

Die UNESCO-
Empfehlung zur

Ethik der Künstlichen Intelligenz: Wegweiser für die Gestaltung unserer Zukunft



unesco

UNESCO-Kommissionen
Deutschland, Niederlande
und Slowenien

Die UNESCO-
Empfehlung zur
**Ethik der Künst-
lichen Intelligenz:**
Wegweiser für
die Gestaltung
unserer Zukunft

Die UNESCO-Empfehlung in Kürze

Die UNESCO-Empfehlung zur Ethik der Künstlichen Intelligenz (KI) ist ein weltweit gültiger Völkerrechtstext, der sicherstellen soll, dass jede Organisation, jedes Unternehmen und jede Einzelperson, die KI entwickelt oder einsetzt, dies auf ethische Weise und in Einklang mit den Menschenrechten tut.

Mit ethischer Entwicklung und Nutzung von KI ist gemeint, dass unsere Interaktionen mit KI-Systemen zu Frieden, Menschenwürde, Nachhaltigkeit und Sicherheit beitragen und diesen Grundprinzipien nicht zuwiderlaufen.

Dazu müssen Kinder und Erwachsene auf der ganzen Welt die Möglichkeit erhalten, angemessenes Wissen

über die technologischen und ethischen Aspekte von KI zu erlangen.

Die Empfehlung der UNESCO ist ein Leitfaden, der Regierungen aufzeigt, wie sie die ethische Perspektive bei der Entwicklung von KI-Gesetzen und -Strategien berücksichtigen sollten.

Die UNESCO veröffentlicht regelmäßig Empfehlungen. Dabei handelt es sich um Instrumente des Völkerrechts, die von allen 193 Mitgliedstaaten der UNESCO verabschiedet wurden.

Nach der Verabschiedung einer Empfehlung müssen die Mitgliedstaaten regelmäßig darüber berichten, wie sie diese umsetzen.

Wegweiser für die Gestaltung unserer Zukunft

Die UNESCO-Empfehlung zur Ethik der Künstlichen Intelligenz:

Warum brauchen wir eine globale Vereinbarung über menschenorientierte KI?

→ Künstliche Intelligenz ist Teil unseres Alltags geworden. Sie hat enormes Potenzial, Gutes zu bewirken und zu nachhaltiger Entwicklung beizutragen. KI kann zum Beispiel in der Landwirtschaft eingesetzt werden, um Ernteerträge vorherzusagen und so dazu beitragen, unsere Ernährung zu sichern. Der Einsatz von KI birgt jedoch auch Risiken. In einigen Fällen wirft die Nutzung von KI-Systemen ethische Fragen auf, weil sie geschlechtsspezifische Diskriminierung verstärken, die Privatsphäre verletzen oder die Manipulation von Medieninhalten erleichtern können. Da alle Länder gleichermaßen mit diesen Fragen konfrontiert sind und die meisten KI-Systeme weltweit eingesetzt werden können, brauchen wir eine grenzübergreifende Perspektive für die ethische Entwicklung und Nutzung von KI. Im November 2021 verabschiedeten die 193 Mitgliedstaaten der UNESCO deshalb den ersten globalen Völkerrechtstext für eine menschenorientierte KI-Entwicklung und -Nutzung: Die Empfehlung zur Ethik der Künstlichen Intelligenz.

Was sind die Hauptziele der Empfehlung?

→ Die Empfehlung bietet Leitplanken für die Entwicklung und den Einsatz von KI-Systemen, damit diese dem Wohl von Mensch und Umwelt dienen und keinen Schaden anrichten. Als weltweit anerkannter Rechtstext formuliert sie ethische Werte und Prinzipien und zeigt, wie diese praktisch umgesetzt werden können. Dazu werden die UNESCO-Mitgliedstaaten aufgefordert, konkrete Maßnahmen in elf verschiedenen Politikfeldern zu ergreifen, darunter Gesundheit, Umwelt und Geschlechtergerechtigkeit. Mit diesem Ansatz zielt die Empfehlung auf den Schutz und die Förderung der Menschenrechte und Grundfreiheiten sowie der Menschenwürde und der Gleichberechtigung ab. In der Empfehlung wird außerdem betont, dass sich ethische Fragen in allen Phasen des Lebenszyklus eines KI-Systems stellen: von der Forschung, dem Design und der Entwicklung des Systems über seine Einführung und Nutzung – einschließlich Wartung, Betrieb, Überwachung und Evaluierung – bis hin zur Beendigung seines Einsatzes.

→ Die Empfehlung wurde von allen 193 UNESCO-Mitgliedsstaaten einstimmig verabschiedet. Sie legt erstmals einen Schwerpunkt auf die besonderen Bedürfnisse und Beiträge von Ländern mit niedrigem und mittlerem Einkommen bei der ethischen Entwicklung und Nutzung von KI. Dabei wird die Gestaltung einer menschenorientierten KI als gesamtgesellschaftliche Aufgabe definiert, die weit über die bloße Regulierung von Technologien hinausgeht. Gleichzeitig betont die Empfehlung, dass jede Regulierung von KI darauf ausgerichtet sein muss, die Menschenrechte zu respektieren und zu fördern.

→ Weiterhin werden in der Empfehlung alle Bereiche des menschlichen Lebens, die durch KI beeinflusst werden (können), umfassend beleuchtet. Dabei werden auch Bereiche in den Blick genommen, die in der bisherigen Diskussion um KI vernachlässigt wurden, **wie zum Beispiel ...**

... die unterschiedlichen Entwicklungsstufen von Ländern,

... die Bedürfnisse marginalisierter/ gefährdeter Gruppen,

... die Interessen des Globalen Südens,

... geschlechtsspezifische Vorurteile und Diskriminierung

... und Nachhaltigkeit.



Empfehlung zur
Ethik der Künstlichen
Intelligenz, verabschiedet
am 23. November 2021
(englischer Originaltext)

Was ist „KI“?

→ Systeme der Künstlichen Intelligenz sind in der Lage, Daten und Informationen auf eine Weise zu verarbeiten, die menschlichen Denkprozessen ähnelt oder diese nachzuahmen scheint. Dies beinhaltet Aspekte des Denkens und Lernens, der Wahrnehmung, Vorhersage, Planung oder Steuerung.

Die Systeme können mit verschiedenen Methoden arbeiten, wie zum Beispiel Machine Learning (maschinelles Lernen) und Machine Reasoning (maschinelles Denken).

Die UNESCO-Empfehlung zur Ethik der Künstlichen Intelligenz:

Wegweiser für die Gestaltung unserer Zukunft

Was sind die wichtigsten Forderungen der Empfehlung?

Um eine ethische Entwicklung und Nutzung von KI zu gewährleisten, definiert die Empfehlung konkrete Rahmenbedingungen in elf Politikfeldern. Die folgenden Beispiele stellen eine Auswahl dieser Rahmenbedingungen dar und veranschaulichen die inhaltliche Bandbreite der Empfehlung:

- 1 Gesetze und Strategien zur Datenregulierung müssen das Recht auf Privatsphäre schützen, auf zentrale Bedenken eingehen (etwa in Zusammenhang mit KI-gestützter Überwachung) und sicherstellen, dass die Rechte des Einzelnen an seinen Daten respektiert werden.
- 2 Die internationale Zusammenarbeit im Bereich der KI muss verstärkt werden, um die drängenden Herausforderungen zu bewältigen und die digitale Kluft zwischen dem Globalen Süden und dem Globalen Norden zu verringern.
- 3 Ethische Folgenabschätzungen müssen während des gesamten Lebenszyklus von KI-Systemen durchgeführt werden, um Schaden von Menschen abzuwenden und um sicherzustellen, dass Menschenrechte respektiert und gefördert werden, wenn KI zum Einsatz kommt.
- 4 Steuerungsmechanismen, wie die Ernennung von KI-Ethik-Beauftragten, müssen geschaffen werden, um potenzielle Schäden durch KI-Systeme abzuwenden bzw. sie zu untersuchen sowie um Wiedergutmachungs- und Abhilfemaßnahmen durchzusetzen.
- 5 Es müssen Anreize für die Entwicklung KI-gestützter Anwendungen zur Überwachung, zum Schutz und zur Wiederherstellung von Ökosystemen geschaffen werden.

- 6 KI muss so entwickelt und eingesetzt werden, dass sie gesellschaftliche Vielfalt (einschließlich kultureller Vielfalt) und Integration bestmöglich fördert. Sie muss Nichtdiskriminierung gewährleisten, Meinungsfreiheit fördern und Geschlechtergerechtigkeit verwirklichen.
- 7 Politische Maßnahmen müssen darauf abzielen, den Bürgerinnen und Bürgern angemessene KI-Kenntnisse zu vermitteln, damit sie mit dieser Technologie umgehen können. Nur so lassen sich die digitalen Gräben in der Gesellschaft verringern.
- 8 KI-Systeme müssen den Zugang zu Informationen und Wissen verbessern.
- 9 Regierungen und private Unternehmen müssen dafür sorgen, dass für Beschäftigte, deren Arbeitsplätze durch neue KI-Technologien und -Systeme bedroht sind, ein fairer Übergang geschaffen wird.
- 10 Die Politik muss Leitlinien für Mensch-Roboter-Interaktion und ihre Auswirkungen auf die zwischenmenschlichen Beziehungen entwickeln. Dabei muss sie insbesondere die psychische und physische Gesundheit der Menschen berücksichtigen.



Aufruf zum Handeln in fünf Schlüsselbereichen

Die Empfehlung gibt den nationalen Regierungen einen klaren ethischen Rahmen für die Entwicklung konkreter politischer Maßnahmen vor. In diesem „Übersetzungsprozess“ der Vorgaben der Empfehlung in nationales Recht kommt sowohl der Zivilgesellschaft als auch dem Privatsektor eine entscheidende Rolle zu. In den folgenden Abschnitten werden fünf der elf Politikfelder der Empfehlung näher beleuchtet. Es handelt sich dabei um Felder, in denen der Einsatz von KI besonders tiefgreifende gesellschaftliche Veränderungen auslöst und dadurch im Alltag für jeden deutlich spürbar wird. Innerhalb dieser Felder werden jeweils konkrete Vorgaben aus der Empfehlung vorgestellt.



● Gesundheit

Im Gesundheitsbereich sind die potenziellen Chancen und Risiken durch KI besonders greifbar. KI revolutioniert bereits heute die Früherkennung von Krankheiten, eröffnet neue Behandlungsmöglichkeiten und kann möglicherweise dazu beitragen, dem zunehmenden Personalmangel im Gesundheits- und Pflegesektor zu begegnen. Der Bereich ist jedoch auch mit wichtigen ethischen Fragen konfrontiert, zum Beispiel rund um den Schutz persönlicher Gesundheitsdaten und den Umgang mit Gehirn-Computer-Schnittstellen. Verwaltungspersonal in Gesundheitseinrichtungen, Patientinnen und Patienten sowie Ärztinnen und Ärzte werden von diesen Fragen direkt betroffen sein.

Regierungen und Akteure des Gesundheitswesens sind daher aufgerufen, folgende Maßnahmen umzusetzen:

- Die Entwicklung von Leitlinien für Mensch-Roboter-Interaktionen und die Untersuchung der Auswirkungen dieser Interaktionen auf zwischenmenschliche Beziehungen; dabei muss ein Fokus auf der psychischen und körperlichen Gesundheit der Menschen liegen, damit Mensch-Roboter-Interaktionen im Einklang mit den Menschenrechten, den Grundfreiheiten, der Förderung der Vielfalt und dem Schutz gefährdeter Personen stehen, insbesondere beim Einsatz von KI-gestützten Systemen in Neurotechnologien und Gehirn-Computer-Schnittstellen;
- Den Einsatz von KI-Systemen im Dienste der menschlichen Gesundheit bei der globalen Überwachung von Krankheitsausbrüchen und dem entsprechenden Datenaustausch;
- Die Förderung von Forschung zu den Auswirkungen der Interaktion von Menschen mit KI-Systemen; besonders berücksichtigt werden sollte dabei, welche psychologischen und kognitiven Auswirkungen diese Systeme auf Kinder und Jugendliche haben können.

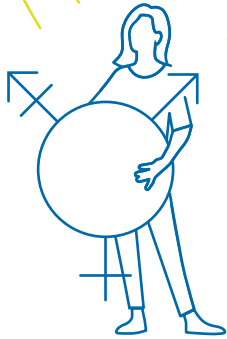
● Kultur

KI hat erhebliche Auswirkungen darauf, wie wir Kultur „produzieren“, wie (viel) Verbraucherinnen und Verbraucher für Kulturgüter bezahlen und inwieweit Kunstschaffende von ihrer Arbeit leben können. Schon heute beeinflussen und entscheiden Algorithmen über unseren „Kulturkonsum“, indem sie beispielsweise mitbestimmen, welche Musik wir hören.

Wichtige Aufgaben für die Regierungen der UNESCO-Mitgliedstaaten sind deshalb:

- Die Zusammenarbeit mit Technologieunternehmen und anderen relevanten Akteuren, um die Vielfalt kultureller Ausdrucksformen und den Zugang zu ihnen zu fördern und gleichzeitig die Sichtbarkeit und Auffindbarkeit lokaler Inhalte sicherzustellen;
- Die Förderung weiterer Forschung zu KI und geistigem Eigentum, um festzustellen, ob und wie Werke, die durch KI oder mit Hilfe von KI geschaffen wurden, durch Rechte des geistigen Eigentums geschützt werden sollten;
- Die Untersuchung der Auswirkungen von KI-Systemen auf die Nuancen menschlicher Sprache und des kulturellen Ausdrucks – insbesondere von Anwendungen, die natürliche Sprache verarbeiten, wie maschinelle Übersetzungstools und Sprachassistenten.





● Geschlechtergerechtigkeit

Durch den Einsatz von KI-Systemen können Vorurteile entstehen oder verstärkt werden. So kann Übersetzungssoftware beispielsweise geschlechtsneutrale Eingaben („the doctor“) in geschlechtsspezifische Ausgaben („der Arzt“) übersetzen. Ähnlich verhält es sich, wenn Unternehmen in einer männerdominierten Branche neue Mitarbeitende mithilfe eines KI-Systems auswählen, das mit den Daten aus den Lebensläufen der vorhandenen Beschäftigten trainiert wurde. Dabei reproduzieren sie mit hoher Wahrscheinlichkeit Vorurteile gegenüber weiblichen Bewerberinnen. Ein wichtiger Grundsatz der Empfehlung ist es daher, diskriminierende oder voreingenommene Ergebnisse von KI-Systemen während ihres gesamten Lebenszyklus zu minimieren und bestenfalls zu vermeiden.

Politisch Verantwortliche, Arbeitgeberinnen und Arbeitgeber in der KI-Forschung und -Entwicklung wie auch andere KI-Akteure sollten daher folgende Schritte einleiten:

- Die Verabschiedung eines Gender-Aktionsplans als Teil der nationalen Digitalpolitik;
- Die Erhöhung der Investitionen in Bereiche wie berufliche Bildung, um beispielsweise die Karrieremöglichkeiten von Frauen in den MINT- und IKT-Fächern zu fördern;
- Die Gewährleistung, dass geschlechtsspezifische Stereotype und diskriminierende Vorurteile nicht in KI-Systeme einfließen (z.B. indem vielfältige, offene und vertrauenswürdige „Gold-Standard-Datensätze“ geschaffen werden).

Die UNESCO-Empfehlung zur Ethik der Künstlichen Intelligenz:

Wegweiser für die Gestaltung unserer Zukunft

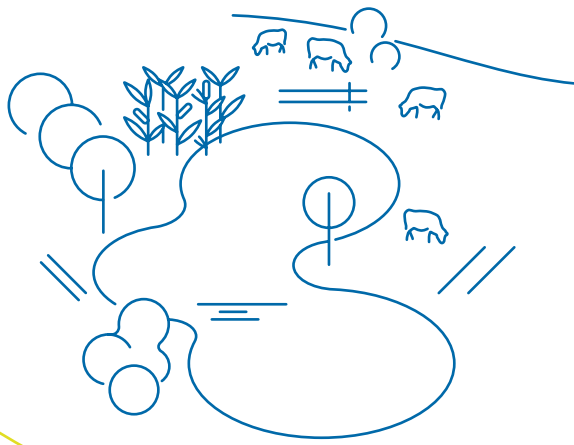


● Bildung und Forschung

Schulen und außerschulische Lernumgebungen spielen eine entscheidende Rolle bei der Sensibilisierung für ethische KI-Entwicklung und -Nutzung. Sie sind ebenfalls bedeutend für die Bemühungen, die digitalen Gräben innerhalb der Gesellschaft zu überwinden. In der Empfehlung wird dem Bildungssektor eine zentrale Rolle bei der Entwicklung von technologischer und ethischer KI-Kompetenz bei Kindern und Erwachsenen weltweit zugesprochen. Gleichzeitig erwirft der Einsatz von E-Learning in Schulen eigene ethische Dilemmata auf. Beispielsweise sollte die Verwendung personenbezogener Daten von Schülerinnen und Schülern den einschlägigen Datenschutzstandards entsprechen. Außerdem muss sorgfältig abgewogen werden, wann und wie KI-gestützte Anwendungen in Schulen eingesetzt werden.

Die nationalen Regierungen, aber auch Arbeitgeberinnen und Arbeitgeber im Bildungssektor, Lehrkräfte und Nichtregierungsorganisationen sind aufgefordert, folgende Maßnahmen zu ergreifen:

- Die Förderung von Sensibilisierungsprogrammen für Entwicklungen im KI-Bereich und für die Auswirkungen von KI-Systemen auf die Menschenrechte von Kindern;
- Die Förderung einer besseren Verbindung zwischen der Vermittlung technischer Fertigkeiten und der Vermittlung von Ethik im KI-Bereich;
- Die Unterstützung von interdisziplinärer KI-Forschung jenseits der Informatik, der Ingenieurwissenschaften und der MINT-Fächer;
- Die Unterstützung wissenschaftlicher Politikberatung zu KI sowie die Förderung eines stärkeren Bewusstseins für die Ethik der KI in der Politik.



● Umwelt und Ökosysteme

Künstliche Intelligenz trägt teilweise bereits zum Schutz der Umwelt bei. So wird KI beispielsweise bei der Analyse von Drohnen- und Satellitenbildern eingesetzt, um biologische Vielfalt zu überwachen und zu erhalten. KI-Systeme können jedoch auch negative Auswirkungen auf die Umwelt haben, zum Beispiel verbraucht die Rechenleistung der Systeme sehr viel Energie.

Die Regierungen der Mitgliedstaaten sind daher aufgerufen, die folgenden Maßnahmen umzusetzen:

- Die Prüfung der direkten und indirekten Umweltauswirkungen von KI-Systemen während ihres gesamten Lebenszyklus und die Überprüfung ihrer Konformität mit Umweltvorschriften;
- Die Gewährleistung, dass KI-gestützte Anwendungen für die Katastrophenvorsorge sowie für die Überwachung, den Schutz und die Regeneration von Ökosystemen entwickelt und eingeführt werden;
- Das Verbot von KI-Systemen, die mit hoher Wahrscheinlichkeit unverhältnismäßig negative Auswirkungen auf die Umwelt haben.

Es ist an der Zeit, eine ethische und menschenorientierte Entwicklung von KI zu unterstützen!

Ziel der UNESCO-Empfehlung ist es, dass KI zum Nutzen der Menschen und der Gesellschaft entwickelt wird. Um sicherzustellen, dass eine auf den Menschen ausgerichtete KI-Entwicklung die Norm wird und bleibt, sind Beiträge aller Beteiligten erforderlich – nicht nur von politisch Verantwortlichen, sondern auch von Organisationen der Zivilgesellschaft, privaten Unternehmen und der breiten Öffentlichkeit. Es bedarf gemeinsamer Anstrengungen, um das derzeit noch offene Möglichkeitsfenster für die Entwicklung ethischer und menschenorientierter KI-Systeme und -Anwendungen zu nutzen. Die Empfehlung der UNESCO bietet hierfür eine wertvolle Grundlage.

Impressum

Herausgeber

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


Niederländische
UNESCO-Nationalkommission
Prins Willem-Alexanderhof 20
2595 BE Den Haag, Niederlande

Slowenische UNESCO-Nationalkommission
Kotnikova 38
1000 Ljubljana, Slowenien

Kontakt

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

Niederländische
UNESCO-Nationalkommission
T +31 (0) 70-3315484
E info@unesco.nl
 www.unesco.nl

Slowenische UNESCO-Nationalkommission
T +386 (0) 1-4784799
E unesco.mizs@gov.si
 [www.gov.si/en/state-authorities/
ministries/ministry-of-higher-education-
science-and-innovation/about-the-ministry/
urad-za-unesco/](http://www.gov.si/en/state-authorities/ministries/ministry-of-higher-education-science-and-innovation/about-the-ministry/urad-za-unesco/)

Redaktion

Anne Diessner, Jeannine Hausmann
(verantwortlich), Maximilian Müngersdorff,
Davor Orlic, Jon Verriet

Übersetzung

Anne Diessner


Stand

April 2023

ISBN

978-3-947675-34-0

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

Gefördert von



Auswärtiges Amt

Wegweiser für die Gestaltung unserer Zukunft

Die UNESCO-Empfehlung zur Ethik der Künstlichen Intelligenz:

mit Unterstützung von



unesco

Centre
Under the auspices
of UNESCO



International Research Centre
of Artificial Intelligence
under the auspices of UNESCO