

UNESCO- Empfehlung zur Ethik der Künstlichen Intelligenz

zweisprachige Publikation
DE / EN



UNESCO- Empfehlung zur Ethik der Künstlichen Intelligenz

Mit den UNESCO-Nationalkommissionen
Luxemburgs, Österreichs und der
Schweiz abgestimmte Übersetzung

Einleitung

Künstliche Intelligenz (KI) ist Teil unseres Alltags. Sie wird von Menschen gemacht und eingesetzt. Sie ist gestaltbar. Daher ist es möglich und nötig, sich weltweit auf Regeln zu verständigen. Solche Regeln sollen sicherstellen, dass KI-Systeme ausschließlich zum Wohle des Menschen eingesetzt werden und die Menschenrechte und Grundfreiheiten nicht verletzen.

Vor diesem Hintergrund hat die UNESCO in den Jahren 2020 und 2021 die Empfehlung zur Ethik der Künstlichen Intelligenz erarbeitet. Das inhaltliche Mandat der UNESCO hat die Konsensfindung erleichtert: Als Organisation der Vereinten Nationen für Bildung, Wissenschaft und Kultur ist sie für die Verhandlung ethischer Fragen rund um den technologischen Fortschritt zuständig und tätig. Die Empfehlung ist der erste global gültige

Völkerrechtstext zu ethischer KI; eine Empfehlung muss anders als eine Konvention nicht in nationales Recht der Mitgliedstaaten übersetzt werden.

Die UNESCO beauftragte 2020 eine Gruppe von Expertinnen und Experten aus aller Welt, einen Textentwurf zu erstellen. Parallel konnten alle 193 UNESCO-Mitgliedstaaten, die Wissenschaft und die Zivilgesellschaft ihre Ansichten zum erhofften Inhalt der Empfehlung zu Protokoll geben. Im Herbst 2020 konnten alle Regierungen den ersten Entwurf kommentieren. Der zweite Entwurf, der diese Kommentare berücksichtigte, wurde im Sommer 2021 von allen Staaten Wort für Wort so lange (etwa 100 Verhandlungsstunden) diskutiert, bis alle zustimmen konnten. Gerade über die Menschenrechte wurde besonders kontrovers

verhandelt. Auf der UNESCO-Generalkonferenz im November 2021 wurde die UNESCO-Empfehlung im Konsens verabschiedet.

Die Empfehlung zeichnet sich aus durch klar definierte ethische Werte und Prinzipien, ihre starke Menschenrechtssprache und konkrete Ziele und Gestaltungsaufgaben für die UNESCO-Mitgliedstaaten. Sie erfüllt somit alle Voraussetzungen für einen globalen Kompass für eine verantwortliche KI-Entwicklung und -Nutzung. Aus Sicht der UNESCO sollen KI-Systeme zu Frieden, Menschenwürde, Nachhaltigkeit und Sicherheit beitragen und diesen Grundprinzipien nicht zuwiderlaufen.

In elf Politikbereichen macht die Empfehlung konkrete Vorgaben für den ethischen Einsatz von KI-Systemen. Dazu zählen bekannte Fragen aus der

Diskussion über KI wie Folgen für die Arbeitswelt, den Gesundheits- oder den Bildungssektor. Gleichzeitig spricht die Empfehlung bisherige „blinde Flecken“ des KI-Rechts an, wie die Interessen des Globalen Südens, geschlechtsspezifische Vorurteile, Diskriminierung oder Fragen der Nachhaltigkeit.

Alle Staaten der Welt können die Empfehlung dazu nutzen, die vielen heute denkbaren Entwicklungspfade künftiger KI-Nutzung in ethische Bahnen zu lenken. Am besten „übersetzen“ sie dazu Vorgaben der Empfehlung in nationales Recht, möglichst in Zusammenarbeit mit der Zivilgesellschaft und den KI-Unternehmen.

Präambel

Die Generalkonferenz der Organisation der Vereinten Nationen für Bildung, Wissenschaft und Kultur (UNESCO), die vom 9. bis zum 24. November 2021 in Paris zu ihrer 41. Sitzung zusammengekommen ist –

In Anerkennung der tiefgreifenden und dynamischen positiven wie negativen Auswirkungen, die Künstliche Intelligenz (KI) auf Gesellschaften, Umwelt, Ökosysteme und das menschliche Leben einschließlich des menschlichen Geistes hat, unter anderem weil ihr Einsatz auf ganz neue Arten das Denken, die Interaktion und die Entscheidungsfindung von Menschen beeinflusst, und Auswirkungen auf die Bildung, die Geistes-, Sozial- und Naturwissenschaften, die Kultur sowie die Kommunikation und Information hat,

Eingedenk dessen, dass die UNESCO kraft ihrer Verfassung das Ziel hat, zu Frieden und Sicherheit beizutragen, indem sie die Zusammenarbeit zwischen den Nationen durch Bildung, Wissenschaft, Kultur, Kommunikation und Information fördert, um die weltweite

Achtung vor Recht und Gerechtigkeit, vor den Menschenrechten und Grundfreiheiten, die den Völkern der Welt bestätigt worden sind, zu stärken,

In der Überzeugung, dass die vorliegende Empfehlung, die in einer globalen Herangehensweise auf der Grundlage des Völkerrechts als ein standardsetzendes Dokument entwickelt wurde, das Menschenwürde, Menschenrechte, Gleichberechtigung der Geschlechter, soziale und wirtschaftliche Gerechtigkeit und Entwicklung, körperliches und geistiges Wohlbefinden, Vielfalt, Verbundenheit, Inklusivität sowie Schutz der Umwelt und der Ökosysteme in den Mittelpunkt stellt, die KI-Technologien in eine verantwortungsvolle Richtung lenken kann,

Geleitet von den Zielen und Grundsätzen der Charta der Vereinten Nationen,

Preamble

The General Conference of the United Nations Educational, Scientific and Cultural Organization (UNESCO), meeting in Paris from 9 to 24 November 2021, at its 41st session,

Recognizing the profound and dynamic positive and negative impacts of artificial intelligence (AI) on societies, environment, ecosystems and human lives, including the human mind, in part because of the new ways in which its use influences human thinking, interaction and decision-making and affects education, human, social and natural sciences, culture, and communication and information,

Recalling that, by the terms of its Constitution, UNESCO seeks to contribute to peace and security by promoting collaboration among nations through education, the sciences, culture, and communication and information, in order to further universal respect for justice, for the rule of law and for the human rights and fundamental freedoms which are affirmed for the peoples of the world,

Convinced that the Recommendation presented here, as a standard-setting instrument developed through a global approach, based on international law, focusing on human dignity and human rights, as well as gender equality, social and economic justice and development, physical and mental well-being, diversity, interconnectedness, inclusiveness, and environmental and ecosystem protection can guide AI technologies in a responsible direction,

Guided by the purposes and principles of the Charter of the United Nations,

In Anbetracht dessen, dass KI-Technologien für die Menschheit sehr nützlich sein können und alle Länder davon profitieren können, andererseits aber auch zu grundsätzlichen ethischen Bedenken führen, beispielsweise im Hinblick auf Verzerrungen, die sie zementieren und verschärfen können und die zu Diskriminierung, Ungleichheit, großen digitalen Zugangsnachteilen, Ausgrenzung und einer Bedrohung kultureller, sozialer und biologischer Vielfalt sowie zu sozialen bzw. wirtschaftlichen Verwerfungen führen können; angesichts der Notwendigkeit, die Funktionsweise der Algorithmen und Daten, mit denen sie trainiert wurden, transparent und verständlich zu machen, sowie in Anbetracht ihrer potenziellen Auswirkungen unter anderem auf die Menschenwürde, auf Menschenrechte und Grundfreiheiten, Gleichberechtigung der Geschlechter, Demokratie, soziale, wirtschaftliche, politische und kulturelle Prozesse, wissenschaftliche und technische Praktiken, Tierwohl, Umwelt und Ökosysteme,

Des Weiteren in der Erkenntnis, dass KI-Technologien bestehende Verwerfungen und Ungleichheiten in der Welt vertiefen können, sowohl innerhalb von Ländern als auch zwischen

Ländern, und dass Gerechtigkeit, Vertrauen und Fairness hochgehalten werden müssen, damit kein Land und niemand zurückgelassen wird, sowohl durch einen fairen Zugang zu KI-Technologien und Nutzung ihrer Vorteile als auch durch Schutz vor ihren negativen Auswirkungen, wobei die unterschiedlichen Gegebenheiten in den einzelnen Ländern anzuerkennen sind und der Wunsch einiger Menschen, sich nicht an allen technologischen Entwicklungen zu beteiligen, zu respektieren ist,

Considering that AI technologies can be of great service to humanity and all countries can benefit from them, but also raise fundamental ethical concerns, for instance regarding the biases they can embed and exacerbate, potentially resulting in discrimination, inequality, digital divides, exclusion and a threat to cultural, social and biological diversity and social or economic divides; the need for transparency and understandability of the workings of algorithms and the data with which they have been trained; and their potential impact on, including but not limited to, human dignity, human rights and fundamental freedoms, gender equality, democracy, social, economic, political and cultural processes, scientific and engineering practices, animal welfare, and the environment and ecosystems,

Also recognizing that AI technologies can deepen existing divides and inequalities in the world, within and between countries, and that justice, trust and fairness must be upheld so that no country and no one should be left behind, either by having fair access to AI technologies and enjoying their benefits or in the protection against their negative implications, while recognizing the

different circumstances of different countries and respecting the desire of some people not to take part in all technological developments,

Im Bewusstsein der Tatsache, dass alle Länder mit einer beschleunigten Nutzung von Informations- und Kommunikationstechnologien sowie KI-Technologien und einem zunehmenden Bedarf an Medien- und Informationskompetenz konfrontiert sind, und dass die digitale Wirtschaft sowohl große gesellschaftliche, wirtschaftliche und ökologische Herausforderungen als auch Chancen für eine gerechte Verteilung der Vorteile bietet, insbesondere für Länder mit niedrigem und mittlerem Einkommen („low and middle income countries“ - LMICs), einschließlich der am wenigsten entwickelten Länder („least developed countries“ - LDCs), Binnenentwicklungsländer („landlocked developed countries“ - LLDCs) und kleiner Inselentwicklungsstaaten („small island developing states“ – SIDS), jedoch nicht auf diese beschränkt, wobei landesspezifische Kulturen, Werte und Kenntnisse anzuerkennen, zu schützen und zu fördern sind, damit nachhaltige digitale Wirtschaftssysteme entwickelt werden können,

Ferner anerkennend, dass KI-Technologien das Potenzial haben, vorteilhaft für die Umwelt und die Ökosysteme zu sein, und dass es zur Ver-

wirklichung dieser Vorteile notwendig ist, mögliche Schäden und negative Auswirkungen auf die Umwelt und die Ökosysteme nicht zu ignorieren, sondern sie anzugehen,

Unter Hinweis darauf, dass die Innovation und Entwicklung durch die Beschäftigung mit Risiken und ethischen Bedenken nicht behindert werden, sondern dadurch vielmehr neue Möglichkeiten eröffnet und ethisch geleitete Forschung und Innovation gefördert werden sollte, die KI-Technologien fest auf dem Boden von Menschenrechten und Grundfreiheiten, Werten und Prinzipien sowie moralischen und ethischen Überlegungen verankern,

Des Weiteren eingedenk dessen, dass die Generalkonferenz der UNESCO im November 2019 auf ihrer 40. Sitzung die Resolution 40 C/37 verabschiedete, mit der sie die Generaldirektorin beauftragte, „ein internationales normsetzendes Instrument zur Ethik der künstlichen Intelligenz (KI) in Form einer Empfehlung auszuarbeiten“, das der Generalkonferenz auf ihrer 41. Tagung im Jahr 2021 vorgelegt werden sollte,

Conscious of the fact that all countries are facing an acceleration in the use of information and communication technologies and AI technologies, as well as an increasing need for media and information literacy, and that the digital economy presents important societal, economic and environmental challenges and opportunities of benefitsharing, especially for low- and middle-income countries (LMICs), including but not limited to least developed countries (LDCs), landlocked developing countries (LLDCs) and small island developing States (SIDS), requiring the recognition, protection and promotion of endogenous cultures, values and knowledge in order to develop sustainable digital economies,

Further recognizing that AI technologies have the potential to be beneficial to the environment and ecosystems, and in order for those benefits to be realized, potential harms to and negative impacts on the environment and ecosystems should not be ignored but instead addressed,

Noting that addressing risks and ethical concerns should not hamper innovation and development but rather provide new opportunities and stimulate ethically-conducted research and innovation that anchor AI technologies in human rights and fundamental freedoms, values and principles, and moral and ethical reflection,

Also recalling that in November 2019, the General Conference of UNESCO, at its 40th session, adopted 40 C/ Resolution 37, by which it mandated the Director-General to prepare an international standard-setting instrument on the ethics of artificial intelligence (AI) in the form of a recommendation”, which is to be submitted to the General Conference at its 41st session in 2021,

In der Erkenntnis, dass die Entwicklung von KI-Technologien mit einer entsprechenden Zunahme der Daten-, Medien- und Informationskompetenz sowie Zugang zu unabhängigen, pluralistischen und vertrauenswürdigen Informationsquellen einhergehen muss, unter anderem weil dies Anstrengungen unterstützt, die Risiken von Falschinformation, Desinformation und Hetze einzudämmen und Schaden durch Missbrauch personenbezogener Daten zu mindern,

Feststellend, dass internationale und nationale rechtliche Regelwerke, die Menschenrechte und Grundfreiheiten, Ethik, die Notwendigkeit des Zugangs zu Daten, Informationen und Wissen, die Freiheit von Forschung und Innovation sowie das Wohl von Mensch, Umwelt und Ökosystemen die Basis für einen normativen Rahmen für KI-Technologien und deren gesellschaftliche Auswirkungen bilden, und dass dieser Rahmen auf der Grundlage eines gemeinsamen Verständnisses und gemeinsamer Ziele einen Zusammenhang zwischen ethischen Werten und Prinzipien und den mit KI-Technologien verbundenen Herausforderungen und Chancen herstellt,

Ebenfalls anerkennend, dass ethische Werte und Prinzipien dadurch, dass sie mit Blick auf die rasante technologische Entwicklung Orientierung bieten, bei der Entwicklung und Umsetzung rechtebasierter politischer Maßnahmen und Rechtsnormen helfen können,

Des Weiteren in der Überzeugung, dass weltweit akzeptierte ethische Standards für KI-Technologien, die vollständig im Einklang mit dem Völkerrecht, insbesondere den Menschenrechten stehen, eine Schlüsselrolle bei der Entwicklung KI-bezogener Normen auf der ganzen Welt spielen können,

Recognizing that the development of AI technologies necessitates a commensurate increase in data, media and information literacy as well as access to independent, pluralistic, trusted sources of information, including as part of efforts to mitigate risks of misinformation, disinformation and hate speech, and harm caused through the misuse of personal data,

Observing that a normative framework for AI technologies and its social implications finds its basis in international and national legal frameworks, human rights and fundamental freedoms, ethics, need for access to data, information and knowledge, the freedom of research and innovation, human and environmental and ecosystem well-being, and connects ethical values and principles to the challenges and opportunities linked to AI technologies, based on common understanding and shared aims,

Also recognizing that ethical values and principles can help develop and implement rights-based policy measures and legal norms, by providing guidance with a view to the fast pace of technological development,

Also convinced that globally accepted ethical standards for AI technologies, in full respect of international law, in particular human rights law, can play a key role in developing AI-related norms across the globe,

Eingedenk der Allgemeinen Erklärung der Menschenrechte (1948), der globalen Menschenrechtspakte einschließlich des Übereinkommens über die Rechtsstellung der Flüchtlinge (1951), des Übereinkommens über die Diskriminierung in Beschäftigung und Beruf (1958), des Internationalen Übereinkommens zur Beseitigung jeder Form von Rassendiskriminierung (1965), des Internationalen Pakts über bürgerliche und politische Rechte (1966), des Internationalen Pakts über wirtschaftliche, soziale und kulturelle Rechte (1966), des Übereinkommens zur Beseitigung jeder Form von Diskriminierung der Frau (1979), des Übereinkommens über die Rechte des Kindes (1989), des Übereinkommens über die Rechte von Menschen mit Behinderungen (2006), des Übereinkommens gegen Diskriminierung im Bildungswesen (1960), des Übereinkommens zum Schutz und zur Förderung der Vielfalt kultureller Ausdrucksformen (2005) sowie aller weiteren einschlägigen internationalen Übereinkünfte, Empfehlungen und Erklärungen,

Ferner unter Hinweis auf die Erklärung der Vereinten Nationen über das Recht auf Entwicklung (1986); die Erklärung über die Verantwortung der

heutigen Generation gegenüber den künftigen Generationen (1997); die Allgemeine Erklärung über Bioethik und Menschenrechte (2005); die Erklärung der Vereinten Nationen über die Rechte indigener Völker (2007); die Resolution der Generalversammlung der Vereinten Nationen zur Umsetzungsprüfung der Ergebnisse des Weltgipfels über die Informationsgesellschaft (A/RES/70/125) (2015); die Resolution der Generalversammlung der Vereinten Nationen zur Transformation unserer Welt: die Agenda 2030 für nachhaltige Entwicklung (A/RES/70/1) (2015); die Empfehlung über die Bewahrung des dokumentarischen Erbes einschließlich des digitalen Erbes und den Zugang zu diesem Erbe (2015); die Erklärung zu ethischen Grundsätzen in Bezug auf den Klimawandel (2017); die Empfehlung über Wissenschaft, Wissenschaftlerinnen und Wissenschaftler (2017); die (2018 vom Internationalen Programm der UNESCO gebilligten) Internet-Universalitätsindikatoren einschließlich der (2015 von der Generalkonferenz der UNESCO gebilligten) R-O-A-M-Prinzipien; die Resolution des Menschenrechtsrats zum „Recht auf Privatsphäre im digitalen Zeitalter“ (A/HRC/RES/-

Bearing in mind the Universal Declaration of Human Rights (1948), the instruments of the international human rights framework, including the Convention Relating to the Status of Refugees (1951), the Discrimination (Employment and Occupation) Convention (1958), the International Convention on the Elimination of All Forms of Racial Discrimination (1965), the International Covenant on Civil and Political Rights (1966), the International Covenant on Economic, Social and Cultural Rights (1966), the Convention on the Elimination of All Forms of Discrimination against Women (1979), the Convention on the Rights of the Child (1989), and the Convention on the Rights of Persons with Disabilities (2006), the Convention against Discrimination in Education (1960), the Convention on the Protection and Promotion of the Diversity of Cultural Expressions (2005), as well as any other relevant international instruments, recommendations and declarations,

Also noting the United Nations Declaration on the Right to Development (1986); the Declaration on the Responsibilities of the Present Generations Towards Future Generations (1997); the Universal Declaration on Bioethics and Human Rights

(2005); the United Nations Declaration on the Rights of Indigenous Peoples (2007); the United Nations General Assembly resolution on the review of the World Summit on the Information Society (A/RES/70/125) (2015); the United Nations General Assembly Resolution on Transforming our world: the 2030 Agenda for Sustainable Development (A/RES/70/1) (2015); the Recommendation Concerning the Preservation of, and Access to, Documentary Heritage Including in Digital Form (2015); the Declaration of Ethical Principles in relation to Climate Change (2017); the Recommendation on Science and Scientific Researchers (2017); the Internet Universality Indicators (endorsed by UNESCO's International Programme for the Development of Communication in 2018), including the ROAM principles (endorsed by UNESCO's General Conference in 2015); the Human Rights Council's resolution on "The right to privacy in the digital age" (A/HRC/RES/42/15) (2019); and the Human Rights Council's resolution on "New and emerging digital technologies and human rights" (A/HRC/RES/41/11) (2019),

42/15) (2019) und die Resolution des Menschenrechtsrats zu „Neuen und aufkommenden digitalen Technologien und Menschenrechten“(A/HRC/RES/41/11) (2019),

Betonend, dass ein besonderes Augenmerk auf LMICs, unter anderem LDCs, LLDCs und SIDS gelegt werden muss, da diese zwar über eigene Kapazitäten verfügen, in der Debatte über die Ethik der KI bislang jedoch unterrepräsentiert sind, was wegen der Vernachlässigung lokalen Wissens, eines kulturellen Pluralismus, von Wertesystemen und des Anspruchs der globalen Fairness im Umgang mit den positiven und negativen Auswirkungen von KI-Technologien bedenklich ist,

Ferner in dem Bewusstsein der Vielzahl bereits existierender nationaler Leitlinien und sonstiger Rahmenwerke und Initiativen, die von relevanten Einrichtungen der Vereinten Nationen, zwischenstaatlichen Organisationen einschließlich regionaler Organisationen, sowie von der Privatwirtschaft, von Berufsorganisationen, nichtstaatlichen Organisationen und der Wissenschaft im Zusammenhang mit der Ethik und der Regulierung von KI-Technologien ausgearbeitet worden sind,

Des Weiteren in der Überzeugung, dass KI-Technologien großen Nutzen mit sich bringen können, dass die Verwirklichung dieser Vorteile jedoch auch zur Verstärkung von Spannungen im Zusammenhang mit Innovation, asymmetrischem Zugang zu Wissen und Technologien einschließlich fehlender digitaler und Alltags-Fähigkeiten führen kann, wodurch die Möglichkeiten der Öffentlichkeit zur aktiven Auseinandersetzung mit KI-bezogenen Themen eingeschränkt werden; weitere Folgen können höhere Barrieren beim Zugang zu Informationen, unzureichende menschliche und institutionelle Kapazitäten, Hindernisse beim Zugang zu technologischen Innovationen und ein Mangel an angemessener physischer und digitaler Infrastruktur sowie an rechtlichen Regelwerken sein, auch im Zusammenhang mit Daten, und überzeugt, dass all diese Probleme angegangen werden sollten,

Emphasizing that specific attention must be paid to LMICs, including but not limited to LDCs, LLDCs and SIDS, as they have their own capacity but have been underrepresented in the AI ethics debate, which raises concerns about neglecting local knowledge, cultural pluralism, value systems and the demands of global fairness to deal with the positive and negative impacts of AI technologies,

Also conscious of the many existing national policies, other frameworks and initiatives elaborated by relevant United Nations entities, intergovernmental organizations, including regional organizations, as well as those by the private sector, professional organizations, non-governmental organizations, and the scientific community, related to the ethics and regulation of AI technologies,

Further convinced that AI technologies can bring important benefits, but that achieving them can also amplify tension around innovation, asymmetric access to knowledge and technologies, including the digital and civic literacy deficit that limits the public's ability to engage in topics related to AI, as well as barriers to access to information and gaps in capacity, human and institutional capacities, barriers to access to technological innovation, and a lack of adequate physical and digital infrastructure and regulatory frameworks, including those related to data, all of which need to be addressed,

Hervorhebend, dass die Stärkung der globalen Zusammenarbeit und Solidarität, auch durch den Multilateralismus, notwendig ist, um den fairen Zugang zu KI-Technologien zu ermöglichen und die Herausforderungen zu bewältigen, die diese in Bezug auf die Vielfalt und die Verbundenheit von Kulturen und ethischen Systemen mit sich bringen, und um potenziellen Missbrauch zu mindern, das volle Potenzial der KI auszuschöpfen, insbesondere im Bereich der Entwicklung, und um sicherzustellen, dass nationale KI-Strategien von ethischen Grundsätzen geleitet sind,

Unter umfassender Berücksichtigung der Tatsache, dass die rasante Entwicklung von KI-Technologien eine Herausforderung für ihre ethische Umsetzung, Lenkung und Kontrolle sowie für die Achtung und den Schutz der kulturellen Vielfalt darstellt und ein disruptives Potenzial im Hinblick auf lokale und regionale ethische Normen und Werte birgt,

1. **Nimmt** die vorliegende Empfehlung zur Ethik der künstlichen Intelligenz am dreiundzwanzigsten November 2021 an;
2. **Dabei empfiehlt sie** den Mitgliedstaaten die Anwendung dieser Empfehlung auf frei-

williger Basis durch geeignete Schritte einschließlich aller notwendigen gesetzgeberischen bzw. sonstigen erforderlichen Maßnahmen in Übereinstimmung mit der verfassungsrechtlichen Praxis und den Organisationsstrukturen eines jeden Staates, um den Prinzipien und Normen der Empfehlung innerhalb ihres Hoheitsgebiets im Einklang mit dem Völkerrecht einschließlich der internationalen Menschenrechtsnormen Wirkung zu verleihen;

3. **Zudem empfiehlt sie** den Mitgliedstaaten die Einbeziehung sämtlicher Interessengruppen einschließlich Wirtschaftsunternehmen, um sicherzustellen, dass sie ihren jeweiligen Rollen bei der Umsetzung dieser Empfehlung gerecht werden; außerdem sollten sie die Empfehlung ins Bewusstsein sämtlicher Behörden, Gremien, Forschungseinrichtungen und akademischen Institutionen sowie der mit KI-Technologien befassten Einrichtungen und Organisationen im öffentlichen, privaten und zivilgesellschaftlichen Bereich rücken, damit die Entwicklung und Nutzung von KI-Technologien sowohl von solider wissenschaftlicher Forschung als auch von ethischer Analyse und Bewertung geleitet ist.

Underlining that the strengthening of global cooperation and solidarity, including through multilateralism, is needed to facilitate fair access to AI technologies and address the challenges that they bring to diversity and interconnectivity of cultures and ethical systems, to mitigate potential misuse, to realize the full potential that AI can bring, especially in the area of development, and to ensure that national AI strategies are guided by ethical principles,

Taking fully into account that the rapid development of AI technologies challenges their ethical implementation and governance, as well as the respect for and protection of cultural diversity, and has the potential to disrupt local and regional ethical standards and values,

1. **Adopts** the present Recommendation on the Ethics of Artificial Intelligence on this twenty-third day of November 2021;
2. **Recommends** that Member States apply on a voluntary basis the provisions of this

Recommendation by taking appropriate steps, including whatever legislative or other measures may be required, in conformity with the constitutional practice and governing structures of each State, to give effect within their jurisdictions to the principles and norms of the Recommendation in conformity with international law, including international human rights law;

3. **Also recommends** that Member States engage all stakeholders, including business enterprises, to ensure that they play their respective roles in the implementation of this Recommendation; and bring the Recommendation to the attention of the authorities, bodies, research and academic organizations, institutions and organizations in public, private and civil society sectors involved in AI technologies, so that the development and use of AI technologies are guided by both sound scientific research as well as ethical analysis and evaluation.

I. Anwendungsbereich



(1) Diese Empfehlung beschäftigt sich mit ethischen Fragen im Zusammenhang mit Künstlicher Intelligenz, soweit sie in das Mandat der UNESCO fallen. Dabei versteht sie Ethik der KI als eine systematische normative Reflexion, basierend auf einem holistischen, umfassenden, multikulturellen und sich weiterentwickelnden Gerüst wechselseitig voneinander abhängiger Werte, Prinzipien und Handlungen. Dieses soll Gesellschaften beim verantwortungsvollen Umgang mit den bekannten und den noch nicht bekannten Auswirkungen von KI-Technologien auf Menschen, Gesellschaften, Umwelt und Ökosysteme Orientierung geben und ihnen eine Grundlage für die Entscheidung bieten, eine KI-Technologie anzunehmen oder abzulehnen. Ethik wird hier verstanden als dynamische Grundlage für die normative Bewertung und Lenkung von KI-Technologien, als ein auf Menschenwürde, Wohlergehen und Schadensverhütung bezogener Kompass, der in der Ethik der Wissenschaft und Technik wurzelt.

I. Scope of application



(1) This Recommendation addresses ethical issues related to the domain of Artificial Intelligence to the extent that they are within UNESCO's mandate. It approaches AI ethics as a systematic normative reflection, based on a holistic, comprehensive, multi-cultural and evolving framework of interdependent values, principles and actions that can guide societies in dealing responsibly with the known and unknown impacts of AI technologies on human beings, societies and the environment and ecosystems, and offers them a basis to accept or reject AI technologies. It considers ethics as a dynamic basis for the normative evaluation and guidance of AI technologies, referring to human dignity, well-being and the prevention of harm as a compass and as rooted in the ethics of science and technology.

(2) Diese Empfehlung bietet keine einzelne Definition von KI, denn eine solche Definition müsste im Laufe der Zeit zwangsläufig an die technologischen Entwicklungen angepasst werden. Es geht vielmehr darum, diejenigen Merkmale von KI-Systemen in den Blick zu nehmen, die von zentraler ethischer Bedeutung sind. Als KI-Systeme im Sinne dieser Empfehlung gelten daher Systeme, die in der Lage sind, Daten und Informationen in einer Weise zu verarbeiten, die intelligentem Verhalten ähnelt, und die typischerweise Aspekte von logischem Schlussfolgern, Lernen, Wahrnehmung, Vorhersage, Planung oder Kontrolle umfassen. Bei diesem Ansatz stehen drei Elemente im Mittelpunkt:

- a) KI-Systeme sind informationsverarbeitende Technologien, die Modelle und Algorithmen solchermaßen zusammenführen, dass sie die Fähigkeit entwickeln zu lernen und kognitive Aufgaben durchzuführen, und die im Ergebnis in materiellen und virtuellen Umgebungen Vorhersagen machen und Entscheidungen treffen können. KI-Systeme sind so konzipiert, dass sie durch Wissensmodellierung und -darstellung sowie durch Nutzung von Daten und Berechnung von Korrelationen mit unterschiedlichen Graden an Autonomie arbeiten. KI-Systeme können verschiedene Methoden einbeziehen, beispielsweise die folgenden, wobei diese Aufzählung nicht erschöpfend ist:
 - i) Maschinelles Lernen, einschließlich Deep Learning („mehrschichtiges Lernen“) und bestärkenden Lernens;
 - ii) Maschinelles Denken („machine reasoning“), einschließlich Planung, Ablaufsteuerung, Wissensdarstellung, logischen Schlussfolgerns, Suche und Optimierung.

(2) This Recommendation does not have the ambition to provide one single definition of AI, since such a definition would need to change over time, in accordance with technological developments. Rather, its ambition is to address those features of AI systems that are of central ethical relevance. Therefore, this Recommendation approaches AI systems as systems which have the capacity to process data and information in a way that resembles intelligent behaviour, and typically includes aspects of reasoning, learning, perception, prediction, planning or control. Three elements have a central place in this approach:

- a) AI systems are information-processing technologies that integrate models and algorithms that produce a capacity to learn and to perform cognitive tasks leading to outcomes such as prediction and decision-making in material and virtual environments. AI systems are designed to operate with varying degrees of autonomy by means of knowledge modelling and representation and by exploiting data and calculating correlations. AI systems may include several methods, such as but not limited to:
 - i) machine learning, including deep learning and reinforcement learning;
 - ii) machine reasoning, including planning, scheduling, knowledge representation and reasoning, search, and optimization.

KI-Systeme können in cyberphysischen Systemen eingesetzt werden, wie dem Internet der Dinge, Robotersystemen, sozialer Robotik und Mensch-Computer-Schnittstellen, die Steuerungs- und Wahrnehmungsprozesse, die Verarbeitung von durch Sensoren erfassten Daten sowie den Betrieb von Aktoren in der Einsatzumgebung von KI-Systemen umfassen.

- b) Ethische Fragen im Zusammenhang mit KI-Systemen betreffen alle Phasen ihres Lebenszyklus, d.h. angefangen bei Forschung, Design und Entwicklung bis hin zu Einsatz und Nutzung, einschließlich Wartung, Betrieb, Handel, Finanzierung, Überwachung und Bewertung, Validierung, Einstellung der Nutzung, Demontage und Außerbetriebnahme. Der Begriff „KI-Akteure“ schließt per Definition all die ein, die in mindestens eine Phase im Lebenszyklus von KI-Systemen einbezogen sind; dabei kann es sich sowohl um natürliche als auch um juristische Personen handeln, wie zum Beispiel Forschende, Programmiererinnen und Programmierer, Ingenieurinnen und Ingenieure, Datenwissenschaftlerinnen und -wissenschaftler, Endnutzerinnen und -nutzer, Wirtschaftsunternehmen, Universitäten sowie öffentliche und private Einrichtungen, um nur einige zu nennen.

AI systems can be used in cyber-physical systems, including the Internet of things, robotic systems, social robotics, and human-computer interfaces, which involve control, perception, the processing of data collected by sensors, and the operation of actuators in the environment in which AI systems work.

- b) Ethical questions regarding AI systems pertain to all stages of the AI system life cycle, understood here to range from research, design and development to deployment and use, including maintenance, operation, trade, financing, monitoring and evaluation, validation, end-of-use, disassembly and termination. In addition, AI actors can be defined as any actor involved in at least one stage of the AI system life cycle, and can refer both to natural and legal persons, such as researchers, programmers, engineers, data scientists, end-users, business enterprises, universities and public and private entities, among others.

c) KI-Systeme werfen neuartige ethische Fragen auf, die sich unter anderem auf deren Auswirkungen auf folgende Bereiche beziehen: Entscheidungsfindung, Beschäftigung und Arbeitswelt, soziale Interaktion, Gesundheitsversorgung, Bildung, Medien, Zugang zu Informationen, digitale Kluft, Schutz personenbezogener Daten und Verbraucherschutz, Umwelt, Demokratie, Rechtsstaatlichkeit, Sicherheit und Polizeiarbeit, Dual-Use (im Sinne einer möglichen zivilen und militärischen Nutzung von KI) sowie Menschenrechte und Grundfreiheiten einschließlich der Meinungsfreiheit, des Schutzes der Privatsphäre und der Nichtdiskriminierung. Darüber hinaus erwachsen neue ethische Herausforderungen aus dem Potenzial von KI-Algorithmen, bereits existierende Verzerrungen zu reproduzieren und zu verstärken und auf diese Weise bestehende Formen von Diskriminierung, Vorurteilen und Stereotypisierung zu verschärfen. Einige dieser Probleme folgen aus der Fähigkeit von KI-Systemen, Aufgaben auszuführen, die bisher nur von Lebewesen ausgeführt werden konnten und in einigen Fällen sogar ausschließlich Menschen vorbehalten waren. Deshalb haben KI-Systeme eine einschneidend neue Rolle in menschlichen Praktiken und in der Gesellschaft, ebenso in deren Beziehung zur Umwelt und zu Ökosystemen, und sie schaffen einen neuen Kontext, in dem Kinder und junge Menschen heranwachsen, ihr Verständnis von der Welt und sich selbst entwickeln, Medien und Informationen kritisch verstehen und lernen, Entscheidungen zu treffen. Langfristig könnten KI-Systeme die Besonderheit menschlicher Erlebnis- und Handlungsfähigkeit infrage stellen und damit weitergehende Fragen aufwerfen, u. a. in Bezug auf das menschliche Selbstverständnis, auf soziale, kulturelle und ökologische Interaktion, auf Autonomie und Handlungsfähigkeit, auf Wert und Würde des Menschen.

c) AI systems raise new types of ethical issues that include, but are not limited to, their impact on decision-making, employment and labour, social interaction, health care, education, media, access to information, digital divide, personal data and consumer protection, environment, democracy, rule of law, security and policing, dual use, and human rights and fundamental freedoms, including freedom of expression, privacy and nondiscrimination. Furthermore, new ethical challenges are created by the potential of AI algorithms to reproduce and reinforce existing biases, and thus to exacerbate already existing forms of discrimination, prejudice and stereotyping. Some of these issues are related to the capacity of AI systems to perform tasks which previously only living beings could do, and which were in some cases even limited to human beings only. These characteristics give AI systems a profound, new role in human practices and society, as well as in their relationship with the environment and ecosystems, creating a new context for children and young people to grow up in, develop an understanding of the world and themselves, critically understand media and information, and learn to make decisions. In the long term, AI systems could challenge humans' special sense of experience and agency, raising additional concerns about, inter alia, human selfunderstanding, social, cultural and environmental interaction, autonomy, agency, worth and dignity.

(3) In dieser Empfehlung wird besonderes Augenmerk auf die breiteren ethischen Implikationen von KI-Systemen im Hinblick auf die Mandatsfelder der UNESCO gelegt, d.h. Bildung, Wissenschaft, Kultur sowie Kommunikation und Information, wie sie in der Vorstudie 2019 zur Ethik der künstlichen Intelligenz der UNESCO-Weltkommission für Wissenschafts- und Technologieethik (COMEST) herausgearbeitet wurden:

- a) Bildung: Denn das Leben in sich digitalisierenden Gesellschaften erfordert angesichts von Arbeitsmarktfolgen, Beschäftigungsfähigkeit und bürgerchaftlicher Beteiligung neue Wege im Bildungswesen, ethische Reflexion, kritisches Denken, verantwortungsvolle Gestaltungspraktiken und neue Kompetenzen;
- b) Wissenschaft im weitesten Sinne und mit allen Disziplinen von den Naturwissenschaften über die Medizin bis hin zu den Sozial- und Geisteswissenschaften: Denn KI-Technologien führen zu neuen Forschungsmöglichkeiten und -ansätzen, verändern unsere Vorstellungen von wissenschaftlichem Verständnis und wissenschaftlicher Erklärung und schaffen neue Grundlagen für die Entscheidungsfindung;
- c) Kulturelle Identität und Vielfalt: Denn KI-Technologien können zwar die Kultur- und Kreativwirtschaft bereichern, aber auch zu zunehmender Konzentration des Angebots an kulturellen Inhalten, Daten, Märkten und Einkommen in den Händen einiger weniger Akteure führen, was sich potenziell negativ auf die Vielfalt und den Pluralismus von Sprachen, Medien, kulturellen Ausdrucksformen, Teilhabe und Gleichberechtigung auswirkt;

(3) This Recommendation pays specific attention to the broader ethical implications of AI systems in relation to the central domains of UNESCO: education, science, culture, and communication and information, as explored in the 2019 Preliminary Study on the Ethics of Artificial Intelligence by the UNESCO World Commission on Ethics of Scientific Knowledge and Technology (COMEST):

- a) Education, because living in digitalizing societies requires new educational practices, ethical reflection, critical thinking, responsible design practices and new skills, given the implications for the labour market, employability and civic participation.
- b) Science, in the broadest sense and including all academic fields from the natural sciences and medical sciences to the social sciences and humanities, as AI technologies bring new research capacities and approaches, have implications for our concepts of scientific understanding and explanation, and create a new basis for decisionmaking.
- c) Cultural identity and diversity, as AI technologies can enrich cultural and creative industries, but can also lead to an increased concentration of supply of cultural content, data, markets and income in the hands of only a few actors, with potential negative implications for the diversity and pluralism of languages, media, cultural expressions, participation and equality.

d) Kommunikation und Information: Denn KI-Technologien spielen eine immer wichtigere Rolle bei der Verarbeitung, Strukturierung und Bereitstellung von Informationen; automatisierter Journalismus und die Algorithmen-basierte Bereitstellung von Nachrichten, Moderation und Kuratierung von Inhalten in sozialen Medien und Suchmaschinen sind nur einige Beispiele, bei denen sich Fragen aufdrängen, unter anderem im Zusammenhang mit dem Zugang zu Informationen, Desinformation, Misinformation, Hass und Hetze, der Entstehung neuer Arten gesellschaftlicher Narrative, Diskriminierung, Meinungsfreiheit, Privatsphäre sowie Medien- und Informationskompetenz.

(4) Diese Empfehlung richtet sich an die Mitgliedstaaten, sowohl in ihrer Eigenschaft als KI-Akteure als auch in ihrer organisatorischen Verantwortlichkeit, rechtliche und regulatorische Rahmenbedingungen für den gesamten Lebenszyklus von KI-Systemen zu entwickeln und unternehmerische Verantwortung zu fördern. Darüber hinaus bietet sie allen KI-Akteuren sowohl aus dem öffentlichen als auch aus dem privaten Bereich eine ethische Orientierung, indem sie als Grundlage für eine ethische Folgenabschätzung entlang des gesamten Lebenszyklus von KI-Systemen dient.

d) Communication and information, as AI technologies play an increasingly important role in the processing, structuring and provision of information; the issues of automated journalism and the algorithmic provision of news and moderation and curation of content on social media and search engines are just a few examples raising issues related to access to information, disinformation, misinformation, hate speech, the emergence of new forms of societal narratives, discrimination, freedom of expression, privacy and media and information literacy, among others.

(4) This Recommendation is addressed to Member States, both as AI actors and as authorities responsible for developing legal and regulatory frameworks throughout the entire AI system life cycle, and for promoting business responsibility. It also provides ethical guidance to all AI actors, including the public and private sectors, by providing a basis for an ethical impact assessment of AI systems throughout their life cycle.

II. Ziele

(5) Diese Empfehlung bietet Grundlagen an, um KI-Systeme zum Wohle der Menschheit, des einzelnen Menschen, der Gesellschaften sowie der Umwelt und der Ökosysteme einzusetzen und Schaden vorzubeugen. Sie zielt außerdem darauf ab, die friedliche Nutzung von KI-Systemen zu fördern.

(6) Ergänzend zu den bereits vielfach in unserer Welt existierenden ethischen Rahmenwerken für KI wurde diese Empfehlung als Text mit weltweit anerkannten Normen geschaffen, der nicht nur Werte und Prinzipien formuliert, sondern durch konkrete Handlungsempfehlungen auch deren praktische Umsetzung in den Blick nimmt und dabei ein besonderes Augenmerk auf Inklusion und Geschlechtergerechtigkeit sowie den Schutz von Umwelt und Ökosystemen legt.

(7) Die Komplexität der ethischen Fragestellungen rund um KI erfordert die Zusammenarbeit einer Vielzahl von Interessengruppen quer zu verschiedenen Ebenen und Sektoren der Gesellschaft, auf weltweiter, regionaler (kontinentaler) und nationaler Ebene; daher soll die vorliegende Empfehlung diese Interessengruppen dazu befähigen, auf der Grundlage eines globalen und interkulturellen Dialogs gemeinsam Verantwortung zu übernehmen.

II. Aims and objectives

(5) This Recommendation aims to provide a basis to make AI systems work for the good of humanity, individuals, societies and the environment and ecosystems, and to prevent harm. It also aims at stimulating the peaceful use of AI systems.

(6) In addition to the existing ethical frameworks regarding AI around the world, this Recommendation aims to bring a globally accepted normative instrument that focuses not only on the articulation of values and principles, but also on their practical realization, via concrete policy recommendations, with a strong emphasis on inclusion issues of gender equality and protection of the environment and ecosystems.

(7) Because the complexity of the ethical issues surrounding AI necessitates the cooperation of multiple stakeholders across the various levels and sectors of international, regional and national communities, this Recommendation aims to enable stakeholders to take shared responsibility based on a global and intercultural dialogue.

(8) Die Ziele dieser Empfehlung sind:

- a) Die Etablierung eines allgemeinen Gerüsts von Werten, Prinzipien und Maßnahmen, das den Staaten bei der Gestaltung völkerrechtskonformer Gesetze, Strategien und sonstiger Instrumente zur KI als Orientierung dient;
- b) Die Bereitstellung von Leitplanken für das Handeln von Einzelpersonen, Gruppen, Gemeinschaften, Institutionen und privatwirtschaftlichen Unternehmen, um die Berücksichtigung ethischer Aspekte in allen Phasen des Lebenszyklus eines KI-Systems zu gewährleisten;
- c) Der Schutz, die Förderung und Achtung der Menschenrechte und Grundfreiheiten, der menschlichen Würde und Gleichberechtigung, einschließlich der Geschlechtergerechtigkeit, Wahrung der Interessen der heutigen und künftiger Generationen, Erhalt der Umwelt, der biologischen Vielfalt und der Ökosysteme sowie Achtung der kulturellen Vielfalt in allen Phasen des Lebenszyklus von KI-Systemen;
- d) Die Förderung eines multidisziplinären und pluralistischen Multistakeholder-Dialogs und einer entsprechenden Konsensbildung über ethische Fragen im Zusammenhang mit KI-Systemen;
- e) Die Förderung des gleichberechtigten Zugangs zu KI-Entwicklungen und -Wissen sowie einer gerechten Aufteilung des Nutzens, wobei Bedürfnisse und Beiträge von LMICs, einschließlich LDCs, LLDCs und SIDS besondere Beachtung finden.

(8) The objectives of this Recommendation are:

- a) to provide a universal framework of values, principles and actions to guide States in the formulation of their legislation, policies or other instruments regarding AI, consistent with international law;
- b) to guide the actions of individuals, groups, communities, institutions and private sector companies to ensure the embedding of ethics in all stages of the AI system life cycle;
- c) to protect, promote and respect human rights and fundamental freedoms, human dignity and equality, including gender equality; to safeguard the interests of present and future generations; to preserve the environment, biodiversity and ecosystems; and to respect cultural diversity in all stages of the AI system life cycle;
- d) to foster multi-stakeholder, multidisciplinary and pluralistic dialogue and consensus building about ethical issues relating to AI systems;
- e) to promote equitable access to developments and knowledge in the field of AI and the sharing of benefits, with particular attention to the needs and contributions of LMICs, including LDCs, LLDCs and SIDS.

III.

Werte und Prinzipien

(9) Die im Folgenden dargelegten Werte und Prinzipien sollen von allen Akteuren im Lebenszyklus eines KI-Systems beachtet und, wo dies erforderlich und angemessen erscheint, durch Änderung bestehender und Ausarbeitung neuer Gesetze, Vorschriften und Unternehmensleitlinien gefördert werden. Dies muss im Einklang mit dem Völkerrecht geschehen, einschließlich der Charta der Vereinten Nationen und den menschenrechtlichen Verpflichtungen der Mitgliedstaaten, sowie mit international vereinbarten sozialen, politischen, ökologischen, bildungspolitischen, wissenschaftlichen und wirtschaftlichen Nachhaltigkeitszielen wie den Zielen der Vereinten Nationen für nachhaltige Entwicklung („Sustainable Development Goals“ – SDGs).

(10) Bei der Gestaltung politischer Maßnahmen und rechtlicher Normen spielen Werte als motivierende Ideale eine bedeutende Rolle. So regt der nachstehend skizzierte Wertekanon zu wünschenswertem Verhalten an und bildet das Fundament für die Prinzipien; die Prinzipien wiederum lassen die Werte, die ihnen zugrunde liegen, deutlicher hervortreten, so dass die Werte leichter in politische Strategien und konkrete Maßnahmen umgesetzt werden können.

III.

Values and principles

(9) The values and principles included below should be respected by all actors in the AI system life cycle, in the first place and, where needed and appropriate, be promoted through amendments to the existing and elaboration of new legislation, regulations and business guidelines. This must comply with international law, including the United Nations Charter and Member States' human rights obligations, and should be in line with internationally agreed social, political, environmental, educational, scientific and economic sustainability objectives, such as the United Nations Sustainable Development Goals (SDGs).

(10) Values play a powerful role as motivating ideals in shaping policy measures and legal norms. While the set of values outlined below thus inspires desirable behaviour and represents the foundations of principles, the principles unpack the values underlying them more concretely so that the values can be more easily operationalized in policy statements and actions.

(11) Wenngleich alle der unten dargelegten Werte und Prinzipien per se erstrebenswert sind, so können diese Werte und Prinzipien in bestimmten praktischen Kontexten durchaus in einem Spannungsverhältnis zueinander stehen. Ein angemessener Umgang mit solchen Spannungsverhältnissen erfordert die Berücksichtigung des Grundsatzes der Verhältnismäßigkeit, die Beachtung der Menschenrechte und Grundfreiheiten und immer eine kontextbezogene Bewertung der jeweiligen Situation. Eine mögliche Einschränkung von Menschenrechten und Grundfreiheiten bedarf in jedem Fall einer rechtlichen Grundlage und muss begründet, notwendig und verhältnismäßig sein und im Einklang mit den völkerrechtlichen Verpflichtungen der Staaten stehen. Eine kluge Handhabung derartiger Szenarien erfordert in aller Regel die Einbindung eines breiten Spektrums geeigneter Interessengruppen, unter Einsatz eines gesellschaftlichen Dialogs, ethischer Überlegungen, Risikoprüfungen und Folgenabschätzungen.

(11) While all the values and principles outlined below are desirable per se, in any practical contexts, there may be tensions between these values and principles. In any given situation, a contextual assessment will be necessary to manage potential tensions, taking into account the principle of proportionality and in compliance with human rights and fundamental freedoms. In all cases, any possible limitations on human rights and fundamental freedoms must have a lawful basis, and be reasonable, necessary and proportionate, and consistent with States' obligations under international law. To navigate such scenarios judiciously will typically require engagement with a broad range of appropriate stakeholders, making use of social dialogue, as well as ethical deliberation, due diligence and impact assessment.

(12) Die Vertrauenswürdigkeit und Integrität des Lebenszyklus von KI-Systemen sind entscheidend dafür, dass KI-Technologien tatsächlich zum Wohle der Menschheit, sowohl des Einzelnen als auch der Gesellschaften, sowie zum Wohle der Umwelt und der Ökosysteme eingesetzt werden und die in dieser Empfehlung dargelegten Werte und Prinzipien verkörpern. Die Gesellschaft braucht gute Gründe für das Vertrauen, dass KI-Systeme dem Einzelnen und der Allgemeinheit Vorteile bringen können, während gleichzeitig angemessene Maßnahmen zur Risikominderung getroffen werden. Eine wesentliche Voraussetzung für die Vertrauenswürdigkeit von KI-Systemen ist, dass sie, soweit angemessen, über ihren gesamten Lebenszyklus einer gründlichen Überwachung durch die jeweiligen Interessengruppen unterzogen werden. Da Vertrauenswürdigkeit durch praktische Umsetzung der Prinzipien aus diesem Dokument entsteht, sind sämtliche der in dieser Empfehlung vorgeschlagenen politischen Maßnahmen auf die Förderung der Vertrauenswürdigkeit in allen Phasen des Lebenszyklus eines KI-Systems ausgerichtet.

(12) The trustworthiness and integrity of the life cycle of AI systems is essential to ensure that AI technologies will work for the good of humanity, individuals, societies and the environment and ecosystems, and embody the values and principles set out in this Recommendation. People should have good reason to trust that AI systems can bring individual and shared benefits, while adequate measures are taken to mitigate risks. An essential requirement for trustworthiness is that, throughout their life cycle, AI systems are subject to thorough monitoring by the relevant stakeholders as appropriate. As trustworthiness is an outcome of the operationalization of the principles in this document, the policy actions proposed in this Recommendation are all directed at promoting trustworthiness in all stages of the AI system life cycle.

III.1 Werte

Achtung, Schutz und Förderung der Menschenrechte, der Grundfreiheiten und der Menschenwürde

(13) Die unantastbare und jedem Menschen innewohnende Würde bildet die Grundlage für das System der universellen, unteilbaren, unveräußerlichen und wechselseitig voneinander abhängigen Menschenrechte und Grundfreiheiten. Die Menschenwürde und -rechte, wie sie im Völkerrecht einschließlich der internationalen Menschenrechtsabkommen verankert sind, zu achten, zu schützen und zu fördern ist daher über den gesamten Lebenszyklus von KI-Systemen von wesentlicher Bedeutung. Die Menschenwürde bezieht sich auf die Anerkennung des gleichen Eigenwerts eines jeden Menschen und die Gleichwertigkeit aller Menschen, ungeachtet ihrer ethnischen Herkunft und Hautfarbe, ihrer Abstammung, ihres Geschlechts, ihres Alters, ihrer Sprache, ihrer Religion und ihrer politischen Überzeugung, ihrer nationalen oder sozialen Herkunft, ihrem wirtschaftlichen oder sozialen Geburtsstatus, einer Behinderung oder sonstiger Merkmale.

(14) Kein Mensch und keine menschliche Gemeinschaft darf in irgendeiner Phase des Lebenszyklus eines KI-Systems geschädigt oder untergeordnet werden, weder in physischer noch in wirtschaftlicher, sozialer, politischer, kultureller oder geistiger Hinsicht. Die Qualität menschlichen Lebens sollte entlang des gesamten Lebenszyklus eines KI-Systems verbessert werden, wobei es einzelnen Personen oder Gruppen überlassen bleiben sollte, „Lebensqualität“ zu definieren, solange es im Rahmen dieser Definition nicht zu einer Verletzung oder einem Missbrauch von Menschenrechten und Grundfreiheiten oder der menschlichen Würde kommt.

III.1 Values

Respect, protection and promotion of human rights and fundamental freedoms and human dignity

(13) The inviolable and inherent dignity of every human constitutes the foundation for the universal, indivisible, inalienable, interdependent and interrelated system of human rights and fundamental freedoms. Therefore, respect, protection and promotion of human dignity and rights as established by international law, including international human rights law, is essential throughout the life cycle of AI systems. Human dignity relates to the recognition of the intrinsic and equal worth of each individual human being, regardless of race, colour, descent, gender, age, language, religion, political opinion, national origin, ethnic origin, social origin, economic or social condition of birth, or disability and any other grounds.

(14) No human being or human community should be harmed or subordinated, whether physically, economically, socially, politically, culturally or mentally during any phase of the life cycle of AI systems. Throughout the life cycle of AI systems, the quality of life of human beings should be enhanced, while the definition of “quality of life” should be left open to individuals or groups, as long as there is no violation or abuse of human rights and fundamental freedoms, or the dignity of humans in terms of this definition.

(15) Menschen können über den gesamten Lebenszyklus von KI-Systemen mit diesen interagieren und von ihnen Unterstützung erhalten, zum Beispiel bei der Betreuung schutzbedürftiger Menschen bzw. Menschen, die sich in gefährdeten Umständen befinden, unter anderem Kinder, ältere Menschen, Menschen mit Behinderungen oder Kranke. Im Rahmen derartiger Interaktionen dürfen Menschen niemals zu Objekten degradiert werden, noch darf ihre Würde auf andere Weise untergraben oder ihre Menschenrechte und Grundfreiheiten verletzt oder missbraucht werden.

(16) Die Menschenrechte und Grundfreiheiten müssen während des gesamten Lebenszyklus von KI-Systemen geachtet, geschützt und gefördert werden. Regierungen, der privatwirtschaftliche Sektor, die Zivilgesellschaft, internationale Organisationen, die technische Fachwelt und die Wissenschaft müssen bei ihren Eingriffen in die Prozesse rund um den Lebenszyklus von KI-Systemen die Menschenrechtsinstrumente und -rahmen beachten. Neue Technologien müssen neue Möglichkeiten zur Förderung der Menschenrechte und zu deren Verteidigung und Verwirklichung bieten und dürfen nicht zu deren Verletzung beitragen.

Günstige Entwicklung von Umwelt und Ökosystemen

(17) Das Gedeihen von Umwelt und Ökosystem sollte während des gesamten Lebenszyklus von KI-Systemen anerkannt, geschützt und gefördert werden. Die Umwelt und die Ökosysteme sind nämlich die existenzielle Voraussetzung dafür, dass die Menschheit und andere Lebewesen Nutzen aus dem KI-Fortschritt KI ziehen können.

(15) Persons may interact with AI systems throughout their life cycle and receive assistance from them, such as care for vulnerable people or people in vulnerable situations, including but not limited to children, older persons, persons with disabilities or the ill. Within such interactions, persons should never be objectified, nor should their dignity be otherwise undermined, or human rights and fundamental freedoms violated or abused.

(16) Human rights and fundamental freedoms must be respected, protected and promoted throughout the life cycle of AI systems. Governments, private sector, civil society, international organizations, technical communities and academia must respect human rights instruments and frameworks in their interventions in the processes surrounding the life cycle of AI systems. New technologies need to provide new means to advocate, defend and exercise human rights and not to infringe them.

Environment and ecosystem flourishing

(17) Environmental and ecosystem flourishing should be recognized, protected and promoted through the life cycle of AI systems. Furthermore, environment and ecosystems are the existential necessity for humanity and other living beings to be able to enjoy the benefits of advances in AI.

(18) Sämtliche Akteure, die in den Lebenszyklus von KI-Systemen einbezogen sind, müssen das geltende Völkerrecht und die innerstaatlichen Rechtsvorschriften, Normen und Verfahren beachten, so zum Beispiel das Vorsorgeprinzip, das auf den Schutz und die Wiederherstellung der Umwelt und der Ökosysteme sowie auf eine nachhaltige Entwicklung ausgelegt ist. Sie sollten die ökologischen Auswirkungen von KI-Systemen eindämmen, unter anderem, jedoch nicht ausschließlich von deren CO₂-Bilanz, um Risikofaktoren für den Klimawandel und die Umwelt zu minimieren und eine nicht nachhaltige Ausbeutung, Nutzung oder Transformation natürlicher Ressourcen zu verhindern, die zur Verschlechterung der Umwelt und zur Zerstörung von Ökosystemen beitragen.

(18) All actors involved in the life cycle of AI systems must comply with applicable international law and domestic legislation, standards and practices, such as precaution, designed for environmental and ecosystem protection and restoration, and sustainable development. They should reduce the environmental impact of AI systems, including but not limited to its carbon footprint, to ensure the minimization of climate change and environmental risk factors, and prevent the unsustainable exploitation, use and transformation of natural resources contributing to the deterioration of the environment and the degradation of ecosystems.

Gewährleistung von Vielfalt und Inklusivität

(19) Achtung, Schutz und Förderung von Vielfalt und Inklusivität sind über den gesamten Lebenszyklus von AI-Systemen im Einklang mit dem Völkerrecht einschließlich der Menschenrechte zu gewährleisten. Dies kann durch Förderung der aktiven Beteiligung eines jeden Einzelnen bzw. sämtlicher Gruppen geschehen, ungeachtet ihrer ethnischen Herkunft und Hautfarbe, ihrer Abstammung, ihres Geschlechts, ihres Alters, ihrer Sprache, ihrer Religion und ihrer politischen Überzeugung, ihrer nationalen oder sozialen Herkunft, ihres wirtschaftlichen oder sozialen Geburtsstatus, einer Behinderung oder sonstiger Merkmale.

(20) Die Wahloptionen von Lebensweisen, Überzeugungen und Meinungen, Ausdrucksformen und persönlichen Erfahrungen, auch hinsichtlich der optionalen Nutzung von KI-Systemen und der Mitgestaltung solcher Architekturen, dürfen in keiner Phase im Lebenszyklus von KI-Systemen eingeschränkt werden.

(21) Darüber hinaus sollten Anstrengungen unternommen werden, auch im Rahmen internationaler Zusammenarbeit, um die Probleme etwa fehlender technischer Infrastruktur, fehlender Bildung und Kompetenzen sowie fehlender Rechtsrahmen, insbesondere in LMICs, LDCs, LLDCs und SIDS, zu überwinden und niemals zum Nachteil von Gemeinschaften auszunutzen.

Ensuring diversity and inclusiveness

(19) Respect, protection and promotion of diversity and inclusiveness should be ensured throughout the life cycle of AI systems, consistent with international law, including human rights law. This may be done by promoting active participation of all individuals or groups regardless of race, colour, descent, gender, age, language, religion, political opinion, national origin, ethnic origin, social origin, economic or social condition of birth, or disability and any other grounds.

(20) The scope of lifestyle choices, beliefs, opinions, expressions or personal experiences, including the optional use of AI systems and the co-design of these architectures should not be restricted during any phase of the life cycle of AI systems.

(21) Furthermore, efforts, including international cooperation, should be made to overcome, and never take advantage of, the lack of necessary technological infrastructure, education and skills, as well as legal frameworks, particularly in LMICs, LDCs, LLDCs and SIDS, affecting communities.

Leben in friedlichen, gerechten und verbundenen Gesellschaften

(22) KI-Akteure sollten die Aufgabe, friedliche und gerechte Gesellschaften zu gewährleisten, in partizipativer und fördernder Weise ausfüllen, beruhend auf einer Zukunft in Verbundenheit zum Wohle aller und im Einklang mit den Menschenrechten und Grundfreiheiten. Der Wert des Lebens in friedlichen und gerechten Gesellschaften verweist auf das Potenzial von KI-Systemen, über ihren gesamten Lebenszyklus zur Verbundenheit aller Lebewesen untereinander und mit der natürlichen Umwelt beizutragen.

(23) Die Vorstellung, dass die Menschen miteinander verbunden sind, beruht auf der Erkenntnis, dass jeder Mensch zu einem größeren Ganzen gehört, das nur dann gedeiht, wenn alle seine Bestandteile sich entfalten können. Um in friedlichen, gerechten und miteinander verbundenen Gesellschaften zu leben, bedarf es einer organischen, unmittelbaren und nicht berechnenden Solidarität, die sich durch ständiges Bemühen um friedliche Beziehungen auszeichnet und darauf ausgerichtet ist, im weitesten Sinne Rücksicht auf die Belange der anderen und der natürlichen Umwelt zu nehmen.

(24) Dieser Wert verlangt nach Förderung von Frieden, Inklusivität und Gerechtigkeit, Gleichheit und Verbundenheit über den gesamten Lebenszyklus von KI-Systemen, insofern als die Prozesse im Lebenszyklus von KI-Systemen niemanden ausgrenzen oder zu Objekten degradieren, die Freiheit und die autonome Entscheidungsfindung nicht beeinträchtigen, die Sicherheit von Menschen und Gemeinschaften nicht gefährden, Individuen und Gruppen nicht spalten oder gegeneinander aufbringen und die Koexistenz zwischen Menschen, anderen Lebewesen und der natürlichen Umwelt nicht bedrohen dürfen.

Living in peaceful, just and interconnected societies

(22) AI actors should play a participative and enabling role to ensure peaceful and just societies, which is based on an interconnected future for the benefit of all, consistent with human rights and fundamental freedoms. The value of living in peaceful and just societies points to the potential of AI systems to contribute throughout their life cycle to the interconnectedness of all living creatures with each other and with the natural environment.

(23) The notion of humans being interconnected is based on the knowledge that every human belongs to a greater whole, which thrives when all its constituent parts are enabled to thrive. Living in peaceful, just and interconnected societies requires an organic, immediate, uncalculated bond of solidarity, characterized by a permanent search for peaceful relations, tending towards care for others and the natural environment in the broadest sense of the term.

(24) This value demands that peace, inclusiveness and justice, equity and interconnectedness should be promoted throughout the life cycle of AI systems, in so far as the processes of the life cycle of AI systems should not segregate, objectify or undermine freedom and autonomous decision-making as well as the safety of human beings and communities, divide and turn individuals and groups against each other, or threaten the coexistence between humans, other living beings and the natural environment.

III.2 Prinzipien

Verhältnismäßigkeit und Schadensverhütung („Do No Harm“)

(25) KI-Technologien gewährleisten nicht notwendigerweise per se das Wohlergehen von Mensch, Umwelt und Ökosystemen – dies sollte zuallererst anerkannt werden. Außerdem darf keiner der Prozesse im Zusammenhang mit dem Lebenszyklus von KI-Systemen über das hinausgehen, was zur Erreichung legitimer Ziele erforderlich ist, und die Prozesse sollten dem jeweiligen Kontext angemessen sein. Es sollte sichergestellt werden, dass im Falle eines denkbaren Schadens für Menschen, Menschenrechte und Grundfreiheiten, Gemeinschaften, die Gesellschaft insgesamt oder für die Umwelt und die Ökosysteme Verfahren zur Risikobewertung greifen und Maßnahmen umgesetzt werden, um das Auftreten eines solchen Schadens zu verhindern.

(26) Die Entscheidung für den Einsatz von KI-Systemen und die Wahl der zu verwendenden KI-Methode sollten von den folgenden Kriterien getragen sein:

- a) Die gewählte KI-Methode muss geeignet und verhältnismäßig sein, um das jeweilige legitime Ziel zu erreichen;
- b) die gewählte KI-Methode darf nicht gegen die in diesem Dokument dargelegten Grundwerte verstoßen, insbesondere darf ihre Anwendung nicht zu einer Verletzung oder einem Missbrauch von Menschenrechten führen; und
- c) die KI-Methode sollte dem Kontext angemessen sein und auf streng wissenschaftlichen Grundlagen beruhen. In Szenarien, bei denen davon auszugehen ist, dass Entscheidungen unumkehrbare oder nur schwer rückgängig zu machende Folgen haben oder in denen möglicherweise Entscheidungen über Leben und Tod fallen, sollte die endgültige Entscheidung von einem Menschen getroffen werden. Insbesondere dürfen KI-Systeme nicht zum Zwecke des Social Scoring (Sozialkreditsystem) oder der Massenüberwachung eingesetzt werden.

III.2 Principles

Proportionality and Do No Harm

(25) It should be recognized that AI technologies do not necessarily, per se, ensure human and environmental and ecosystem flourishing. Furthermore, none of the processes related to the AI system life cycle shall exceed what is necessary to achieve legitimate aims or objectives and should be appropriate to the context. In the event of possible occurrence of any harm to human beings, human rights and fundamental freedoms, communities and society at large or the environment and ecosystems, the implementation of procedures for risk assessment and the adoption of measures in order to preclude the occurrence of such harm should be ensured.

(26) The choice to use AI systems and which AI method to use should be justified in the following ways: (a) the AI method chosen should be appropriate and proportional to achieve a given legitimate aim; (b) the AI method chosen should not infringe upon the foundational values captured in this document, in particular, its use must not violate or abuse human rights; and (c) the AI method should be appropriate to the context and should be based on rigorous scientific foundations. In scenarios where decisions are understood to have an impact that is irreversible or difficult to reverse or may involve life and death decisions, final human determination should apply. In particular, AI systems should not be used for social scoring or mass surveillance purposes.

Sicherheit und Schutz

(27) Um Sicherheit und Schutz von Mensch, Umwelt und Ökosystemen zu gewährleisten, sollten über den gesamten Lebenszyklus von KI-Systemen Maßnahmen zur Identifizierung, Vermeidung und Beseitigung von ungewollten Schäden (Sicherheitsrisiken) und Anfälligkeiten gegen Angriffe (Schwachstellen) ergriffen werden. Sichere und gegen Angriffe geschützte KI wird durch die Entwicklung tragfähiger Rahmenbedingungen für den Datenzugang ermöglicht, die die Privatsphäre schützen und dafür sorgen, dass KI-Systeme durch Verwendung hochwertiger Daten besser trainiert und validiert werden können.

Safety and security

(27) Unwanted harms (safety risks), as well as vulnerabilities to attack (security risks) should be avoided and should be addressed, prevented and eliminated throughout the life cycle of AI systems to ensure human, environmental and ecosystem safety and security. Safe and secure AI will be enabled by the development of sustainable, privacy-protective data access frameworks that foster better training and validation of AI models utilizing quality data.

Fairness und Nichtdiskriminierung

(28) KI-Akteure sollten die soziale Gerechtigkeit fördern und das Prinzip von Fairness und Gerechtigkeit im Einklang mit dem Völkerrecht in jeglicher Hinsicht wahren. Dies ist nur durch einen inklusiven Ansatz möglich, der sicherstellt, dass alle Menschen von KI-Technologien profitieren können und ihr Nutzen für alle zugänglich ist, wobei die spezifischen Bedürfnisse verschiedener Altersgruppen, Kulturen und Sprachgruppen, die Belange von Menschen mit Behinderungen, Mädchen und Frauen sowie von benachteiligten, marginalisierten und schutzbedürftigen Menschen bzw. Menschen in gefährdeten Umständen zu berücksichtigen sind. Die Mitgliedstaaten sollten sich dafür einsetzen, dass alle Menschen einschließlich lokaler Gemeinschaften einen inklusiven Zugang zu KI-Systemen mit lokal relevanten Inhalten und Diensten erhalten, wobei die Mehrsprachigkeit und die kulturelle Vielfalt zu berücksichtigen sind. Die Mitgliedstaaten sollten darauf hinarbeiten, digitale Gräben zu überwinden, und inklusiven Zugang zu und Beteiligung an der Entwicklung von KI sicherzustellen. Auf der nationalen Ebene sollten die Mitgliedstaaten im Hinblick auf den Zugang zu KI-Systemen und die Beteiligung am Lebenszyklus dieser Systeme sowohl die Chancengerechtigkeit zwischen ländlichen und städtischen Gebieten als auch zwischen allen Menschen fördern, ungeachtet ihrer ethnischen Herkunft und Hautfarbe, ihrer Abstammung, ihres Geschlechts, ihres Alters, ihrer Sprache, ihrer Religion und ihrer politischen Überzeugung, ihrer nationalen oder sozialen Herkunft, ihrem wirtschaftlichen oder sozialen Geburtsstatus, einer Behinderung oder sonstiger Merkmale. Auf der internationalen Ebene haben die technologisch am weitesten fortgeschrittenen Länder eine solidarische Verantwortung gegenüber den am wenigsten fortgeschrittenen Ländern und sollten sicherstellen, dass der Nutzen der KI-Systeme gerecht geteilt wird, damit die Tatsache, dass

Fairness and non-discrimination

(28) AI actors should promote social justice and safeguard fairness and non-discrimination of any kind in compliance with international law. This implies an inclusive approach to ensuring that the benefits of AI technologies are available and accessible to all, taking into consideration the specific needs of different age groups, cultural systems, different language groups, persons with disabilities, girls and women, and disadvantaged, marginalized and vulnerable people or people in vulnerable situations. Member States should work to promote inclusive access for all, including local communities, to AI systems with locally relevant content and services, and with respect for multilingualism and cultural diversity. Member States should work to tackle digital divides and ensure inclusive access to and participation in the development of AI. At the national level, Member States should promote equity between rural and urban areas, and among all persons regardless of race, colour, descent, gender, age, language, religion, political opinion, national origin, ethnic origin, social origin, economic or social condition of birth, or disability and any other grounds, in terms of access to and participation in the AI system life cycle. At the international level, the most technologically advanced countries have a responsibility of solidarity with the least advanced to ensure that the benefits of AI technologies are shared such that access to and participation in the AI system life cycle for the latter contributes to a fairer world order with regard to information, communication, culture, education, research and socio-economic and political stability.

Zugang zu den KI-Technologie und Beteiligung an deren Lebenszyklus zu einer gerechteren Weltordnung in Bezug auf Information, Kommunikation, Kultur, Bildung, Forschung sowie sozioökonomischer und politischer Stabilität beitragen.

(29) Zur Gewährleistung der Fairness von KI-Systemen sollten KI-Akteure alle möglichen Anstrengungen unternehmen, um diskriminierende oder von Verzerrungen geprägte Anwendungen und Ergebnisse über den gesamten Lebenszyklus eines KI-Systems auf ein Minimum zu reduzieren und dafür zu sorgen, dass diese nicht fortbestehen oder gefestigt werden. Gegen Diskriminierung und vorurteilsbehaftete algorithmische Festlegungen sollten wirksame Rechtsbehelfe zur Verfügung stehen.

(30) Darüber hinaus müssen digitale Gräben und Wissensgefälle sowohl innerhalb von Ländern als auch zwischen verschiedenen Ländern über den gesamten Lebenszyklus eines KI-Systems angegangen werden, auch hinsichtlich des Zugangs und der Qualität des Zugangs zu Technologie und Daten im Einklang mit den einschlägigen nationalen, regionalen und internationalen Rechtsrahmen, sowie hinsichtlich der Vernetzungsmöglichkeiten, der Kenntnisse und Fertigkeiten sowie einer aussagekräftigen Beteiligung der betroffenen Gemeinschaften, damit eine Gleichbehandlung aller Menschen gewährleistet ist.

(29) AI actors should make all reasonable efforts to minimize and avoid reinforcing or perpetuating discriminatory or biased applications and outcomes throughout the life cycle of the AI system to ensure fairness of such systems. Effective remedy should be available against discrimination and biased algorithmic determination.

(30) Furthermore, digital and knowledge divides within and between countries need to be addressed throughout an AI system life cycle, including in terms of access and quality of access to technology and data, in accordance with relevant national, regional and international legal frameworks, as well as in terms of connectivity, knowledge and skills and meaningful participation of the affected communities, such that every person is treated equitably.

Nachhaltigkeit

(31) Die Entwicklung nachhaltiger Gesellschaften ist abhängig von der Erreichung einer Reihe komplexer Ziele in einem Kontinuum menschlicher, sozialer, kultureller, wirtschaftlicher und ökologischer Dimensionen. Je nachdem, wie KI-Technologien in Ländern mit unterschiedlichem Entwicklungsstand eingesetzt werden, kann ihr Aufkommen für die Nachhaltigkeitsziele entweder förderlich sein oder ihre Verwirklichung behindern. Die kontinuierliche Bewertung der menschenbezogenen, sozialen, kulturellen, wirtschaftlichen und ökologischen Folgen von KI-Technologien sollte daher im vollen Bewusstsein der Auswirkungen von KI-Technologien auf die Nachhaltigkeit erfolgen, die sich in einer Reihe von sich stetig weiterentwickelnden Zielen in unterschiedlichen Dimensionen manifestiert, wie sie gegenwärtig beispielsweise in den Zielen für nachhaltige Entwicklung (SDGs) der Vereinten Nationen vereinbart sind.

Sustainability

(31) The development of sustainable societies relies on the achievement of a complex set of objectives on a continuum of human, social, cultural, economic and environmental dimensions. The advent of AI technologies can either benefit sustainability objectives or hinder their realization, depending on how they are applied across countries with varying levels of development. The continuous assessment of the human, social, cultural, economic and environmental impact of AI technologies should therefore be carried out with full cognizance of the implications of AI technologies for sustainability as a set of constantly evolving goals across a range of dimensions, such as currently identified in the Sustainable Development Goals (SDGs) of the United Nations.

Recht auf Privatsphäre und Datenschutz

(32) Das für den Schutz der Menschenwürde, der menschlichen Autonomie und der menschlichen Handlungsfähigkeit wesentliche Recht auf Privatsphäre muss über den gesamten Lebenszyklus von KI-Systemen geachtet, geschützt und gefördert werden. Es ist wichtig, dass Daten für KI-Systeme auf eine Art und Weise gesammelt, verwendet, weitergegeben, gespeichert und gelöscht werden, die mit dem Völkerrecht und den in dieser Empfehlung dargelegten Werten und Prinzipien im Einklang steht und gleichzeitig die einschlägigen nationalen, regionalen (heißt hier: auf Ebene von Kontinenten) und internationalen Rechtsrahmen beachtet.

(33) Angemessene Datenschutzrahmen und entsprechende Durchsetzungsmechanismen sollten auf Basis eines alle Interessengruppen einbeziehenden Multistakeholder-Ansatzes auf nationaler oder internationaler Ebene geschaffen, durch Rechtssysteme geschützt und über den gesamten Lebenszyklus von KI-Systemen gewährleistet werden. Die Datenschutzrahmen und alle damit zusammenhängenden Mechanismen sollten sich an den internationalen Datenschutzprinzipien und -normen für die Sammlung, Verwendung und Offenlegung personenbezogener Daten und die Ausübung der Rechte betroffener Personen orientieren; dabei ist sicherzustellen, dass die Verarbeitung personenbezogener Daten einen legitimen Zweck verfolgt und auf einer gültigen Rechtsgrundlage beruht, einschließlich einer Einwilligungserklärung.

Right to privacy, and data protection

(32) Privacy, a right essential to the protection of human dignity, human autonomy and human agency, must be respected, protected and promoted throughout the life cycle of AI systems. It is important that data for AI systems be collected, used, shared, archived and deleted in ways that are consistent with international law and in line with the values and principles set forth in this Recommendation, while respecting relevant national, regional and international legal frameworks.

(33) Adequate data protection frameworks and governance mechanisms should be established in a multi-stakeholder approach at the national or international level, protected by judicial systems, and ensured throughout the life cycle of AI systems. Data protection frameworks and any related mechanisms should take reference from international data protection principles and standards concerning the collection, use and disclosure of personal data and exercise of their rights by data subjects while ensuring a legitimate aim and a valid legal basis for the processing of personal data, including informed consent.

(34) Algorithmische Systeme erfordern angemessene Folgenabschätzungen zum Schutz der Privatsphäre, die auch gesellschaftliche und ethische Überlegungen zu ihrer Nutzung und eine innovative Anwendung des Konzepts des „eingebauten Datenschutzes“ („privacy by design approach“) umfassen. KI-Akteure müssen sicherstellen, dass sie für die Gestaltung und Umsetzung von KI-Systemen in einer Weise in der Verantwortung bleiben, die den Schutz personenbezogener Daten über den gesamten Lebenszyklus des KI-Systems gewährleistet.

(34) Algorithmic systems require adequate privacy impact assessments, which also include societal and ethical considerations of their use and an innovative use of the privacy by design approach. AI actors need to ensure that they are accountable for the design and implementation of AI systems in such a way as to ensure that personal information is protected throughout the life cycle of the AI system.

Menschliche Aufsicht und Entscheidung

(35) Die Mitgliedstaaten sollten sicherstellen, dass es immer möglich ist, die ethische und rechtliche Verantwortung für jede Phase im Lebenszyklus von KI-Systemen natürlichen Personen oder bestehenden juristischen Personen zuzuordnen; Gleiches gilt für Rechtsbehelfe im Zusammenhang mit KI-Systemen. Menschliche Aufsicht bezieht sich somit nicht nur auf die individuelle Aufsicht durch Menschen, sondern gegebenenfalls auch auf eine gesamtheitliche öffentliche Aufsicht.

(36) Auch wenn Menschen sich manchmal aus Effizienzgründen dazu entschließen mögen, sich auf KI-Systeme zu verlassen, so muss die Entscheidung über die Abgabe der Kontrolle innerhalb eines begrenzten Kontextes stets Menschen vorbehalten bleiben, denn Menschen können zwar beim Treffen von Entscheidungen und bei ihren Handlungen auf KI-Systeme zurückgreifen, aber ein KI-System kann niemals die letztendliche menschliche Verantwortung und Rechenschaftspflicht ersetzen. Entscheidungen über Leben und Tod dürfen grundsätzlich nicht KI-Systemen übertragen werden.

Human oversight and determination

(35) Member States should ensure that it is always possible to attribute ethical and legal responsibility for any stage of the life cycle of AI systems, as well as in cases of remedy related to AI systems, to physical persons or to existing legal entities. Human oversight refers thus not only to individual human oversight, but to inclusive public oversight, as appropriate.

(36) It may be the case that sometimes humans would choose to rely on AI systems for reasons of efficacy, but the decision to cede control in limited contexts remains that of humans, as humans can resort to AI systems in decision-making and acting, but an AI system can never replace ultimate human responsibility and accountability. As a rule, life and death decisions should not be ceded to AI systems.

Transparenz und Erklärbarkeit

(37) Transparenz und Erklärbarkeit von KI-Systemen sind häufig wesentliche Voraussetzungen für die Wahrung, Verteidigung und Förderung von Menschenrechten, Grundfreiheiten und ethischen Grundsätzen. Transparenz ist notwendig, damit einschlägige nationale und internationale Haftungsregelungen wirksam greifen können. Fehlende Transparenz kann auch dazu führen, dass Entscheidungen, die auf von KI-Systemen erzeugten Ergebnissen beruhen, nicht wirksam angefochten werden können, womit das Recht auf ein faires Verfahren und einen wirksamen Rechtsbehelf verletzt sein kann und ein rechtmäßiger Einsatz solcher Systeme nur in eingeschränkten Bereichen möglich ist.

Transparency and explainability

(37) The transparency and explainability of AI systems are often essential preconditions to ensure the respect, protection and promotion of human rights, fundamental freedoms and ethical principles. Transparency is necessary for relevant national and international liability regimes to work effectively. A lack of transparency could also undermine the possibility of effectively challenging decisions based on outcomes produced by AI systems and may thereby infringe the right to a fair trial and effective remedy, and limits the areas in which these systems can be legally used.

(38) Wenngleich Anstrengungen unternommen werden müssen, um die Transparenz und Erklärbarkeit von KI-Systemen, einschließlich solcher mit extraterritorialen Auswirkungen, über ihren gesamten Lebenszyklus zu erhöhen, um die Voraussetzungen für eine demokratische Lenkung und Kontrolle zu schaffen, so muss der Grad an Transparenz und Erklärbarkeit doch immer in einem angemessenen Verhältnis zum Kontext und den jeweiligen Auswirkungen stehen, denn unter Umständen muss zwischen Transparenz und Erklärbarkeit und anderen Prinzipien wie Privatsphäre und Sicherheit abgewogen werden. Betroffene sollten umfassend informiert werden, wenn eine Entscheidung auf KI-Algorithmen beruht bzw. auf deren Grundlage getroffen wird, gerade wenn es um Entscheidungen geht, die sich negativ auf ihre Sicherheit oder ihre Menschenrechte auswirken; man sollte ihnen unter diesen Umständen ermöglichen, von dem betreffenden KI-Akteur oder von öffentlichen Einrichtungen erläuternde Informationen zu verlangen. Außerdem sollten die Betroffenen sich über die Gründe informieren können, die zu einer Entscheidung geführt haben, durch die ihre Rechte und Freiheiten beeinträchtigt werden, und sie sollten die Möglichkeit haben, sich an einen benannten Mitarbeiter bzw. eine benannte Mitarbeiterin des privatwirtschaftlichen Unternehmens oder der öffentlichen Einrichtung zu wenden, der bzw. die eine solche Entscheidung überprüfen und korrigieren kann. Wenn die Bereitstellung eines Produkts oder die Erbringung einer Dienstleistung unmittelbar durch ein KI-System oder mit dessen Hilfe erfolgt, sollten die KI-Akteure die Nutzerinnen und Nutzer ordnungsgemäß und rechtzeitig darüber informieren.

(38) While efforts need to be made to increase transparency and explainability of AI systems, including those with extra-territorial impact, throughout their life cycle to support democratic governance, the level of transparency and explainability should always be appropriate to the context and impact, as there may be a need to balance between transparency and explainability and other principles such as privacy, safety and security. People should be fully informed when a decision is informed by or is made on the basis of AI algorithms, including when it affects their safety or human rights, and in those circumstances should have the opportunity to request explanatory information from the relevant AI actor or public sector institutions. In addition, individuals should be able to access the reasons for a decision affecting their rights and freedoms, and have the option of making submissions to a designated staff member of the private sector company or public sector institution able to review and correct the decision. AI actors should inform users when a product or service is provided directly or with the assistance of AI systems in a proper and timely manner.

(39) Aus soziotechnischer Sicht trägt mehr Transparenz zu friedlicheren, gerechteren, demokratischeren und inklusiveren Gesellschaften bei. Sie ermöglicht öffentliche Kontrolle, die Korruption und Diskriminierung eindämmen und auch dazu beitragen kann, negative Auswirkungen auf Menschenrechte zu erkennen und ihnen vorzubeugen. Transparenz zielt darauf ab, den jeweiligen Adressaten geeignete Informationen zur Verfügung zu stellen, um ihnen Verständnis zu ermöglichen und um Vertrauen zu fördern. Im Falle von KI-Systemen bedeutet Transparenz auch, Menschen in die Lage zu versetzen, die Umsetzung eines KI-Systems in jeder einzelnen Stufe zu verstehen - in Abhängigkeit vom jeweiligen Kontext und der Sensibilität des Systems. Transparenz kann auch heißen, dass Einblick in die Faktoren gewährt wird, die eine bestimmte Vorhersage oder Entscheidung beeinflussen, und dass Auskunft darüber gegeben wird, ob angemessene Garantien (wie Maßnahmen zur Gewährleistung von Sicherheit oder Fairness) bestehen oder nicht. In Fällen, in denen ein großes Risiko negativer Auswirkungen auf Menschenrechte besteht, kann es im Sinne der Transparenz auch erforderlich sein, den Code bzw. die Datensätze offenzulegen.

(39) From a socio-technical lens, greater transparency contributes to more peaceful, just, democratic and inclusive societies. It allows for public scrutiny that can decrease corruption and discrimination, and can also help detect and prevent negative impacts on human rights. Transparency aims at providing appropriate information to the respective addressees to enable their understanding and foster trust. Specific to the AI system, transparency can enable people to understand how each stage of an AI system is put in place, appropriate to the context and sensitivity of the AI system. It may also include insight into factors that affect a specific prediction or decision, and whether or not appropriate assurances (such as safety or fairness measures) are in place. In cases of serious threats of adverse human rights impacts, transparency may also require the sharing of code or datasets.

(40) Erklärbarkeit bedeutet, Ergebnisse von KI-Systemen verständlich zu machen und Einblick in sie zu gewähren. Die Erklärbarkeit von KI-Systemen bezieht sich auch auf die Verständlichkeit von Eingabe, Ausgabe und Funktionsweise eines jeden einzelnen algorithmischen Bausteins und seines Beitrags zum Ergebnis des Systems. Die Erklärbarkeit hängt also eng mit der Transparenz zusammen, da die Ergebnisse und die Teilprozesse, die zu den Ergebnissen führen, kontextangemessen verständlich und rückverfolgbar sein sollten. KI-Akteure sollten sich dazu verpflichten, dafür zu sorgen, dass die entwickelten Algorithmen erklärbar sind. Im Falle von KI-Anwendungen, die sich in einer Weise auf den Endnutzer oder die Endnutzerin auswirken, die nicht nur vorübergehender Natur, nicht leicht rückgängig zu machen oder aus anderen Gründen nicht risikoarm ist, sollte sichergestellt werden, dass für jede Entscheidung, die zu der betreffenden Maßnahme geführt hat, eine nachvollziehbare Erklärung geliefert wird; nur so kann das Ergebnis als transparent gelten.

(41) Transparenz und Erklärbarkeit stehen in engem Zusammenhang mit angemessenen Maßnahmen im Rahmen der Verantwortlichkeit und Rechenschaftspflicht sowie mit der Vertrauenswürdigkeit von KI-Systemen.

(40) Explainability refers to making intelligible and providing insight into the outcome of AI systems. The explainability of AI systems also refers to the understandability of the input, output and the functioning of each algorithmic building block and how it contributes to the outcome of the systems. Thus, explainability is closely related to transparency, as outcomes and sub-processes leading to outcomes should aim to be understandable and traceable, appropriate to the context. AI actors should commit to ensuring that the algorithms developed are explainable. In the case of AI applications that impact the end user in a way that is not temporary, easily reversible or otherwise low risk, it should be ensured that the meaningful explanation is provided with any decision that resulted in the action taken in order for the outcome to be considered transparent.

(41) Transparency and explainability relate closely to adequate responsibility and accountability measures, as well as to the trustworthiness of AI systems.

Verantwortlichkeit und Rechenschaftspflicht

(42) Über den gesamten Lebenszyklus von KI-Systemen müssen die KI-Akteure und die Mitgliedstaaten die Menschenrechte und Grundfreiheiten achten, schützen und fördern und sollten sich auch für den Schutz der Umwelt und der Ökosysteme einsetzen, in Wahrnehmung ihrer jeweiligen ethischen und rechtlichen Verantwortung entsprechend dem nationalen Recht und dem Völkerrecht, insbesondere den Menschenrechtsverpflichtungen der Mitgliedstaaten sowie im Einklang mit ethischen Leitlinien, auch in Bezug auf die KI-Akteure, die in ihrem jeweiligen Hoheitsgebiet und unter ihrer tatsächlichen Kontrolle agieren. Die ethische Verantwortung und die Haftung für Entscheidungen und Handlungen, die in irgendeiner Weise auf einem KI-System beruhen, sollten letztlich immer den KI-Akteuren zurechenbar sein, entsprechend ihrer Rolle im Lebenszyklus des KI-Systems.

(43) Um die Rechenschaftspflicht für KI-Systeme und ihre Auswirkungen über ihren gesamten Lebenszyklus zu gewährleisten, sollten geeignete Mechanismen für Aufsicht, Folgenabschätzung, Überprüfung und Sorgfaltspflicht („due diligence“) entwickelt werden, auch zum Schutz von Personen, die Missstände melden („whistle-blower“). Sowohl die technische als auch die institutionelle Ausgestaltung sollten die Überprüfbarkeit und Rückverfolgbarkeit (der Funktionsweise) von KI-Systemen gewährleisten, insbesondere um Konflikten mit Menschenrechtsnormen und -standards sowie Bedrohungen für das Wohl von Umwelt und Ökosystemen entgegenzuwirken.

Responsibility and accountability

(42) AI actors and Member States should respect, protect and promote human rights and fundamental freedoms, and should also promote the protection of the environment and ecosystems, assuming their respective ethical and legal responsibility, in accordance with national and international law, in particular Member States' human rights obligations, and ethical guidance throughout the life cycle of AI systems, including with respect to AI actors within their effective territory and control. The ethical responsibility and liability for the decisions and actions based in any way on an AI system should always ultimately be attributable to AI actors corresponding to their role in the life cycle of the AI system.

(43) Appropriate oversight, impact assessment, audit and due diligence mechanisms, including whistle-blowers' protection, should be developed to ensure accountability for AI systems and their impact throughout their life cycle. Both technical and institutional designs should ensure auditability and traceability of (the working of) AI systems in particular to address any conflicts with human rights norms and standards and threats to environmental and ecosystem well-being.

Bewusstsein und Kompetenzen

(44) Das Bewusstsein der Öffentlichkeit für KI-Technologien und den Wert von Daten sowie deren Verständnis sollten durch offene und zugängliche Bildung, zivilgesellschaftliches Engagement, Schulungsmaßnahmen zum Erwerb digitaler Kenntnisse und zur KI-Ethik sowie zum Aufbau von Medien- und Informationskompetenz gefördert werden; dabei sollten Regierungen, zwischenstaatliche Organisationen, die Zivilgesellschaft, die Wissenschaft, die Medien und Führungspersönlichkeiten in informellen Gruppen sowie die Privatwirtschaft zusammenarbeiten. Sie sollten der bestehenden sprachlichen, sozialen und kulturellen Vielfalt Rechnung tragen, um eine wirksame Beteiligung der breiten Öffentlichkeit zu gewährleisten, damit alle Mitglieder der Gesellschaft fundierte Entscheidungen über die Nutzung von KI-Systemen treffen können und vor ungebührlicher Einflussnahme geschützt sind.

(45) Das Lernen über die Auswirkungen von KI-Systemen sollte auch das Lernen über, durch und für die Menschenrechte und Grundfreiheiten umfassen; dies bedeutet, dass die Herangehensweise an KI-Systeme sowie das Verständnis dieser Systeme von ihren Auswirkungen auf die Menschenrechte und den Zugang zu Rechten sowie ihren Folgen für die Umwelt und Ökosysteme ausgehen sollten.

Awareness and literacy

(44) Public awareness and understanding of AI technologies and the value of data should be promoted through open and accessible education, civic engagement, digital skills and AI ethics training, media and information literacy and training led jointly by governments, intergovernmental organizations, civil society, academia, the media, community leaders and the private sector, and considering the existing linguistic, social and cultural diversity, to ensure effective public participation so that all members of society can take informed decisions about their use of AI systems and be protected from undue influence.

(45) Learning about the impact of AI systems should include learning about, through and for human rights and fundamental freedoms, meaning that the approach and understanding of AI systems should be grounded by their impact on human rights and access to rights, as well as on the environment and ecosystems.

Flexible Organisation und Zusammenarbeit unter Einbeziehung aller Interessengruppen

(46) Bei der Nutzung von Daten sind das Völkerrecht und die nationale Souveränität zu achten. Das bedeutet, dass die Staaten im Einklang mit dem Völkerrecht die in ihrem Hoheitsgebiet erzeugten oder durch ihr Hoheitsgebiet fließenden Daten regeln und Maßnahmen für eine wirksame Regulierung der Daten, einschließlich des Datenschutzes, auf der Grundlage der Achtung des Rechts auf Privatsphäre im Einklang mit dem Völkerrecht und anderen Menschenrechtsnormen und -standards ergreifen können.

(47) Eine inklusive Organisation von KI-Entscheidungen erfordert die Beteiligung verschiedener Interessengruppen über den gesamten Lebenszyklus eines KI-Systems, damit die Vorteile von allen geteilt werden können und zu einer nachhaltigen Entwicklung beitragen. Zu diesen Interessengruppen gehören unter anderem Regierungen, zwischenstaatliche Organisationen, die technische Fachwelt, die Zivilgesellschaft, Forschende und die Wissenschaft, Medien, Bildungseinrichtungen, politische Entscheidungsträgerinnen und Entscheidungsträger, privatwirtschaftliche Unternehmen, Menschenrechtsorganisationen und Organisationen, die für Gleichstellung oder Antidiskriminierung eintreten, sowie Gruppen für Jugendliche und Kinder. Um Zusammenarbeit zu erleichtern, sollte auf offene Standards und Interoperabilität geachtet werden. Es sollten Maßnahmen ergriffen werden, die dem technologischen Wandel und der Entstehung neuer Interessengruppen Rechnung tragen und es marginalisierten Gruppen, Gemeinschaften und Einzelpersonen ermöglichen, sich wirkungsvoll einzubringen, gegebenenfalls auch unter Respektierung der Selbstverwaltung ihrer Daten wie im Falle indigener Völker.

Multi-stakeholder and adaptive governance and collaboration

(46) International law and national sovereignty must be respected in the use of data. That means that States, complying with international law, can regulate the data generated within or passing through their territories, and take measures towards effective regulation of data, including data protection, based on respect for the right to privacy in accordance with international law and other human rights norms and standards.

(47) Participation of different stakeholders throughout the AI system life cycle is necessary for inclusive approaches to AI governance, enabling the benefits to be shared by all, and to contribute to sustainable development. Stakeholders include but are not limited to governments, intergovernmental organizations, the technical community, civil society, researchers and academia, media, education, policy-makers, private sector companies, human rights institutions and equality bodies, anti-discrimination monitoring bodies, and groups for youth and children. The adoption of open standards and interoperability to facilitate collaboration should be in place. Measures should be adopted to take into account shifts in technologies, the emergence of new groups of stakeholders, and to allow for meaningful participation by marginalized groups, communities and individuals and, where relevant, in the case of Indigenous Peoples, respect for the self-governance of their data.

IV. Politische Handlungs- bereiche

(48) Mit den politischen Maßnahmen, die in den nachfolgend aufgeführten Politikbereichen beschrieben sind, sollen die in dieser Empfehlung dargelegten Werte und Prinzipien in die Praxis umgesetzt werden. Die Hauptaufgabe für die Mitgliedstaaten besteht darin, wirksame Maßnahmen, zum Beispiel politische Rahmenbedingungen oder Mechanismen, einzuführen und dafür zu sorgen, dass auch andere Interessengruppen wie privatwirtschaftliche Unternehmen, Hochschulen und Forschungseinrichtungen sowie die Zivilgesellschaft diese beachten, indem die Mitgliedstaaten unter anderem alle Interessengruppen dazu anregen, Instrumente zur Bewertung der Auswirkungen auf die Menschenrechte, die Rechtsstaatlichkeit, die Demokratie und ethische Folgen sowie zur Ausübung von Sorgfaltspflichten („due diligence“) im Einklang mit Orientierungshilfen wie den Leitprinzipien der Vereinten Nationen für Wirtschaft und Menschenrechte zu entwickeln. Der Prozess zur Entwicklung solcher Maßnahmen bzw. Mechanismen sollte alle Interessengruppen einbeziehen und die Gegebenheiten und Prioritäten der einzelnen Mitgliedstaaten berücksichtigen. Dabei kann die UNESCO als Partner fungieren und die Mitgliedstaaten bei der Entwicklung, Überwachung und Evaluierung politischer Mechanismen unterstützen.

IV. Areas of policy action

(48) The policy actions described in the following policy areas operationalize the values and principles set out in this Recommendation. The main action is for Member States to put in place effective measures, including, for example, policy frameworks or mechanisms, and to ensure that other stakeholders, such as private sector companies, academic and research institutions, and civil society adhere to them by, among other actions, encouraging all stakeholders to develop human rights, rule of law, democracy, and ethical impact assessment and due diligence tools in line with guidance including the United Nations Guiding Principles on Business and Human Rights. The process for developing such policies or mechanisms should be inclusive of all stakeholders and should take into account the circumstances and priorities of each Member State. UNESCO can be a partner and support Member States in the development as well as monitoring and evaluation of policy mechanisms.

(49) Aus Sicht der UNESCO ist zu berücksichtigen, dass ihre Mitgliedstaaten in unterschiedlichem Grad fähig sind zur Umsetzung dieser Empfehlung, zum Beispiel in wissenschaftlicher, technischer, wirtschaftlicher, bildungspolitischer, rechtlicher, regulatorischer, infrastruktureller, gesellschaftlicher, oder kultureller Hinsicht. Dieser „Grad der Umsetzungsfähigkeit“ muss als dynamische Größe verstanden werden. Um die wirksame Umsetzung dieser Empfehlung zu ermöglichen, wird die UNESCO daher (1) eine Methodik zur Bewertung dieses Grades der Umsetzungsfähigkeit („readiness assessment methodology“) entwickeln, der interessierten Mitgliedstaaten dabei helfen soll, zu ermitteln, wie weit sie zu bestimmten Zeitpunkten auf ihrem Weg zur Umsetzung entlang eines Kontinuums von Dimensionen fortgeschritten sind, und (2) interessierte Mitgliedstaaten bei der Entwicklung einer UNESCO-Methodik für die ethische Folgenabschätzung („Ethical Impact Assessment“ – EIA) von KI-Technologien, beim Austausch bewährter Verfahren und Bewertungsleitlinien sowie bei sonstigen Mechanismen und analytischen Arbeiten unterstützen.

(49) UNESCO recognizes that Member States will be at different stages of readiness to implement this Recommendation, in terms of scientific, technological, economic, educational, legal, regulatory, infrastructural, societal, cultural and other dimensions. It is noted that “readiness” here is a dynamic status. In order to enable the effective implementation of this Recommendation, UNESCO will therefore: (1) develop a readiness assessment methodology to assist interested Member States in identifying their status at specific moments of their readiness trajectory along a continuum of dimensions; and (2) ensure support for interested Member States in terms of developing a UNESCO methodology for Ethical Impact Assessment (EIA) of AI technologies, sharing of best practices, assessment guidelines and other mechanisms and analytical work.

Politikbereich 1: Ethische Folgenabschätzung

(50) Die Mitgliedstaaten sind gehalten, Rahmen für Folgenabschätzungen, beispielsweise ethische Folgenabschätzungen, einzuführen, um Nutzen sowie bedenkliche Aspekte und Risiken von KI-Systemen zu erkennen und zu bewerten, und sie sollten neben weiteren Sicherheitsmechanismen geeignete Maßnahmen zur Risikoprävention, -minderung und -überwachung ergreifen. Mit derartigen Folgenabschätzungen sollen die Auswirkungen auf die Menschenrechte und Grundfreiheiten ermittelt werden, insbesondere, aber nicht ausschließlich in Bezug auf die Rechte marginalisierter und besonders schutzbedürftiger Menschen und Menschen in gefährdeten Umständen, und die Auswirkungen auf Arbeitsrechte, auf die Umwelt und die Ökosysteme sowie ethische und soziale Implikationen; außerdem soll im Einklang mit den in dieser Empfehlung dargelegten Werten und Prinzipien die Einbindung der Bürgerinnen und Bürger gefördert werden.

Policy area 1: Ethical Impact Assessment

(50) Member States should introduce frameworks for impact assessments, such as ethical impact assessment, to identify and assess benefits, concerns and risks of AI systems, as well as appropriate risk prevention, mitigation and monitoring measures, among other assurance mechanisms. Such impact assessments should identify impacts on human rights and fundamental freedoms, in particular but not limited to the rights of marginalized and vulnerable people or people in vulnerable situations, labour rights, the environment and ecosystems and ethical and social implications, and facilitate citizen participation in line with the values and principles set forth in this Recommendation.

(51) Die Mitgliedstaaten und privatwirtschaftliche Unternehmen sollten Sorgfalts- und Aufsichtsmechanismen entwickeln, um negative Auswirkungen von KI-Systemen auf die Achtung der Menschenrechte, die Rechtsstaatlichkeit und die Inklusivität von Gesellschaften zu erkennen, zu verhindern, abzumildern und Rechenschaft über ihren Umgang damit abzulegen. Die Mitgliedstaaten sollten auch in der Lage sein, die sozioökonomischen Auswirkungen von KI-Systemen auf Armut zu bewerten und sicherstellen, dass sich die Gräben zwischen armen und reichen Menschen sowie die digitale Kluft zwischen und innerhalb von Ländern durch den massiven Einsatz von KI-Technologien in der Gegenwart und der Zukunft nicht vergrößert. Zu diesem Zweck sollten insbesondere durchsetzbare Transparenzprotokolle für den Zugang zu Informationen, auch Informationen von öffentlichem Interesse, die sich in Privatbesitz befinden, implementiert werden. Mitgliedstaaten, privatwirtschaftliche Unternehmen und die Zivilgesellschaft sollten die soziologischen und psychologischen Auswirkungen von KI-gestützten Empfehlungen auf die Entscheidungsautonomie von Menschen untersuchen. KI-Systeme, die als potenziell riskant für die Menschenrechte erkannt wurden, sollten von KI-Akteuren im Rahmen der ethischen Folgenabschätzung auf breiter Basis getestet werden, falls erforderlich auch unter realen Bedingungen, bevor sie auf den Markt gebracht werden.

(51) Member States and private sector companies should develop due diligence and oversight mechanisms to identify, prevent, mitigate and account for how they address the impact of AI systems on the respect for human rights, rule of law and inclusive societies. Member States should also be able to assess the socioeconomic impact of AI systems on poverty and ensure that the gap between people living in wealth and poverty, as well as the digital divide among and within countries, are not increased with the massive adoption of AI technologies at present and in the future. In order to do this, in particular, enforceable transparency protocols should be implemented, corresponding to the access to information, including information of public interest held by private entities. Member States, private sector companies and civil society should investigate the sociological and psychological effects of AI-based recommendations on humans in their decision-making autonomy. AI systems identified as potential risks to human rights should be broadly tested by AI actors, including in real-world conditions if needed, as part of the Ethical Impact Assessment, before releasing them in the market.

(52) Zur ethischen Folgenabschätzung gehört auch, dass Mitgliedstaaten und Wirtschaftsunternehmen geeignete Maßnahmen ergreifen, um alle Phasen im Lebenszyklus eines KI-Systems zu beobachten, einschließlich der Funktionsweise der für die Entscheidungsfindung verwendeten Algorithmen, sowie der Daten und der am Prozess beteiligten KI-Akteure, insbesondere wenn es um öffentlich angebotene Dienste geht und in Fällen, in denen eine direkte Interaktion mit dem Endnutzer oder der Endnutzerin erforderlich ist. Die ethischen Aspekte bei der Bewertung von KI-Systemen umfassen auch die menschenrechtlichen Verpflichtungen der Mitgliedstaaten.

(53) Die Regierungen sollten einen Regulierungsrahmen schaffen, der insbesondere für Behörden Verfahren zur Durchführung ethischer Folgenabschätzungen für KI-Systeme vorsieht, um Folgen vorherzusagen, Risiken zu mindern, schädliche Folgen zu vermeiden, die Beteiligung der Bürgerinnen und Bürger zu erleichtern und gesellschaftliche Herausforderungen anzugehen. Im Rahmen dieser Abschätzung sollten auch geeignete Aufsichtsmechanismen, unter anderem zur Gewährleistung von Überprüfbarkeit, Rückverfolgbarkeit und Erklärbarkeit, eingesetzt werden, die eine Bewertung von Algorithmen, Daten und Entwurfsprozessen ermöglichen; außerdem sollten KI-Systeme auch extern überprüft werden. Ethische Folgenabschätzungen sollten transparent sein und, wo dies angebracht ist, der Öffentlichkeit zugänglich gemacht werden. Solche Abschätzungen sollten außerdem viele Disziplinen, Kulturen und Interessengruppen einbeziehen, sie sollten pluralistisch und inklusiv sein. Gerade Behörden sollten verpflichtet werden, die von ihnen implementierten bzw. eingesetzten KI-Systeme durch die Einführung geeigneter Mechanismen und Instrumente zu überwachen.

(52) Member States and business enterprises should implement appropriate measures to monitor all phases of an AI system life cycle, including the functioning of algorithms used for decision-making, the data, as well as AI actors involved in the process, especially in public services and where direct end-user interaction is needed, as part of ethical impact assessment. Member States' human rights law obligations should form part of the ethical aspects of AI system assessments.

(53) Governments should adopt a regulatory framework that sets out a procedure, particularly for public authorities, to carry out ethical impact assessments on AI systems to predict consequences, mitigate risks, avoid harmful consequences, facilitate citizen participation and address societal challenges. The assessment should also establish appropriate oversight mechanisms, including auditability, traceability and explainability, which enable the assessment of algorithms, data and design processes, as well as include external review of AI systems. Ethical impact assessments should be transparent and open to the public, where appropriate. Such assessments should also be multidisciplinary, multi-stakeholder, multicultural, pluralistic and inclusive. The public authorities should be required to monitor the AI systems implemented and/or deployed by those authorities by introducing appropriate mechanisms and tools.

Politikbereich 2: Ethische Organisationsformen und Verantwortung

(54) Die Mitgliedstaaten sollten sicherstellen, dass die organisatorischen Mechanismen für KI inklusiv, transparent, multidisziplinär, multilateral (dies beinhaltet im Falle von Schäden auch die Möglichkeit abhilfeschaffender Maßnahmen und Rechtsmittel über Grenzen hinweg) und multistakeholderbasiert sind. Die organisatorischen Mechanismen sollten insbesondere Aspekte der Antizipation, des wirksamen Rechtsschutzes, der Überwachung von Auswirkungen, der Rechtsdurchsetzung und der Wiedergutmachung umfassen.

(55) Die Mitgliedstaaten sollten sicherstellen, dass durch KI-Systeme verursachte Schäden untersucht werden und Wiedergutmachung gewährt wird, indem sie wirkungsvolle Mechanismen zur Rechtsdurchsetzung und zum Rechtsschutz einführen, um zu gewährleisten, dass die Menschenrechte und Grundfreiheiten sowie die rechtsstaatlichen Prinzipien in der digitalen wie in der physischen Welt geachtet werden. Im Rahmen derartiger Mechanismen sollten Abhilfemaßnahmen sowohl von privatwirtschaftlichen als auch von öffentlich-rechtlichen Unternehmen gewährt werden. Zu diesem Zweck sollten die Überprüfbarkeit und die Rückverfolgbarkeit von KI-Systemen gefördert werden. Darüber hinaus sollten die Mitgliedstaaten ihre institutionellen Kapazitäten stärken, um dieser Verpflichtung nachzukommen, und mit Forschenden und anderen Akteuren zusammenarbeiten, um den potenziell böswilligen Einsatz von KI-Systemen zu untersuchen, zu unterbinden und die Folgen abzumildern.

Policy area 2: Ethical governance and stewardship

(54) Member States should ensure that AI governance mechanisms are inclusive, transparent, multidisciplinary, multilateral (this includes the possibility of mitigation and redress of harm across borders) and multi-stakeholder. In particular, governance should include aspects of anticipation, and effective protection, monitoring of impact, enforcement and redress.

(55) Member States should ensure that harms caused through AI systems are investigated and redressed, by enacting strong enforcement mechanisms and remedial actions, to make certain that human rights and fundamental freedoms and the rule of law are respected in the digital world and in the physical world. Such mechanisms and actions should include remediation mechanisms provided by private and public sector companies. The auditability and traceability of AI systems should be promoted to this end. In addition, Member States should strengthen their institutional capacities to deliver on this commitment and should collaborate with researchers and other stakeholders to investigate, prevent and mitigate any potentially malicious uses of AI systems.

(56) Die Mitgliedstaaten sind aufgerufen, je nach Sensibilität des Anwendungsbereichs und der erwarteten Auswirkungen auf die Menschenrechte, die Umwelt und die Ökosysteme sowie unter Berücksichtigung anderer in dieser Empfehlung dargelegter ethischer Erwägungen nationale und kontinentweite („regionale“) KI-Strategien zu entwickeln und Formen der weichen Organisation („soft governance“) in Betracht zu ziehen, wie etwa einen Zertifizierungsmechanismus für KI-Systeme und die gegenseitige Anerkennung ihrer Zertifizierung. Ein solcher Mechanismus könnte mehrstufige Überprüfungen der Systeme, der Daten und der Einhaltung ethischer Richtlinien und Verfahrenserfordernisse umfassen. Jedoch sollten dadurch weder Innovation behindert noch kleine und mittlere Unternehmen oder Start-ups, die Zivilgesellschaft sowie Forschungs- und Wissenschaftsorganisationen durch übermäßigen Verwaltungsaufwand belastet werden. Im Rahmen solcher Mechanismen sollte auch eine regelmäßige Beobachtung von KI-Systemen stattfinden, um die Robustheit der Systeme, ihre dauerhafte Integrität und die Einhaltung ethischer Standards über ihren gesamten Lebenszyklus hinweg sicherzustellen; dies bedeutet, dass gegebenenfalls Re-Zertifizierungen erforderlich sind.

(57) Die Mitgliedstaaten und ihre Behörden sollten eine transparente Selbstbewertung bestehender und geplanter KI-Systeme vornehmen, insbesondere zur Beurteilung der Angemessenheit des Einsatzes von KI; falls diese bejaht wird, sollte eine weitere Bewertung zur Ermittlung der geeigneten Methode erfolgen, sowie eine Beurteilung der Frage, ob ein solcher Einsatz zu Verstößen gegen die menschenrechtlichen Verpflichtungen der Mitgliedstaaten oder zu Missbrauch von Menschenrechten führen würde; in diesem Fall ist der Einsatz zu verbieten.

(56) Member States are encouraged to develop national and regional AI strategies and to consider forms of soft governance such as a certification mechanism for AI systems and the mutual recognition of their certification, according to the sensitivity of the application domain and expected impact on human rights, the environment and ecosystems, and other ethical considerations set forth in this Recommendation. Such a mechanism might include different levels of audit of systems, data, and adherence to ethical guidelines and to procedural requirements in view of ethical aspects. At the same time, such a mechanism should not hinder innovation or disadvantage small and medium enterprises or start-ups, civil society as well as research and science organizations, as a result of an excessive administrative burden. These mechanisms should also include a regular monitoring component to ensure system robustness and continued integrity and adherence to ethical guidelines over the entire life cycle of the AI system, requiring re-certification if necessary.

(57) Member States and public authorities should carry out transparent self-assessment of existing and proposed AI systems, which, in particular, should include the assessment of whether the adoption of AI is appropriate and, if so, should include further assessment to determine what the appropriate method is, as well as assessment as to whether such adoption would result in violations or abuses of Member States' human rights law obligations, and if that is the case, prohibit its use.

(58) Die Mitgliedstaaten sollten öffentlichen Stellen, privatwirtschaftlichen Unternehmen und Organisationen der Zivilgesellschaft nahelegen, unterschiedliche Interessengruppen in ihre KI-bezogene Organisation einzubinden und die Einsetzung eines bzw. einer unabhängigen KI-Ethikbeauftragten oder einen sonstigen Mechanismus in Erwägung zu ziehen, um die Aktivitäten im Bereich der ethischen Folgenabschätzung, der Überprüfung und kontinuierlichen Überwachung zu beaufsichtigen und eine ethische Orientierung beim Einsatz von KI-Systemen zu gewährleisten. Die Mitgliedstaaten, privatwirtschaftliche Unternehmen und Organisationen der Zivilgesellschaft sind aufgerufen, mit Unterstützung der UNESCO ein Netzwerk unabhängiger KI-Ethikbeauftragter aufzubauen, um diesen Prozess auf nationaler, kontinentweiter und globaler Ebene zu unterstützen.

(59) Die Mitgliedstaaten sollten auf nationaler Ebene den Aufbau eines digitalen Ökosystems für eine ethische und inklusive Entwicklung von KI-Systemen und den Zugang zu diesem digitalen Ökosystem fördern, auch um Lücken beim Zugang zum Lebenszyklus von KI-Systemen zu schließen; parallel dazu sollten sie zur internationalen Zusammenarbeit beitragen. Zu einem solchen Ökosystem gehören insbesondere digitale Technologien und Infrastrukturen sowie gegebenenfalls Mechanismen für den Austausch von KI-Wissen.

(58) Member States should encourage public entities, private sector companies and civil society organizations to involve different stakeholders in their AI governance and to consider adding the role of an independent AI Ethics Officer or some other mechanism to oversee ethical impact assessment, auditing and continuous monitoring efforts and ensure ethical guidance of AI systems. Member States, private sector companies and civil society organizations, with the support of UNESCO, are encouraged to create a network of independent AI Ethics Officers to give support to this process at national, regional and international levels.

(59) Member States should foster the development of, and access to, a digital ecosystem for ethical and inclusive development of AI systems at the national level, including to address gaps in access to the AI system life cycle, while contributing to international collaboration. Such an ecosystem includes, in particular, digital technologies and infrastructure, and mechanisms for sharing AI knowledge, as appropriate.

(60) Die Mitgliedstaaten sollten in Zusammenarbeit mit internationalen Organisationen, transnationalen Unternehmen, akademischen Einrichtungen und der Zivilgesellschaft Mechanismen einführen, die eine aktive Beteiligung aller Mitgliedstaaten, insbesondere der LMICs, und ganz besonders der LDCs, LLDCs und SIDS, an den internationalen Diskussionen über KI-bezogene Organisationsformen sicherstellen. Dies kann beispielsweise durch Bereitstellung von finanziellen Mitteln geschehen, um zu gewährleisten, dass alle Weltregionen die gleichen Möglichkeiten zur Beteiligung haben, oder durch andere Maßnahmen. Um eine inklusive Zusammensetzung von KI-Foren sicherzustellen, sollten die Mitgliedstaaten außerdem dafür sorgen, dass KI-Akteure, insbesondere solche aus LMICs, und ganz besonders aus LDCs, LLDCs und SIDS, ungehindert aus ihrem jeweiligen Mitgliedstaat ausreisen und in andere Mitgliedstaaten einreisen können, um an solchen Foren teilzunehmen.

(61) Änderungen an bestehenden und die Formulierung neuer nationaler Rechtsvorschriften zu KI-Systemen müssen den menschenrechtlichen Verpflichtungen der Mitgliedstaaten Rechnung tragen und die Menschenrechte und Grundfreiheiten über den gesamten Lebenszyklus von KI-Systemen fördern. Diese Förderung sollte auch in Form von organisatorischen Initiativen, Beispielen für gute Zusammenarbeit bei KI-Systemen und, im Zuge der Weiterentwicklung von KI-Technologien, in Form der Ausarbeitung nationaler und internationaler technischer und methodischer Leitlinien erfolgen. Bei ihrem Vorgehen im Zusammenhang mit KI-Systemen müssen die unterschiedlichen Sektoren, auch der Privatsektor, die Menschenrechte und Grundfreiheiten achten, schützen und fördern, indem sie sowohl bestehende als auch neue Instrumente in Verbindung mit dieser Empfehlung nutzen.

(60) Member States should establish mechanisms, in collaboration with international organizations, transnational corporations, academic institutions and civil society, to ensure the active participation of all Member States, especially LMICs, in particular LDCs, LLDCs and SIDS, in international discussions concerning AI governance. This can be through the provision of funds, ensuring equal regional participation, or any other mechanisms. Furthermore, in order to ensure the inclusiveness of AI fora, Member States should facilitate the travel of AI actors in and out of their territory, especially from LMICs, in particular LDCs, LLDCs and SIDS, for the purpose of participating in these fora.

(61) Amendments to the existing or elaboration of new national legislation addressing AI systems must comply with Member States' human rights law obligations and promote human rights and fundamental freedoms throughout the AI system life cycle. Promotion thereof should also take the form of governance initiatives, good exemplars of collaborative practices regarding AI systems, and national and international technical and methodological guidelines as AI technologies advance. Diverse sectors, including the private sector, in their practices regarding AI systems must respect, protect and promote human rights and fundamental freedoms using existing and new instruments in combination with this Recommendation.

(62) Mitgliedstaaten, die KI-Systeme für menschenrechtssensible Anwendungsfälle wie Strafverfolgung, Sozialfürsorge, Beschäftigung, Medien und Informationsangebote, Gesundheitsversorgung oder das unabhängige Justizsystem einführen, sollten Mechanismen zur Beobachtung der sozialen und wirtschaftlichen Auswirkungen solcher Systeme durch geeignete Aufsichtsbehörden wie zum Beispiel unabhängige Datenschutzbehörden, Einrichtungen der Branchenaufsicht und öffentliche Aufsichtsorgane etablieren.

(63) Die Mitgliedstaaten sollten die Kapazitäten ihrer Justiz stärken, Entscheidungen in Bezug auf KI-Systeme nach rechtsstaatlichen Prinzipien und im Einklang mit dem Völkerrecht und internationalen Standards zu treffen, auch bei der Nutzung von KI in ihren Entscheidungsprozessen; dabei ist sicherzustellen, dass der Grundsatz der Aufsicht durch Menschen gewahrt bleibt. Werden KI-Systeme von der Justiz eingesetzt, so sind ausreichende Sicherungen vorzusehen, damit unter anderem der Schutz der grundlegenden Menschenrechte, die Rechtsstaatlichkeit, die richterliche Unabhängigkeit sowie der Grundsatz der menschlichen Kontrolle und Aufsicht gewährleistet sind und eine vertrauenswürdige, gemeinwohlorientierte und menschenzentrierte Entwicklung und Nutzung von KI-Systemen in der Justiz sichergestellt ist.

(62) Member States that acquire AI systems for human rights-sensitive use cases, such as law enforcement, welfare, employment, media and information providers, health care and the independent judiciary system should provide mechanisms to monitor the social and economic impact of such systems by appropriate oversight authorities, including independent data protection authorities, sectoral oversight and public bodies responsible for oversight.

(63) Member States should enhance the capacity of the judiciary to make decisions related to AI systems as per the rule of law and in line with international law and standards, including in the use of AI systems in their deliberations, while ensuring that the principle of human oversight is upheld. In case AI systems are used by the judiciary, sufficient safeguards are needed to guarantee inter alia the protection of fundamental human rights, the rule of law, judicial independence as well as the principle of human oversight, and to ensure a trustworthy, public interest-oriented and human-centric development and use of AI systems in the judiciary.

(64) Die Mitgliedstaaten sollten dafür sorgen, dass Regierungen und multilaterale Organisationen eine führende Rolle bei der Gewährleistung von Schutz und Sicherheit von KI-Systemen spielen, wobei eine Vielzahl von Interessengruppen einbezogen werden sollte. Konkret sollten die Mitgliedstaaten, internationale Organisationen und andere relevante Gremien internationale Standards entwickeln, die mess- und überprüfbare Sicherheits- und Transparenzniveaus beschreiben, so dass Systeme objektiv bewertet werden können, und um festzustellen, inwieweit diese den Standards entsprechen. Zudem sollten die Mitgliedstaaten wie die Unternehmen strategische Forschung zu potenziellen Sicherheitsrisiken und Schwachstellen von KI-Technologien dauerhaft unterstützen und Forschung zu Transparenz und Erklärbarkeit, Inklusion und Kompetenzbildung fördern, indem sie dafür in verschiedenen Bereichen und für unterschiedliche Niveaus, beispielsweise Fachsprache und Gemeinsprache, zusätzliche Mittel bereitstellen.

(65) Die Mitgliedstaaten sollten Konzepte und Vorgaben umsetzen, die sicherstellen, dass das Handeln von KI-Akteuren über den gesamten Lebenszyklus von KI-Systemen mit internationalen Menschenrechtsrahmen, -normen und -grundsätzen vereinbar ist. Dabei sind die bestehende kulturelle und soziale Vielfalt einschließlich lokaler Bräuche und religiöser Traditionen gebührend zu berücksichtigen, es gilt jedoch der Vorrang der universellen Menschenrechte.

(64) Member States should ensure that governments and multilateral organizations play a leading role in ensuring the safety and security of AI systems, with multistakeholder participation. Specifically, Member States, international organizations and other relevant bodies should develop international standards that describe measurable, testable levels of safety and transparency, so that systems can be objectively assessed and levels of compliance determined. Furthermore, Member States and business enterprises should continuously support strategic research on potential safety and security risks of AI technologies and should encourage research into transparency and explainability, inclusion and literacy by putting additional funding into those areas for different domains and at different levels, such as technical and natural language.

(65) Member States should implement policies to ensure that the actions of AI actors are consistent with international human rights law, standards and principles throughout the life cycle of AI systems, while taking into full consideration the current cultural and social diversities, including local customs and religious traditions, with due regard to the precedence and universality of human rights.

(66) Die Mitgliedstaaten sollten Mechanismen einsetzen, die von KI-Akteuren verlangen, jede Art von Stereotypisierung – unabhängig davon, ob diese konzeptionell angelegt oder durch Nachlässigkeit entstanden ist – in den Ergebnissen von KI-Systemen und -Daten offenzulegen und zu bekämpfen, und die sicherstellen, dass die Trainingsdatensätze für KI-Systeme nicht kulturelle, wirtschaftliche oder soziale Ungleichheiten und Vorurteile, die Verbreitung von Desinformation und Misinformation sowie die Behinderung der Meinungsfreiheit und den Zugang zu Informationen begünstigen. Besonderes Augenmerk sollte dabei auf Weltregionen gelegt werden, in denen es nur wenige Daten gibt.

(67) Die Mitgliedstaaten sollten Konzepte und Vorgaben umsetzen, die dafür sorgen, dass Vielfalt und Inklusion in den KI-Entwicklungsteams und Trainingsdatensätzen gemäß der Zusammensetzung ihrer Bevölkerung besser abgebildet werden, und die den gleichberechtigten Zugang zu KI-Technologien und ihrem Nutzen sicherstellen, insbesondere für marginalisierte Gruppen und sowohl in ländlichen als auch in städtischen Regionen.

(66) Member States should put in place mechanisms to require AI actors to disclose and combat any kind of stereotyping in the outcomes of AI systems and data, whether by design or by negligence, and to ensure that training data sets for AI systems do not foster cultural, economic or social inequalities, prejudice, the spreading of disinformation and misinformation, and disruption of freedom of expression and access to information. Particular attention should be given to regions where the data are scarce.

(67) Member States should implement policies to promote and increase diversity and inclusiveness that reflect their populations in AI development teams and training datasets, and to ensure equal access to AI technologies and their benefits, particularly for marginalized groups, both from rural and urban zones.

(68) Die Mitgliedstaaten sollten Regulierungsrahmen entwickeln bzw. vorhandene überprüfen und gegebenenfalls anpassen, um Rechenschaftspflicht und Verantwortung für den Inhalt und die Resultate von KI-Systemen in den verschiedenen Phasen ihres Lebenszyklus sicherzustellen. Soweit erforderlich, sollten die Mitgliedstaaten Haftungsregelungen einführen oder die Auslegung bestehender Regelungen klären, damit die Verantwortung für die Ergebnisse und die Arbeitsweise von KI-Systemen eindeutig zugewiesen werden kann. Darüber hinaus müssen die Mitgliedstaaten bei der Entwicklung von Rechtsrahmen insbesondere beachten, dass die letztendliche Verantwortung und Rechenschaftspflicht immer bei natürlichen oder juristischen Personen liegen muss und dass einem KI-System keine eigene Rechtspersönlichkeit zugestanden werden darf. Um dies zu gewährleisten, sollten solche Regulierungsrahmen den Grundsatz der Aufsicht durch Menschen betonen und einen umfassenden Ansatz etablieren, der KI-Akteure und die beteiligten technologischen Prozesse in den verschiedenen Phasen des Lebenszyklus von KI-Systemen in den Mittelpunkt stellt.

(68) Member States should develop, review and adapt, as appropriate, regulatory frameworks to achieve accountability and responsibility for the content and outcomes of AI systems at the different phases of their life cycle. Member States should, where necessary, introduce liability frameworks or clarify the interpretation of existing frameworks to ensure the attribution of accountability for the outcomes and the functioning of AI systems. Furthermore, when developing regulatory frameworks, Member States should, in particular, take into account that ultimate responsibility and accountability must always lie with natural or legal persons and that AI systems should not be given legal personality themselves. To ensure this, such regulatory frameworks should be consistent with the principle of human oversight and establish a comprehensive approach focused on AI actors and the technological processes involved across the different stages of the AI system life cycle.

(69) Bei der Ausarbeitung neuer Normen, falls sie noch nicht existieren, oder der Anpassung bestehender Rechtsrahmen sollten die Mitgliedstaaten sämtliche KI-Akteure einbeziehen (u. a., jedoch nicht ausschließlich aus den Bereichen Forschung, Zivilgesellschaft, Justiz und Versicherungswesen, außerdem Investorinnen und Investoren, Herstellerinnen und Hersteller, Ingenieurinnen und Ingenieure, Anwältinnen und Anwälte sowie Nutzerinnen und Nutzer). Günstigenfalls gehen aus solchen Normen bewährte Verfahren, Gesetze und Vorschriften hervor. Die Mitgliedstaaten werden ferner ermutigt, mit Mechanismen wie „Strategie-Prototypen“ und ordnungspolitischen Reallaboren zu arbeiten, um die Entwicklung von Gesetzen, Vorschriften und Strategien und deren regelmäßige Überprüfung zu beschleunigen, damit sie mit der raschen Entwicklung neuer Technologien Schritt halten, und um zu gewährleisten, dass Gesetze und Vorschriften vor ihrem formalen Beschluss zunächst in einem sicheren Umfeld getestet werden können. Die Mitgliedstaaten sollten Regierungen auf Landes- bzw. kommunaler Ebene bei der Entwicklung lokaler Strategien, Vorschriften und Gesetze unterstützen, die mit nationalen und internationalen Rechtsrahmen vereinbar sind.

(70) Die Mitgliedstaaten sollten klare Anforderungen an die Transparenz und Erklärbarkeit von KI-Systemen stellen, um die Vertrauenswürdigkeit von KI-Systemen über ihren gesamten Lebenszyklus zu gewährleisten. Diese Anforderungen sollten sich auch auf die Gestaltung und Umsetzung von Wirkmechanismen beziehen, die der Natur des Anwendungsbereichs, der beabsichtigten Verwendung, der Zielgruppe und der Realisierbarkeit eines jeden KI-Systems Rechnung tragen.

(69) In order to establish norms where these do not exist, or to adapt the existing legal frameworks, Member States should involve all AI actors (including, but not limited to, researchers, representatives of civil society and law enforcement, insurers, investors, manufacturers, engineers, lawyers and users). The norms can mature into best practices, laws and regulations. Member States are further encouraged to use mechanisms such as policy prototypes and regulatory sandboxes to accelerate the development of laws, regulations and policies, including regular reviews thereof, in line with the rapid development of new technologies and ensure that laws and regulations can be tested in a safe environment before being officially adopted. Member States should support local governments in the development of local policies, regulations and laws in line with national and international legal frameworks.

(70) Member States should set clear requirements for AI system transparency and explainability so as to help ensure the trustworthiness of the full AI system life cycle. Such requirements should involve the design and implementation of impact mechanisms that take into consideration the nature of application domain, intended use, target audience and feasibility of each particular AI system.

Politikbereich 3: Datenpolitik

(71) Die Mitgliedstaaten sollten Strategien zum Datenqualitätsmanagement entwickeln, die eine kontinuierliche Bewertung der Qualität der Trainingsdaten für KI-Systeme gewährleisten, einschließlich der Angemessenheit der Datenerhebungs- und -auswahlverfahren, geeigneter Datensicherheits- und -schutzmaßnahmen sowie Feedback-Mechanismen, um aus Fehlern zu lernen und bewährte Verfahren zwischen allen KI-Akteuren auszutauschen.

(72) Die Mitgliedstaaten sollten geeignete Sicherungen zum Schutz des Rechts auf Privatsphäre im Einklang mit dem Völkerrecht einführen und dabei auch auf Bedenken hinsichtlich möglicher Überwachung der Bevölkerung eingehen. Unter anderem sollten sie rechtliche Regelwerke einführen bzw. durchsetzen, die einen angemessenen Schutz im Einklang mit dem Völkerrecht bieten. Sie sind gehalten, alle KI-Akteure, auch privatwirtschaftliche Unternehmen, nachdrücklich dazu aufzufordern, die bestehenden internationalen Standards zu befolgen und im Rahmen ethischer Folgenabschätzungen, die die weiteren sozioökonomischen Auswirkungen der beabsichtigten Datenverarbeitung berücksichtigen, insbesondere angemessene Privatsphäre-Folgenabschätzungen durchzuführen, und in ihren Systemen das Konzept der eingebauten Privatsphäre („privacy by design“) umzusetzen. Die Privatsphäre sollte über den gesamten Lebenszyklus von KI-Systemen geachtet, geschützt und gefördert werden.

Policy area 3: Data policy

(71) Member States should work to develop data governance strategies that ensure the continual evaluation of the quality of training data for AI systems including the adequacy of the data collection and selection processes, proper data security and protection measures, as well as feedback mechanisms to learn from mistakes and share best practices among all AI actors.

(72) Member States should put in place appropriate safeguards to protect the right to privacy in accordance with international law, including addressing concerns such as surveillance. Member States should, among others, adopt or enforce legislative frameworks that provide appropriate protection, compliant with international law. Member States should strongly encourage all AI actors, including business enterprises, to follow existing international standards and, in particular, to carry out adequate privacy impact assessments, as part of ethical impact assessments, which take into account the wider socio-economic impact of the intended data processing, and to apply privacy by design in their systems. Privacy should be respected, protected and promoted throughout the life cycle of AI systems.

(73) Die Mitgliedstaaten sollten sicherstellen, dass jeder Mensch die Rechte an seinen persönliche Daten behält und diese durch einen entsprechenden Rahmen geschützt werden, der insbesondere Folgendes vorsieht: Transparenz; angemessene Garantien hinsichtlich der Verarbeitung sensibler Daten; ein angemessenes Datenschutzniveau; wirksame und sinnvolle Systeme und Mechanismen zur Gewährleistung der Rechenschaftspflicht; uneingeschränkte Ausübung der den Betroffenen zustehenden Rechte einschließlich der Möglichkeit, auf ihre persönlichen Daten in KI-Systemen zuzugreifen und sie zu löschen, soweit dem nicht bestimmte Umstände im Einklang mit dem Völkerrecht entgegenstehen; ein angemessenes Schutzniveau unter vollständiger Einhaltung der Datenschutzvorschriften, wenn Daten zu kommerziellen Zwecken verwendet werden, zum Beispiel zur Ermöglichung sehr gezielter Werbung oder bei grenzüberschreitender Übermittlung; sowie eine wirksame unabhängige Aufsicht im Rahmen einer operationalisierten Datenpolitik, die sicherstellt, dass jede betroffene Person die Kontrolle über ihre persönlichen Daten behält, und der die Vorteile eines freien internationalen Informationsflusses begünstigt, einschließlich des Zugangs zu Daten.

(73) Member States should ensure that individuals retain rights over their personal data and are protected by a framework, which notably foresees: transparency; appropriate safeguards for the processing of sensitive data; an appropriate level of data protection; effective and meaningful accountability schemes and mechanisms; the full enjoyment of the data subjects' rights and the ability to access and erase their personal data in AI systems, except for certain circumstances in compliance with international law; an appropriate level of protection in full compliance with data protection legislation where data are being used for commercial purposes such as enabling micro-targeted advertising, transferred crossborder; and an effective independent oversight as part of a data governance mechanism which keeps individuals in control of their personal data and fosters the benefits of a free flow of information internationally, including access to data.

(74) Die Mitgliedstaaten sollten Datenschutzrichtlinien oder gleichwertige Regelwerke erstellen oder bestehende Regelungsrahmen verstärken, um den vollständigen Schutz personenbezogener und sensibler Daten zu gewährleisten, die im Falle ihrer Offenlegung zu gravierenden Schäden, Verletzungen oder Härten für Einzelpersonen führen können. Beispiele hierfür sind Daten über Gesetzesverstöße, Strafverfahren und Verurteilungen und damit zusammenhängende Sicherungsmaßnahmen, biometrische, genetische und Gesundheitsdaten sowie personenbezogene Daten zu ethnischer Herkunft, Hautfarbe, Abstammung, Geschlecht, Alter, Sprache, Religion, politischer Überzeugung, nationaler oder sozialer Herkunft, wirtschaftlichem oder sozialem Geburtsstatus, Behinderung oder sonstigen Merkmalen.

(75) Die Mitgliedstaaten sollten sich für offene Daten einsetzen. In dieser Hinsicht sollten sie eine Überprüfung ihrer Richtlinien und Rechtsrahmen in Betracht ziehen, auch hinsichtlich des Zugangs zu Informationen und einer offenen Verwaltung, damit den KI-spezifischen Anforderungen Rechnung getragen wird und Mechanismen wie beispielsweise offene Datenrepositorien für öffentlich finanzierte oder in öffentlichem Besitz befindliche Daten und Quellcodes sowie Datentreuhandmodelle gefördert werden, um eine sichere, faire, legale und ethische geteilte Nutzung von Daten zu unterstützen.

(74) Member States should establish their data policies or equivalent frameworks, or reinforce existing ones, to ensure full security for personal data and sensitive data, which, if disclosed, may cause exceptional damage, injury or hardship to individuals. Examples include data relating to offences, criminal proceedings and convictions, and related security measures; biometric, genetic and health data; and -personal data such as that relating to race, colour, descent, gender, age, language, religion, political opinion, national origin, ethnic origin, social origin, economic or social condition of birth, or disability and any other characteristics.

(75) Member States should promote open data. In this regard, Member States should consider reviewing their policies and regulatory frameworks, including on access to information and open government to reflect AI-specific requirements and promoting mechanisms, such as open repositories for publicly funded or publicly held data and source code and data trusts, to support the safe, fair, legal and ethical sharing of data, among others.

(76) Die Mitgliedstaaten sollten dafür sorgen und sich dafür einsetzen, dass für Training, Entwicklung und Nutzung von KI-Systemen qualitativ hochwertige und belastbare Datensätze verwendet werden, und die Erhebung und Nutzung dieser Daten sorgsam überwachen. Soweit möglich und praktikabel könnten sie in diesem Zusammenhang auch in die Erstellung von Datensätzen investieren, die einem Goldstandard entsprechen, einschließlich offener und vertrauenswürdiger Datensätze, die divers sind und auf einer gültigen Rechtsgrundlage basieren, wozu auch die Zustimmung der Betroffenen gehört, soweit diese gesetzlich vorgeschrieben ist. Die Mitgliedstaaten sollten dazu aufrufen, Standards für die Kommentierung von Datensätzen zu schaffen, einschließlich einer Aufschlüsselung von Daten nach Geschlecht und anderen Merkmalen, so dass leicht festgestellt werden kann, wie ein Datensatz erhoben wurde und welche Eigenschaften er hat.

(77) Wie im Bericht des High-level Panel on Digital Cooperation (Hochrangiges Gremium für digitale Zusammenarbeit) des Generalsekretärs der Vereinten Nationen vorgeschlagen, sollten die Mitgliedstaaten mit Unterstützung der Vereinten Nationen und der UNESCO dort, wo dies sinnvoll ist, einen Digital-Commons-Ansatz für Daten verfolgen, die Interoperabilität von Instrumenten und Datensätzen sowie Schnittstellen von Systemen verbessern, in denen Daten gehostet werden, sowie privatwirtschaftliche Unternehmen dazu anhalten, die von ihnen erhobenen Daten, soweit angemessen, mit sämtlichen Interessengruppen etwa für Forschungs-, Innovations- oder Gemeinwohlzwecke auszutauschen. Außerdem sollten sie öffentliche und private Bemühungen zur Schaffung kollaborativer Plattformen für den Austausch hochwertiger Daten in vertrauenswürdigen und gesicherten Datenräumen unterstützen.

(76) Member States should promote and facilitate the use of quality and robust datasets for training, development and use of AI systems, and exercise vigilance in overseeing their collection and use. This could, if possible and feasible, include investing in the creation of gold standard datasets, including open and trustworthy datasets, which are diverse, constructed on a valid legal basis, including consent of data subjects, when required by law. Standards for annotating datasets should be encouraged, including disaggregating data on gender and other bases, so it can easily be determined how a dataset is gathered and what properties it has.

(77) Member States, as also suggested in the report of the United Nations Secretary-General's High-level Panel on Digital Cooperation, with the support of the United Nations and UNESCO, should adopt a digital commons approach to data where appropriate, increase interoperability of tools and datasets and interfaces of systems hosting data, and encourage private sector companies to share the data they collect with all stakeholders, as appropriate, for research, innovation or public benefits. They should also promote public and private efforts to create collaborative platforms to share quality data in trusted and secured data spaces.

Politikbereich 4: Entwicklungs- und internationale Zusammenarbeit

(78) Mitgliedstaaten und transnationale Unternehmen sollten der KI-Ethik eine hohe Priorität einräumen, indem sie Diskussionen über KI-bezogene ethische Fragen in die einschlägigen Foren der internationalen, zwischenstaatlichen und multistakeholderbasierten Zusammenarbeit aufnehmen.

(79) Die Mitgliedstaaten sollten sicherstellen, dass der Einsatz von KI in Bereichen der Entwicklungszusammenarbeit wie beispielsweise Bildung, Wissenschaft, Kultur, Kommunikation und Information, Gesundheitsversorgung, Landwirtschaft und Nahrungsmittelversorgung, Umwelt, Bewirtschaftung natürlicher Ressourcen und von Infrastruktur sowie Wirtschaftsplanung und -wachstum den in dieser Empfehlung dargelegten Werten und Grundsätzen entspricht.

(80) Die Mitgliedstaaten sollten über internationale Organisationen Plattformen für die internationale Zusammenarbeit zu KI für Entwicklungszwecke bereitstellen, unter anderem indem sie Expertise, Finanzmittel, Daten, Fachwissen und Infrastruktur beitragen und die Zusammenarbeit zwischen verschiedenen Interessengruppen zur Bewältigung schwieriger Entwicklungsprobleme erleichtern, insbesondere für LMICs und ganz besonders für LDCs, LLDCs und SIDS.

(81) Die Mitgliedstaaten sollten sich für internationale Zusammenarbeit in der KI-Forschung und -Innovation einsetzen; dies schließt auch den Einsatz für Forschungs- und Innovationszentren und -netzwerke ein, die eine stärkere Beteiligung und Führungsrolle von Forschenden aus LMICs und anderen Ländern einschließlich LDCs, LLDCs und SIDS fördern.

Policy area 4: Development and international cooperation

(78) Member States and transnational corporations should prioritize AI ethics by including discussions of AI-related ethical issues into relevant international, intergovernmental and multi-stakeholder fora.

(79) Member States should ensure that the use of AI in areas of development such as education, science, culture, communication and information, health care, agriculture and food supply, environment, natural resource and infrastructure management, economic planning and growth, among others, adheres to the values and principles set forth in this Recommendation.

(80) Member States should work through international organizations to provide platforms for international cooperation on AI for development, including by contributing expertise, funding, data, domain knowledge, infrastructure, and facilitating multi-stakeholder collaboration to tackle challenging development problems, especially for LMICs, in particular LDCs, LLDCs and SIDS.

(81) Member States should work to promote international collaboration on AI research and innovation, including research and innovation centres and networks that promote greater participation and leadership of researchers from LMICs and other countries, including LDCs, LLDCs and SIDS.

(82) Die Mitgliedstaaten sollten internationale Organisationen und Forschungseinrichtungen sowie transnationale Unternehmen einbinden, um Forschung im Bereich der KI-Ethik zu fördern, auf deren Grundlage eine ethische Nutzung von KI-Systemen durch öffentliche und private Einrichtungen basieren kann; in diesem Rahmen sollte auch dazu geforscht werden, inwieweit spezifische ethische Leitplanken in bestimmten Kulturen und Kontexten anwendbar sind und welche Möglichkeiten es gibt, technologisch praktikable Lösungen im Einklang mit diesen Leitplanken zu entwickeln.

(83) Die Mitgliedstaaten sollten die internationale Zusammenarbeit im Bereich der KI fördern, um Unterscheidungen nach Weltregion und technischem Entwicklungsstand zu überwinden. Technischer Austausch und Beratungen zwischen den Mitgliedstaaten und ihren Bevölkerungen, zwischen öffentlichem und privatem Sektor sowie zwischen den technisch am weitesten und den am wenigsten fortgeschrittenen Ländern sollten unter voller Achtung des Völkerrechts stattfinden.

(82) Member States should promote AI ethics research by engaging international organizations and research institutions, as well as transnational corporations, that can be a basis for the ethical use of AI systems by public and private entities, including research into the applicability of specific ethical frameworks in specific cultures and contexts, and the possibilities to develop technologically feasible solutions in line with these frameworks.

(83) Member States should encourage international cooperation and collaboration in the field of AI to bridge geo-technological lines. Technological exchanges and consultations should take place between Member States and their populations, between the public and private sectors, and between and among the most and least technologically advanced countries in full respect of international law.

Politikbereich 5: Umwelt und Ökosysteme

(84) Die Mitgliedstaaten und privatwirtschaftliche Unternehmen sollten die unmittelbaren und mittelbaren ökologischen Auswirkungen von KI-Systemen, unter anderem, jedoch nicht ausschließlich hinsichtlich ihrer CO₂-Bilanz und ihres Energieverbrauchs, sowie der ökologischen Auswirkungen der Rohstoffgewinnung für die Herstellung von KI-Technologien über ihren gesamten Lebenszyklus untersuchen, und sie sollten die Auswirkungen von KI-Systemen und Dateninfrastrukturen auf die Umwelt verringern. Die Mitgliedstaaten sollten sicherstellen, dass sich alle KI-Akteure an Umweltvorschriften und -verfahren halten.

(85) Die Mitgliedstaaten sollten Anreize schaffen, soweit dies notwendig und sinnvoll erscheint, um zu erreichen, dass rechthebasierte und ethische KI-gestützte Lösungen entwickelt und eingeführt werden, die der Stärkung der Widerstandsfähigkeit gegen Katastrophen, dem Schutz, der Beobachtung und Wiederherstellung der Umwelt und der Ökosysteme sowie dem Erhalt des Planeten dienen. Während des gesamten Lebenszyklus dieser KI-Systeme sollten lokale und indigene Gemeinschaften einbezogen und Konzepte wie Kreislaufwirtschaft sowie nachhaltige Verbrauchs- und Produktionsmuster unterstützt werden. Hier einige Beispiele, wozu KI-Systeme, soweit erforderlich und angemessen, eingesetzt werden könnten:

- a) Unterstützung von Schutz, Beobachtung und Bewirtschaftung natürlicher Ressourcen;
- b) Unterstützung der Vorhersage, Vorbeugung, Kontrolle und Minderung klimabedingter Probleme;
- c) Unterstützung einer effizienteren und nachhaltigeren Lebensmittel-Erzeugung;

Policy area 5: Environment and ecosystems

(84) Member States and business enterprises should assess the direct and indirect environmental impact throughout the AI system life cycle, including, but not limited to, its carbon footprint, energy consumption and the environmental impact of raw material extraction for supporting the manufacturing of AI technologies, and reduce the environmental impact of AI systems and data infrastructures. Member States should ensure compliance of all AI actors with environmental law, policies and practices.

(85) Member States should introduce incentives, when needed and appropriate, to ensure the development and adoption of rights-based and ethical AI-powered solutions for disaster risk resilience; the monitoring, protection and regeneration of the environment and ecosystems; and the preservation of the planet. These AI systems should involve the participation of local and indigenous communities throughout the life cycle of AI systems and should support circular economy type approaches and sustainable consumption and production patterns. Some examples include using AI systems, when needed and appropriate, to:

- a) Support the protection, monitoring and management of natural resources.
- b) Support the prediction, prevention, control and mitigation of climate-related problems.
- c) Support a more efficient and sustainable food ecosystem.

- d) Unterstützung eines beschleunigten Zugangs zu nachhaltiger Energie und deren Einführung auf breiter Basis;
- e) Ermöglichung und Förderung der durchgehenden Verankerung nachhaltiger Infrastrukturen, nachhaltiger Geschäftsmodelle und eines nachhaltigen Finanzsystems für eine nachhaltige Entwicklung;
- f) Erkennung von umweltschädlichen Stoffen oder Vorhersage des Grades der Umweltbelastung zur Unterstützung relevanter Interessengruppen bei der Identifizierung, Planung und Durchführung gezielter Maßnahmen zur Vorbeugung und Verringerung von Verschmutzung und dass man ihr ausgesetzt ist.

(86) In Anbetracht des potenziell daten- bzw. ressourcenintensiven Charakters einiger KI-Methoden und ihrer jeweiligen Auswirkungen auf die Umwelt sollten die Mitgliedstaaten bei der Auswahl dieser Methoden darauf achten, dass die KI-Akteure im Einklang mit dem Grundsatz der Verhältnismäßigkeit daten-, energie- und ressourceneffizienten KI-Methoden den Vorzug geben. Es sollten Anforderungen entwickelt werden, die verlangen, dass mittels geeigneter Nachweise gezeigt werden muss, dass eine KI-Anwendung die beabsichtigte Wirkung erzielen wird, oder dass die mit einer KI-Anwendung verbundenen Sicherheitsvorkehrungen ihren Einsatz rechtfertigen können. Ist dies nicht möglich, sollte nach dem Vorsorgeprinzip gehandelt, und in Fällen, in denen unverhältnismäßig negative Auswirkungen auf die Umwelt zu erwarten sind, auf den Einsatz von KI verzichtet werden.

- d) Support the acceleration of access to and mass adoption of sustainable energy.
- e) Enable and promote the mainstreaming of sustainable infrastructure, sustainable business models and sustainable finance for sustainable development.
- f) Detect pollutants or predict levels of pollution and thus help relevant stakeholders identify, plan and put in place targeted interventions to prevent and reduce pollution and exposure.

(86) When choosing AI methods, given the potential data-intensive or resource-intensive character of some of them and the respective impact on the environment, Member States should ensure that AI actors, in line with the principle of proportionality, favour data, energy and resource-efficient AI methods. Requirements should be developed to ensure that appropriate evidence is available to show that an AI application will have the intended effect, or that safeguards accompanying an AI application can support the justification for its use. If this cannot be done, the precautionary principle must be favoured, and in instances where there are disproportionate negative impacts on the environment, AI should not be used.

Politikbereich 6: Gleichstellung der Geschlechter

(87) Die Mitgliedstaaten sollten sicherstellen, dass das Potenzial digitaler Technologien und künstlicher Intelligenz, zur Gleichstellung der Geschlechter beizutragen, in vollem Umfang ausgeschöpft wird. Sie müssen gewährleisten, dass die Menschenrechte und Grundfreiheiten von Mädchen und Frauen sowie ihre Sicherheit und Unversehrtheit in keiner Phase des Lebenszyklus von KI-Systemen verletzt werden. Darüber hinaus sollte die ethische Folgenabschätzung eine übergreifende Geschlechterperspektive beinhalten.

(88) Die Mitgliedstaaten sollten zweckgebundene Mittel aus ihren öffentlichen Haushalten für die Finanzierung von geschlechtergerechten Programmen bereitstellen, dafür sorgen, dass ihre nationale Digitalpolitik einen Aktionsplan für die Gleichstellung der Geschlechter vorsieht, und geeignete Strategien entwickeln, beispielsweise im Bereich der beruflichen Bildung, die Mädchen und Frauen fördern und sicherstellen, dass diese in der durch die KI vorangetriebenen digitalen Wirtschaft nicht zurückgelassen werden. Es sollten gezielte Investitionen in spezielle Programme und geschlechterspezifische Sprache erwogen und umgesetzt werden, um eine stärkere Beteiligung von Frauen und Mädchen in Disziplinen wie Mathematik, Informatik, Naturwissenschaft und Technik (MINT), einschließlich der Informations- und Kommunikationstechnologien (IKT) zu erreichen, sie besser auf den Arbeitsmarkt vorzubereiten, ihre gleichberechtigten Karriere- und Aufstiegsmöglichkeiten zu verbessern und die berufliche Weiterentwicklung von Frauen und Mädchen zu fördern.

Policy area 6: Gender

(87) Member States should ensure that the potential for digital technologies and artificial intelligence to contribute to achieving gender equality is fully maximized, and must ensure that the human rights and fundamental freedoms of girls and women, and their safety and integrity are not violated at any stage of the AI system life cycle. Moreover, Ethical Impact Assessment should include a transversal gender perspective.

(88) Member States should have dedicated funds from their public budgets linked to financing gender-responsive schemes, ensure that national digital policies include a gender action plan, and develop relevant policies, for example, on labour education, targeted at supporting girls and women to make sure they are not left out of the digital economy powered by AI. Special investment in providing targeted programmes and genderspecific language, to increase the opportunities of girls' and women's participation in science, technology, engineering, and mathematics (STEM), including information and communication technologies (ICT) disciplines, preparedness, employability, equal career development and professional growth of girls and women, should be considered and implemented.

(89) Die Mitgliedstaaten sollten sicherstellen, dass das Potenzial von KI-Systemen zur Förderung der tatsächlichen Gleichstellung der Geschlechter ausgeschöpft wird. Sie sollten dafür sorgen, dass diese Technologien die in der analogen Welt teils bestehenden großen geschlechtsbedingten Unterschiede nicht weiter verschärfen, sondern vielmehr diese Unterschiede beseitigen. Zu diesen Unterschieden gehören: Einkommensunterschiede zwischen Frauen und Männern, ihre ungleiche Vertretung in bestimmten Berufen und Tätigkeitsfeldern, der geringe Anteil von Frauen in den obersten Führungs- und Vorstandsetagen oder in KI-Forschungsteams, das Bildungsgefälle, das Gefälle beim digitalen Zugang und Zugang zu KI sowie bei der Annahme und Nutzung dieser Technologien, die Kaufkraftlücke sowie die ungleiche Verteilung von unbezahlter Arbeit und Betreuungsaufgaben in unserer Gesellschaft.

(90) Die Mitgliedstaaten müssen dafür sorgen, dass geschlechtsspezifische Stereotype und diskriminierende Verzerrungen nicht in KI-Systeme übernommen werden, sondern dass diese identifiziert und proaktiv beseitigt werden. Es müssen Anstrengungen unternommen werden, um zu verhindern, dass sich die negativen Auswirkungen der technologischen Kluft bei der Gleichstellung der Geschlechter und bei der Bekämpfung von Gewalt gegen Frauen und Mädchen und unterrepräsentierte Gruppen in Form von Belästigung, Mobbing oder Menschenhandel, auch im Online-Bereich, verstärken.

(89) Member States should ensure that the potential of AI systems to advance the achievement of gender equality is realized. They should ensure that these technologies do not exacerbate the already wide gender gaps existing in several fields in the analogue world, and instead eliminate those gaps. These gaps include: the gender wage gap; the unequal representation in certain professions and activities; the lack of representation at top management positions, boards of directors, or research teams in the AI field; the education gap; the digital and AI access, adoption, usage and affordability gap; and the unequal distribution of unpaid work and of the caring responsibilities in our societies.

(90) Member States should ensure that gender stereotyping and discriminatory biases are not translated into AI systems, and instead identify and proactively redress these. Efforts are necessary to avoid the compounding negative effect of technological divides in achieving gender equality and avoiding violence such as harassment, bullying or trafficking of girls and women and under-represented groups, including in the online domain.

(91) Die Mitgliedstaaten sollten Unternehmertum, Beteiligung und Engagement von Frauen in allen Phasen des Lebenszyklus von KI-Systemen fördern, indem sie unter anderem wirtschaftliche, regulatorische und sonstige Anreize schaffen und unterstützende Mechanismen und Maßnahmen initiieren und fördern, die auf eine ausgewogene Beteiligung von Frauen und Männern an der akademischen KI-Forschung sowie auf eine ausgewogene Vertretung der Geschlechter in Führungspositionen, Unternehmensvorständen und Forschungsteams von Digital- und KI-Unternehmen abzielen. Die Mitgliedstaaten sollten dafür sorgen, dass öffentliche Mittel (für Innovation, Forschung und Technologien) in inklusive Programme und Unternehmen mit einem ausgewogenen Geschlechterverhältnis fließen, und Anreize schaffen, auch private Mittel für Fördermaßnahmen nach dem Grundsatz der „positiven Diskriminierung“ einzusetzen. Sie sollten Konzepte für ein benachteiligungs- und belästigungsfreies Umfeld entwickeln und durchsetzen, flankiert von der Anregung, bewährte Verfahren zur Förderung der Vielfalt über den gesamten Lebenszyklus von KI-Systemen auszutauschen.

(92) Die Mitgliedstaaten sollten die Geschlechterdiversität in der akademischen und der industriellen KI-Forschung fördern, indem sie Mädchen und Frauen Anreize für den Einstieg in dieses Fachgebiet bieten, Mechanismen zur Bekämpfung von Geschlechterstereotypen und Benachteiligungen in der Welt der KI-Forschung umsetzen und akademische und private Einrichtungen dazu anregen, bewährte Verfahren zur Förderung der Geschlechterdiversität auszutauschen.

(93) Die UNESCO kann dabei unterstützen, ein Best Practice Repository aufzubauen, in dem bewährte Verfahren zur Förderung der Beteiligung von Mädchen, Frauen und unterrepräsentierten Gruppen in allen Phasen des Lebenszyklus von KI-Systemen gespeichert werden.

(91) Member States should encourage female entrepreneurship, participation and engagement in all stages of an AI system life cycle by offering and promoting economic, regulatory incentives, among other incentives and support schemes, as well as policies that aim at a balanced gender participation in AI research in academia, gender representation on digital and AI companies' top management positions, boards of directors and research teams. Member States should ensure that public funds (for innovation, research and technologies) are channelled to inclusive programmes and companies, with clear gender representation, and that private funds are similarly encouraged through affirmative action principles. Policies on harassment-free environments should be developed and enforced, together with the encouragement of the transfer of best practices on how to promote diversity throughout the AI system life cycle.

(92) Member States should promote gender diversity in AI research in academia and industry by offering incentives to girls and women to enter the field, putting in place mechanisms to fight gender stereotyping and harassment within the AI research community, and encouraging academic and private entities to share best practices on how to enhance gender diversity.

(93) UNESCO can help form a repository of best practices for incentivizing the participation of girls, women and under-represented groups in all stages of the AI system life cycle.

Politikbereich 7: Kultur

(94) Die Mitgliedstaaten sollten KI-Systeme geeignet dazu nutzen, um das materielle, dokumentarische und immaterielle Kulturerbe einschließlich gefährdeter und indigener Sprachen und indigenen Wissens zu erhalten, zu bereichern, zu verstehen, zu fördern, zu verwalten und zugänglich zu machen, beispielsweise durch Einführung oder Aktualisierung von Bildungsprogrammen im Zusammenhang mit der Anwendung von KI-Systemen in diesen Bereichen und durch Gewährleistung eines partizipativen Ansatzes, der sich an Institutionen und die Öffentlichkeit richtet.

(95) Die Mitgliedstaaten sollten die kulturellen Auswirkungen von KI-Systemen, insbesondere von Anwendungen zur Verarbeitung natürlicher Sprache („natural language processing“ – NLP) wie maschinelle Übersetzung und Sprachassistenten, auf die Nuancen menschlicher Sprache und des menschlichen Ausdrucks zu untersuchen und zu thematisieren. Solche Untersuchungen sollten einen Beitrag zur Ausgestaltung und Umsetzung von Strategien leisten, um den größtmöglichen Nutzen aus diesen Systemen zu ziehen, indem sie kulturelle Unterschiede überbrücken und Menschen ein besseres Verständnis ermöglichen, wobei aber auch auf die negativen Auswirkungen eingegangen wird, wie zum Beispiel eine zurückgehende Nutzung, die zum Aussterben gefährdeter Sprachen, lokaler Dialekte und tonaler und kultureller Variationen im Zusammenhang mit der menschlichen Sprache und Ausdrucksweise führen könnte.

Policy area 7: Culture

(94) Member States are encouraged to incorporate AI systems, where appropriate, in the preservation, enrichment, understanding, promotion, management and accessibility of tangible, documentary and intangible cultural heritage, including endangered languages as well as indigenous languages and knowledges, for example by introducing or updating educational programmes related to the application of AI systems in these areas, where appropriate, and by ensuring a participatory approach, targeted at institutions and the public.

(95) Member States are encouraged to examine and address the cultural impact of AI systems, especially natural language processing (NLP) applications such as automated translation and voice assistants, on the nuances of human language and expression. Such assessments should provide input for the design and implementation of strategies that maximize the benefits from these systems by bridging cultural gaps and increasing human understanding, as well as addressing the negative implications such as the reduction of use, which could lead to the disappearance of endangered languages, local dialects, and tonal and cultural variations associated with human language and expression.

(96) Die Mitgliedstaaten sollten die Weiterbildung von Künstlerinnen und Künstlern und Kreativschaffenden zu digitalen Themen und KI-Technologien fördern, damit diese die Eignung dieser Technologien für den Einsatz in ihrem Beruf beurteilen und zur Entwicklung und Umsetzung geeigneter KI-Technologien beitragen können, denn KI-Technologien werden bereits bei der Erstellung, Produktion, Verteilung bzw. Ausstrahlung und beim Konsum einer Vielzahl von Kulturgütern und -dienstleistungen eingesetzt; dabei muss die Bedeutung des Erhalts des kulturellen Erbes, der Vielfalt und der künstlerischen Freiheit im Blick behalten werden.

(97) Die Mitgliedstaaten sollten lokal verankerte Kulturbetriebe und kleine und mittlere Unternehmen, die im Kulturbereich tätig sind, für KI-Tools sensibilisieren und sie ermutigen, diese zu bewerten, um der Konzentrationsgefahr auf dem Kulturmarkt entgegenzuwirken.

(98) Die Mitgliedstaaten sollten Technologieunternehmen und andere Interessengruppen einbeziehen, um ein vielfältiges Angebot an kulturellen Ausdrucksformen und vielfältige Zugangsmöglichkeiten zu diesen Ausdrucksformen zu fördern und insbesondere dafür zu sorgen, dass die Sichtbarkeit und Auffindbarkeit lokal verankerter Inhalte durch algorithmische Empfehlungen verbessert wird.

(96) Member States should promote AI education and digital training for artists and creative professionals to assess the suitability of AI technologies for use in their profession, and contribute to the design and implementation of suitable AI technologies, as AI technologies are being used to create, produce, distribute, broadcast and consume a variety of cultural goods and services, bearing in mind the importance of preserving cultural heritage, diversity and artistic freedom.

(97) Member States should promote awareness and evaluation of AI tools among local cultural industries and small and medium enterprises working in the field of culture, to avoid the risk of concentration in the cultural market.

(98) Member States should engage technology companies and other stakeholders to promote a diverse supply of and plural access to cultural expressions, and in particular to ensure that algorithmic recommendation enhances the visibility and discoverability of local content.

(99) Die Mitgliedstaaten sollten neue Forschung an der Schnittstelle zwischen KI und geistigem Eigentum fördern, um zum Beispiel zu ermitteln, ob und wie die mittels KI-Technologien geschaffenen Werke durch geistige Eigentumsrechte geschützt werden sollten. Die Mitgliedstaaten sollten auch untersuchen, wie sich KI-Technologien auf die Rechte und Interessen von Inhabern geistigen Eigentums auswirken, deren Werke für Forschungs-, Entwicklungs-, Trainings- und Umsetzungszwecke von KI-Anwendungen verwendet werden.

(100) Die Mitgliedstaaten sollten Museen, Galerien, Bibliotheken und Archive auf nationaler Ebene dazu anregen, KI-Systeme zu nutzen, um auf ihre Sammlungen aufmerksam zu machen und ihre Bibliotheken, Datenbanken und Wissensbasen zu erweitern und gleichzeitig für ihre Nutzerinnen und Nutzer zugänglich zu machen.

(99) Member States should foster new research at the intersection between AI and intellectual property (IP), for example to determine whether or how to protect with IP rights the works created by means of AI technologies. Member States should also assess how AI technologies are affecting the rights or interests of IP owners, whose works are used to research, develop, train or implement AI applications.

(100) Member States should encourage museums, galleries, libraries and archives at the national level to use AI systems to highlight their collections and enhance their libraries, databases and knowledge base, while also providing access to their users.

Politikbereich 8: Bildung und Forschung

(101) Die Mitgliedstaaten sollten mit internationalen Organisationen, Bildungseinrichtungen sowie privaten und nichtstaatlichen Einrichtungen zusammenarbeiten, um der Öffentlichkeit in allen Ländern und auf allen Bildungsebenen angemessene KI-Kompetenzen zu vermitteln und Menschen zu befähigen, die digitale Kluft und Ungleichheiten beim digitalen Zugang, die sich infolge der breiten Einführung von KI-Systemen ergeben, zu verringern.

(102) Die Mitgliedstaaten sollten den Erwerb von Fähigkeiten fördern, die für das Erlernen des Umgangs mit KI unerlässlich sind, beispielsweise Lesen und Schreiben, Rechnen, Programmieren und digitale Kenntnisse, Medien- und Informationskompetenz sowie kritisches und kreatives Denken, Teamarbeit, Kommunikation, sozioemotionale Fähigkeiten und Kenntnisse in der KI-Ethik, insbesondere in Ländern und Regionen oder Gebieten innerhalb von Ländern, in denen es bei der Vermittlung dieser Fähigkeiten noch erhebliche Lücken gibt.

(103) Die Mitgliedstaaten sollten Programme zur allgemeinen Sensibilisierung in Bezug auf KI-Entwicklungen fördern, auch in Bezug auf Daten, Möglichkeiten und Herausforderungen von KI-Technologien, die Auswirkungen und Implikationen von KI-Systemen auf die Menschenrechte, auch auf die Rechte von Kindern. Solche Programme sollten sowohl für Fachgruppen als auch für Laien zugänglich sein.

Policy area 8: Education and research

(101) Member States should work with international organizations, educational institutions and private and non-governmental entities to provide adequate AI literacy education to the public on all levels in all countries in order to empower people and reduce the digital divides and digital access inequalities resulting from the wide adoption of AI systems.

(102) Member States should promote the acquisition of “prerequisite skills” for AI education, such as basic literacy, numeracy, coding and digital skills, and media and information literacy, as well as critical and creative thinking, teamwork, communication, socio-emotional and AI ethics skills, especially in countries and in regions or areas within countries where there are notable gaps in the education of these skills.

(103) Member States should promote general awareness programmes about AI developments, including on data and the opportunities and challenges brought about by AI technologies, the impact of AI systems on human rights and their implications, including children’s rights. These programmes should be accessible to nontechnical as well as technical groups.

(104) Die Mitgliedstaaten sollten Forschungsinitiativen ermöglichen, unter anderem zum verantwortungsvollen und ethischen Einsatz von KI-Technologien im Unterricht, in der Lehrerbildung und beim E-Learning, um die Möglichkeiten in diesem Bereich zu verbessern und die Herausforderungen und Risiken zu verringern. Diese Initiativen sollten hinsichtlich der Qualität der Bildung und der Auswirkungen des Einsatzes von KI-Technologien auf Lernende und Lehrende angemessen evaluiert werden. Die Mitgliedstaaten sollten auch darauf hinwirken, dass KI-Technologien die Handlungsfähigkeit von Lernenden und Lehrenden stärken und ihnen eine bessere Lehr- und Lernerfahrung ermöglichen. Dabei ist zu bedenken, dass zwischenmenschliche und soziale Aspekte und der Wert traditioneller Unterrichtsformen für die Beziehung zwischen Lehrenden und Lernenden und zwischen Lernenden untereinander von entscheidender Bedeutung sind und bei der Diskussion über die Einführung von KI-Technologien im Bildungsbereich berücksichtigt werden müssen. KI-Systeme, die beim Lernen eingesetzt werden, sollten in Bezug auf das Monitoring, die Bewertung von Fähigkeiten oder die Vorhersage von Verhaltensweisen der Lernenden strengen Anforderungen unterliegen. KI sollte den Lernprozess unterstützen, ohne kognitive Fähigkeiten einzuschränken und ohne sensible Informationen zu extrahieren, im Einklang mit einschlägigen Datenschutzstandards zum Schutz personenbezogener Daten. Die während der Interaktion des Lernenden mit dem KI-System zum Wissenserwerb erhobenen Daten dürfen nicht missbraucht, zweckentfremdet oder gesetzeswidrig genutzt werden, auch nicht für kommerzielle Zwecke.

(104) Member States should encourage research initiatives on the responsible and ethical use of AI technologies in teaching, teacher training and e-learning, among other issues, to enhance opportunities and mitigate the challenges and risks involved in this area. The initiatives should be accompanied by an adequate assessment of the quality of education and impact on students and teachers of the use of AI technologies. Member States should also ensure that AI technologies empower students and teachers and enhance their experience, bearing in mind that relational and social aspects and the value of traditional forms of education are vital in teacher-student and student-student relationships and should be considered when discussing the adoption of AI technologies in education. AI systems used in learning should be subject to strict requirements when it comes to the monitoring, assessment of abilities, or prediction of the learners' behaviours. AI should support the learning process without reducing cognitive abilities and without extracting sensitive information, in compliance with relevant personal data protection standards. The data handed over to acquire knowledge collected during the learner's interactions with the AI system must not be subject to misuse, misappropriation or criminal exploitation, including for commercial purposes.

(105) Die Mitgliedstaaten sollten sich dafür einsetzen, dass Mädchen und Frauen, Menschen unterschiedlicher ethnischer Herkunft und verschiedener Kulturen, Menschen mit Behinderungen, marginalisierte und besonders schutzbedürftige Menschen bzw. Menschen in gefährdeten Lebenslagen, Minderheiten und alle anderen Menschen, die noch nicht in vollem Umfang von einer digitalen Inklusion profitieren, auf allen Ebenen an KI-Bildungsprogrammen teilnehmen und dabei auch Führungsrollen übernehmen; ebenso sollten die Mitgliedstaaten diesbezüglich bewährte Verfahren beobachten und deren Austausch mit anderen Mitgliedstaaten fördern.

(106) Die Mitgliedstaaten sollten im Einklang mit ihren nationalen Bildungsprogrammen und -traditionen Lehrpläne für KI-Ethik für sämtliche Bildungsniveaus entwickeln und sich dafür einsetzen, dass die Vermittlung technischer KI-Fähigkeiten und die Behandlung humanistischer, ethischer und sozialer Aspekte der KI-Bildung miteinander verknüpft werden. Es sollten Onlinekurse und digitale Ressourcen für die KI-Ethikausbildung in lokalen Sprachen, auch in indigenen Sprachen, entwickelt werden; dabei ist der Diversität der Kontexte Rechnung zu tragen, insbesondere sollten Formate vorgesehen werden, die auch für Menschen mit Behinderungen zugänglich sind.

(105) Member States should promote the participation and leadership of girls and women, diverse ethnicities and cultures, persons with disabilities, marginalized and vulnerable people or people in vulnerable situations, minorities and all persons not enjoying the full benefits of digital inclusion, in AI education programmes at all levels, as well as the monitoring and sharing of best practices in this regard with other Member States.

(106) Member States should develop, in accordance with their national education programmes and traditions, AI ethics curricula for all levels, and promote cross-collaboration between AI technical skills education and humanistic, ethical and social aspects of AI education. Online courses and digital resources of AI ethics education should be developed in local languages, including indigenous languages, and take into account the diversity of environments, especially ensuring accessibility of formats for persons with disabilities.

(107) Die Mitgliedstaaten sollten KI-Forschung, insbesondere Forschung zur Ethik der KI, fördern und unterstützen, beispielsweise durch Investitionen in diese Forschung oder durch die Schaffung von Anreizen für den öffentlichen und den privaten Sektor, in diesen Bereich zu investieren, in der Erkenntnis, dass Forschung im Sinne der Förderung des Völkerrechts und der in dieser Empfehlung dargelegten Werte und Grundsätze wesentlich zur Weiterentwicklung und Verbesserung der KI-Technologien beiträgt. Außerdem sollten die Mitgliedstaaten bewährte Verfahren aus der Forschung und von Unternehmen, die KI auf ethische Weise entwickeln, öffentlich bekannt machen und die Zusammenarbeit mit solchen Forschenden und Unternehmen fördern.

(108) Die Mitgliedstaaten sollten dafür sorgen, dass KI-Forschende in Forschungsethik geschult werden, und von ihnen verlangen, dass sie ethische Erwägungen in ihre Entwürfe, Produkte und Veröffentlichungen einfließen lassen, insbesondere bei der Analyse der von ihnen verwendeten Datensätze, der Art und Weise der Datenkommentierung und in Bezug auf die Qualität und die Tragweite der Ergebnisse möglicher Anwendungen.

(109) Die Mitgliedstaaten sollten privatwirtschaftliche Unternehmen dazu anhalten, der Wissenschaft für Forschungszwecke Zugang zu ihren Daten zu ermöglichen, insbesondere in LMICs und ganz besonders in den LDCs, LLDCs und SIDS. Dabei sind die einschlägigen Rechtsnormen für den Schutz der Privatsphäre und den Datenschutz einzuhalten.

(107) Member States should promote and support AI research, notably AI ethics research, including for example through investing in such research or by creating incentives for the public and private sectors to invest in this area, recognizing that research contributes significantly to the further development and improvement of AI technologies with a view to promoting international law and the values and principles set forth in this Recommendation. Member States should also publicly promote the best practices of, and cooperation with, researchers and companies who develop AI in an ethical manner.

(108) Member States should ensure that AI researchers are trained in research ethics and require them to include ethical considerations in their designs, products and publications, especially in the analyses of the datasets they use, how they are annotated, and the quality and scope of the results with possible applications.

(109) Member States should encourage private sector companies to facilitate the access of the scientific community to their data for research, especially in LMICs, in particular LDCs, LLDCs and SIDS. This access should conform to relevant privacy and data protection standards.

(110) Um eine kritische Auswertung der KI-Forschung und eine angemessene Kontrolle hinsichtlich potenziellen Missbrauchs oder nachteiliger Auswirkungen zu gewährleisten, sollten die Mitgliedstaaten sicherstellen, dass künftige Entwicklungen im Bereich der KI-Technologien auf der Grundlage fundierter und unabhängiger wissenschaftlicher Forschung erfolgen, außerdem sollten sie interdisziplinäre KI-Forschung fördern, die neben Mathematik, Informatik, Naturwissenschaft und Technik (MINT) auch andere Disziplinen einbezieht, beispielsweise Kulturwissenschaften, Pädagogik, Ethik, internationale Beziehungen, Recht, Sprachwissenschaft, Philosophie, Politikwissenschaft, Soziologie und Psychologie.

(111) Da KI-Technologien große Chancen auf wissenschaftlichen Fortschritt in Theorie und Praxis bieten, insbesondere in den traditionell modellgestützten Disziplinen, sollten die Mitgliedstaaten sich bemühen, das Bewusstsein der Wissenschaftsgemeinschaft sowohl für den Nutzen als auch für die Grenzen und Risiken ihrer Anwendung zu schärfen; dazu gehören auch Anstrengungen, durch die sichergestellt werden kann, dass die aus datengestützten Ansätzen, Modellen und Behandlungen gezogenen Schlussfolgerungen solide und fundiert sind. Darüber hinaus sollten die Mitgliedstaaten die Bedeutung der Wissenschaft für die Gestaltung von Politik und die Sensibilisierung für die Stärken und Schwächen von KI-Technologien stärken.

(110) To ensure a critical evaluation of AI research and proper monitoring of potential misuses or adverse effects, Member States should ensure that any future developments with regards to AI technologies should be based on rigorous and independent scientific research, and promote interdisciplinary AI research by including disciplines other than science, technology, engineering and mathematics (STEM), such as cultural studies, education, ethics, international relations, law, linguistics, philosophy, political science, sociology and psychology.

(111) Recognizing that AI technologies present great opportunities to help advance scientific knowledge and practice, especially in traditionally model-driven disciplines, Member States should encourage scientific communities to be aware of the benefits, limits and risks of their use; this includes attempting to ensure that conclusions drawn from data-driven approaches, models and treatments are robust and sound. Furthermore, Member States should welcome and support the role of the scientific community in contributing to policy and in cultivating awareness of the strengths and weaknesses of AI technologies.

Politikbereich 9: Kommunikation und Information

(112) Die Mitgliedstaaten sollten KI-Systeme nutzen, um die Zugänglichkeit von Informationen und Wissen zu verbessern. In diesem Rahmen können sie etwa Forschende, die Wissenschaft, Journalistinnen und Journalisten, die breite Öffentlichkeit sowie Entwicklerinnen und Entwickler unterstützen, um die Freiheit der Meinungsäußerung, die Wissenschaftsfreiheit zu stärken, Informationen leichter zugänglich zu machen und darauf hinzuwirken, dass offizielle Daten und Informationen proaktiv offengelegt werden.

(113) Die Mitgliedstaaten sollten dafür sorgen, dass KI-Akteure im Zusammenhang mit automatischer Erstellung, Moderation und Kuratierung von Inhalten die Meinungsfreiheit und die Zugänglichkeit von Informationen gewährleisten und fördern. Sie sollten geeignete Rahmenbedingungen schaffen, gegebenenfalls auch durch entsprechende Regulierung, um bei Online-Kommunikations- und -Informationsanbietern Transparenz herzustellen und zu gewährleisten, dass Nutzerinnen und Nutzer Zugang zu einer Vielfalt von Meinungen haben, unverzüglich über die Gründe für eine etwaige Entfernung oder sonstige Behandlung von Inhalten benachrichtigt werden und mit Rechtsmitteln dagegen vorgehen können.

Policy area 9: Communication and information

(112) Member States should use AI systems to improve access to information and knowledge. This can include support to researchers, academia, journalists, the general public and developers, to enhance freedom of expression, academic and scientific freedoms, access to information, and increased proactive disclosure of official data and information.

(113) Member States should ensure that AI actors respect and promote freedom of expression as well as access to information with regard to automated content generation, moderation and curation. Appropriate frameworks, including regulation, should enable transparency of online communication and information operators and ensure users have access to a diversity of viewpoints, as well as processes for prompt notification to the users on the reasons for removal or other treatment of content, and appeal mechanisms that allow users to seek redress.

(114) Die Mitgliedstaaten sollten in den Aufbau digitaler Kompetenzen sowie Medien- und Informationskompetenzen investieren und deren Erwerb unterstützen, um kritisches Denken zu fördern und die Kompetenzen zu stärken, die erforderlich sind, um die Arbeitsweise und die Auswirkungen von KI-Systemen zu verstehen und dadurch Desinformation, Misinformation sowie Hass und Hetze entgegenzuwirken oder diese zu entschärfen. Im Rahmen dieser Bemühungen sollte auch ein besseres Verständnis sowohl der positiven als auch der potenziell schädlichen Auswirkungen von Empfehlungsdiensten vermittelt werden, damit diese besser beurteilt werden können.

(115) Die Mitgliedstaaten sollten ein förderliches Umfeld für die Medien schaffen, das ihnen die Rechte und Ressourcen gewährt, die sie brauchen, um wirksam über nützliche und schädliche Aspekte von KI-Systemen berichten zu können, und sie sollten darauf hinwirken, dass die Medien KI-Systeme bei ihrer Arbeit auf ethische Weise einsetzen.

(114) Member States should invest in and promote digital and media and information literacy skills to strengthen critical thinking and competencies needed to understand the use and implication of AI systems, in order to mitigate and counter disinformation, misinformation and hate speech. A better understanding and evaluation of both the positive and potentially harmful effects of recommender systems should be part of those efforts.

(115) Member States should create enabling environments for media to have the rights and resources to effectively report on the benefits and harms of AI systems, and also encourage media to make ethical use of AI systems in their operations.

Politikbereich 10: Wirtschaft und Arbeitswelt

(116) Die Mitgliedstaaten sollten die Auswirkungen von KI-Systemen auf die Arbeitsmärkte und die Folgen für die Anforderungen an Bildung in allen Ländern und ganz besonders in Ländern mit arbeitsintensiver Wirtschaft untersuchen und sich damit auseinandersetzen. So kann es etwa erforderlich sein, auf allen Bildungsebenen die Vermittlung eines breiteren Spektrums an Schlüsselqualifikationen und interdisziplinären Fähigkeiten einzuführen, um sowohl den heutigen Arbeitskräften als auch künftigen Generationen eine faire Chance zu geben, auf einem sich rasch verändernden Markt Arbeit zu finden, und um sicherzustellen, dass sie sich der ethischen Aspekte von KI-Systemen bewusst sind. Fähigkeiten wie „Lernen lernen“, Kommunikation, kritisches Denken, Teamarbeit, Einfühlungsvermögen und die Fähigkeit, sein Wissen auch auf andere Bereiche zu übertragen, sollten ebenso gelehrt werden wie Fachwissen und technische Fertigkeiten, neben Aufgaben, für die eine geringere Qualifizierung erforderlich ist. Klarheit darüber zu schaffen, welche Fähigkeiten gefragt sind, und Lehrpläne entsprechend zu aktualisieren ist von entscheidender Bedeutung.

Policy area 10: Economy and labour

(116) Member States should assess and address the impact of AI systems on labour markets and its implications for education requirements, in all countries and with special emphasis on countries where the economy is labourintensive. This can include the introduction of a wider range of “core” and interdisciplinary skills at all education levels to provide current workers and new generations a fair chance of finding jobs in a rapidly changing market, and to ensure their awareness of the ethical aspects of AI systems. Skills such as “learning how to learn”, communication, critical thinking, teamwork, empathy, and the ability to transfer one’s knowledge across domains, should be taught alongside specialist, technical skills, as well as low-skilled tasks. Being transparent about what skills are in demand and updating curricula around these are key.

(117) Die Mitgliedstaaten sollten Kooperationsvereinbarungen zwischen Regierungen, akademischen Einrichtungen, Einrichtungen der beruflichen Aus- und Weiterbildung, der Industrie, Arbeitnehmerorganisationen und der Zivilgesellschaft unterstützen, um die Kluft zwischen Anforderungen und Qualifikationen zu überbrücken und Ausbildungsprogramme auf die Arbeit der Zukunft und die Bedürfnisse der Industrie, einschließlich kleiner und mittelständischer Unternehmen, zuzuschneiden. Projektbasierte Lehr- und Lernansätze für KI, in deren Rahmen auch Partnerschaften zwischen öffentlichen Einrichtungen, privatwirtschaftlichen Unternehmen, Hochschulen und Forschungszentren möglich sind, sollten gefördert werden.

(117) Member States should support collaboration agreements among governments, academic institutions, vocational education and training institutions, industry, workers' organizations and civil society to bridge the gap of skillset requirements to align training programmes and strategies with the implications of the future of work and the needs of industry, including small and medium enterprises. Project-based teaching and learning approaches for AI should be promoted, allowing for partnerships between public institutions, private sector companies, universities and research centres.

(118) Die Mitgliedstaaten sollten mit privatwirtschaftlichen Unternehmen, zivilgesellschaftlichen Organisationen und anderen Interessengruppen einschließlich Erwerbstätiger und Gewerkschaften zusammenarbeiten, um den Übergang für Arbeitnehmerinnen und Arbeitnehmer, für die besondere Risiken bestehen, verträglich zu gestalten. Dazu sollten auch Umschulungs- und Weiterqualifizierungsprogramme aufgelegt, wirksame Mechanismen zur Weiterbeschäftigung während dieser Übergangszeiten gefunden und Lösungen in Form eines „Sicherheitsnetzes“ für diejenigen eruiert werden, die nicht umgeschult werden können. Die Mitgliedstaaten sollten Programme entwickeln und umsetzen, um zu den erkannten Herausforderungen zu forschen und diese anzugehen, wobei sie unter anderem Weiterqualifizierungs- und Umschulungsmaßnahmen, verbesserte soziale Absicherung, proaktive industriepolitische Strategien und Interventionen, Steuervergünstigungen und neue Besteuerungsformen vorsehen könnten. Die Mitgliedstaaten sollten ausreichende öffentliche Mittel für diese Programme zur Verfügung stellen. Relevante Regulierung, beispielsweise Steuersysteme, sollten sorgfältig überprüft und bei Bedarf geändert werden, um die Folgen von Arbeitslosigkeit abzufedern, die durch KI-getriebene Automatisierung entsteht.

(119) Die Mitgliedstaaten sollten Forschende dazu anregen und dabei unterstützen, die Folgen von KI-Systemen für die lokale Arbeitswelt zu analysieren, um kommende Entwicklungen und Herausforderungen zu antizipieren. Diese Studien sollten interdisziplinär angelegt sein und die Auswirkungen von KI-Systemen auf wirtschaftliche, soziale und geografische Faktoren sowie auf die Interaktion zwischen Mensch und Maschine und die Beziehungen von Menschen untereinander untersuchen, damit sie Hinweise auf optimale Verfahren bei Umschulungs- und Umstrukturierungsmaßnahmen liefern können.

(118) Member States should work with private sector companies, civil society organizations and other stakeholders, including workers and unions to ensure a fair transition for at-risk employees. This includes putting in place upskilling and reskilling programmes, finding effective mechanisms of retaining employees during those transition periods, and exploring “safety net” programmes for those who cannot be retrained. Member States should develop and implement programmes to research and address the challenges identified that could include upskilling and reskilling, enhanced social protection, proactive industry policies and interventions, tax benefits, new taxation forms, among others. Member States should ensure that there is sufficient public funding to support these programmes. Relevant regulations, such as tax regimes, should be carefully examined and changed if needed to counteract the consequences of unemployment caused by AI-based automation.

(119) Member States should encourage and support researchers to analyse the impact of AI systems on the local labour environment in order to anticipate future trends and challenges. These studies should have an interdisciplinary approach and investigate the impact of AI systems on economic, social and geographic sectors, as well as on human-robot interactions and humanhuman relationships, in order to advise on reskilling and redeployment best practices.

(120) Die Mitgliedstaaten sollten geeignete Schritte unternehmen, um wettbewerbsorientierte Märkte und Verbraucherschutz zu gewährleisten, und dabei mögliche Maßnahmen und Mechanismen auf nationaler, kontinentweiter und internationaler Ebene in Betracht ziehen, um den Missbrauch marktbeherrschender Stellungen bzw. von Monopolisierung im Zusammenhang mit KI-Systemen über ihren gesamten Lebenszyklus zu verhindern, unabhängig davon, ob es dabei um Daten, Forschung, Technologie oder Märkte geht. Die Mitgliedstaaten sollten daraus resultierenden Ungleichheiten vorbeugen, relevante Märkte bewerten und wettbewerbsorientierte Märkte fördern. Dabei müssen LMICs, insbesondere LDCs, LLDCs und SIDS, gebührend berücksichtigt werden, da sie missbräuchlicher Ausnutzung von marktbeherrschenden Stellungen besonders schutzlos ausgeliefert sind, unter anderem aufgrund fehlender Infrastruktur, mangelnder personeller Kapazitäten und mangelhafter Rechtsrahmen. KI-Akteure, die KI-Systeme in Ländern entwickeln, die ethische Standards für KI aufgestellt oder angenommen haben, sollten sich auch an diese Standards halten, wenn sie diese Produkte in Länder exportieren, in denen es solche Standards möglicherweise nicht gibt, oder wenn sie die Produkte in solchen Ländern entwickeln oder einsetzen; dabei müssen sie geltende völkerrechtliche Verpflichtungen sowie die nationalen Rechtsvorschriften, Standards und Verfahren dieser Länder beachten.

(120) Member States should take appropriate steps to ensure competitive markets and consumer protection, considering possible measures and mechanisms at national, regional and international levels, to prevent abuse of dominant market positions, including by monopolies, in relation to AI systems throughout their life cycle, whether these are data, research, technology, or market. Member States should prevent the resulting inequalities, assess relevant markets and promote competitive markets. Due consideration should be given to LMICs, in particular LDCs, LLDCs and SIDS, which are more exposed and vulnerable to the possibility of abuses of market dominance as a result of a lack of infrastructure, human capacity and regulations, among other factors. AI actors developing AI systems in countries which have established or adopted ethical standards on AI should respect these standards when exporting these products, developing or applying their AI systems in countries where such standards may not exist, while respecting applicable international law and domestic legislation, standards and practices of these countries.

Politikbereich 11: Gesundheit und soziales Wohlergehen

(121) Die Mitgliedstaaten sollten möglichst wirksame KI-Systeme einsetzen, um die Gesundheit der Menschen zu verbessern und ihr Recht auf Leben zu schützen, auch durch Eindämmung von Seuchenausbrüchen; gleichzeitig sollten sie internationale Solidarität stärken und aufrechterhalten, damit globale Gesundheitsrisiken und Unwägbarkeiten bewältigt werden können; dabei müssen sie sicherstellen, dass der Einsatz von KI-Systemen in ihrem Gesundheitswesen mit ihren völkerrechtlichen und menschenrechtlichen Verpflichtungen vereinbar ist. Die Mitgliedstaaten sollten darauf achten, dass die an KI-Systemen in der Gesundheitsversorgung beteiligten Akteure die Bedeutung der Beziehungen von Patientinnen und Patienten zu ihren Familien und zum medizinischen und Pflegepersonal berücksichtigen.

(122) Die Mitgliedstaaten sollten sicherstellen, dass Entwicklung und Einsatz von KI-Systemen im Bereich der Gesundheit im Allgemeinen und besonders im Bereich der psychischen Gesundheit, unter angemessener Berücksichtigung von Kindern und Jugendlichen, dahingehend geregelt werden, dass die Systeme sicher, wirksam, effizient, wissenschaftlich und medizinisch geprüft sein müssen und evidenzbasierte Innovation und medizinischen Fortschritt ermöglichen. Darüber hinaus werden die Mitgliedstaaten nachdrücklich aufgefordert, Patientinnen und Patienten und sie vertretende Personen aktiv in alle relevanten Schritte der Systementwicklung im Bereich der digitalen Gesundheitsmaßnahmen einzubeziehen.

Policy area 11: Health and social well-being

(121) Member States should endeavour to employ effective AI systems for improving human health and protecting the right to life, including mitigating disease outbreaks, while building and maintaining international solidarity to tackle global health risks and uncertainties, and ensure that their deployment of AI systems in health care be consistent with international law and their human rights law obligations. Member States should ensure that actors involved in health care AI systems take into consideration the importance of a patient's relationships with their family and with health care staff.

(122) Member States should ensure that the development and deployment of AI systems related to health in general and mental health in particular, paying due attention to children and youth, is regulated to the effect that they are safe, effective, efficient, scientifically and medically proven and enable evidence-based innovation and medical progress. Moreover, in the related area of digital health interventions, Member States are strongly encouraged to actively involve patients and their representatives in all relevant steps of the development of the system.

(123) Bei der Regulierung von KI-gestützten Lösungen zur Vorhersage, Erkennung und Behandlung im Rahmen der Gesundheitsversorgung sollten die Mitgliedstaaten ihr besonderes Augenmerk darauf richten,

- a) Aufsicht zu gewährleisten, damit Verzerrungen auf ein Minimum reduziert und gemildert werden;
- b) sicherzustellen, dass medizinische Fachkräfte, Patientinnen und Patienten, Pflegende und diejenigen, die Dienste in Anspruch nehmen, bei allen relevanten Schritten der Algorithmenentwicklung als „Fachexpertinnen und -experten“ einbezogen werden;
- c) angesichts der denkbaren Notwendigkeit medizinischer Überwachung den Schutz der Privatsphäre angemessen zu berücksichtigen und sicherzustellen, dass alle einschlägigen nationalen und internationalen Datenschutzanforderungen erfüllt werden;
- d) für wirksame Mechanismen zu sorgen, die gewährleisten, dass Personen, deren personenbezogene Daten analysiert werden, über die Verwendung und Analyse ihrer Daten informiert sind und ihre Einwilligung dazu erteilen, ohne dass dadurch der Zugang zur Gesundheitsversorgung behindert wird;
- e) zu gewährleisten, dass menschliche Betreuung und endgültige Entscheidungen über Diagnose und Behandlung immer durch Menschen erfolgen, wobei anerkannt werden sollte, dass KI-Systeme ihre Arbeit unterstützen können;
- f) dafür zu sorgen, dass KI-Systeme vor ihrem klinischen Einsatz durch eine Forschungsethikkommission geprüft werden, wo dies erforderlich ist.

(123) Member States should pay particular attention in regulating prediction, detection and treatment solutions for health care in AI applications by:

- a) ensuring oversight to minimize and mitigate bias;
- b) ensuring that the professional, the patient, caregiver or service user is included as a “domain expert” in the team in all relevant steps when developing the algorithms;
- c) paying due attention to privacy because of the potential need for being medically monitored and ensuring that all relevant national and international data protection requirements are met;
- d) ensuring effective mechanisms so that those whose personal data is being analysed are aware of and provide informed consent for the use and analysis of their data, without preventing access to health care;
- e) ensuring the human care and final decision of diagnosis and treatment are taken always by humans while acknowledging that AI systems can also assist in their work;
- f) ensuring, where necessary, the review of AI systems by an ethical research committee prior to clinical use.

(124) Die Mitgliedstaaten sollten Forschung zu den potenziell schädlichen Auswirkungen von KI-Systemen auf die psychische Gesundheit etablieren, beispielsweise zu verstärktem Auftreten von Depressionen, Angststörungen und sozialer Isolation, die unter anderem zu Suchterkrankungen führen und Menschenhandel, Radikalisierung und Misinformation begünstigen können, und sie sollten regulierende Maßnahmen treffen.

(125) Im Hinblick auf die künftige Entwicklung von Robotern sollten die Mitgliedstaaten auf der Grundlage von Forschungsergebnissen Leitlinien für die Interaktion zwischen Menschen und Robotern entwickeln, die die Auswirkungen auf die Beziehungen von Menschen untereinander im Blick haben, wobei ein besonderes Augenmerk auf die psychische und physische Gesundheit der Menschen zu legen ist. Besondere Aufmerksamkeit muss hierbei Robotern zukommen, die in der Gesundheitsversorgung, bei der Pflege älterer Menschen oder Menschen mit Behinderungen sowie im Bildungsbereich eingesetzt werden. Gleiches gilt für Roboter für Kinder, Spielzeugroboter, Chatbots und Begleitroboter für Kinder und Erwachsene. Darüber hinaus sollten KI-Technologien auch dafür eingesetzt werden, Sicherheit und Ergonomie beim Umgang mit Robotern zu verbessern, auch in Arbeitsumgebungen, in denen Menschen und Roboter zusammenarbeiten. Insbesondere muss im Auge behalten werden, dass KI auch zu Manipulationszwecken und zum Missbrauch von menschlichen kognitiven Verzerrungen eingesetzt werden kann.

(124) Member States should establish research on the effects and regulation of potential harms to mental health related to AI systems, such as higher degrees of depression, anxiety, social isolation, developing addiction, trafficking, radicalization and misinformation, among others.

(125) Member States should develop guidelines for humanrobot interactions and their impact on human-human relationships, based on research and directed at the future development of robots, and with special attention to the mental and physical health of human beings. Particular attention should be given to the use of robots in health care and the care for older persons and persons with disabilities, in education, and robots for use by children, toy robots, chatbots and companion robots for children and adults. Furthermore, assistance of AI technologies should be applied to increase the safety and ergonomic use of robots, including in a human-robot working environment. Special attention should be paid to the possibility of using AI to manipulate and abuse human cognitive biases.

(126) Die Mitgliedstaaten müssen sicherstellen, dass bei der Interaktion zwischen Menschen und Robotern dieselben Werte und Grundsätze beachtet werden, die auch für alle anderen KI-Systeme gelten, wie die Menschenrechte und Grundfreiheiten, die Förderung der Vielfalt und der Schutz besonders schutzbedürftiger Menschen bzw. Menschen, die sich in gefährdeten Umständen befinden. Zur Wahrung der menschlichen Würde und Autonomie müssen ethische Aspekte auch im Zusammenhang mit KI-gestützten Systemen für Neurotechnologien und Gehirn-Computer-Schnittstellen berücksichtigt werden.

(127) Die Mitgliedstaaten müssen sicherstellen, dass es für Nutzerinnen und Nutzer immer leicht ersichtlich bleibt, ob sie mit einem lebenden Wesen oder mit einem KI-System interagieren, das menschliche oder tierische Eigenschaften imitiert, und dass sie eine solche Interaktion wirksam ablehnen und menschliche Betreuung verlangen können.

(128) Die Mitgliedstaaten sollten Maßnahmen ergreifen, um für die Anthropomorphisierung von KI-Systemen und Technologien zu sensibilisieren, die menschliche Emotionen erkennen und nachahmen, auch hinsichtlich der Sprache, in der man über sie spricht, und sie sollten die Erscheinungsformen, ethischen Implikationen und möglichen Grenzen einer solchen Anthropomorphisierung bewerten, insbesondere im Kontext der Interaktion zwischen Robotern und Menschen und ganz besonders, wenn Kinder beteiligt sind.

(126) Member States should ensure that human-robot interactions comply with the same values and principles that apply to any other AI systems, including human rights and fundamental freedoms, the promotion of diversity, and the protection of vulnerable people or people in vulnerable situations. Ethical questions related to AI-powered systems for neurotechnologies and braincomputer interfaces should be considered in order to preserve human dignity and autonomy.

(127) Member States should ensure that users can easily identify whether they are interacting with a living being, or with an AI system imitating human or animal characteristics, and can effectively refuse such interaction and request human intervention.

(128) Member States should implement policies to raise awareness about the anthropomorphization of AI technologies and technologies that recognize and mimic human emotions, including in the language used to mention them, and assess the manifestations, ethical implications and possible limitations of such anthropomorphization, in particular in the context of robot-human interaction and especially when children are involved.

(129) Die Mitgliedstaaten sollten sich für Zusammenarbeit in der Forschung einsetzen und sie fördern, um die Folgen dauerhafter Interaktion zwischen Menschen und KI-Systemen zu erforschen, mit einem besonderen Schwerpunkt auf den psychologischen und kognitiven Auswirkungen, die solche Systeme auf Kinder und Jugendliche haben können. Dies sollte unter Heranziehung einer Vielzahl von Normen, Prinzipien, Protokollen und fachlichen Herangehensweisen geschehen, wobei Änderungen von Verhaltensweisen und Gewohnheiten bewertet und dadurch verursachte kulturelle und gesellschaftliche Auswirkungen sorgfältig untersucht werden müssen. Darüber hinaus sollten die Mitgliedstaaten darauf hinwirken, dass zu den Folgen von KI-Technologien für die Leistungsfähigkeit von Gesundheitssystemen und zu deren Ergebnissen für die Gesundheit geforscht wird.

(130) Die Mitgliedstaaten und sämtliche Interessengruppen sollten Mechanismen einführen, um Kinder und Jugendliche in Gesprächen, Debatten und Entscheidungsprozessen in Bezug auf die Auswirkungen von KI-Systemen auf ihr Leben und ihre Zukunft einzubeziehen und ihnen eine Stimme zu geben.

(129) Member States should encourage and promote collaborative research into the effects of long-term interaction of people with AI systems, paying particular attention to the psychological and cognitive impact that these systems can have on children and young people. This should be done using multiple norms, principles, protocols, disciplinary approaches, and assessment of the modification of behaviours and habits, as well as careful evaluation of the downstream cultural and societal impacts. Furthermore, Member States should encourage research on the effect of AI technologies on health system performance and health outcomes.

(130) Member States, as well as all stakeholders, should put in place mechanisms to meaningfully engage children and young people in conversations, debates and decisionmaking with regard to the impact of AI systems on their lives and futures.

V. Überwachung und Evaluierung

(131) Die Mitgliedstaaten sollten Strategien, Programme und Mechanismen im Zusammenhang mit der KI-Ethik entsprechend ihren jeweiligen Rahmenbedingungen, Regierungsverfasstheit und verfassungsrechtlichen Bestimmungen glaubwürdig und transparent überwachen und evaluieren; dabei sollten sowohl quantitative als auch qualitative Methoden zu Anwendung kommen. Zur Unterstützung der Mitgliedstaaten kann die UNESCO folgende Beiträge leisten:

- a) Entwicklung einer UNESCO-Methodik für die ethische Folgenabschätzung („Ethical Impact Assessment“ – EIA) von KI-Technologien auf der Grundlage streng wissenschaftlicher Forschung und verankert in internationalen Menschenrechtsnormen, die Erarbeitung eines Leitfadens für die Umsetzung dieser Methodik über den gesamten Lebenszyklus von KI-Systemen sowie von Materialien für den Kapazitätsaufbau, um die Mitgliedstaaten in ihren Bemühungen zu unterstützen, Regierungsbeamtinnen und -beamte, politische Entscheidungsträgerinnen und -träger und andere relevante KI-Akteure in der EIA-Methodik zu schulen;
- b) Entwicklung einer UNESCO-Methodik, die die Mitgliedstaaten dabei unterstützt zu beurteilen, wie weit sie auf ihrem Weg zur Umsetzung („readiness assessment methodology“) entlang eines Kontinuums von Dimensionen zu bestimmten Zeitpunkten fortgeschritten sind;

V. Monitoring and evaluation

(131) Member States should, according to their specific conditions, governing structures and constitutional provisions, credibly and transparently monitor and evaluate policies, programmes and mechanisms related to ethics of AI, using a combination of quantitative and qualitative approaches. To support Member States, UNESCO can contribute by:

- a) developing a UNESCO methodology for Ethical Impact Assessment (EIA) of AI technologies based on rigorous scientific research and grounded in international human rights law, guidance for its implementation in all stages of the AI system life cycle, and capacity-building materials to support Member States' efforts to train government officials, policy-makers and other relevant AI actors on EIA methodology;
- b) developing a UNESCO readiness assessment methodology to assist Member States in identifying their status at specific moments of their readiness trajectory along a continuum of dimensions;

- c) Entwicklung einer UNESCO-Methodik zur Ex-ante- und Ex-post-Evaluierung der Wirksamkeit und Effizienz der Strategien zur KI-Ethik und den Impulssetzungen im Hinblick auf die Erreichung definierter Ziele;
- d) Stärkung forschungs- und evidenzbasierter Untersuchungen von Strategien zu KI-Ethik und die diesbezügliche Berichterstattung;
- e) Sammlung und Verbreitung von Informationen über Fortschritte und Innovationen sowie von Forschungsberichten, wissenschaftlichen Veröffentlichungen, Daten und Statistiken zu KI-Ethik-Strategien, auch im Rahmen bestehender Initiativen, um voneinander zu lernen, den Austausch bewährter Verfahren zu fördern und die Umsetzung dieser Empfehlung voranzutreiben.

(132) Bei den Verfahren zur Überwachung und Evaluierung sollte eine breite Beteiligung aller Interessengruppen gewährleistet sein, einschließlich besonders schutzbedürftiger Menschen und Menschen in gefährdeten Umständen, jedoch nicht auf diese beschränkt. Dabei muss der sozialen, kulturellen und geschlechtlichen Vielfalt Rechnung getragen werden, um Lernprozesse zu verbessern und die Beziehungen zwischen Ergebnissen, Entscheidungsfindungsprozessen, Transparenz und Rechenschaftspflicht für die Ergebnisse zu stärken.

- c) developing a UNESCO methodology to evaluate ex ante and ex post the effectiveness and efficiency of the policies for AI ethics and incentives against defined objectives;
- d) strengthening the research- and evidence-based analysis of and reporting on policies regarding AI ethics;
- e) collecting and disseminating progress, innovations, research reports, scientific publications, data and statistics regarding policies for AI ethics, including through existing initiatives, to support sharing best practices and mutual learning, and to advance the implementation of this Recommendation.

(132) Processes for monitoring and evaluation should ensure broad participation of all stakeholders, including, but not limited to, vulnerable people or people in vulnerable situations. Social, cultural and gender diversity should be ensured, with a view to improving learning processes and strengthening the connections between findings, decision-making, transparency and accountability for results.

(133) Im Interesse der Förderung bewährter Strategien und Verfahren im Zusammenhang mit der Ethik der künstlichen Intelligenz sollten geeignete Instrumente und Indikatoren entwickelt werden, um deren Wirksamkeit und Effizienz im Hinblick auf vereinbarte Standards, Prioritäten und Ziele zu bewerten, einschließlich spezifischer Ziele für Personen, die benachteiligten und marginalisierten Bevölkerungsgruppen angehören, sowie für besonders schutzbedürftige Menschen und Menschen in gefährdeten Umständen, und um die Auswirkungen von KI-Systemen auf individueller und gesellschaftlicher Ebene beurteilen zu können. Die Überwachung und Bewertung der Auswirkungen von KI-Systemen und damit in Zusammenhang stehenden KI-Ethikstrategien und -verfahren sollte kontinuierlich und systematisch in einer Weise erfolgen, die den jeweiligen Risiken angemessen ist. Dies sollte auf der Grundlage international vereinbarter Rahmen geschehen, und die Evaluierung privater wie auch öffentlicher Institutionen, Anbieter und Programme beinhalten, wobei auch Selbstevaluierungen, Tracer-Methoden und die Entwicklung von Indikatorreihen einzubeziehen sind. Die Datenerhebung und -verarbeitung muss im Einklang mit dem Völkerrecht, nationalen Rechtsvorschriften zum Datenschutz und zum Schutz der Privatsphäre sowie mit den in dieser Empfehlung dargelegten Werten und Prinzipien erfolgen.

(134) Insbesondere bieten sich bestimmte Mechanismen für die Mitgliedstaaten zur Überwachung und Evaluierung an, wie Ethikkommissionen, Beobachtungsstellen für KI-Ethik, Repositorien für menschenrechtskonforme und ethische Entwicklung von KI-Systemen oder Beiträge zu bestehenden Initiativen durch Thematisierung der Einhaltung ethischer Prinzipien in allen Zuständigkeitsbereichen der UNESCO, einen Mechanismus für den Erfahrungsaustausch, KI-Reallabore und einen Leitfaden für alle KI-Akteure, anhand dessen sie feststellen können, inwieweit sie in diesem Dokument aufgeführten Handlungsempfehlungen beachten.

(133) In the interests of promoting best policies and practices related to ethics of AI, appropriate tools and indicators should be developed for assessing the effectiveness and efficiency thereof against agreed standards, priorities and targets, including specific targets for persons belonging to disadvantaged, marginalized populations, and vulnerable people or people in vulnerable situations, as well as the impact of AI systems at individual and societal levels. The monitoring and assessment of the impact of AI systems and related AI ethics policies and practices should be carried out continuously in a systematic way proportionate to the relevant risks. This should be based on internationally agreed frameworks and involve evaluations of private and public institutions, providers and programmes, including self-evaluations, as well as tracer studies and the development of sets of indicators. Data collection and processing should be conducted in accordance with international law, national legislation on data protection and data privacy, and the values and principles outlined in this Recommendation.

(134) In particular, Member States may wish to consider possible mechanisms for monitoring and evaluation, such as an ethics commission, AI ethics observatory, repository covering human rights-compliant and ethical development of AI systems, or contributions to existing initiatives by addressing adherence to ethical principles across UNESCO's areas of competence, an experiencesharing mechanism, AI regulatory sandboxes, and an assessment guide for all AI actors to evaluate their adherence to policy recommendations mentioned in this document.

VI. Nutzung und gewinnbringender Einsatz der vorliegenden Empfehlung

(135) Die Mitgliedstaaten und alle anderen in dieser Empfehlung genannten Interessengruppen sind gehalten, die hier dargelegten ethischen Werte, Prinzipien und Normen im Zusammenhang mit KI zu respektieren, zu fördern und zu schützen und alle realisierbaren Schritte zu unternehmen, um die Handlungsempfehlungen in die Tat umzusetzen.

(136) Die Mitgliedstaaten sollten sich bemühen, ihr eigenes Tätigwerden im Hinblick auf diese Empfehlung zu erweitern und auszubauen, indem sie mit allen relevanten nationalen und internationalen staatlichen und nichtstaatlichen Organisationen sowie mit transnationalen Unternehmen und wissenschaftlichen Organisationen zusammenarbeiten, deren Aktivitäten in den Anwendungsbereich und unter die Ziele dieser Empfehlung fallen. Die Entwicklung einer UNESCO-Methodik zur ethischen Folgenabschätzung und die Einrichtung nationaler KI-Ethikkommissionen können dabei wichtige Instrumente sein.

VI. Utilization and exploitation of the present Recommendation

(135) Member States and all other stakeholders as identified in this Recommendation should respect, promote and protect the ethical values, principles and standards regarding AI that are identified in this Recommendation, and should take all feasible steps to give effect to its policy recommendations.

(136) Member States should strive to extend and complement their own action in respect of this Recommendation, by cooperating with all relevant national and international governmental and non-governmental organizations, as well as transnational corporations and scientific organizations, whose activities fall within the scope and objectives of this Recommendation. The development of a UNESCO Ethical Impact Assessment methodology and the establishment of national commissions for the ethics of AI can be important instruments for this.

VII. Einsatz für die vorliegende Empfehlung

(137) Von allen Organisationen der Vereinten Nationen ist es primär die Aufgabe der UNESCO, sich für diese Empfehlung einzusetzen und sie bekannt zu machen; sie wird dabei mit anderen betroffenen Organisationen der Vereinten Nationen zusammenarbeiten; dabei respektiert sie deren Mandat und vermeidet Doppelarbeit.

(138) Die UNESCO, einschließlich ihrer Gremien wie der Weltkommission für Ethik in Wissenschaft und Technologie (COMEST), des Internationalen Ausschusses für Bioethik (IBC) und des Zwischenstaatlichen Ausschusses für Bioethik (IGBC), wird auch mit anderen internationalen, regionalen und subregionalen staatlichen und nichtstaatlichen Organisationen zusammenarbeiten.

(139) Wenngleich das Mandat zur Förderung und zum Schutz innerhalb der UNESCO in die Zuständigkeit der Regierungen und zwischenstaatlichen Gremien fällt, ist auch die Zivilgesellschaft ein wichtiger Akteur, wenn es darum geht, für die Interessen des öffentlichen Sektors einzutreten; daher muss die UNESCO deren Legitimität sicherstellen und dafür werben.

VII. Promotion of the present Recommendation

(137) UNESCO has the vocation to be the principal United Nations agency to promote and disseminate this Recommendation, and accordingly will work in collaboration with other relevant United Nations entities, while respecting their mandate and avoiding duplication of work.

(138) UNESCO, including its bodies, such as the World Commission on the Ethics of Scientific Knowledge and Technology (COMEST), the International Bioethics Committee (IBC) and the Intergovernmental Bioethics Committee (IGBC), will also work in collaboration with other international, regional and sub-regional governmental and non-governmental organizations.

(139) Even though, within UNESCO, the mandate to promote and protect falls within the authority of governments and intergovernmental bodies, civil society will be an important actor to advocate for the public sector's interests and therefore UNESCO needs to ensure and promote its legitimacy.

VIII. Schluss- bestimmungen

(140) Diese Empfehlung muss als Ganzes betrachtet werden; die zugrunde liegenden Werte und Prinzipien sind als sich ergänzend und zueinander in Wechselbeziehung stehend zu verstehen.

(141) Kein Teil dieser Empfehlung kann so ausgelegt werden, dass er an die Stelle völkerrechtlicher Verpflichtungen oder Rechte von Staaten tritt, zu Änderungen daran führt oder sie anderweitig beeinträchtigt, und nichts darin kann als Billigung einer Tätigkeit oder Handlung eines Staates oder eines sonstigen politischen, wirtschaftlichen oder sozialen Akteurs, einer Gruppe oder einer Person verstanden werden, die den Menschenrechten und Grundfreiheiten, der Menschenwürde und dem sorgsamem Umgang mit der Umwelt und Ökosystemen, sowohl lebender als auch unbelebter Natur, zuwiderläuft.

VIII. Final provisions

(140) This Recommendation needs to be understood as a whole, and the foundational values and principles are to be understood as complementary and interrelated.

(141) Nothing in this Recommendation may be interpreted as replacing, altering or otherwise prejudicing States' obligations or rights under international law, or as approval for any State, other political, economic or social actor, group or person to engage in any activity or perform any act contrary to human rights, fundamental freedoms, human dignity and concern for the environment and ecosystems, both living and non-living.

Impressum

Herausgeberin

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

Kontakt

Deutsche UNESCO-Kommission e.V.
T +49 (0) 228 60 497-165
E zukunft@unesco.de
🌐 www.unesco.de

Übersetzung

Silke Rasche-Walther

ISBN

978-3-947675-38-8

Gefördert durch:



Auswärtiges Amt

Copyright

Die Texte dieser Publikation sind lizenziert unter der Creative-Commons-Lizenz Namensnennung – Weitergabe unter gleichen Bedingungen 4.0 International (CC-BY-NC-SA 4.0).

🌐 www.creativecommons.org/licenses/by-sa/4.0/

Design und Layout

Panatom

