



Empfehlung des Rats zu künstlicher Intelligenz

**OECD-Rechts-
instrumente**



Dies ist keine amtliche Übersetzung. Obwohl die größtmöglichen Anstrengungen unternommen wurden, um die Übereinstimmung mit den Originaltexten zu gewährleisten, sind der englische und der französische Text die einzigen amtlichen Fassungen, die auf der OECD-Website <https://legalinstruments.oecd.org> zur Verfügung stehen.

DER RAT,

IM HINBLICK AUF Artikel 5b) des Übereinkommens über die Organisation für wirtschaftliche Zusammenarbeit und Entwicklung vom 14. Dezember 1960;

IM HINBLICK AUF die OECD-Leitsätze für multinationale Unternehmen [[OECD/LEGAL/0144](#)]; die Empfehlung des Rats über Leitlinien für den Schutz des Persönlichkeitsbereichs und den grenzüberschreitenden Verkehr personenbezogener Daten (Recommendation of the Council concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data) [[OECD/LEGAL/0188](#)]; die Empfehlung des Rats über die Leitlinien für die Kryptographiepolitik (Recommendation of the Council concerning Guidelines for Cryptography Policy) [[OECD/LEGAL/0289](#)]; die Empfehlung des Rats für einen verbesserten Zugang zu und eine effektivere Nutzung von Informationen des öffentlichen Sektors (Recommendation of the Council on Enhanced Access and More Effective Use of Public Sector Information) [[OECD/LEGAL/0362](#)]; die Empfehlung des Rats über Risikomanagement im Bereich der digitalen Sicherheit für wirtschaftlichen und sozialen Wohlstand (Recommendation of the Council on Digital Security Risk Management for Economic and Social Prosperity) [[OECD/LEGAL/0415](#)]; die Empfehlung des Rats über den Verbraucherschutz im elektronischen Geschäftsverkehr (Recommendation of the Council on Consumer Protection in E-commerce) [[OECD/LEGAL/0422](#)]; die Erklärung zur digitalen Wirtschaft: Innovation, Wachstum und sozialer Wohlstand (Declaration on the Digital Economy: Innovation, Growth and Social Prosperity) (Erklärung von Cancún) [[OECD/LEGAL/0426](#)]; die Erklärung zur Stärkung von KMU und unternehmerischer Initiative für Produktivität und inklusives Wachstum (Declaration on Strengthening SMEs and Entrepreneurship for Productivity and Inclusive Growth) [[OECD/LEGAL/0439](#)]; sowie die auf der Tagung der für Arbeit und Beschäftigung zuständigen Ministerinnen und Minister der OECD-Länder verabschiedete Erklärung des Rats der OECD auf Ministerebene von 2016 zur Erhöhung der Resilienz und der Inklusivität der Arbeitsmärkte (2016 Ministerial Statement on Building more Resilient and Inclusive Labour Markets);

IM HINBLICK AUF die in der Agenda 2030 für nachhaltige Entwicklung aufgeführten und von der Generalversammlung der Vereinten Nationen verabschiedeten Ziele für nachhaltige Entwicklung (A/RES/70/1) sowie die 1948 verabschiedete Allgemeine Erklärung der Menschenrechte;

IM HINBLICK AUF die derzeit in anderen internationalen Regierungs- und Nichtregierungsforen durchgeführten wichtigen Arbeiten im Bereich der künstlichen Intelligenz (im Folgenden als „KI“ bezeichnet);

IN DER ERKENNTNIS, dass KI tiefgreifende, weitreichende und globale Auswirkungen hat, die Gesellschaften, Wirtschaftszweige und die Arbeitswelt grundlegend verändern, und sich diese Entwicklung in Zukunft voraussichtlich noch verstärken wird;

IN DER ERKENNTNIS, dass KI das Potenzial besitzt, die Wohlfahrt und das Wohlergehen der Menschen zu verbessern, zu einer positiven nachhaltigen globalen Wirtschaftstätigkeit beizutragen, die Innovationstätigkeit und die Produktivität zu steigern und einen Beitrag zur Bewältigung dringender globaler Herausforderungen zu leisten;

IN DER ERKENNTNIS, dass diese Transformationen gleichzeitig sowohl innerhalb als auch auf Ebene der

Gesellschaften und Volkswirtschaften völlig unterschiedliche Effekte haben können, insbesondere im Hinblick auf wirtschaftliche Veränderungen, den Wettbewerb, Umbrüche auf dem Arbeitsmarkt, Ungleichheiten sowie die Auswirkungen auf die Demokratie und die Menschenrechte, den Schutz der Privatsphäre und den Datenschutz sowie die digitale Sicherheit;

IN DER ERKENNTNIS, dass Vertrauen eine wichtige Triebkraft des digitalen Wandels ist; dass die Vertrauenswürdigkeit von KI-Systemen ein entscheidender Faktor für die Verbreitung und Übernahme von KI ist, auch wenn die künftigen Arten und Auswirkungen von KI-Anwendungen u.U. nur schwer vorherzusehen sind; und dass eine fundierte, gesamtgesellschaftliche öffentliche Debatte notwendig ist, um das positive Potenzial der Technologie auszuschöpfen und zugleich die damit verbundenen Risiken zu begrenzen;

UNTER HINWEIS DARAUF, dass bestimmte bestehende nationale und internationale rechtliche, regulatorische und politische Regelwerke bereits für KI relevant sind, einschließlich derjenigen mit Bezug auf Menschenrechte, Verbraucherschutz und Schutz personenbezogener Daten, geistige Eigentumsrechte, verantwortungsbewusstes unternehmerisches Handeln und Wettbewerb, und in Anerkennung der Tatsache, dass einige Regelwerke möglicherweise im Hinblick auf ihre Eignung überprüft und neue Ansätze entwickelt werden müssen;

IN DER ERKENNTNIS, dass angesichts der raschen Entwicklung und Einführung von KI ein stabiles Politikumfeld für alle Stakeholder, unabhängig von deren jeweiliger Rolle und dem jeweiligen Kontext, erforderlich ist, das einen menschenzentrierten Ansatz für vertrauenswürdige KI fördert, die Forschung unterstützt und wirtschaftliche Innovationsanreize bewahrt;

IN DER ERWÄGUNG, dass es von entscheidender Bedeutung ist, die Chancen, die KI-Anwendungen bieten, zu nutzen und die daraus resultierenden Herausforderungen anzugehen, und die Stakeholder zu befähigen sich einzubringen, um die Einführung einer vertrauenswürdigen KI in der Gesellschaft zu fördern und dafür zu sorgen, dass die Vertrauenswürdigkeit von KI ein Wettbewerbsfaktor auf dem Weltmarkt wird;

Auf Vorschlag des Ausschusses für digitale Wirtschaft;

I. **KOMMT ÜBEREIN**, dass die folgenden Begriffe für die Zwecke dieser Empfehlung folgendermaßen zu verstehen sind:

- *KI-System*: Ein KI-System ist ein maschinenbasiertes System, das für bestimmte von Menschen definierte Ziele Voraussagen machen, Empfehlungen abgeben oder Entscheidungen treffen kann, die das reale oder virtuelle Umfeld beeinflussen. KI-Systeme können mit einem unterschiedlichen Grad an Autonomie ausgestattet sein.
- *Lebenszyklus eines KI-Systems*: Der Lebenszyklus eines KI-Systems besteht aus folgenden Phasen: *i)* „Design, Daten und Modelle“, wobei es sich um eine kontextabhängige Abfolge handelt, die Planung und Design, Datensammlung und -verarbeitung sowie Modellierung umfasst; *ii)* „Verifizierung und Validierung“; *iii)* „Einführung“ sowie *iv)* „Betrieb und Monitoring“. Diese Phasen finden häufig iterativ statt und sind nicht zwangsläufig sequenziell.

Die Entscheidung, ein KI-System außer Betrieb zu nehmen, kann zu jedem Zeitpunkt der Betriebs- oder Monitoringphase erfolgen.

- *KI-Wissen*: KI-Wissen bezieht sich auf die Kompetenzen und Ressourcen, die erforderlich sind, um den Lebenszyklus des KI-Systems zu verstehen und daran mitzuwirken, wie Daten, Codes, Algorithmen, Modelle, Forschung, Know-how, Ausbildungsprogramme, Governance, Prozesse und empfehlenswerte Praktiken.
- *KI-Akteure*: KI-Akteure sind Akteure, die im Lebenszyklus des KI-Systems eine aktive Rolle spielen, einschließlich Organisationen und Personen, die KI einführen oder anwenden.
- *Stakeholder*: Stakeholder umfassen alle Organisationen und Personen, die direkt oder indirekt an KI-Systemen mitwirken oder davon betroffen sind. KI-Akteure sind eine Untergruppe der Stakeholder.

Abschnitt 1: Grundsätze einer verantwortungsvollen Steuerung vertrauenswürdiger KI

II. **EMPFIEHLT**, dass die Mitglieder und Nichtmitglieder, die dieser Empfehlung zustimmen (im Folgenden die „zustimmenden Länder“ genannt) die folgenden für alle Stakeholder relevanten Grundsätze einer verantwortungsvollen Steuerung vertrauenswürdiger KI fördern und umsetzen.

III. **FORDERT** alle KI-Akteure **AUF**, die folgenden Grundsätze einer verantwortungsvollen Steuerung vertrauenswürdiger KI entsprechend ihrer jeweiligen Rolle zu fördern und umzusetzen.

IV. **UNTERSTREICHT**, dass die folgenden Grundsätze komplementär sind und als Ganzes betrachtet werden sollten.

1.1. Inklusives Wachstum, nachhaltige Entwicklung und Lebensqualität

Die Stakeholder sollten sich proaktiv für eine verantwortungsvolle Steuerung vertrauenswürdiger KI einsetzen, die darauf abzielt, einen Nutzen für die Menschen und den Planeten zu erzielen – wie eine Verbesserung der menschlichen Fähigkeiten und eine Steigerung der Kreativität, eine stärkere Einbindung unterrepräsentierter Bevölkerungsgruppen, eine Verringerung der wirtschaftlichen, sozialen, geschlechtsspezifischen und sonstigen Ungleichheiten sowie den Schutz der natürlichen Umwelt – und damit ein inklusives Wachstum, eine nachhaltige Entwicklung und die Lebensqualität zu fördern.

1.2. Menschenzentrierte Werte und Fairness

- a) Die KI-Akteure sollten während des gesamten Lebenszyklus eines KI-Systems das Prinzip der Rechtsstaatlichkeit, die Menschenrechte und demokratische Werte achten. Hierzu zählen u.a. Freiheit, Würde und Selbstbestimmung, Schutz der Privatsphäre und Datenschutz, Nichtdiskriminierung und Gleichbehandlung, Vielfalt, Fairness, soziale Gerechtigkeit und international anerkannte Arbeitsrechte.
- b) Zu diesem Zweck sollten die KI-Akteure Mechanismen und Schutzmaßnahmen wie die Möglichkeit menschlicher Entscheidungen einführen, die dem Kontext angemessen sind und dem neuesten Stand der Technik entsprechen.

1.3. Transparenz und Erklärbarkeit

Die KI-Akteure sollten sich in Bezug auf KI-Systeme zu Transparenz und verantwortungsvoller Offenlegung verpflichten. Zu diesem Zweck sollten sie sachdienliche, dem Kontext angemessene und dem neuesten Stand der Technik entsprechende Informationen bereitstellen, um:

- i. ein allgemeines Verständnis von KI-Systemen zu fördern,
- ii. den Stakeholdern ihre Interaktionen mit KI-Systemen bewusst zu machen, einschließlich jener am Arbeitsplatz,
- iii. die von einem KI-System Betroffenen zu befähigen, das Ergebnis zu verstehen, und
- iv. die Betroffenen, denen durch ein KI-System Nachteile entstehen, zu befähigen, das jeweilige Ergebnis auf der Grundlage von klaren und leicht verständlichen Informationen über die einzelnen Faktoren und die der jeweiligen Vorhersage, Empfehlung oder Entscheidung zugrunde liegende Logik zu hinterfragen.

1.4. Robustheit und Sicherheit

- a) KI-Systeme sollten während ihres gesamten Lebenszyklus robust und sicher sein, sodass sie bei normaler, vorhersehbarer bestimmungsgemäßer oder missbräuchlicher Verwendung bzw. unter sonstigen ungünstigen Bedingungen angemessen funktionieren und kein unzumutbares Sicherheitsrisiko darstellen.

- b) Zu diesem Zweck sollten KI-Akteure die Rückverfolgbarkeit gewährleisten – u.a. in Bezug auf Datensätze, Prozesse und die im Lebenszyklus eines KI-Systems getroffenen Entscheidungen –, um eine Analyse der Ergebnisse und eingeholten Antworten des KI-Systems zu ermöglichen, die dem Kontext angemessen ist und dem neuesten Stand der Technik entspricht.
- c) Die KI-Akteure sollten in allen Phasen des Lebenszyklus eines KI-Systems, basierend auf ihren Rollen, dem Kontext und ihrer Handlungsfähigkeit, stets einen systematischen Risikomanagementansatz verfolgen, um den Risiken im Zusammenhang mit KI-Systemen zu begegnen, einschließlich Risiken in Bezug auf den Schutz der Privatsphäre, die digitale Sicherheit und Verzerrungen (Bias).

1.5. Rechenschaftspflicht

Die KI-Akteure sollten für den einwandfreien Betrieb der KI-Systeme und die Achtung der oben genannten Grundsätze rechenschaftspflichtig sein; die Rechenschaftslegung sollte auf ihren Rollen und dem jeweiligen Kontext basieren und dem neuesten Stand der Technik entsprechen.

Abschnitt 2: Einzelstaatliche Maßnahmen und internationale Zusammenarbeit für vertrauenswürdige KI

V. **EMPFIEHLT**, dass die zustimmenden Länder die folgenden Empfehlungen im Rahmen ihrer nationalen Politik und ihrer internationalen Zusammenarbeit im Einklang mit den in Abschnitt 1 genannten Grundsätzen unter besonderer Berücksichtigung kleiner und mittlerer Unternehmen (KMU) umsetzen.

2.1. In KI-Forschung und -Entwicklung investieren

- a) Die Länder sollten im Hinblick auf Forschung und Entwicklung, einschließlich interdisziplinärer Anstrengungen, langfristige öffentliche Investitionen in Betracht ziehen und Anreize für private Investitionen schaffen, um Innovationen im Bereich einer vertrauenswürdigen KI zu fördern, deren Fokus auf schwierigen technischen Fragen sowie auf sozialen, rechtlichen und ethischen Implikationen bzw. Politikproblemen im Zusammenhang mit KI liegt.
- b) Die Länder sollten außerdem im Hinblick auf offene Datensätze, die repräsentativ sind und dem Schutz der Privatsphäre sowie dem Datenschutz gerecht werden, öffentliche Investitionen in Betracht ziehen und Anreize für private Investitionen schaffen, um ein KI-Forschungs- und -Entwicklungsumfeld ohne unangemessenen Bias zu fördern und die Interoperabilität sowie die Anwendung von Standards zu verbessern.

2.2. Ein digitales Ökosystem für KI fördern

Die Länder sollten die Entwicklung eines digitalen Ökosystems für vertrauenswürdige KI und den Zugang dazu fördern. Ein solches Ökosystem umfasst insbesondere die digitalen Technologien und die digitale Infrastruktur sowie die Mechanismen für einen angemessenen Austausch von KI-Wissen. Diesbezüglich sollten die Länder die Förderung von Mechanismen wie Datentreuhandmodellen in Betracht ziehen, um den

sicheren, fairen, legalen und ethisch korrekten Austausch von Daten zu stärken.

2.3. Ein für KI günstiges Politikumfeld schaffen

- a) Die Länder sollten ein Politikumfeld begünstigen, das einen zügigen Übergang vertrauenswürdiger KI-Systeme von der Forschungs- und Entwicklungsphase hin zur Errichtungs- und Betriebsphase fördert. Zu diesem Zweck sollten sie den Einsatz von Experimenten in Betracht ziehen, um ein kontrolliertes Umfeld zu schaffen, in dem KI-Systeme getestet und gegebenenfalls skaliert werden können.
- b) Die Länder sollten ihre für KI-Systeme geltenden Politik- und Regulierungsrahmen bzw. Evaluierungsmechanismen überprüfen und gegebenenfalls anpassen, um Innovationen und Wettbewerb in Bezug auf KI zu fördern.

2.4. Die Kompetenzen der Menschen stärken und sich auf den Wandel des Arbeitsmarkts vorbereiten

- a) Die Länder sollten eng mit den Stakeholdern zusammenarbeiten, um sich für den Wandel der Arbeitswelt und der Gesellschaft zu wappnen. Sie sollten die Menschen befähigen, KI-Systeme des gesamten Anwendungsspektrums effektiv zu nutzen und mit ihnen zu interagieren, u.a. indem sie sie mit den erforderlichen Kompetenzen ausstatten.
- b) Die Länder sollten, u.a. im Rahmen eines sozialen Dialogs, Maßnahmen ergreifen, die im Zuge der Einführung von KI eine für Arbeitskräfte faire Umstellung gewährleisten, beispielsweise durch Fort- und Weiterbildungsprogramme im Lauf des Arbeitslebens, Unterstützungsmaßnahmen für entlassene Arbeitskräfte und die Eröffnung neuer Arbeitsmarktchancen.
- c) Die Länder sollten darüber hinaus eng mit den betroffenen Akteuren zusammenarbeiten, um einen verantwortungsvollen Einsatz von KI in der Arbeitswelt zu fördern, die Sicherheit der Arbeitskräfte und die Beschäftigungsqualität zu erhöhen, unternehmerische Initiative und Produktivität zu stärken und sich zu bemühen, eine faire Teilhabe breiter Bevölkerungsschichten am Nutzen von KI zu gewährleisten.

2.5. Internationale Zusammenarbeit für vertrauenswürdige KI

- a) Die Länder, einschließlich der Entwicklungsländer, sollten aktiv zusammenarbeiten, auch mit den Stakeholdern, um diese Grundsätze zu stärken und bei der verantwortungsvollen Steuerung vertrauenswürdiger KI Fortschritte zu erzielen.
- b) Die Länder sollten im Rahmen der OECD und anderer globaler und regionaler Foren zusammenarbeiten, um den Austausch von KI-Wissen angemessen zu fördern. Sie sollten internationale, sektorübergreifende und offene Multi-Stakeholder-Initiativen unterstützen, um auf lange Sicht eine Expertise im KI-Bereich zu entwickeln.
- c) Die Länder sollten die Ausarbeitung konsensbasierter, globaler technischer Multi-Stakeholder-Standards für interoperable und vertrauenswürdige KI fördern.

- d) Die Länder sollten darüber hinaus die Entwicklung international vergleichbarer Messgrößen zur Evaluierung der KI-Forschung, -Entwicklung und -Einführung anregen und diese selbst nutzen, und eine Evidenzbasis aufbauen, um die Fortschritte zu evaluieren, die bei der Umsetzung dieser Grundsätze erzielt wurden.

VI. **ERSUCHT** den Generalsekretär und die zustimmenden Länder, diese Empfehlung zu verbreiten.

VII. **ERSUCHT** die nicht zustimmenden Länder, diese Empfehlung gebührend zur Kenntnis zu nehmen und ihr zuzustimmen.

VIII. **BEAUFTRAGT** den Ausschuss für digitale Wirtschaft:

- a) aufbauend auf dieser Empfehlung seine wichtige Arbeit im Bereich künstliche Intelligenz fortzusetzen und dabei den Arbeiten anderer internationaler Foren Rechnung zu tragen, und den Rahmen zur Bewertung der evidenzbasierten KI-Politik weiterzuentwickeln;
- b) weitere praktische Leitlinien zur Umsetzung dieser Empfehlung zu erarbeiten bzw. zu erneuern und dem Rat bis spätestens Ende Dezember 2019 über die erzielten Fortschritte Bericht zu erstatten;
- c) ein Forum für den Austausch von Informationen über KI-Maßnahmen und -Aktivitäten, einschließlich der Erfahrungen mit der Umsetzung dieser Empfehlung, bereitzustellen und einen interdisziplinären Dialog zwischen verschiedenen Stakeholdern zu fördern, um das Vertrauen in KI zu stärken und deren Einführung voranzutreiben;
- d) in Zusammenarbeit mit anderen einschlägigen Ausschüssen die Umsetzung dieser Empfehlung zu überwachen und dem Rat spätestens fünf Jahre nach deren Verabschiedung und im Anschluss daran regelmäßig darüber Bericht zu erstatten.