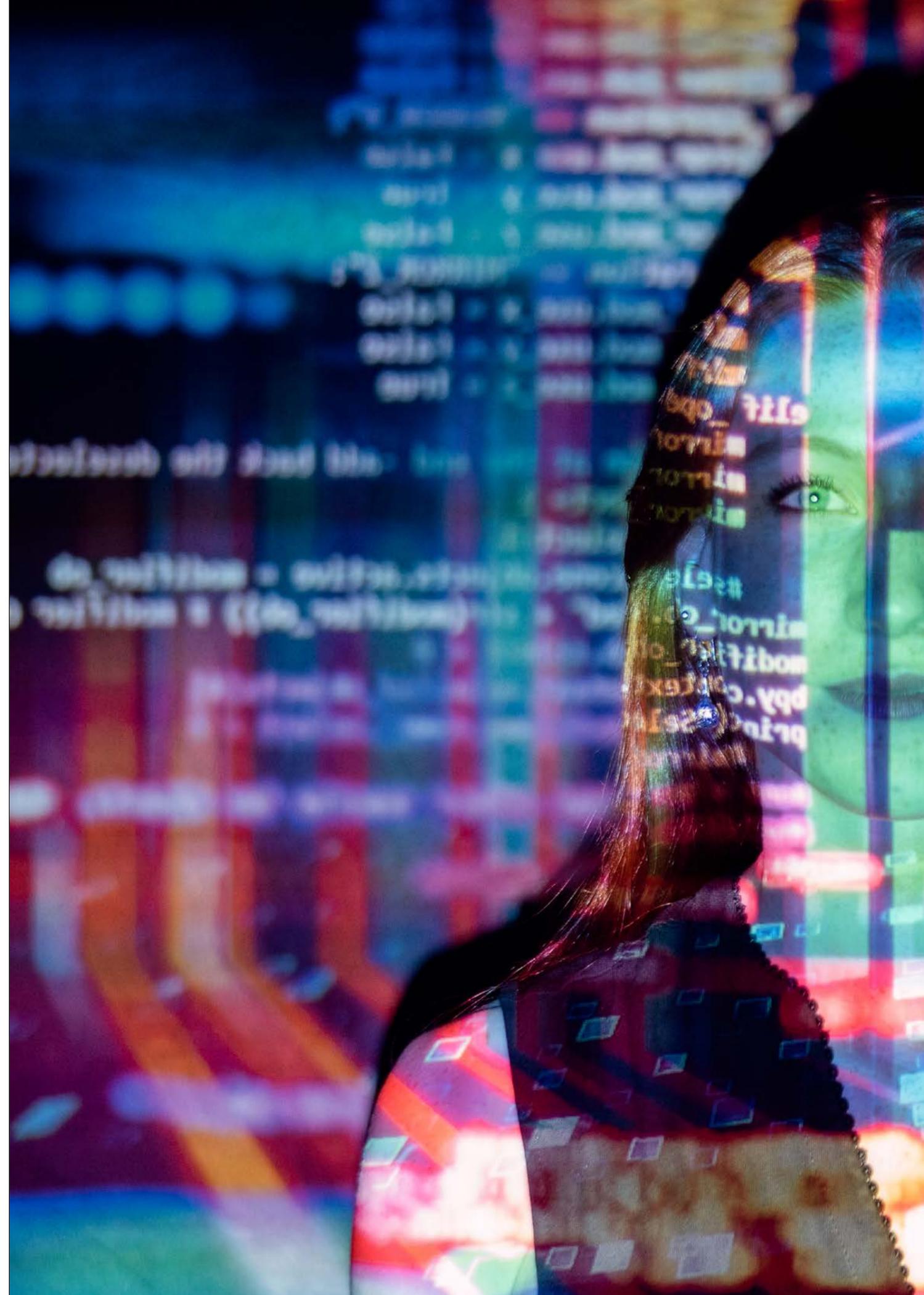
61 So der Niederländische Wissenschaftliche Rat (WRR) in der englischen Zusammenfassung seines Berichts „Opgave AI. De nieuwe systeemtechnologie“, siehe WRR, Mission AI, The New System Technology, 11. November 2021, [https://www.wrr.nl/binaries/wrr/documenten/rapporten/2021/11/11/opgave-ai-de-nieuwe-systeemtechnologie/Summary+WRRreport\\_Mission+AI\\_The+New+System+Technology\\_R105.pdf](https://www.wrr.nl/binaries/wrr/documenten/rapporten/2021/11/11/opgave-ai-de-nieuwe-systeemtechnologie/Summary+WRRreport_Mission+AI_The+New+System+Technology_R105.pdf), S.39.

Ein deutlicher Mehrwert der UNESCO-Empfehlung lässt sich auch dahingehend konstatieren, dass in dieser auf die kulturellen Aspekte hingewiesen wird, die auch bei dem Einsatz von KI-Systemen berührt sein können, und eben auch auf die besonderen wirtschaftlichen Dynamiken, die bei LMICs anders, potenziell deutlich vulnerabler gelagert sein können. Besonders hervorzuheben ist in diesem Zusammenhang, dass auch ausdrücklich auf die substanziellen Beiträge hingewiesen wird, die LMICs zu einer faireren KI-Nutzung und besseren KI-Regulierung leisten.<sup>62</sup> Dass die UNESCO-Empfehlung zudem einen Akzent auf die Bedeutung und Schutzwürdigkeit einzelner besonderer, zum Teil marginalisierter Gruppen legt, kann ein wichtiger Impuls sein, um ein hinreichendes Schutzniveau durch taugliche regulative Instrumente und Mechanismen zu gewährleisten – insbesondere auch im Rahmen des Gesetzgebungsprozesses der EU.

Zu den unter II. schon angesprochenen besonderen Stärken der UNESCO-Empfehlung gehören indes der Legitimitätsfördernde Aushandlungsprozess mit Multi-Stakeholder-Struktur und ihr Charakter als erste globale Regeln für eine ethisch abgesicherte Nutzung von KI (bei 193 beteiligten Staaten). Bedeutende Mehrwerte liegen, besonders im Vergleich zu den anderen Empfehlungen, Studien und Rechtsakten in den konkreten Bezügen zu den elf Politikfeldern, darunter Bildung und Wissenschaft, Kommunikation, Gesundheit sowie Umwelt. Bemerkenswert ist auch, dass in der Empfehlung aus ethischen Regeln (die zu großen Teilen menschenrechtlich abgestützt sind) sinnvolle Forderungen wie Impact Assessments für die KI-Nutzung abgeleitet werden.

62 Siehe beispielsweise Jason Edward Lewis (Hg.), Indigenous Protocol and Artificial Intelligence Position Paper, Honolulu, Hawaii: The Initiative for Indigenous Futures and the Canadian Institute for Advanced Research (CIFAR), 15. April 2020, <https://doi.org/10.11573/spectrum.library.concordia.ca.00986506>.

Bedingungen in Deutschland  
für die Implementierung  
der UNESCO KI-Ethik-Empfehlung



# III. Bedingungen in Deutschland für die Implementierung der UNESCO KI-Ethik-Empfehlung

## III.1. Einführung

Auch auf nationaler Ebene erfährt KI sowohl im öffentlichen wie auch im politischen Diskurs große Aufmerksamkeit – insbesondere im Hinblick auf die gesetzliche Regulierung und Förderung von KI. Mit Blick auf die folgenden Politikfelder der Empfehlung beantworten die kommenden Seiten, welche Bedingungen die Implementierung der Empfehlung fördern.

- Ethical Impact Assessment (Policy Area 1)
- Ethical Governance and Stewardship (Policy Area 2)
- Data Policy (Policy Area 3)
- Development and International Cooperation (Policy Area 4)
- Gender (Policy Area 6)

Zur Einschätzung der Bedingungen für die Umsetzung der Empfehlung in Deutschland werden folgende Fragen mit Blick auf die angeführten Politikfelder adressiert:

Inwieweit gibt es in Deutschland und auf Ebene der EU bereits (politische und regulatorische) Prozesse, Instrumente und Governance-Strukturen, die KI in diesen Politikfeldern gemäß der Empfehlung adressieren (Gesetze, Plattformen, Finanzierungsprogramme, Gremien, Prüfmechanismen, Standards etc.) – wie also sind Deutschland und die EU hier grundsätzlich aufgestellt? Hinsichtlich welcher Inhalte der Empfehlung besteht in Deutschland noch besonders großer Regulierungs- bzw. Handlungsbedarf? Ist dieser Regulierungs- bzw. Handlungsbedarf, falls er besteht, auch schon in anderen Kontexten (z. B. durch die Enquete-Kommission) angesprochen worden?

Daran schließt sich dann jeweils ein kurzer Abgleich mit Regeln oder Initiativen im Globalen Süden (insb. Afrika) an.

## III.2. Ethical Impact Assessment (Policy Area 1)

Im Folgenden sollen die Inhalte der Empfehlung bezüglich ethischer Folgen von KI-Systemen und ihre derzeitige Umsetzung untersucht werden. Dabei wurden primär der Vorschlag für den KI-Rechtsakt der EU (EU-AI)<sup>63</sup> und die nationale

63 Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für Künstliche Intelligenz und zur Änderung bestimmter Rechtsakte der Union, 2021/0106 (COD), 2021. Zitiert als EU-AI.

KI-Strategie der deutschen Bundesregierung<sup>64</sup> (NKIS) in den Blick genommen. Ergänzend wurden die Ethik-Leitlinien für eine vertrauenswürdige KI der Hochrangigen Expertengruppe für Künstliche Intelligenz der EU-Kommission<sup>65</sup> (HEG-KI), der Bericht der Enquete-Kommission Künstliche Intelligenz des deutschen Bundestages<sup>66</sup> (EKKI) und das Gutachten der Datenethikkommission der Bundesregierung<sup>67</sup> (DEK) untersucht.

### Nummer 50 der UNESCO-Empfehlung: Ethische Folgenabschätzungen

Nr. 50 der Empfehlung fordert Mitgliedstaaten auf, Rahmenbedingungen für Impact Assessments (Folgenabschätzungen) einzuführen, z. B. eine „ethische Folgenabschätzung“, die das Ziel hat, die Vorteile, Bedenken und Risiken von KI-Systemen zu ermitteln und zu bewerten, sowie geeignete Maßnahmen zur Risikoprävention, -minderung und -überwachung neben anderen Sicherheitsmechanismen vorzuschlagen. Der EU-AI fokussiert dagegen in den §§ 30 ff. auf Einstufungen einer KI als Hochrisiko-KI, je nachdem, welche Gefahren zu erwarten sind. Auch ethische Aspekte wie die Auswirkungen auf Grundrechte fließen hier ein. Die DEK empfiehlt auf nationaler Ebene, dass der Staat an der Entwicklung ethischer Standards für KI beteiligt sein sollte. Auf diesen Standards sollten Folgenabschätzungen basieren.<sup>68</sup> Die DEK empfiehlt ebenfalls, „Akteure der Zivilgesellschaft, Datenschutzbehörden, Verbraucherschützer oder Vertreter von Betroffenenorganisationen“<sup>69</sup> am Prozess der Standardisierung zu beteiligen. Ähnlich formuliert es die EKKI. Demnach steht die Bundesregierung in der Pflicht, die Grundlagen für eine Folgenabschätzung von

64 Strategie Künstliche Intelligenz der Bundesregierung, Fortschreibung 2020, 2. Dezember 2020, [https://www.ki-strategie-deutschland.de/files/downloads/201201\\_Fortschreibung\\_KI-Strategie.pdf](https://www.ki-strategie-deutschland.de/files/downloads/201201_Fortschreibung_KI-Strategie.pdf). Zitiert als NKIS.

65 Hochrangige Expertengruppe für Künstliche Intelligenz, Ethik-Leitlinien für eine vertrauenswürdige KI, 8. April 2019, <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>. Zitiert als HEG-KI.

66 Bericht der Enquete-Kommission Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale, BT-DrS 19/23700, 28. Oktober 2020, <https://dserver.bundestag.de/btd/19/237/1923700.pdf>. Zitiert als EKKI.

67 Gutachten der Datenethikkommission der Bundesregierung, Oktober 2019, [https://www.bmjv.de/Shared-Docs/Downloads/DE/Themen/Fokusthemen/Gutachten\\_DEK\\_DE.pdf?\\_\\_blob=publicationFile&v=5](https://www.bmjv.de/Shared-Docs/Downloads/DE/Themen/Fokusthemen/Gutachten_DEK_DE.pdf?__blob=publicationFile&v=5). Zitiert als DEK.

68 DEK, S. 76.

69 Ebenda.

KI zu schaffen.<sup>70</sup> Die Nr. 50 der UNESCO-Empfehlung ist somit weitgehend aufgegriffen und muss nur mehr in der Praxis flächendeckend umgesetzt werden.

### Nummer 51 der UNESCO-Empfehlung: Sorgfaltspflichten und Aufsichtsmechanismen

Nr. 51 der Empfehlung fordert Mitgliedstaaten und Unternehmen des Privatsektors auf, Sorgfalts- und Aufsichtsmechanismen zu entwickeln, um die Auswirkungen von KI-Systemen auf die Achtung der Menschenrechte, die Rechtsstaatlichkeit und integrative Gesellschaften zu ermitteln, etwaige negative Konsequenzen zu verhindern bzw. abzumildern und Rechenschaft darüber abzulegen, wie sie damit umgehen.

Der EU-AI sieht gem. Art. 9 auch ein abgestuftes Risikomanagementsystem für KI vor. Dieses umfasst auch Maßnahmen zur Sicherung von Grundrechten. Tests müssen gem. Art. 9 Abs. 6 „geeignet sein, die Zweckbestimmung des KI-Systems zu erfüllen“. Ähnlich findet es sich in der NKIS. Diese fordert eine angemessene Kontrollstruktur und Überprüfbarkeit von KI-Anwendungen.<sup>71</sup> Die HEG-KI betont, dass „menschliche Werte im Mittelpunkt der Entwicklung, Einführung, Nutzung und Überwachung der KI-Systeme stehen“ müssen. Dies diene auch der Achtung der Grundrechte.<sup>72</sup>

Im Gegensatz zur UNESCO-Empfehlung sehen die nationale KI-Strategie und der geplante AI-Rechtsakt der EU keine verpflichtenden Maßnahmen zur Analyse des sozioökonomischen Einflusses von KI auf Armut vor. Die HEG-KI und die EKKI erkennen die Möglichkeit von KI, die Kluft zwischen Arm und Reich zu überbrücken<sup>73</sup> und wirtschaftliche Chancengleichheit zu fördern.<sup>74</sup> Konkrete Maßnahmen zur Erforschung solcher Möglichkeiten finden jedoch keine Erwähnung.

Anspruch auf Herausgabe von Transparenz-Protokollen: Die HEG-KI sieht Transparenz als eines der Kriterien, die eine vertrauenswürdige KI erfüllen muss.<sup>75</sup> Eine angemessene Dokumentation müsse Ex-Post-Kontrollen durch Behörden ermöglichen.<sup>76</sup> Gem. Art. 23 des EU-AI können die zuständigen Behörden die Protokolle nach Art. 20 herausverlangen, wenn dies zur Übereinstimmung der KI mit den Pflichten der Richtlinie notwendig ist. Einen Herausgabeanspruch für andere Personen oder Organisationen sieht der EU-AI nicht vor.

70 EKKI, S. 121.

71 NKIS, S. 25.

72 HEG-KI, S. 48, Rn. 153.

73 Ebenda, S. 11.

74 EKKI, S. 12.

75 HEG-KI, S. 10.

76 Ebenda, S. 29.

Dagegen betont die UNESCO-Empfehlung, dass der Zugang zu Informationen, einschließlich Informationen von öffentlichem Interesse, die sich im Besitz privater Einrichtungen befinden, gesichert sein muss. Dies dient nicht nur dazu, den Einfluss des sozioökonomischen Status auf KI-Nutzung zu bewerten, sondern auch sicherzustellen, dass die Kluft zwischen Menschen, die in Reichtum und Armut leben, sowie die digitale Kluft zwischen und innerhalb von Ländern durch den massiven Einsatz von KI-Technologien jetzt und in Zukunft nicht noch größer wird. Die Mitgliedstaaten, Privatunternehmen und die Zivilgesellschaft sollten die soziologischen und psychologischen Auswirkungen von KI-gestützten Empfehlungen auf die Entscheidungsautonomie des Menschen untersuchen. KI-Systeme, die als potenzielle Risiken für die Menschenrechte identifiziert werden, sollten von KI-Akteur\*innen im Rahmen der ethischen Folgenabschätzung auf breiter Basis getestet werden, gegebenenfalls auch unter realen Bedingungen, bevor sie auf den Markt gebracht werden.

Die Nationale KI-Strategie erkennt Transparenz ebenfalls als einen wichtigen Baustein für KI an.<sup>77</sup> Weiter als dieses Bekenntnis geht das Gutachten der EKKI. Demnach könnte Transparenz auch mittels regulatorischer Maßnahmen durchgesetzt werden, um Gefahren einzudämmen.<sup>78</sup> Auch müssten sich Personen gegen Diskriminierungen von KI zur Wehr setzen können. Dies erfordere einen Anspruch auf Transparenz.<sup>79</sup> Ähnlich sieht es die DEK, welche einen Anspruch auf Transparenz der verwendeten Daten, je nach Risikokategorie der verwendeten KI, für notwendig erachtet.<sup>80</sup>

Soziologische und psychologische Effekte von KI-basierten Empfehlungssystemen müssen erforscht werden: Der EU-AI verbietet in Art. 5 Abs. 1 lit. a zwar jedes KI-System, „das Techniken der unterschwellig Beeinflussung außerhalb des Bewusstseins einer Person einsetzt, um das Verhalten einer Person in einer Weise wesentlich zu beeinflussen, die dieser Person oder einer anderen Person einen physischen oder psychischen Schaden zufügt oder zufügen kann“. Was jedoch einen psychischen Schaden zufügt, wird im Richtlinienvorschlag nicht beantwortet. Auch enthält er keine Pflicht zur Erforschung dieses Feldes. Auch die HEG-KI beschäftigt sich nicht mit den Folgen von Empfehlungsalgorithmen.

Die EKKI sieht in der interdisziplinären Erforschung der Auswirkungen von Empfehlungsalgorithmen eine dringende Aufgabe.<sup>81</sup> Die DEK stimmt mit dieser Annahme überein, beschränkt dies jedoch nicht auf die Auswirkungen von Empfehlungsalgorithmen. Die Auswirkungen von jeder Form von KI müssten weiter erforscht werden.<sup>82</sup>

77 NKIS, S. 25.

78 EKKI, S. 464.

79 Ebenda, S. 64.

80 DEK, S. 186.

81 EKKI, S. 461.

82 DEK, S. 174.



Der EU-AI sieht in Art. 5 umfassende Testpflichten über den gesamten Lebenszyklus von KI vor. Dies umfasst auch Tests vor der Einführung der KI in den Markt. Gem. Art. 6 müssen die Testverfahren geeignet sein, „die Zweckbestimmung des KI-Systems zu erfüllen“. Damit ist ein Test in realer Umgebung nicht zwingend erforderlich. Die nationale KI-Strategie erkennt die Wichtigkeit von Tests an, insbesondere von Tests in Reallaboren.<sup>83</sup> Die Bundesregierung unterstützt etwa digitale Testfelder.<sup>84</sup>

#### Nummer 52 der UNESCO-Empfehlung: Überwachung aller Phasen des KI-Lebenszyklus

Nr. 52 der Empfehlung sieht vor, dass Mitgliedstaaten und Unternehmen als Teil der ethischen Folgenabschätzung geeignete Maßnahmen ergreifen sollen, um alle Phasen des Lebenszyklus eines KI-Systems zu überwachen, einschließlich der Funktionsweise der für die Entscheidungsfindung verwendeten Algorithmen, der Daten sowie der am Prozess beteiligten KI-Akteur\*innen, insbesondere im öffentlichen Dienst und dort, wo eine direkte Interaktion mit Endnutzer\*innen erforderlich ist. Die menschenrechtlichen Verpflichtungen der Mitgliedstaaten sollten Teil der ethischen Aspekte der Bewertung von KI-Systemen sein.

Der EU-AI sieht in Art. 9 Abs. 2 vor, dass das Risikomanagement von Hochrisiko-KI als kontinuierlicher Prozess während des gesamten Lebenszyklus eines KI-Systems zu verstehen ist. Auch die HEG-KI empfiehlt eine solch fortlaufende Überwachung, unter anderem durch Behörden.<sup>85</sup>

Gem. Art. 7 Abs. 1 lit. b des EU-AI ist die Verletzung von Grundrechten ein Aspekt der Risikobewertung von KI. Die nationale KI-Strategie erklärt den Schutz von Grundrechten zum Ziel aller Normen und Standards für KI. Die EKKI sieht das Grundgesetz sowie die Grundrechtecharta der Europäischen Union als Maßstab für die politische Gestaltung von KI.<sup>86</sup> Daher sei die Förderung und Regulierung von KI an diesen Rechten und Werten zu orientieren.<sup>87</sup> Die DEK unterstützt den eingeschlagenen Weg, dass sich die KI-Strategie Europas strikt an „europäischen Werten und Grundrechten, wie (...) insbesondere auch (...) der Charta der Grundrechte der Europäischen Union und (...) der Konvention zum Schutz der Menschenrechte und Grundfreiheiten des Europarats“ orientiere.<sup>88</sup> Dies umfasse auch ein Verbot von KI, wenn diese in einer Weise in „den Kern der Privatsphäre und die Integrität der Persönlichkeit“ eingreift, die mit den Grundrechten nicht vereinbar sei.<sup>89</sup>

85 HEG-KI, S.29.  
86 EKKI, S.83.  
87 Ebenda, S.137.  
88 DEK, S.226.  
89 Ebenda, S.97.

#### Nummer 53 der UNESCO-Empfehlung: Rechtsrahmen für Folgenabschätzungen

Nr. 53 der Empfehlung legt Regierungen nahe, einen Rechtsrahmen zu verabschieden, der ein Verfahren festlegt, insbesondere für Behörden, um ethische Folgenabschätzungen für KI-Systeme durchzuführen, um Folgen vorherzusagen, Risiken zu mindern, schädliche Folgen zu vermeiden, die Beteiligung der Bürger\*innen zu erleichtern und gesellschaftliche Herausforderungen anzugehen. Im Rahmen der Bewertung sollten auch geeignete Aufsichtsmechanismen, einschließlich Überprüfbarkeit, Rückverfolgbarkeit und Erklärbarkeit, geschaffen werden, die eine Bewertung von Algorithmen, Daten und Entwurfsprozessen ermöglichen und eine externe Überprüfung von KI-Systemen beinhalten. Ethische Folgenabschätzungen sollten transparent sein und gegebenenfalls der Öffentlichkeit zugänglich gemacht werden. Solche Abschätzungen sollten auch multidisziplinär, von mehreren Interessengruppen getragen sowie multikulturell, pluralistisch und integrativ sein. Die Behörden sollten verpflichtet werden, die von ihnen eingeführten und/oder genutzten KI-Systeme durch die Einführung geeigneter Mechanismen und Instrumente zu überwachen.

Der EU-AI regelt in Art. 64 Abs. 1 die Befugnisse der Marktüberwachungsbehörden. Diese erhalten uneingeschränkter Zugang zu Trainings-, Validierungs- und Testdatensätzen. Bei speziellen Grundrechtsverletzungen sieht Art. 64. Abs. 3 besondere Maßnahmen vor. Die Bewertung einer KI als Risikosystem obliegt ebenfalls den Behörden und erfolgt nach den in Titel III Kapitel 1 festgelegten Vorgaben.<sup>90</sup>

Die Nationale KI-Strategie sieht die Förderung und Weiterentwicklung von praktischen Ansätzen zum Risikomanagement vor.<sup>91</sup> Außerdem sollten Aufsicht und Durchsetzung von Vorgaben primär den Aufsichtsbehörden zugewiesen werden. Diese hätten bereits Kompetenz auf diesem Gebiet.<sup>92</sup> Ethik-Folgenabschätzungen sollten transparent und, wo angemessen, öffentlich verfügbar sein.

Weder die Empfehlung der HEG-KI noch der EU-AI sehen vor, dass die Ethische Folgenabschätzung veröffentlicht wird. Nach der NKIS gehört „ein risikoadäquates Maß an Transparenz“ zu den Anforderungen an KI.<sup>93</sup> Spezielle Pflichten zur Veröffentlichung von Ethik-Folgenabschätzungen sieht diese jedoch nicht vor.

90 EU-AI, Art. 30 ff.  
91 NKIS, S.64.  
92 Ebenda.  
93 Ebenda, S.25.

#### Fazit

Es besteht noch bedeutender Nachholbedarf seitens des deutschen Gesetzgebers und anderer KI-Akteur\*innen, wenn es um eine Umsetzung der UNESCO-Empfehlung geht. Besonders vielversprechend ist es, die bestehenden Ethik-Empfehlungen und Dokumente von OECD, Europarat und EU gedanklich in ihrer Gesamtheit zu erfassen und die jeweiligen Mehrwerte herauszudestillieren. Die Empfehlung der UNESCO bietet einen globalen und ganzheitlich Rahmen, der politikfeldübergreifend wichtige Maßnahmen vorschlägt. Der kommende Rechtsakt der EU hingegen hat den Vorteil, für eine kleine Gruppe von Staaten verbindlich zu sein. In einem nächsten Schritt wären daher in einer vergleichenden Sichtung die jeweiligen Mehrwerte zu identifizieren, wobei die UNESCO-Empfehlung als thematisch breiteste den besten Rahmen bietet. Nur ein Beispiel: Bis auf die UNESCO-Empfehlung enthalten die untersuchten Policy-Dokumente keine konkreten Pläne zur Erforschung der soziokulturellen Auswirkungen von KI auf Armut.

Obwohl im Entwurf der EU-AI-Richtlinie nicht zu finden, empfehlen andere Institutionen neben der UNESCO den Anspruch auf Herausgabe von Transparenzprotokollen für Nutzer\*innen. Ebenfalls nicht auf europäischer, sondern nur auf nationaler Ebene wird die Erforschung der soziologischen und psychologischen Folgen von Empfehlungsalgorithmen gefordert. Auch fehlt es auf europäischer Ebene an einer Pflicht zum Test der KI in realer Umgebung. Deutschland könnte hier in der Umsetzung der Empfehlung vorangehen und Transparenzkontrolle von KI-Anbieter\*innen anfordern. Auch die digitale Kluft innerhalb Deutschlands zwischen formal Gebildeten und formal weniger Gebildeten, zwischen Arm und Reich, ist zu überwinden. Eine Ethikfolgenabschätzung für Algorithmenutzung ließe sich etwa über das Büro für Technikfolgenabschätzung beim Deutschen Bundestag verankern.

Ein entsprechender klarer Rechtsrahmen fehlt noch, unter anderem weil auf die entsprechende Festlegung im Rechtsakt der EU zur KI gewartet wurde. Der Rechtsrahmen muss sehr sensibel verschiedene Ziele verfolgen können, nämlich sowohl effektive Überwachungsmechanismen einzuführen, als auch diese zu überwachen. Besonders Behörden sollten verpflichtet werden, die eigene verwendete KI zu überwachen: Die DEK empfiehlt eine Pflicht für Behörden, die verwendete KI zu überwachen<sup>94</sup> und eine Ansprechperson für den Einsatz von KI zu benennen.<sup>95</sup> Die EKKI empfiehlt ebenfalls, dass Behörden die von ihnen eingesetzte KI überwachen und dafür die Kompetenzen und Ressourcen der Behörden gestärkt werden.<sup>96</sup> Diese Punkte werden in der Empfehlung auch vorgebracht, aber noch um einige Anforderungen erweitert.

94 DEK, S.198.  
95 Ebenda, S.29.  
96 EKKI, S.194.

83 NKIS, S.21.  
84 Ebenda, S.22.

Ein kurzer vergleichender Blick nach Indien: Die größte Demokratie der Welt ist eines der wenigen Entwicklungs- und Schwellenländer mit eigener KI-Strategie.<sup>97</sup> Diese sieht unter anderem vor, dass KI einer Ethik-Folgenabschätzung zu unterziehen ist. Dies soll insbesondere Diskriminierung durch KI verhindern bzw. begrenzen. Ebenso wird die Einführung eines Ethikrates für KI empfohlen.<sup>98</sup> Eine weniger konkrete Initiative im Sinne eines ethischen Perspektivenwechsels ist die Fokussierung auf buddhistische Werte bei der ethischen Folgenabschätzung von KI.<sup>99</sup>

### III.3. Ethical Governance and Stewardship (Policy Area 2)

Die in den Empfehlungen 57 und 58 angeregten Selbstbewertungen durch öffentliche Stellen im Hinblick auf Ihren KI-Einsatz sind in der bisherigen Form in Deutschland lediglich geplant, eine etablierte Behördenpraxis besteht noch nicht.<sup>100</sup> Im Hinblick auf die Sicherung eines insbesondere in den Punkten 68–70 der Empfehlung umrissenen, auf rechenschaftspflichtige und verantwortungsvolle KI gerichteten Ordnungsrahmens besteht in Deutschland derzeit noch kein allgemeiner regulatorischer Rahmen. Vorhanden sind dagegen partielle bzw. bereichsspezifische Regelungen des Medienstaatsvertrages, etwa die Kennzeichnungspflicht für sogenannte Social Bots (§ 18 Abs. 3 MStV)<sup>101</sup> und die medienrechtlichen Transparenzanforderungen für algorithmische Empfehlungssysteme in §§ 93 ff. MStV.<sup>102</sup> Die in Punkt 56 der Empfehlung nahegelegte nationale KI-Strategie wurde von der Großen Koalition ursprünglich im November 2018

verabschiedet<sup>103</sup> und zuletzt mit einem Fortschreibungsbericht zur Strategie von Dezember 2020<sup>104</sup> aktualisiert, um den „Auf- und Ausbau der Qualitätsinfrastruktur auf der Basis eines angemessenen Ordnungsrahmens zu einem System für sichere und vertrauenswürdige KI zu stärken“ und „die zivilgesellschaftliche Vernetzung und die Einbeziehung in die Entwicklung und Nutzung von gemeinwohlorientierter KI zu unterstützen“. Im Hinblick auf die von Empfehlung 56 insbesondere anvisierten Formen der Soft Governance von KI sind zwar bereits Normungsinitiativen entlang der sogenannten Normungsroadmap KI<sup>105</sup> angestoßen. Diese basieren derzeit aber auf einer primär technischen Perspektive und beziehen die von der Empfehlung vorgesehenen mehrstufigen Überprüfungssysteme unter Einbeziehung ethischer Standards und prozeduraler Anhörungs- und Überprüfungspflichten gegenüber Betroffenen allenfalls unvollständig mit ein. Die Empfehlung kann daher zukünftig als Anstoß für weitergehende praktische Initiativen dienen, etwa für Experimente mit institutionalisierten Nutzer\*innen oder einer Betroffenenmitbestimmung im Kontext der KI-Governance.

In vergleichbarer Weise zeichnet sich auch auf europäischer Ebene eine Überbetonung der technischen Standardsetzung gegenüber teilhabeorientierteren Governancemodellen im Kontext der KI-Regulierung ab. Der im April dieses Jahres von der Kommission vorgelegte Entwurf eines Europäischen KI-Gesetzes<sup>106</sup> lässt die auch im Hinblick auf die von den Empfehlungen 61–63 berührten Fragen der Vereinbarkeit des Einsatzes von KI mit Grund- und Menschenrechten und dazu erforderlicher Teilhabemodelle im Unklaren.<sup>107</sup> Es erscheint wahrscheinlich, dass aufgrund des Anreizregimes des KI-Gesetzes Normungsbehörden, insbesondere das Europäische Komitee für elektrotechnische Normung, eine starke praktische Rolle in der „On-the-ground“-Regelsetzung im Anwendungsbereich des KI-Gesetzes erhalten.<sup>108</sup>

103 Bundesregierung, KI-Strategie, 2018, [www.ki-strategie-deutschland.de](http://www.ki-strategie-deutschland.de).

104 NKIS.

105 DIN, Normungsroadmap KI, <https://www.din.de/de/forschung-und-innovation/themen/kuenstliche-intelligenz/fahrplan-festlegen>.

106 Europäische Kommission, „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 COM(2021) 206 final 2021/0106 (COD), 2021, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206>.

107 Für eine umfangreiche gemeinsame Stellungnahmen zahlreicher NGOs hierzu vgl. European Digital Rights, <https://edri.org/wp-content/uploads/2021/01/EDRI-open-letter-AI-red-lines.pdf>, 12. Januar 2021.

108 Michael Veale/Frederik Zuiderveen Borgesius, „Demystifying the Draft EU Artificial Intelligence Act — Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach“, Computer Law Review International 22.4 (2021), S. 14.

Normungsverfahren sind strukturell für von KI-Produkten und –Dienstleistungen Betroffene und NGOs eher unzugänglich. Damit geht die Gefahr der Überbetonung von Industrieinteressen einher.

Dennoch bildet der Entwurf für das europäische KI-Gesetz einige der Grundvoraussetzungen im Hinblick auf die Sicherung menschenrechtskonformer und teilhabeorientierter KI im Sinne der Empfehlung ab. Der (wahrscheinliche) zukünftige Ordnungsrahmen sieht insbesondere ein Verfahren zur Meldung schwerwiegender Vorfälle und Fehlfunktionen vor (Punkt 62) und statuiert risikobasierte Transparenzpflichten (vgl. Punkt 52). Im Hinblick auf die praktischen Umsetzungsbedingungen ist noch unklar, ob der einzurichtende Ausschuss für KI hinreichend ausgestattet wird, um seine Rolle effektiv auszufüllen und öffentliche wie Individualinteressen der Betroffenen zu schützen.<sup>109</sup> Außerhalb des EU-Entwurfs ist die Förderung der auf KI beruhenden Innovation eng mit weiteren Initiativen, insbesondere dem Daten-Governance-Gesetz,<sup>110</sup> der Richtlinie über offene Daten<sup>111</sup> und anderen Initiativen im Rahmen der EU-Strategie für Daten<sup>112</sup> verknüpft, die insgesamt das Ziel verfolgen, vertrauenswürdige Mechanismen und Dienste für die Weiterverwendung, das Teilen und die Zusammenführung von Daten zu schaffen.

Die von Punkt 64 der Empfehlung in den Fokus genommene Integration internationaler Organisationen in die Entwicklung von KI-Strategien ist in Deutschland noch ausbaufähig. So umfasste die KI-Strategie der Bundesregierung bisher lediglich eine geplante Untersuchung der Machbarkeit und möglicher Elemente eines völkerrechtlichen Rechtsrahmens für die Entwicklung, Gestaltung und Anwendung von KI, aber keine klaren Teilhabemodell.<sup>113</sup> Zur von Punkt 67 in den Blick genommenen inklusiven Zusammensetzung von Entwicklungsteams und Datensätzen gibt es bisher keine soliden Vorschläge. Im Kontext des EU-AI sollen lediglich die freiwilligen Verhaltenskodizes gem. Art. 69, welche die Kommission und der geplante europäische KI-Ausschuss fördern sollen, dahingehende Anforderungen enthalten.

109 Europäischer Datenschutzausschuss und Europäischer Datenschutzbeauftragter, Gemeinsame Stellungnahme zum Gesetz über Künstliche Intelligenz, 18. Juni 2021, [https://edpb.europa.eu/system/files/2021-06/edpb-edps\\_joint\\_opinion\\_ai\\_regulation\\_en.pdf](https://edpb.europa.eu/system/files/2021-06/edpb-edps_joint_opinion_ai_regulation_en.pdf).

110 Vorschlag für eine Verordnung über europäische Daten-Governance (Data-Governance-Gesetz), COM(2020) 767, 25. November 2020, <https://eur-lex.europa.eu/legal-content/DE/ALL/?uri=CELEX:52020PC0767>.

111 Richtlinie (EU) 2019/1024 des Europäischen Parlaments und des Rates vom 20. Juni 2019 über offene Daten und die Weiterverwendung von Informationen des öffentlichen Sektors (PE/28/2019/REV/1, ABl. L 172 vom 26.6.2019, S. 56).

112 Europäische Kommission, „Eine europäische Datenstrategie“, COM(2020) 66 final, 19. Februar 2020, [https://ec.europa.eu/info/sites/default/files/communication-european-strategy-data-19feb2020\\_de.pdf](https://ec.europa.eu/info/sites/default/files/communication-european-strategy-data-19feb2020_de.pdf).

113 NKIS, S. 33.

### Fazit

Der zur Zeit in Deutschland vorherrschende Ansatz der „Stewardship“ von KI-Prozessen ist im Kontext der Diversitätssicherung wie auch im Hinblick auf die Beteiligung von Interessenträger\*innen an der Konzeption und Entwicklung von KI-Systemen noch deutlich ausbaufähig. Angezeigt erscheint gerade im Hinblick auf Letzteres (und auch angesichts der UNESCO-Empfehlung) eine politische und rechtliche Festlegung, welche Teilhabemodelle erprobt werden sollen. Angesichts begrenzter Ressourcen der Nachprüfung von KI-Entscheidungen müssen insoweit Experimente mit institutionellen Strukturen unternommen werden.

Diese Beobachtung deckt sich in Teilen mit Vorschlägen der Enquete-Kommission zum Ausbau von Transfer- und Kooperationsmechanismen zwischen Wissenschaft, Wirtschaft, Politik und Zivilgesellschaft im Bereich der KI-Forschung. Diese schlug etwa vor, im Rahmen von Citizen Sciences, Reallaboren und anderen inklusiven Ansätzen verschiedene Partizipationsmodelle zu erproben<sup>114</sup> und im Bereich „Arbeit“ adaptierte Formen bisheriger Arbeitnehmer\*innenmitbestimmung in KI-Governancesysteme zu integrieren.<sup>115</sup> Derartige teilhabeorientierte institutionelle und prozedurale Modelle für Betroffenenmitbestimmung sollten grundsätzlich über den Bereich „Arbeit“ hinaus erprobt werden. Dies erscheint insbesondere angesichts der derzeitigen Überbetonung von Normungsverfahren in Regulierungsansätzen (s.o.), die zwar einen epistemischen regulatorischen Mehrwert aufweisen können, wegen ihres geringen Legitimationsniveaus aber weitergehende Maßnahmen zur Beteiligung von Betroffenen an der KI-Governance nicht ersetzen können, angeraten.<sup>116</sup>

Bei einem Vergleich der Empfehlungen in den oben genannten Politikbereichen zu bestehenden Regularien oder Initiativen in Afrika zeigt sich, dass dort erst wenige formalisierte nationale KI-Strategien eingeführt wurden.<sup>117</sup> Unabhängig davon haben aber einige Länder Taskforces und Gremien eingesetzt, die sich zur KI-Politik im hier besprochenen Politikfeld positioniert haben, etwa Ugandas „Expert National Task

114 EKKI, S. 102.

115 Ebenda, S. 317f., S. 320f.

116 Zur Analyse des EU-AI-Vorschlags in dieser Hinsicht eindringlich Michael Veale/Frederik Zuiderveen Borgesius, „Demystifying the Draft EU Artificial Intelligence Act — Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach“, Computer Law Review International 22.4 (2021), S. 15.

117 Stand März 2021 soll Mauritius das einzige Land in Afrika mit einer „voll formalisierten“ nationalen KI-Strategie gewesen sein, vgl. Arthur Gwagwa/Patti Kachidza/Kathleen Siminyu/Matthew Smith, Responsible Artificial Intelligence in Sub-Saharan Africa: Landscape and general state of play, [https://ircai.org/wp-content/uploads/2021/03/AI4D\\_Report\\_Responsible\\_AI\\_in\\_SSA.pdf](https://ircai.org/wp-content/uploads/2021/03/AI4D_Report_Responsible_AI_in_SSA.pdf), 2021, S. 16 f., mit Verweis auf A. Sey, „Sub-Saharan Africa: Regional Analysis“, in: Oxford Insights, & International Development Research Centre (IDRC). Government artificial intelligence readiness index, 2020, <https://www.oxfordinsights.com/government-ai-readiness-index-2020>.

Force on the Fourth Industrial Revolution“ (4IR),<sup>118</sup> in deren Abschlussbericht<sup>119</sup> die Realisierung des Nutzens von KI als Fortschrittswerkzeug die Hauptrolle gegenüber potenziellen Gefahren einnimmt. Auch die durch KI in diesem Politikfeld aufgeworfenen Fragen nach Partizipation und Vernetzung unterschiedlicher KI-Akteur\*innen und den passenden institutionellen Strukturen finden im Globalen Süden Ausdruck: Der Abschlussbericht der ugandischen 4IR verweist etwa auf Initiativen für „Centers of Excellence“ des Nationalen Verbands der Softwareunternehmen in Indien, die potenzielle Stakeholder zur Entwicklung bzw. Nutzung bestimmter KI-Anwendungen vernetzen. Dieses Modell will Uganda auf den Agrarsektor übertragen, um dort über eine Public-Private-Partnership Verbesserungen zu erreichen.<sup>120</sup>

### III.4. Data Policy (Policy Area 3)

Die deutsche Bundesregierung hat im November 2018 ihre Strategie zur Künstlichen Intelligenz verabschiedet.<sup>121</sup> Auf der Website „www.ki-strategie-deutschland.de“ sind die Strategie, Maßnahmen und aktuelle Entwicklungen abrufbar. Hier wird explizit „Datennutzung, Datensicherheit, Recht und Ethik“ als eines der Handlungsfelder aufgelistet. Das Strategie-Papier soll einen „Überblick über Ziele, Grundlagen und Handlungsfelder für eine Nationale Strategie Künstliche Intelligenz“ liefern. Unter Punkt 3.8 „Daten verfügbar machen und Nutzung erleichtern“ werden einige Themenfelder gemäß der Empfehlung adressiert. So soll insbesondere die Qualität der Daten erhöht und der Zugang zu Daten erleichtert werden.<sup>122</sup> Hierbei soll das Recht auf informationelle Selbstbestimmung und datenschutzrechtliche Vorschriften stets gewahrt werden.<sup>123</sup> Auch qualitativ hochwertige Trainingsdatensätze sollen gefördert werden.<sup>124</sup> Zudem soll ein Austausch mit Datenschutzaufsichtsbehörden zur Entwicklung von KI-Systemen stattfinden.<sup>125</sup>

118 Xinhuanet, „Uganda Prepares to Harness Opportunities of 4th Industrial Revolution“, 4. September 2019, [http://www.xinhuanet.com/english/2019-04/09/c\\_137963317.htm](http://www.xinhuanet.com/english/2019-04/09/c_137963317.htm).  
 119 Vgl. 4IR Task Force, „Uganda’s National 4IR Strategy: A continental 4IR hub that enables a smart and connected Ugandan Society“, 2021, <https://ict.go.ug/wp-content/uploads/2020/10/Executive-Summary-Ugandas-National-4IR-Strategy.pdf>.  
 120 Ebenda, S.14, mit Verweisen auf die Einrichtungen der indischen National Association of Software and Service Companies (NASSCOM), vgl. <https://coe-dsai.nasscom.in> und <https://www.coe-iot.com>.  
 121 Bundesministerium für Wirtschaft und Energie, Strategie Künstliche Intelligenz der Bundesregierung, 16. November 2018, <https://www.bmwi.de/Redaktion/DE/Publikationen/Technologie/strategie-kuenstliche-intelligenz-der-bundesregierung.html>.  
 122 Ebenda, S.33.  
 123 Ebenda.  
 124 Ebenda, S.35.  
 125 Ebenda, S.37.

Im November 2019 wurde ein Zwischenbericht veröffentlicht, der über die einzelnen Maßnahmen zur Umsetzung der Strategie berichtet.<sup>126</sup> Hier wurde bezüglich des Themenfeldes „Data Policy“ eine Dateninfrastruktur „GAIA-X“ eingerichtet, der regelmäßige Austausch zu KI-Datenschutzfragen mit Datenschutzaufsichtsstellen begonnen und ein Open-Data-Portal geschaffen.<sup>127</sup> Zudem wurde in der Zwischenzeit die Datenethik-Kommission eingesetzt, die im Oktober 2019 ihr Gutachten zur Strategie übergeben hat.<sup>128</sup> In diesem Gutachten werden die Achtung der Rechte beteiligter Personen, Nutzen und Teilen von Daten, Datenqualität, Informationssicherheit und Transparenz betont.<sup>129</sup> Außerdem werden unter anderem Datenrechte und korrespondierende Datenpflichten<sup>130</sup>, Anforderungen an die Nutzung personenbezogener Daten<sup>131</sup> und der Zugang zu personenbezogenen Daten<sup>132</sup> ausführlich beschrieben. Im Dezember 2020 wurde die Fortschreibung der KI-Strategie veröffentlicht, welche die weiteren Schritte zur Umsetzung der Strategie formuliert.<sup>133</sup> Im Abschlussbericht der Enquete-Kommission Künstliche Intelligenz vom Oktober 2020 wird ausführlich auf KI und Daten eingegangen.<sup>134</sup>

Des Weiteren gibt es die Plattform „Lernende Systeme“.<sup>135</sup> Im Rahmen des Schwerpunktes „Datenökosysteme“ wurde hier eine Broschüre veröffentlicht, die als wichtige Leitlinien Zugang zu Daten und Qualität der Daten definiert.<sup>136</sup> Außerdem haben das Deutsche Institut für Normung (DIN) und die Deutsche Kommission Elektrotechnik Elektronik Informationstechnik (DKE) zusammen mit dem Bundesministerium für Wirtschaft und Energie (BMWi) eine Roadmap zu Normen und Standards der Künstlichen Intelligenz veröffentlicht.<sup>137</sup> Unter dem Schwerpunktthema „IT-Sicherheit bei KI-Systemen“

126 Bundesregierung, Zwischenbericht ein Jahr KI-Strategie, abrufbar unter „Downloads“ auf der Website [www.ki-strategie-deutschland.de](http://www.ki-strategie-deutschland.de).  
 127 Ebenda, S.7.  
 128 DEK.  
 129 Ebenda, S.81 ff.  
 130 Ebenda, S.85 ff.  
 131 Ebenda, S.95 ff.  
 132 Ebenda, S.124 ff.  
 133 NKIS.  
 134 EKKI.  
 135 Siehe Homepage Plattform Lernende Systeme, <https://www.plattform-lernende-systeme.de/startseite.html>.  
 136 Lernende Systeme, Von Daten zu Wertschöpfung, Juli 2020, [https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/PLS\\_Booklet\\_Datenoekosysteme.pdf](https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/PLS_Booklet_Datenoekosysteme.pdf).  
 137 DIN und DKE, Deutsche Normungsroadmap Künstliche Intelligenz, November 2020, <https://www.din.de/resource/blob/772438/6b5ac6680543eff9fe-372603514be3e6/normungsroadmap-ki-data.pdf>.

werden Datenschutz<sup>138</sup> und sichere Daten<sup>139</sup> beschrieben. Der KI-Bundesverband ist ein weiteres Gremium, das sich auch regelmäßig in Positionspapieren und Pressemitteilungen mit „Data Policy“ im Anwendungsbereich Künstlicher Intelligenz befasst.<sup>140</sup>

Bezüglich Förderprogrammen ist der Aufbau von KI-Servicezentren hervorzuheben.<sup>141</sup> Hier soll der Zugang zu Datensätzen ermöglicht und gleichzeitig die Datensouveränität gewährleistet werden. Eine vollständige Auflistung der Förderprogramme ist unter „www.ki-strategie-deutschland.de“ auffindbar.<sup>142</sup>

Auf europäischer Ebene wurde im April 2018 eine Agenda der Europäischen Kommission veröffentlicht, die unter anderem den Zugang zu Daten und einen geeigneten ethischen und rechtlichen Rahmen betont.<sup>143</sup> Im Frühjahr 2020 wurde daraufhin das „Weißbuch zur Künstlichen Intelligenz – Ein europäisches Konzept für Exzellenz und Vertrauen“<sup>144</sup> vorgestellt. Darin wird auch auf die Risiken für den Schutz personenbezogener Daten und die Sicherheit eingegangen.<sup>145</sup> Zudem behandelt im Entwurf zur Regulierung von KI-Systemen Art. 10 Daten und Daten-Governance und Art. 17 beschreibt das Qualitätsmanagementsystem.<sup>146</sup> Die Wahrung des Datenschutzes wird durchwegs erwähnt.<sup>147</sup> Außerdem hat die High-Level Expert Group on AI in einem Papier im Jahr 2019 auch Empfehlungen zu „Policy“ bezüglich Datenökonomie, Schutz der betroffenen Personen und Datenzugang veröffentlicht.<sup>148</sup> Die von dieser Gruppe aufgestellten Ethik-Leitlinien enthalten ebenfalls Anforderungen zu technischer Sicherheit, Schutz der Privatsphäre, Transparenz und

138 Ebenda, S.101.  
 139 Ebenda, S.105 ff.  
 140 Positionspapiere und Stellungnahmen des KI Bundesverbandes, siehe <https://ki-verband.de>.  
 141 Bundesministerium für Bildung und Forschung, Richtlinie zur Förderung zum Aufbau von KI-Servicezentren, 1. Oktober 2021, <https://www.bmbf.de/bmbf/shareddocs/bekanntmachungen/de/2021/10/2021-10-01-Bekanntmachung-KI.html>.  
 142 Auch die Länder haben KI-Strategien mit „Data Policy“-Schwerpunkten veröffentlicht. Auf ihre Darstellung wird an dieser Stelle verzichtet. Sie sind jedoch unter [www.ki-strategie-deutschland.de](http://www.ki-strategie-deutschland.de) abrufbar.  
 143 Europäische Kommission, Mitteilung der Kommission – Künstliche Intelligenz für Europa, COM(2018) 237 final, 25. April 2018, <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=COM%3A2018%3A237%3AFIN>.  
 144 Europäische Kommission, Weissbuch zur Künstlichen Intelligenz – ein europäisches Konzept für Exzellenz und Vertrauen, COM(2020) 65 final, 19. Februar 2020, [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_de.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_de.pdf).  
 145 Ebenda, S.12 ff.  
 146 Europäische Kommission, 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, COM(2021) 206 final, 21. April 2021, <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206>.  
 147 Ebenda, Erwägungsgrund 36.  
 148 High-Level Expert Group on AI, Policy and Investment Recommendations for Trustworthy AI, 26. Juni 2019, <https://digital-strategy.ec.europa.eu/en/library/policy-and-investment-recommendations-trustworthy-artificial-intelligence>.

Fairness.<sup>149</sup> Darüber hinaus findet man auf der Homepage der Europäischen KI-Allianz einen Überblick zu einschlägigen Dokumenten und Blogbeiträgen von Stakeholdern.<sup>150</sup>

### Fazit

Handlungsbedarf besteht in Deutschland bei der Abstimmung der datenschutzrechtlichen Vorschriften mit KI-Strategien. Dies insbesondere mit Blick auf Punkt 74 des Empfehlungsvorschlags, welcher empfiehlt, dass bereits bestehende Datenschutzvorschriften verstärkt werden, um personenbezogene Daten und darunter besonders sensible Daten zu schützen. Zudem könnten Wissenschaft und Praxis intensiver miteinander Ergebnisse weiterentwickeln und nicht vorwiegend parallel forschen. Dies wäre im Sinne des Punktes 77 des Empfehlungsvorschlags zur Schaffung von Kooperationsplattformen für den Austausch hochwertiger Daten bedeutsam. Diese Punkte werden in oben genannten Dokumenten meist thematisiert. Schlussendlich wird insbesondere das Teilen von Daten durch private Unternehmen ausschlaggebend für KI sein. Einige stehen dieser Idee aufgrund des datenbasierten Geschäftsmodells skeptisch gegenüber. Hier können Unternehmen gemäß Punkt 77 des Empfehlungsvorschlags noch mehr ermutigt werden, die von ihnen gesammelten Daten zu teilen, um Forschung und Innovation zu fördern.

Die Initiative „FAIR Forward – Künstliche Intelligenz für alle mit Afrika und Asien“ adressiert den Zugang zu Trainingsdaten und Rahmenbedingungen für einen besseren Datenschutz.<sup>151</sup> Die „Global South – AI4COVID“-Initiative beschäftigt sich auch mit Rahmenbedingungen für die Datenverarbeitung, jedoch mit dem Schwerpunkt Covid-19.<sup>152</sup> Zudem thematisiert die Initiative „Artificial Intelligence for Development in Africa (AI4D Africa)“ Rahmenbedingungen für Dateneinsatz bei KI.<sup>153</sup> Der Fokus dieser Initiativen liegt vor allem darauf, dass der Globale Süden bei KI nicht abgehängt wird. Zudem soll von den Ländern selbst und nicht nur für die Länder KI entwickelt werden. Gemäß „OECD.AI“ haben nur die wenigsten afrikanischen Staaten Initiativen zu KI begründet.<sup>154</sup>

149 HEG-KI.  
 150 Siehe European AI Alliance, <https://futurium.ec.europa.eu/en/european-ai-alliance/pages/about?language=en>.  
 151 Siehe FAIR Forward, Toolkit Digitalisierung, <https://toolkit-digitalisierung.de/fair-forward-ki-fuer-alle-mit-afrika-und-asien>.  
 152 Siehe Global South AI4COVID Program, <https://covid-south.ai/about>.  
 153 Siehe AI4D Africa, <https://africa.ai4d.ai/about-ai4d>.  
 154 Siehe OECD.AI Policy Observatory, <https://oecd.ai/en/dashboards>.

### III.5. Development and International Cooperation (Policy Area 4)

Das Politikfeld 4 ist den Themen Entwicklung und internationale Kooperation gewidmet. Empfehlung Nr. 78 richtet sich an Mitgliedstaaten und transnationale Unternehmen und fordert diese auf, der KI-Ethik Vorrang einzuräumen, indem sie Diskussionen über KI-bezogene ethische Fragen in die einschlägigen internationalen, zwischenstaatlichen und Multi-Stakeholder-Foren aufnehmen. Teilweise ist das schon geschehen. In der Strategie zur Künstlichen Intelligenz der Bundesregierung aus dem Jahr 2018 wird unter Punkt 3.11 „Nationale und internationale Vernetzung“ als eigenständiges Handlungsfeld genannt.<sup>155</sup> Hierunter wird die Zusammenarbeit und der Austausch mit EU-Institutionen sowie anderen Wirtschaftsräumen beschrieben.<sup>156</sup> Zudem soll ein Teil des Handlungsfeldes der Aufbau von Kapazitäten und Wissen zu KI in Entwicklungsländern sein.<sup>157</sup> Im Zwischenbericht zur KI-Strategie wird die Begleitung und Unterstützung der Bundesregierung auf europäischer Ebene dargelegt.<sup>158</sup> Außerdem wird die aktive Arbeit in internationalen Wirtschaftsräumen durch G7-, G20- und OECD-Beteiligung betont.<sup>159</sup> Die Errichtung internationaler Zukunftslabore und die Initiative „FAIR Forward – Künstliche Intelligenz für alle mit Afrika und Asien“ zur Nutzung von KI, die den Globalen Süden einbezieht, werden als weitere bereits vollzogene Maßnahmen aufgelistet.<sup>160</sup>

Auch in der Fortschreibung der KI-Strategie wird die Unterstützung bei der Schaffung von Rahmenbedingungen für KI im Globalen Süden genannt.<sup>161</sup> Das entspricht Empfehlung Nr. 79, die Mitgliedstaaten dazu auffordert, dass der Einsatz von KI in Entwicklungsbereichen wie Bildung, Wissenschaft, Kultur, Kommunikation und Information, Gesundheitsfürsorge, Landwirtschaft und Nahrungsmittelversorgung, Umwelt, Management natürlicher Ressourcen und Infrastruktur, Wirtschaftsplanung und -wachstum den in dieser Empfehlung dargelegten Werten und Grundsätzen gerecht wird. Im Gutachten der Datenethikkommission wird die Wahl eines europäischen Wegs begrüßt, der sich von anderen Wertesystemen, Kulturen und Gesellschaftsmodellen unterscheidet.<sup>162</sup> Es wird betont, diese Erkenntnisse auf globaler Ebene im außereuropäischen Kontext noch

155 Bundesministerium für Wirtschaft und Energie, Strategie Künstliche Intelligenz der Bundesregierung, 16. November 2018, <https://www.bmwi.de/Redaktion/DE/Publikationen/Technologie/strategie-kuenstliche-intelligenz-der-bundesregierung.html>.

156 Ebenda, S. 42 f.

157 Ebenda, S. 43.

158 Bundesregierung, Zwischenbericht ein Jahr KI-Strategie, November 2019, S. 8, abrufbar unter „Downloads“ auf der Website [www.ki-strategie-deutschland.de](http://www.ki-strategie-deutschland.de).

159 Ebenda.

160 Ebenda, S. 9.

161 NKIS.

162 DEK.

mehr einzubringen.<sup>163</sup> Im Abschlussbericht der Enquete-Kommission wird als eine der zentralen Handlungsempfehlungen die Zusammenarbeit zwischen Forschung, Wirtschaft und Gesellschaft erwähnt.<sup>164</sup>

Eine Arbeitsgruppe der Plattform „Lernende Systeme“ beschäftigt sich mit dem Entwicklungsfeld „Gesundheit & Pflege“.<sup>165</sup> Zudem wird „KI in der Medizin“ in der deutschen Normungsroadmap ein Kapitel gewidmet.<sup>166</sup> Ansonsten hat der KI-Bundesverband im Februar 2021 ein Positionspapier mit dem Titel „Wie Künstliche Intelligenz Klimaschutz und Nachhaltigkeit fördern kann“ veröffentlicht.<sup>167</sup> Darin werden unter anderem die Schaffung von Nachhaltigkeits-Plattformen zur interdisziplinären Kooperation sowie die Entwicklungsfelder Energie und Landwirtschaft thematisiert.<sup>168</sup> Ferner legt das Förderprogramm „Digital GreenTech – Umwelttechnik trifft Digitalisierung“ ebenfalls das Hauptaugenmerk auf die Anwendungsfelder Ressourceneffizienz, nachhaltige Landwirtschaft und Geotechnologie.<sup>169</sup>

Verstärkt muss sich Deutschland indes für die KI-Forschungskooperationen einsetzen, besonders mit den am wenigsten entwickelten Ländern (LDCs) (Empfehlung 81 und 82), einschließlich Forschungs- und Innovationszentren und -netzen, die eine stärkere Beteiligung und Führungsrolle von Forscher\*innen des Globalen Südens erwarten lassen.

Im „Weißbuch zur Künstlichen Intelligenz – Ein europäisches Konzept für Exzellenz und Vertrauen“ wird als einer der wichtigsten Bausteine die Kooperation zwischen öffentlichem und Privatsektor gefordert.<sup>170</sup> Überdies sieht die High-Level Expert Group on AI als Erkenntnis ihres Berichts die Notwendigkeit von Multi-Stakeholder-Kooperationen zur Implementierung der genannten Empfehlungen vor.<sup>171</sup> Zudem nennt sie als Beispiele für vertrauenswürdige KI anhand ethischer Leitlinien den Einsatz in den Entwick-

163 Ebenda.

164 EKKI, S. 107.

165 Siehe Plattform Lernende Systeme, <https://www.plattform-lernende-systeme.de/startseite.html>.

166 DIN und DKE, Deutsche Normungsroadmap Künstliche Intelligenz, November 2020, <https://www.din.de/resource/blob/772438/6b5ac6680543eff9fe372603514be3e6/normungsroadmap-ki-data.pdf>, S. 135 ff.

167 KI Bundesverband, Wie Künstliche Intelligenz Klimaschutz und Nachhaltigkeit fördern kann, 22. Februar 2021, <https://ki-verband.de/wp-content/uploads/2021/02/KIBV-Klima-Positionspapier-1.pdf>.

168 Ebenda.

169 Siehe FONa, [https://www.fona.de/de/massnahmen/foerdermassnahmen/DigitalGreenTech.php?utm\\_source=hootsuite](https://www.fona.de/de/massnahmen/foerdermassnahmen/DigitalGreenTech.php?utm_source=hootsuite).

170 Europäische Kommission, Weissbuch zur Künstlichen Intelligenz – Ein europäisches Konzept für Exzellenz und Vertrauen, COM(2020) 65 final, 19. Februar 2020, [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_de.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_de.pdf), S. 3.

171 High-Level Expert Group on AI, Policy and Investment Recommendations for Trustworthy AI, 26. Juni 2019, <https://digital-strategy.ec.europa.eu/en/library/policy-and-investment-recommendations-trustworthy-artificial-intelligence>, S. 48.

lungsfeldern Klimaschutz und nachhaltige Infrastruktur, Gesundheit und Bildung.<sup>172</sup> Zusätzlich findet man auf der Homepage der Europäischen KI-Allianz einen Überblick zu einschlägigen Dokumenten und Blogbeiträgen von Stakeholdern.<sup>173</sup>

#### Fazit

In Deutschland besteht bezüglich der Policy Area „Development and International Cooperation“ noch Handlungsbedarf im Bereich der Privatunternehmen. Zumeist werden Politik und Wissenschaft als Kooperationspartner beschrieben, jedoch wird der private Sektor insbesondere bezüglich Entwicklungsarbeit im Globalen Süden kaum erwähnt. Dies wäre gemäß Punkt 82 der UNESCO-Empfehlung dringend notwendig, um unter anderem KI für spezifische Kulturen und Kontexte zu entwickeln. Die Zusammenarbeit richtet ihren Fokus generell auf den europäischen Kontinent, was insbesondere durch die Erwähnung des „europäischen Weges“ deutlich wird. Einzelne Initiativen bzw. Strategien erwähnen die globale Zusammenarbeit insbesondere mit Schwellen- und Entwicklungsländern. Jedoch sollte hier der Fokus auch auf die Einbeziehung von forschenden Personen aus den Entwicklungsländern selbst liegen. Dies wäre im Sinne des Punktes 81 der Empfehlung. Zudem besteht

172 HEG-KI, S. 42 ff.

173 Siehe European AI Alliance, <https://futurium.ec.europa.eu/en/european-ai-alliance/pages/about?language=en>.

noch Regulierungs- bzw. Handlungsbedarf betreffend der Überwindung geotechnologischer Grenzen, wie es Punkt 83 betont.

Die bereits oben erwähnte Initiative „FAIR Forward – Künstliche Intelligenz für alle mit Afrika und Asien“ ist ein Beispiel für die aktive Einbeziehung des Globalen Südens.<sup>174</sup> Die „Global South – AI4COVID“-Initiative unterstützt interdisziplinäre Forschung zu COVID-19 in Entwicklungsländern.<sup>175</sup> Zudem legt die Initiative „Artificial Intelligence for Development in Africa (AI4D Africa)“ den Schwerpunkt auf die Ziele für nachhaltige Entwicklung (Verbesserung der Lebensmittelkette, Bildung, Gesundheit und Klimawandel).<sup>176</sup> Während im europäischen Kontext der Wirtschaftsstandort im Fokus steht, geht es im Globalen Süden bei KI vermehrt um eine nachhaltige Entwicklung auf ökologischer und sozialer Ebene. Die Kooperation wird meist von Organisationen und Ländern initiiert, die nicht Teil des Globalen Südens sind.

Ein wichtiges Commitment findet sich am Ende des Abschnitts in Empfehlung Nr. 83, die daran erinnert, dass jede Kooperation fußend auf ethischen Werten und „unter uneingeschränkter Achtung des Völkerrechts“ stattfindet.

174 Siehe FAIR Forward, Toolkit Digitalisierung, <https://toolkit-digitalisierung.de/fair-forward-ki-fuer-alle-mit-afrika-und-asien>.

175 Siehe Global South AI4COVID Program, <https://covid-south.ai/about>.

176 Siehe AI4D Africa, <https://africa.ai4d.ai/about-ai4d>.



### III.6. Gender (Policy Area 6)

In Deutschland und auf Ebene der EU gibt es im Politikfeld Gender (und Diversity) eher wenige (politische und regulatorische) Prozesse, Instrumente und Governance-Strukturen, die KI aktiv adressieren. In der KI-Strategie der deutschen Bundesregierung von November 2018<sup>177</sup> taucht der Begriff Gender genau einmal auf, im Bereich der Risikofolgenabschätzung von KI im Kontext Erwerbsarbeit. Implizit wird dieses Politikfeld unter der Prämisse der Nichtdiskriminierung mitverhandelt. So wird innerhalb der Strategie an verschiedenen Stellen an das allgemeine Diskriminierungsverbot erinnert.<sup>178</sup> „KI-basierte Entscheidungen, Dienstleistungen und Produkte“ sollen, so die Strategie, mit Blick auf mögliche unzulässige Diskriminierungen überprüfbar gemacht werden.<sup>179</sup> Zudem soll durch „Transparenz, Nachvollziehbarkeit und Überprüfbarkeit der KI-Systeme“ ein „effektiver Schutz gegen Verzerrungen, Diskriminierungen, Manipulationen oder sonstige missbräuchliche Nutzungen insbesondere beim Einsatz von algorithmenbasierten Prognose- und Entscheidungssystemen“ erwirkt werden,<sup>180</sup> wobei explizit auch die „Diskriminierung von benachteiligten Bevölkerungsgruppen in Entwicklungsländern“ genannt wird.<sup>181</sup> Zur Sicherung dieser Gebote „prüft die Bundesregierung die Einrichtung bzw. den Ausbau von staatlichen Stellen und privaten Prüfinstitutionen zur Kontrolle algorithmischer Entscheidungen“.<sup>182</sup> Dazu gehören sowohl das Etablieren von Auditierungsstandards und Standards für (Technik-)Folgenabschätzungen wie das Schaffen einer Struktur, in der alle Elemente des KI-Prozesses transparent und nachvollziehbar zur Prüfung grundrechtlicher Kriterien offengelegt werden.<sup>183</sup>

Empfehlung Nr. 87 verpflichtet die Mitgliedstaaten überdies sicherzustellen, dass das Potenzial digitaler Technologien und Künstlicher Intelligenz zur Verwirklichung der Gleichstellung der Geschlechter beiträgt und in vollem Umfang ausgeschöpft wird. Mitgliedstaaten müssen gewährleisten, dass die Menschenrechte und Grundfreiheiten von Mädchen und Frauen sowie ihre Sicherheit und Integrität in keiner Phase des Lebenszyklus von KI-Systemen verletzt werden. Darüber hinaus sollte die ethische Folgenabschätzung – zu dieser schon vorher – eine transversale Geschlechterperspektive beinhalten.

177 Bundesministerium für Wirtschaft und Energie, Strategie Künstliche Intelligenz der Bundesregierung, 16. November 2018, <https://www.bmwi.de/Redaktion/DE/Publikationen/Technologie/strategie-kuenstliche-intelligenz-der-bundesregierung.html>.

178 Vgl. ebenda, S. 39.

179 Vgl. ebenda, S. 38.

180 Vgl. ebenda, S. 39.

181 Vgl. ebenda.

182 Vgl. ebenda, S. 40.

183 Vgl. ebenda.

Auf der Website „www.ki-strategie-deutschland.de“ sind die Strategie, Maßnahmen und aktuelle Entwicklungen abrufbar. Auch hier gilt Nichtdiskriminierung, Gender & Diversity nicht als eigenes Handlungsfeld, findet mit Ausnahme der Förderung von KI-Nachwuchs-Wissenschaftlerinnen kaum Beachtung. Auch der im November 2019 veröffentlichte Zwischenbericht zur Umsetzung der Strategie<sup>184</sup> enthält keinerlei Aussagen zum Policy-Feld Gender. Auch deshalb, weil weder die Antidiskriminierungsstelle des Bundes<sup>185</sup> noch das für Gender zuständige Bundesministerium für Familie, Senioren, Frauen und Jugend (BMFSFJ)<sup>186</sup> von Beginn an aktiv in die Umsetzung der Strategie eingebunden waren. Das BMFSFJ koordiniert seit Kürzerem drei Projekte in Richtung Gleichstellung am Arbeitsmarkt unter dem Dach der KI-Strategie. Die Fortschreibung der KI-Strategie aus 2020 verbleibt ebenfalls weitgehend im Ungefähren.<sup>187</sup> Dafür weist immerhin die zwischenzeitlich gegründete Datenethikkommission in ihrem Gutachten zur Strategie auf die Gefahren der Diskriminierung durch Algorithmen und die fehlenden regulatorischen Ansätze in Richtung

184 Bundesregierung, Zwischenbericht ein Jahr KI-Strategie, November 2019, abrufbar unter „Downloads“ auf der Website [www.ki-strategie-deutschland.de](http://www.ki-strategie-deutschland.de). Während die vom BMBF geförderte Plattform Lernende Systeme schon 2019 eine Studie zum Thema herausgab: Künstliche Intelligenz und Diskriminierung, <https://www.plattform-lernende-systeme.de/publikationen-details/kuenstliche-intelligenz-und-diskriminierung-herausforderungen-und-loesungsansaeetze.html>; ebenso wie das vom BMAS geförderte KI-Observatorium die Themen Geschlecht und Diskriminierung in den Blick nimmt: <https://www.ki-observatorium.de/indikatoren#tab-638>.

185 Antidiskriminierungsstelle, <https://www.antidiskriminierungsstelle.de/DE/startseite/startseite-node.html> – Diese haben das Thema durchaus im Blick und 2019 eine Studie zum Thema Diskriminierung und Algorithmen veröffentlicht: Carsten Orwat, Diskriminierungsrisiken durch Verwendung von Algorithmen, [https://www.antidiskriminierungsstelle.de/SharedDocs/downloads/DE/publikationen/Expertenstudie\\_diskriminierungsrisiken\\_durch\\_verwendung\\_von\\_algorithmen.pdf?\\_\\_blob=publication-file&v=3](https://www.antidiskriminierungsstelle.de/SharedDocs/downloads/DE/publikationen/Expertenstudie_diskriminierungsrisiken_durch_verwendung_von_algorithmen.pdf?__blob=publication-file&v=3). Hier fordern sie im Rahmen ihrer Schussfolgerungen unter anderem: „Die im GG und AGG festgelegten Kataloge geschützter Merkmale sind dahingehend zu prüfen, ob neue Analysemethoden, vor allem mit Algorithmen der künstlichen Intelligenz, zur automatisierten Identifizierung von Persönlichkeitsmerkmalen ihre Erweiterung erfordern. Dabei werden auch Merkmale identifizierbar und für Differenzierungen zugänglich, die dazu genutzt werden können, die Angewiesenen auf ein Gut, eine Ressource oder eine Position zu ermitteln und zu nutzen, um strukturelle Überlegenheit aufzubauen oder zu erhöhen. Die noch weitgehend unbekanntes Zusammenhänge zwischen technisch machbaren und potenziell gefährdeten Schutzziele sollten erforscht und ihre Legitimität sollte gesellschaftlich beurteilt werden.“ (S. 112) Weitere Empfehlungen in Richtung Regulierung folgen ab S. 135 ff.

186 BMFSFJ, <https://www.bmfsfj.de/bmfsfj>. Auch sie fördern zum Beispiel Diversität in der KI-Entwicklung wie z. B. die Initiative <https://www.jumpp.de/diversitaet-ki>; dieses Thema wird auch im Dritten Gleichstellungsbericht der Bundesregierung von 2021 stark thematisiert (siehe z. B. S. 17 ff.) <https://www.bmfsfj.de/resource/blob/184544/665a7070dbc68f9984fe968dc05fd139/dritter-gleichstellungsbericht-bundestagsdrucksache-data.pdf>. Auch hier findet sich durchaus Kritik an der Ausrichtung der KI-Strategie: „Andere digitalisierungsbezogene Strategien wie die ‚Strategie Künstliche Intelligenz‘ der Bundesregierung (2018) oder die Digitalstrategie des BMBF ‚Digitale Zukunft: Lernen. Forschen. Wissen‘ (2019) lassen einen geschlechter- und frauenpolitischen Ansatz vermissen (...). Die im Zuge der Gutachtererstellung geführten Gespräche mit weiteren Digitalgremien bestätigen diese Feststellung des Deutschen Frauenrates.“

187 NKIS.

einer echten erweiterten Regulierung und Rechtsdurchsetzung hin.<sup>188</sup> In eine ganz ähnliche Richtung argumentiert der Abschlussbericht der Enquete-Kommission Künstliche Intelligenz,<sup>189</sup> der in den Handlungsempfehlungen sehr deutlich fordert, „KI-gestützte Entscheidungen regelmäßig auf Diskriminierungsfreiheit (zu) überprüfen“.<sup>190</sup>

Auf europäischer Ebene wurde im April 2018 eine erste Agenda zur Künstlichen Intelligenz veröffentlicht.<sup>191</sup> In ihr findet das Thema Diversität und Nichtdiskriminierung vor allem im Rahmen der Entwicklung von KI Platz, wird aber durchaus weiter – proaktiver – verstanden: „Bei der Entwicklung von KI müssen mehr Frauen und Menschen mit unterschiedlichen Hintergründen, einschließlich Menschen mit Behinderungen, eingebunden werden, angefangen bei einer inklusiven KI-Aus- und -Weiterbildung, um sicherzustellen, dass KI nichtdiskriminierend und integrativ ist.“ Damit wird im KI-Kontext auf EU-Ebene Gender nicht nur mit Blick auf das Ergebnis (Diskriminierung) und eine einfache Lösung (Gleichbehandlung) betrachtet, sondern breiter: als Komplex verschiedener, intersektionaler Diskriminierungen mit unterschiedlichen rechtlichen, gesellschaftlichen und bildungspolitischen Anknüpfungspunkten.

Der Grundsatz der Nichtdiskriminierung wird unter dem Dach der Grund- und Menschenrechte in den Kontext Schutz der Privatsphäre und Verbraucherschutz eingeordnet. Neben einer Regulierung in Richtung Transparenz, Nachvollziehbarkeit und Rechenschaftspflicht von Algorithmen wird hier zudem verstärkt Forschung gefordert, um tatsächliche Diskriminierungsmuster aufzudecken und wissenschaftlich nachweisen zu können. An diesen ersten Wurf knüpft auch das im Frühjahr 2020 veröffentlichte „Weißbuch zur Künstlichen Intelligenz – Ein

188 DEK.

189 EKKI.

190 Ebenda, S. 195

191 Europäische Kommission, Mitteilung der Kommission – Künstliche Intelligenz für Europa, COM(2018) 237 final, 25. April 2018, <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=COM%3A2018%3A237%3AFIN>; mehr Hintergrundinformationen zum Komplex auch hier: [https://ec.europa.eu/commission/presscorner/detail/de/QANDA\\_21\\_1683#2](https://ec.europa.eu/commission/presscorner/detail/de/QANDA_21_1683#2).

europäisches Konzept für Exzellenz und Vertrauen“<sup>192</sup> an. Hier werden die Gefahren der Diskriminierung durch KI mit Blick auf Geschlecht schon im ersten Absatz erwähnt. Dieses Bewusstsein zieht sich, wenn auch teils oberflächlich, durch das gesamte „Weißbuch“: „Infolge der Nutzung von KI können die Werte, auf denen die EU gründet, beeinträchtigt und Grundrechte verletzt werden. Dies gilt auch für das Recht auf freie Meinungsäußerung, die Versammlungsfreiheit, die Achtung der Menschenwürde, die Nichtdiskriminierung ungeachtet des Geschlechts, der Rasse oder der ethnischen Herkunft, der religiösen Überzeugung oder Weltanschauung, einer Behinderung, des Alters oder der sexuellen Ausrichtung, den Schutz personenbezogener Daten und des Privatlebens, das Recht auf einen wirksamen gerichtlichen Rechtsbehelf und ein faires Verfahren sowie den Verbraucherschutz.“<sup>193</sup> Explizit wird anerkannt, dass Diskriminierung aufgrund von Geschlecht oder Rasse im Rahmen von Software zur Vorhersage der Rückfälligkeit von Straftätern bereits nachgewiesen werden konnte, ebenso wie im Bereich der Gesichtserkennung.<sup>194</sup>

Eine Studie im Auftrag des Europarates, auf die immer wieder verwiesen wird, zeigt darüber hinaus eine große Zahl von Grundrechten auf, die durch den Einsatz von KI beeinträchtigt werden könnten.<sup>195</sup> Die Kommission unterstützt außerdem die Kernforderung nach „Vielfalt,

192 Europäische Kommission, Weißbuch zur Künstlichen Intelligenz – Ein europäisches Konzept für Exzellenz und Vertrauen, COM(2020) 65 final, 19. Februar 2020, [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_de.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_de.pdf); das Weißbuch wird von der Bundesregierung weitestgehend unterstützt: [https://www.ki-strategie-deutschland.de/files/downloads/Stellungnahme\\_BReg\\_Weissbuch\\_KI.pdf](https://www.ki-strategie-deutschland.de/files/downloads/Stellungnahme_BReg_Weissbuch_KI.pdf) – von zivilgesellschaftlicher Seite dagegen häufig kritisiert. Vielfach deshalb, weil biometrische Massenüberwachungstechnologien eben nicht ausdrücklich verboten werden, ebenso bleibt unklar, welche gesetzgeberischen Grenzen für KI-basierte Systeme gezogen werden sollten, die gegen Grundrechte verstoßen. Auch wird nicht wirklich deutlich, auf welcher Grundlage die derzeitige Liste der risikoreichen Anwendungen erstellt wurde und was als valider Grund für eine künftige Risikoeinstufung gelten darf. Vgl. z. B. AlgorithmWatch, AlgorithmWatch's response to the European Commission's proposed regulation on Artificial Intelligence – A major step with major gaps, 22. April 2021, <https://algorithmwatch.org/en/response-to-eu-ai-regulation-proposal-2021>; rote Linien für die Anwendung von KI, die Grundrechte bedroht, fordern auch eine Vielzahl weiterer Akteure\*innen. Vgl. zB EDRI, Civil society calls for AI red lines in the European Union's Artificial Intelligence proposal, 12. Januar 2021, <https://edri.org/our-work/civil-society-call-for-ai-red-lines-in-the-european-unions-artificial-intelligence-proposal/>; auch die Einschätzung des Deutschen Juristinnenbundes, Stellungnahme: 21-14, 30. Juni 2021, <https://www.djb.de/presse/stellungnahmen/detail/st21-14> ist in diesem Zusammenhang sehr aufschlussreich.

193 Ebenda, S. 12 ff.

194 Ebenda, S. 14.

195 Council of Europe, Algorithms and Human Rights – Study on the human rights dimensions of automated data processing techniques (in particular algorithms) and possible regulatory implications, März 2018, <https://rm.coe.int/algorithms-and-human-rights-en-rev16807956b5>.

Nichtdiskriminierung und Fairness“.<sup>196</sup> Sie stellt überdies klar, dass Entwickler\*innen und Nutzer\*innen von KI bereits den geltenden europäischen Rechtsvorschriften über Grundrechte unterliegen, weist aber darauf hin, dass diese mit Blick auf KI gegebenenfalls nachgearbeitet werden müssen.<sup>197</sup> Es gilt: Den Schwerpunkt ihres Regulierungsrahmens legt die EU auf das Vermeiden „schwerster Schäden“, wozu explizit das Brechen mit Grundrechten auch und vor allem mit Blick auf den Grundsatz der Nichtdiskriminierung zählt.<sup>198</sup> Im Bereich der Trainingsdaten für KI-Systeme zum Beispiel empfiehlt das „Weißbuch“ das Etablieren von Auflagen, die Entwickler\*innen darauf verpflichten, „umfassende Datensätze zu verwenden, die ausreichend repräsentativ sind.“<sup>199</sup> Hinzu kommt das verpflichtende Dokumentieren der „für Programmierung und Training verwendeten Methoden (...) unter Vermeidung von Verzerrungen, die zu verbotenen Diskriminierungen führen könnten.“<sup>200</sup>

Darüber hinaus müssen Mitgliedstaaten aber, so Empfehlung Nr. 88, zweckgebundene Mittel aus ihren öffentlichen Haushalten für die Finanzierung von geschlechtergerechten Programmen bereitstellen und sicherstellen, dass die nationale Digitalpolitik einen Aktionsplan für die Gleichstellung der Geschlechter enthält.

Die UNESCO-Empfehlung<sup>201</sup> weist in ihren Paragrafen 87–93 zum Politikfeld Gender viele Ähnlichkeiten mit den von der EU avisierten und teils bereits national auf den Weg gebrachten Standards auf, geht aber viel weiter. Dazu gehören der besondere Schutz der Grund- und Menschenrechte sowie marginalisierter Gruppen unter dem Gebot der Nichtdiskriminierung ebenso wie die Förderung von Diversität im Sektor KI(-Entwicklung), vor allem was den Faktor Repräsentation und Relevanz angeht. Diskriminierende Biases, tradiert durch Technik, sollen unterbunden und das Erstarren von inklusiver, diverser KI-Entwicklung gefördert werden. Die UNESCO legt in ihrer Empfehlung einen sehr starken Fokus auf den Faktor Repräsentation und bleibt doch mit allen Forderungen, Maßnahmen und Handlungsempfehlungen zunächst an der Oberfläche.

196 Europäische Kommission, Weißbuch zur Künstlichen Intelligenz – Ein europäisches Konzept für Exzellenz und Vertrauen, COM(2020) 65 final, 2020, [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_de.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_de.pdf), S.12; und auch in der deutschen Normungsrroadmap Künstliche Intelligenz taucht dieser Passus nachdrücklich auf: <https://www.din.de/resource/blob/772438/6b5ac6680543eff-9fe372603514be3e6/normungsrroadmap-ki-data.pdf>, S. 57.

197 Europäische Kommission, Weißbuch zur Künstlichen Intelligenz – Ein europäisches Konzept für Exzellenz und Vertrauen, COM(2020) 65 final, 2020, [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_de.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_de.pdf), S. 11 ff.

198 Ebenda, S. 13.

199 Ebenda, S. 23.

200 Ebenda, S. 24.

201 Vgl. UNESCO, Draft text of the Recommendation on the Ethics of Artificial Intelligence, SHS/IGM-AIETHICS/2021/JUN/3 Rev.2, 25. Juni 2021, <https://unesdoc.unesco.org/ark:/48223/pf0000377897>, S.20 f., Policy Area 6: Gender.

Die UNESCO-Empfehlung legt merklich einen Schwerpunkt auf die Wahrung von Gender Equality bei der KI-Entwicklung und -Nutzung; hier besteht für Deutschland noch großer Handlungsbedarf. Sowohl die Gleichbehandlung der Geschlechter wie auch der Faktor Diversität in KI-Entwickler\*innen-Teams hat in Deutschland noch viel Luft nach oben (UNESCO Punkt 19). Obwohl die Förderung von Frauen und zunehmend auch diverser Menschen im Bereich der Naturwissenschaften und MINT schon seit Jahren immer wieder auf Bundesebene artikuliert und durch verschiedene Förderprojekte im Bildungs- und Berufsbildungsbereich unterstützt wird (UNESCO Forderungen 87–89, 91–93), ist Deutschland im Bereich der Gleichstellung der Geschlechter in der Entwicklung und Beforschung von KI sowie im Aufbau interdisziplinärer Teams noch immer schlecht aufgestellt.

So sind in Deutschland nur 16% der KI-Fachleute weiblich (weltweit sind es 22%), Frauen verdienen in Deutschland im KI-Sektor im Schnitt 25% weniger als ihre männlichen Kollegen.<sup>202</sup> Um dies zu verändern, braucht es neben dem Aufheben von geschlechterstereotyper frühkindlicher Erziehung und Bildung auch explizit Unterstützung für Frauen und diverse Menschen, um nachhaltig und erfolgreich in den vorwiegend weißen, männlichen Technologie-Entwicklungsbereich vorstoßen und sich dort etablieren zu können. Auch extrinsische Incentives sollten hier vermehrt kommuniziert werden: Diversere Teams ermöglichen nachweislich diversere und auch erfolgreichere Produkte. Zudem braucht es mehr Sichtbarkeit von Frauen in der KI als Vorbilder, beispielsweise durch die Verbreitung von Erfolgsgeschichten.<sup>203</sup>

Neben barrierefreien Bildungsangeboten für jedes Alter braucht es eine politisch getriebene, ernsthafte strukturelle Veränderung der (Wirtschafts-)Logik von IT- & KI-Unternehmen.<sup>204</sup> Die derzeitige Schiefelage in der Repräsentation wird nicht nur über Bildung zu reparieren sein, vielmehr braucht es starke Anreize zu strukturellen und systemischen Veränderungen innerhalb der Organisationskulturen der Unternehmen und ein Umdenken der KI-Entwicklung in intersektionaler Perspektive, um verschiedene Ausschlusssysteme sichtbar zu machen und zu überwinden.<sup>205</sup>

202 Vgl. Hedda Nier, „Wie weiblich ist die IT?“, statista, 20. März 2018, <https://de.statista.com/infografik/13283/frauen-inder-tech-branche>.

203 Vgl. HIIG AI Lab, Positionspapier zum Roundtable „KI und Frauen“, Dezember 2020, [https://www.hiig.de/wp-content/uploads/2020/12/Positionspapier-KI-und-Frauen-WEB\\_V2.pdf](https://www.hiig.de/wp-content/uploads/2020/12/Positionspapier-KI-und-Frauen-WEB_V2.pdf), S. 7 ff.

204 Ebenda, S. 11.

205 Vgl. ebenda, S. 12.

## Fazit

Mitgliedstaaten sollten dafür sorgen, dass KI die in der analogen Welt in verschiedenen Bereichen bereits bestehenden großen geschlechtsspezifischen Unterschiede nicht noch weiter verschärft, sondern diese Unterschiede vielmehr beseitigt. Zu diesen Diskrepanzen gehören: das geschlechtsspezifische Lohngefälle, die ungleiche Vertretung in bestimmten Berufen und Tätigkeiten, die fehlende Vertretung in Führungspositionen, Aufsichtsräten oder Forschungsteams im Bereich der KI, das Bildungsgefälle, das Gefälle beim Zugang zu digitalen Inhalten, bei der ungleichen Verteilung unbezahlter Arbeit und von Betreuungsaufgaben. Die Mitgliedstaaten sollen auch, so die Forderung Nr. 90, sicherstellen, dass geschlechtsspezifische Stereotype und diskriminierende Vorurteile nicht in KI-Systeme einfließen, sondern dass diese identifiziert und proaktiv beseitigt werden. Wie eine Expert\*innen-gruppe zuletzt formulierte: „KI-Entwicklung muss intersektional gedacht werden, um manifeste Ausschlusssysteme zu durchbrechen und starre Kategoriensysteme sichtbar zu machen. Dazu bedarf es nicht zuletzt einer intensiven Auseinandersetzung mit der Frage nach Zugang zu materiellen wie immateriellen Ressourcen, Zugangschancen und Teilhabemöglichkeiten, Chancen auf Bildung und Beteiligung.“<sup>206</sup>

Laut OECD-Übersicht gibt es Initiativen zu den Bereichen KI und Gender in sechs afrikanischen Ländern.<sup>207</sup> Natürlich gibt es einzelne, teils auch von Deutschland geförderte Initiativen, die das Thema treiben, so zum Beispiel „FAIR Forward – Künstliche Intelligenz für alle mit Afrika und Asien“<sup>208</sup>, diese behandeln das Thema Gender, Diversity und Nichtdiskriminierung aber nur am Rande. In der Initiative „Artificial Intelligence for Development in Africa“ (AI4D Africa) wird das Fördern von inklusiver, diversitätssensibler KI durchaus stärker in den Fokus gestellt, aber noch nicht in einen regulatorischen Rahmen gegossen.<sup>209</sup>

Mit Blick auf die Verpflichtung (Nr. 90), dass die Gleichstellung der Geschlechter in keiner Phase des Lebenszyklus von KI-Systemen verletzt werden darf, steht Deutschland auch noch ganz am Anfang: Zu sichern, dass geschlechtsspezifische Stereotype und diskriminierende Vorurteile nicht in KI-Systeme einfließen, sondern dass diese identifiziert und proaktiv beseitigt werden, wird zu substanziellen Änderungen im staatlichen Umgang mit KI-basierten Entscheidungsfindungsprozessen führen.

Gleichbehandlung hat auch mit Datenpolitik zu tun, da viele Daten mit Genderdimension oder jene, die sich auf sexuelle Orientierung beziehen, sehr sensibel sind. In Punkt 74 der Empfehlung

206 Ebenda.

207 Vgl. OECD, Dashboards, <https://oecd.ai/en/dashboards>.

208 FAIR Forward, Toolkit Digitalisierung, <https://toolkit-digitalisierung.de/fair-forward-ki-fuer-alle-mit-afrika-und-asien>.

209 Vgl. Artificial Intelligence for Development in Africa, <https://africa.ai4d.ai/blog/ai4d-afrika-launch-driving-inclusive-ai-innovation-in-afrika>.

bekennen sich die Staaten dazu, sensible Daten besonders zu schützen. Dazu gehören: Daten über Straftaten, Strafverfahren und Verurteilungen sowie damit zusammenhängende Sicherheitsmaßnahmen, biometrische, genetische und Gesundheitsdaten sowie personenbezogene Daten wie Daten über Rasse, Hautfarbe, Abstammung, Geschlecht, Alter, Sprache, Religion, politische Meinung, nationale Herkunft, ethnische Herkunft, soziale Herkunft, wirtschaftliche oder soziale Umstände der Geburt oder Behinderung und andere Merkmale.

Auch im Bereich „biometrische Fernidentifikation“ wird mindestens implizit deutlich, wie hochsensibel diese Technologien anderen Grundrechten gegenüberstehen, und dass ein Verletzen ebendieser Grundrechte dringend verhindert werden müsse.<sup>210</sup> Noch konkreter wird hier der Entwurf von EU-Parlament und -Rat als Vorschlag für die Verordnung von harmonisierten Vorschriften für Künstliche Intelligenz (Gesetz über Künstliche Intelligenz).<sup>211</sup> Hier wird in Abs. 13, 15, 17, 33, 35–39, 44, 45 und 47 sehr detailliert aufgegriffen, was im „Weißbuch“ schon angelegt ist, nämlich das unbedingte Verhindern der Verstärkung von (impliziten) Biases durch in hochsensiblen Bereichen eingesetzte Künstliche Intelligenz, die mit nicht intersektional-diskriminierungsfreien Datensets trainiert wurde.<sup>212</sup>

Überwachungstechnologien gefährden überall auf der Welt vor allem ohnehin marginalisierte Gruppen. Ihr Einsatz und die Bedingungen ihres Einsatzes müssen im Bereich der Strafverfolgung, des Sozialstaates sowie in allen anderen gesellschaftlichen Bereichen stets kritisch hinterfragt und fortlaufend unabhängig, umfassend und intersektional vor dem Hintergrund postkolonialer Machtverhältnisse und unter dem Dach der Grund- und Menschenrechte evaluiert werden.<sup>213</sup>

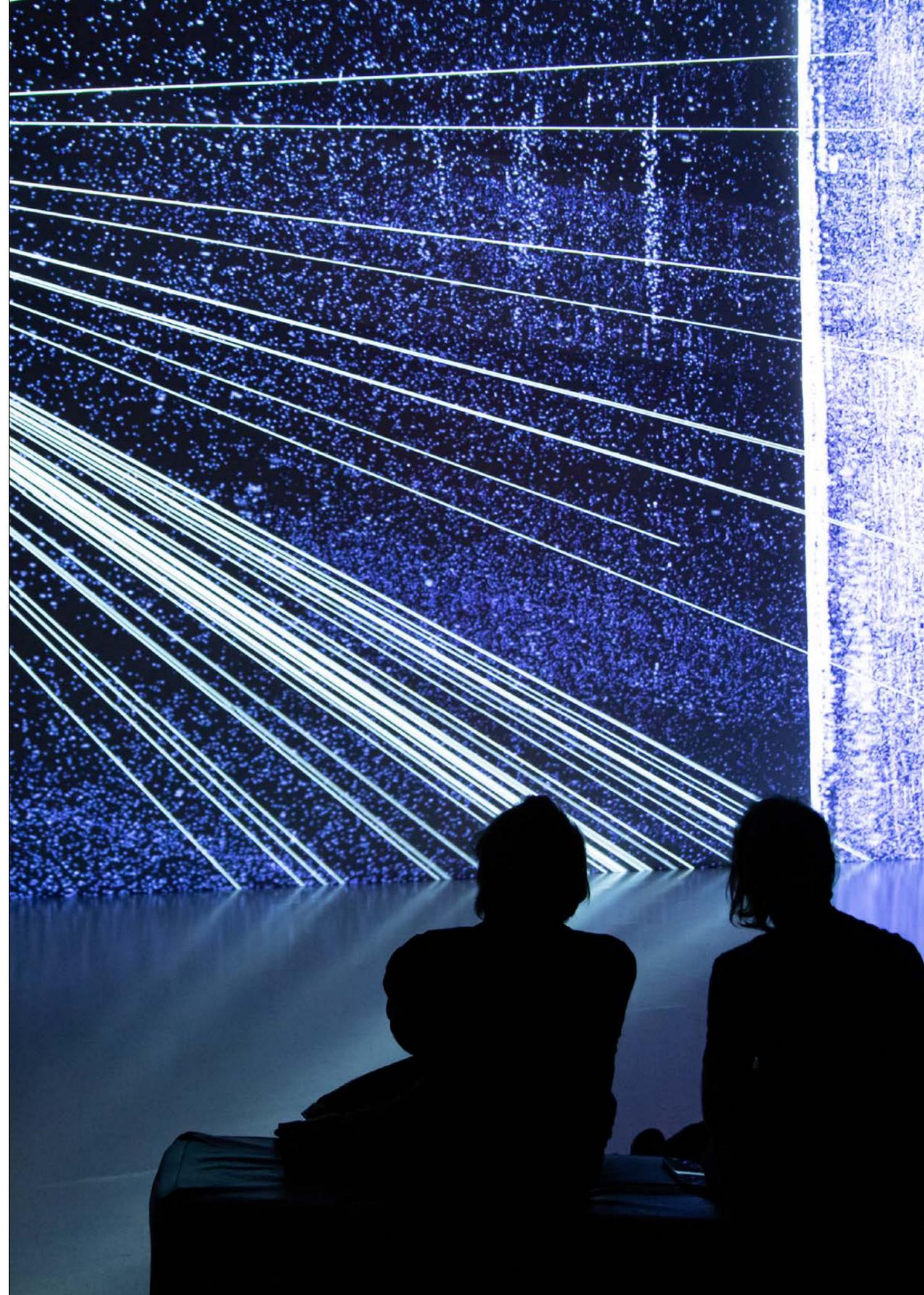
210 Ebenda, S. 25.

211 Europäische Kommission, 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, COM(2021) 206 final, 21. April 2021, <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206>, mehr Hintergründe und Zeitläufe auch hier: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM:2021:205:FIN>.

212 Ebenda; ähnliche Herangehensweisen mit Blick auf das Verhindern von Biases in Datensätzen und KI, die diversere Besetzung von Entwickler\*innen-Teams und den Faktor Transparenz finden wir auch bei der High-Level Expert Group on AI, z.B. hier: Policy and Investment Recommendations for Trustworthy AI, 26. Juni 2019, <https://digital-strategy.ec.europa.eu/en/library/policy-and-investment-recommendations-trustworthy-artificial-intelligence> und hier: Ethics guidelines for trustworthy AI, 8. April 2019, <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>: „Diversity, non-discrimination and fairness: Unfair bias must be avoided, as it could have multiple negative implications, from the marginalization of vulnerable groups, to the exacerbation of prejudice and discrimination. Fostering diversity, AI systems should be accessible to all, regardless of any disability, and involve relevant stakeholders throughout their entire life circle.“ Sowie hier: European AI Alliance, <https://futurium.ec.europa.eu/en/european-ai-alliance/pages/about?language=en>.

213 Vgl. netzforma\* e.V., Wenn KI – dann feministisch, Berlin 2020, [https://netzforma.org/wp-content/uploads/2021/01/2020\\_wenn-ki-dann-feministisch\\_netzforma.pdf](https://netzforma.org/wp-content/uploads/2021/01/2020_wenn-ki-dann-feministisch_netzforma.pdf), Kapitel Überwachung, S. 15–30 sowie S. 206 ff.

Fazit



## IV. Fazit

Während die meisten ethischen Leitlinien zu Künstlicher Intelligenz allgemeine Prinzipien propagieren und nicht auf praktische Umsetzung abzielen, liegt mit der KI-Ethik-Empfehlung der UNESCO ein erster, in einem offenen und inklusiven Prozess ausverhandelter Völkerrechtstext vor, der ethische Prinzipien in konkrete politische Gestaltungsaufgaben übersetzt. Die Empfehlung umfasst weltweit akzeptierte ethische Standards für KI-Technologien unter voller Beachtung der Menschenrechte.

Bemerkenswert ist, dass sich alle 193 Mitgliedstaaten der UNESCO zu der Empfehlung bekannt haben. Der vorangegangene zweijährige Entstehungsprozess, bei dem auch nichtstaatliche Akteur\*innen ihre Stimmen einbringen konnten, stellt einen bedeutenden Mehrwert an Legitimation dar. Inhaltlich sticht der ganzheitliche Ansatz der Empfehlung positiv hervor. Einerseits wird deutlich, dass die Empfehlung KI-Ethik als Gesellschaftsaufgabe versteht, für deren Erfüllung es gilt, Bürger\*innen für Potenziale und Risiken von KI zu sensibilisieren. Andererseits definiert die Empfehlung KI-Ethik auch als unmittelbar anwendbare Normenquelle: Sehr plausibel ist daher die Formulierung von politischen Gestaltungsaufgaben für KI-Entwicklung und -Nutzung in elf Politikfeldern, darunter Bildung und Wissenschaft, Kommunikation, Gesundheit sowie Umwelt. Nicht abstrakte Forderungen, sondern konkrete Verbesserungen stehen im Mittelpunkt.

Die Empfehlung nutzt die Vorteile der Ethik gegenüber dem Recht, insbesondere mit Blick auf die Entwicklung und Umsetzung von politischen Maßnahmen und bei der Präzisierung von Rechtsnormen, wie den gleichbehandlungsrechtlichen Verpflichtungen, die auf KI-Anwendungen ausgeweitet werden. Durch die Bezugnahme auf Ethik sowie den weichen Rechtsstatus einer Empfehlung konnten sowohl eine starke Menschenrechtssprache als auch der Fokus auf die praktische Umsetzung im Dokument verankert werden. Dies wäre bei der Zielsetzung eines „harten“ globalen KI-Rechts nicht möglich gewesen, da vielfach normative Ordnungen erst im Entstehen sind. Bemerkenswert ist auch die Berücksichtigung bisheriger „blind spots“ des KI-Rechts, wie etwa Auswirkungen von KI auf die Bereiche Gender, Nachhaltigkeit, Entwicklungspolitik oder Bildung. Weiterhin definiert die Empfehlung Verfahrens- und Transparenzpflichten: Konkret werden etwa

**„Jetzt sind die Mitgliedstaaten gefragt, ihrer völkerrechtlichen Verpflichtung nachzukommen, indem sie die politischen Gestaltungsaufgaben der Empfehlung mit Leben füllen.“**

ethische Impact Assessments für die Entwicklung und Nutzung von KI-Systemen gefordert, die wiederum in einem multistakeholderbasierten, offenen und responsiven Prozess stattfinden sollen.

Ein deutlicher Mehrwert der UNESCO-Empfehlung lässt sich auch dahingehend konstatieren, dass in dieser auf die kulturellen Aspekte hingewiesen wird, die beim Einsatz von KI-Systemen berührt sein können. Hierzu zählen potenziell positive Effekte wie der Erhalt bedrohter Sprachen, aber auch mögliche negative Auswirkungen wie die Schaffung neuer Eintrittsbarrieren auf digitalen Kulturplattformen. Ferner legt die Empfehlung erstmals einen Fokus auf die besonderen Dynamiken, mit denen Länder des Globalen Südens bei der Nutzung und Entwicklung von KI-Systemen konfrontiert sein können. Dabei wird nicht nur auf die potenziell höhere Vulnerabilität von Entwicklungsländern hingewiesen, sondern auch deren mögliche Beiträge für eine fairere KI-Nutzung und bessere KI-Regulierung hervorgehoben. Indem die UNESCO-Empfehlung zudem einen Akzent auf die Bedeutung und Schutzwürdigkeit bestimmter marginalisierter Gruppen legt, kann ein wichtiger Impuls in Richtung des Gesetzgebungsprozesses der EU gesetzt werden, um dort ein hinreichendes Schutzniveau durch taugliche regulative Instrumente und Mechanismen zu gewährleisten.

Die KI-Ethik-Empfehlung der UNESCO ist der erste globale Konsens für eine ethische Entwicklung und Nutzung Künstlicher Intelligenz. Sie eröffnet damit ein einzigartiges Möglichkeitsfenster, KI-Entwicklungspfade ethisch einzuhegen und somit im Sinne der Menschen(rechte) aktiv zu gestalten. Jetzt sind die Mitgliedstaaten gefragt, ihrer völkerrechtlichen Verpflichtung nachzukommen, indem sie die politischen Gestaltungsaufgaben der Empfehlung mit Leben füllen. Die Gesetze, Programme und Finanzierungsinstrumente, die nun auf Grundlage der Empfehlung verabschiedet werden, müssen in der Rückschau der Gradmesser für deren Erfolg sein.

# Impressum

## Herausgeberin

---

Deutsche UNESCO-Kommission e.V.  
Martin-Luther-Allee 42  
53175 Bonn

## Vertretungsberechtigte:

Prof. Dr. Maria Böhmer (Präsidentin)  
Prof. Dr. Christoph Wulf (1. Vizepräsident)  
Prof. Dr. Hartwig Carsten Lüdtko  
(2. Vizepräsident)  
Dr. Roman Luckscheiter (Generalsekretär)  
Dr. Lutz Möller (Besonderer Vertreter  
gem. § 30 BGB)

Rechtsform: Eingetragener Verein (Satzung)

Vereinssitz: Bonn, Eintragung im Vereins-  
register des Amtsgerichts – Registergericht –  
Bonn, Registernummer: VR 4827

## Redaktion

---

Jeannine Hausmann (verantwortlich),  
Dr. Maximilian Müngersdorff

## Gestaltung

---

Panatom, Berlin

## Copyright

---

Die Texte dieser Publikation sind unter der  
Creative Commons-Lizenz Namensnennung-  
Nicht-kommerziell 3.0 Deutschland (CC BY-NC  
3.0 DE) lizenziert. <https://creativecommons.org/licenses/by-sa/4.0/deed.de>. Die Fotos sind  
von der Lizenz ausgenommen.

## ISBN

---

978-3-947675-26-5

## Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet  
diese Publikation in der Deutschen National-  
bibliografie; detaillierte bibliografische Daten  
sind im Internet über <http://dnb.d-nb.de>  
abrufbar.

## Autor\*innen

---

Matthias C. Kettemann ist Professor für Innovation,  
Theorie und Recht am Institut für Theorie und  
Zukunft des Rechts der Universität Innsbruck,  
Forschungsprogrammleiter am Leibniz-Institut  
für Medienforschung | Hans-Bredow-Institut  
(HBI), Leiter der Forschungsgruppe „Globaler  
Konstitutionalismus und das Internet“ am  
Humboldt-Institut für Internet und Gesellschaft,  
Berlin, und der Forschungsgruppe „Platform and  
Content Governance“ am Sustainable Computing  
Lab der Wirtschaftsuniversität Wien, Co-Leiter  
der Sektion Internetworkrecht am Max-Planck-  
Institut für ausländisches öffentliches Recht  
und Völkerrecht, Heidelberg, und assoziiertes  
Mitglied des Forschungszentrums Gesellschaft-  
licher Zusammenhalt – Teilinstitut Hamburg,  
des Forschungsverbands „Normative Ordnungen“  
der Goethe-Universität Frankfurt am Main und  
des Forschungsnetzwerks „Human Factor in  
Digital Transformation“ der Karl-Franzens-  
Universität Graz.

Philipp Jaud ist wissenschaftlicher Mitarbeiter  
und Lisa-Maria Riedl studentische Mitarbeiterin  
am Institut für Theorie und Zukunft des Rechts  
der Universität Innsbruck.

Martin Fertmann, Vincent Hofmann, Katharina  
Mosene und Johannes Schmees sind wissen-  
schaftliche Mitarbeiter\*innen (Junior Researcher)  
am Leibniz-Institut für Medienforschung |  
Hans-Bredow-Institut.

## Fotografien

---

S. 10 © Martin Hieslmair  
S. 15 © Gerd Altmann  
S. 19 © iStock, metamorworks  
S. 21 © Adi Goldstein  
S. 24 © Martin Wagenhan  
S. 29 © This is Engineering  
S. 32 © iStock, Peter Howell  
S. 45 © vog.photo

