



Ethik-Briefing

Leitfaden für eine verantwortungsvolle Entwicklung
und Anwendung von KI-Systemen

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

 **acatech**
DEUTSCHE AKADEMIE DER
TECHNIKWISSENSCHAFTEN

WHITEPAPER

Jessica Heesen et al.
AG IT-Sicherheit, Privacy,
Recht und Ethik

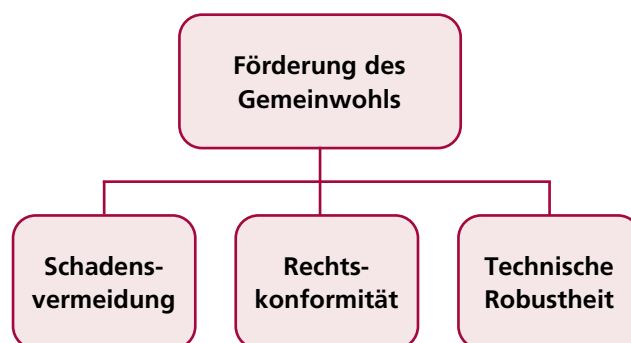
Inhalt

Zusammenfassung	3
1. Zielsetzung	5
2. Kriterien für eine verantwortungsvolle Entwicklung und Anwendung von KI-Systemen.....	7
2.1 Kategorien von Anforderungen	7
2.2 Konkrete Anforderungen an KI-Systeme	9
3. Anwendbarkeit in Unternehmen.....	24
3.1 Bedeutung ethischer Kriterien für ausgewählte Unternehmen.....	24
3.2 Praxisbeispiele.....	25
4. Fazit und Ausblick	27
Literatur.....	28
Über dieses Whitepaper.....	31

Zusammenfassung

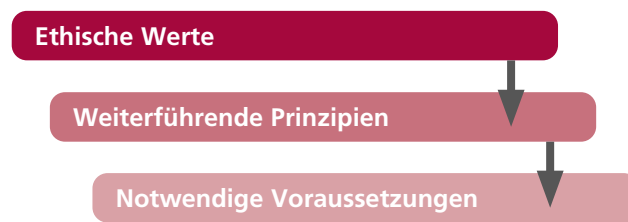
Ethische Werte und Prinzipien spielen für viele Menschen auch im wirtschaftlichen Kontext eine große Rolle. Dies gilt ganz besonders für die Entwicklung von Künstlicher Intelligenz (KI). Auch wenn die Möglichkeiten, die der Einsatz von KI-Systemen in den verschiedenen Anwendungsbereichen bringt, von vielen positiv bewertet werden, bestehen bezüglich dieser technologischen Innovationen auch noch einige Bedenken. Mangelndes Vertrauen ist ein Hauptgrund für die fehlende Akzeptanz von KI-Systemen. Die Arbeitsgruppe IT-Sicherheit, Privacy, Recht und Ethik der Plattform Lernende Systeme will diesem Problem begegnen und gibt in einem Leitfaden einen Überblick, welche Herausforderungen bei der verantwortungsvollen Entwicklung und Anwendung von KI-Systemen beachtet werden müssen und welche Maßnahmen genutzt werden können, um einen ethisch reflektierten Technikentwicklungsprozess umzusetzen (Kapitel 1). Der Leitfaden dient als Orientierung für Entwickelnde, Anbietende, Nutzende und Betroffene von KI-Systemen.

Die Entwicklung und der Einsatz von KI-Systemen sollten immer das Ziel verfolgen, zur Förderung des Gemeinwohls beizutragen. Daraus ergibt sich, dass KI-Systeme so eingesetzt werden sollen, dass Schaden für Einzelpersonen, die Gemeinschaft und die Umwelt vermieden wird, die Rechtskonformität von KI-Systemen in der Praxis von Entwickelnden, Anbietenden und Nutzenden gewährleistet ist und das eingesetzte KI-System die notwendige technische Robustheit erfüllt, dass es zu keinem Zeitpunkt ein unannehmbares Sicherheitsrisiko darstellt.



Die Autorinnen und Autoren des Whitepapers definieren dazu Kriterien für eine verantwortungsvolle Entwicklung und Anwendung von KI-Systemen sowie konkrete Anforderungen an KI-Systeme (Kapitel 2). Diese Kriterien lassen sich in drei übergeordnete Kategorien ethischer Werte aufteilen: Selbstbestimmung, Gerechtigkeit und Schutz der Privatsphäre. An diese Werte schließen weiterführende ethische Prinzipien sowie notwendige (technische und nicht-technische) Voraussetzungen zur Realisierung an.

Selbstbestimmung begründet als grundlegendes Prinzip demokratischer Gesellschaften die Forderung, dass KI-Systeme menschliche Akteure bei der Entscheidungsfindung unterstützen sollen. Notwendige Voraussetzungen zur Realisierung sind Transparenz



und Nachvollziehbarkeit sowie offene Schnittstellen und Systemoperabilität in Bezug auf Selbstbestimmung der oder des Einzelnen auf dem Markt. Das **Prinzip der Gerechtigkeit fordert Gleichheit vor dem Gesetz, Diskriminierungsfreiheit und Chancengerechtigkeit.** Fairness in der Verteilungsgerechtigkeit und Nachhaltigkeit als intergenerationelle Gerechtigkeit sind konkrete Anforderungen. Zur Realisierung der Gerechtigkeit von KI-Systemen sind Rechenschaftspflichten notwendige Voraussetzungen. **Privatsphäre** bezeichnet einen Raum, der sich von der Öffentlichkeit abgrenzt und der geschützt ist vor äußeren Eingriffen. Mit dem zunehmenden Einsatz von KI wird der Schutz der Privatsphäre eine Herausforderung. Datenminimierung und Transparenz bei Datenerhebung und -verarbeitung gelten als zentrale Voraussetzungen zum Schutz der Privatsphäre.

Diese abstrakten ethischen Werte können durch konkrete Maßnahmen realisiert werden, wozu von den Expertinnen und Experten konkrete Anforderungen an bestimmte Akteursgruppen definiert werden (Kapitel 2). Der Leitfaden schlägt dazu Maßnahmen für Entwickelnde und Nutzende vor, etwa die Entwicklung von KI-Systemen, die erklärbare und nachvollziehbare Ergebnisse liefern, sowie die Etablierung von Kontrollmechanismen.

Die Diskussion ethischer Werte und Handlungsoptionen soll einen Beitrag zur Realisierung eines ethisch fundierten Technologie-, Entwicklungs- und Anwendungsprozesses leisten. Die Rückkopplung in die Praxis von Unternehmen in Deutschland ist hierbei eine wichtige Zielsetzung. Im Zentrum steht dabei die Frage, wie die theoretischen Werte in der Praxis realisierbar sind. Viele Unternehmen sind sich der Bedeutung ethischer Werte zur Realisierung eines verantwortungsvollen Technologie- und Anwendungsprozesses nicht nur bewusst, sie haben auch bereits Maßnahmen ergriffen, wie die Einhaltung dieser Werte gewährleistet werden kann. Dazu werden im Whitepaper Praxisbeispiele vorgestellt, mit deren Hilfe exemplarisch ein Ansatz zur Realisierung von ethischen Werten aufgezeigt wird (Kapitel 3).

1. Zielsetzung

Ethische Werte und Prinzipien spielen für viele Menschen auch im wirtschaftlichen Kontext eine große Rolle. Fast 75 Prozent aller Befragten einer Studie des Bundesverbandes Digitale Wirtschaft (BVDW) von September 2019 geben an, dass für sie die Einhaltung ethischer Grundsätze in der Produktentwicklung einen hohen oder sehr hohen Stellenwert einnehmen sollte. Und für immerhin über die Hälfte der befragten Unternehmen hat Ethik bereits einen hohen oder sehr hohen Stellenwert (BVDW 2019). Dieser Stellenwert gilt ganz besonders auch für die Entwicklung von Künstlicher Intelligenz (KI).

Auch wenn die Möglichkeiten, die der Einsatz von KI-Systemen in den verschiedenen Anwendungsbereichen bringt, von vielen positiv bewertet werden, bestehen in der Bevölkerung insgesamt noch Bedenken bezüglich der für viele noch weitgehend unbekannten Technologie. So würden fast 60 Prozent der Befragten einer Studie des VdTÜV auf die Möglichkeit verzichten, in einem Smarthome mit KI-Systemen zu wohnen, unter anderem, weil sie einen Verlust ihrer Autonomie befürchten (VdTÜV 2020). Mangelndes Vertrauen ist ein Hauptgrund für die fehlende Akzeptanz von Technologien, auch von KI-Systemen. Das Weißbuch der Europäischen Union zu Künstlicher Intelligenz und ebenso die Stellungnahme der Bundesregierung greifen den Aspekt „Vertrauen“ auf und widmen ihm jeweils ein eigenes Kapitel (Europäische Kommission 2020, Bundesregierung 2020). Um auf bestehende Befürchtungen bezüglich KI zu reagieren und Vertrauen in die neue Technologie zu schaffen, ist ein ethisch reflektierter Entwicklungsprozess und ein ethisch basierter Einsatz von KI wichtig.

Mit dem Ethik-Briefing, einem Leitfaden für eine verantwortungsvolle Entwicklung und Anwendung von KI, greift die Plattform Lernende Systeme den in der KI-Strategie der Bundesregierung formulierten „ethics by, in and for design“-Ansatz auf und leistet einen Beitrag zu dessen Umsetzung. Dabei werden europäische Ansätze, wie beispielsweise die Ergebnisse der High Level Expert Group on Artificial Intelligence der EU-Kommission (AI HLEG), ebenso beachtet wie weitere (internationale) Initiativen. Hierzu zählen unter anderem die 2019 von der OECD veröffentlichten Empfehlungen zur Künstlichen Intelligenz (OECD 2019).

Der Leitfaden soll Orientierung geben

- welche ethischen und gesellschaftlichen **Herausforderungen** bei Entwicklung und Anwendung von KI-Systemen beachtet werden müssen und wie diese konkret definiert werden können;
- welche technischen und nicht-technischen **Maßnahmen** genutzt werden können, um einen ethisch reflektierten Technikentwicklungsprozess umzusetzen;

- wie ein entsprechendes Bewusstsein und die vorgeschlagenen Maßnahmen in **Unternehmensprozesse** integriert werden können;
- welche **sektorspezifischen** Herausforderungen und Maßnahmen beachtet werden müssen (anhand der Anwendungsszenarien);
- inwieweit die **Vorschläge in der Wirtschaft umgesetzt werden können**.

Dieser Leitfaden soll als Orientierung für Entwickelnde, Anbietende, Nutzende und Betroffene dienen, die einen ethisch reflektierten Entwicklungs- und Anwendungsprozess von KI-Systemen verfolgen. Ein ethisch reflektiertes Design und ein ebensolcher Einsatz sind notwendig, aber nicht immer hinreichend, um alle Risiken, die mit KI für Anwendungen einhergehen, auszuschließen. So kann es durch menschliches oder technisches Versagen sowie durch Missbrauch der KI-Systeme zu nicht intendierten Folgen kommen, die trotz eines umsichtigen Entwicklungsprozesses nicht abzusehen sind.

Die Frage nach ethischen Kriterien für die Anwendung und den Einsatz eines KI-Systems steht in engem Zusammenhang mit der Diskussion um Regulierung und Zertifizierung im Allgemeinen sowie Kritikalität von KI-Systemen im Besonderen (Heesen et al. 2020). Vorliegender Leitfaden gibt in einem ersten Schritt eine Orientierungshilfe für die maßgeblichen Werte, die bei der Entwicklung und Anwendung eines KI-Systems beachtet werden sollten. Allgemein lässt sich die Faustformel aufstellen: **Je höher die Risiken durch den Einsatz eines Systems sind, desto mehr Gewicht soll die Einhaltung ethischer Prinzipien bereits im Entwicklungsprozess haben. Gleichwohl sollten diese im Optimalfall auch bei Anwendungen mit geringem Schadenspotenzial beachtet werden.**

2. Kriterien für eine verantwortungsvolle Entwicklung und Anwendung von KI-Systemen

Ziel des Papiers ist es, Kriterien für eine verantwortungsvolle Entwicklung und Anwendung von KI-Systemen zu definieren. Diese KI-Systeme sollten vertrauenswürdig, nützlich und wirtschaftlich einsetzbar sein. Hierbei dienen bereits existierende Richtlinien, wie beispielsweise die der HLEG der Europäischen Union oder der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) als Orientierung.

2.1 Kategorien von Anforderungen

Die Anforderungen an einen ethisch reflektierten Technologieentwicklungsprozess sowie einen wertbasierten Einsatz der KI-Systeme lassen sich in drei übergeordnete Kategorien ethischer Werte aufteilen: Selbstbestimmung, Gerechtigkeit und Schutz der Privatheit (siehe Tabelle).¹ An diese Werte schließen weiterführende ethische Prinzipien, die die übergeordneten Werte ausdifferenzieren, sowie notwendige (technische und nicht-technische) Voraussetzungen zur Realisierung der Werte an.

Selbstbestimmung, Gerechtigkeit und Schutz der Privatsphäre stehen als ethische Werte in einem relativen Verhältnis zueinander, der Anspruch auf gleichwertige Umsetzbarkeit der drei Kategorien ist nicht gegeben. Denn zur Realisierung des einen kann es nötig sein, hinsichtlich eines anderen Wertes Abstriche machen zu müssen. So kann beispielsweise die Selbstbestimmung des Einzelnen zeitweise hinter das Prinzip der Gleichheit zurücktreten, wenn Einschränkungen durch kollektive Schutzmaßnahmen zur Seuchenbekämpfung dringlich notwendig sind, um eine gleiche Gesundheitsversorgung für alle zu gewährleisten.

Unabhängig von diesen Kategorien sollten KI-Systeme immer mit dem Ziel entwickelt und eingesetzt werden, zur Förderung des Gemeinwohls beizutragen. Dabei muss ein ausgewogenes Gleichgewicht zwischen partikularen Interessen und dem Wohl der Gemeinschaft gefunden werden.

Vor den Ausführungen zu den spezifischen ethischen Anforderungen an KI-Systeme sollen drei vorgelagerte Basisannahmen genannt werden:

- KI-Systeme müssen so eingesetzt werden, dass Schaden für Einzelpersonen, Gemeinschaften oder beispielsweise auch die Umwelt vermieden oder in gerechtfertigter Abwägung eines kleinen Schadens gegen einen größeren Nutzen minimiert wird.

¹ Zu den Diskussionen um verschiedene Wertkategorien zur Entwicklung von KI siehe z. B. Floridi 2018 oder Fjeld et al. 2020.

Besonders zu beachten ist die **Vermeidung von Schäden** an Gesundheit und Leben durch KI-Systeme.

- Die **Rechtskonformität** von KI-Systemen und der Praxis von Entwickelnden, Anbietenden und Nutzenden ist selbstverständliche Voraussetzung und Bestandteil ethisch-normativer Anforderungen. Dies gilt für alle jeweils bestehenden Gesetze.
- Das eingesetzte KI-System muss die notwendige **technische Robustheit** gewährleisten, das heißt so zuverlässig sein, dass es zu keinem Zeitpunkt ein unannehmbares Sicherheitsrisiko darstellt (OECD 2019).

Auf Grundlage dieser Basisannahmen und zur weiterführenden Integration ethischer Zwecke in die Entwicklung von KI stehen die folgenden Wertsetzungen im Vordergrund:

Wert: Selbstbestimmung

Prinzipien

- Förderung der Autonomie
- Verantwortungswahrnehmung

Notwendige Voraussetzungen

- Erklärbarkeit
- Nachvollziehbarkeit
- Vielfalt des Marktangebots
- Offene Schnittstellen und Systemoperabilität

Wert: Gerechtigkeit

Prinzipien

- Gleichheit
- Diskriminierungsfreiheit
- Diversität und Vielfalt
- Fairer Zugang zu den Vorteilen von KI
- Nachhaltigkeit

Notwendige Voraussetzungen

- Auskunftspflichten im Entwicklungs- und Anwendungszyklus
- Zurechenbarkeit von Verantwortung und Rechenschaft

Wert: Schutz der Privatheit und der Persönlichkeit

Prinzipien

- Privatsphäre als Rückzug aus der Öffentlichkeit
- Anonymität als Privatheit in der Öffentlichkeit
- Informationelle Selbstbestimmung
- Integrität der persönlichen Identität

Notwendige Voraussetzungen

- Transparenz als zentrale Forderung von Datenschutz
- Minimierung der Datenerhebung und -verarbeitung, insbesondere personenbezogene Daten
- Privacy by Design

2.2 Konkrete Anforderungen an KI-Systeme

Die vorgestellten Kategorien von Anforderungen an KI-Systeme umfassen viele Einzelaspekte. Die vorangestellte Tabelle gibt einen Überblick, welche einzelnen Elemente sie jeweils konkret beschreiben.

Das folgende Kapitel beinhaltet eine kurze Definition der jeweiligen Anforderungen inklusive illustrierender Beispiele sowie Möglichkeiten der praktischen Umsetzung für Entwickelnde, Anbietende, Nutzende und Betroffene von KI-Systemen. Dieser Leitfaden fasst dabei **Entwickelnde** als Unternehmen (und deren Mitarbeiterinnen und Mitarbeiter) auf, die KI-Systeme entwickeln. Die aus ethischer Sicht relevanten Aspekte einer Technikentwicklung beginnen nicht erst im Rahmen der Arbeit der entwickelnden Ingenieurinnen und Ingenieure. Bereits die strategische Ausrichtung eines Unternehmens und die hinter der Technologieentwicklung stehenden Konzepte zur Kommerzialisierung sind wesentliche Elemente für ein gelingendes ethisches Gesamtkonzept. **Anbietende** im Sinne dieses Papiers sind Unternehmen (und deren Mitarbeiterinnen und Mitarbeiter), die KI-Systeme in ihren Produkten einsetzen, ohne selbst für die Entwicklung der KI-Systeme verantwortlich zu zeichnen. **Nutzende** sind Unternehmen oder Institutionen der öffentlichen Hand (und deren Mitarbeiterinnen und Mitarbeiter), die KI-Systeme direkt als Software oder Produkte mit KI-Systemen im Kontakt mit Kundinnen und Kunden bzw. Bürgerinnen und Bürgern einsetzen. Nutzende können auch direkt Bürgerinnen und Bürger sein. Die Kategorie der **Betroffenen** bezeichnet schließlich die Menschen, die direkt von Auswirkungen der KI-Systeme betroffen sind, und schließt damit sowohl die Kundschaft als auch die Angestellten eines Unternehmens ein.

Entwickelnde, Anbietende und Nutzende stellen die Hauptadressaten dieses Papiers dar. Alle vier Gruppen sprechen aber jeweils unterschiedliche Aspekte im Umgang mit KI-Systemen an, gleichwohl ist die Unterscheidung nicht trennscharf. Ebenso ist die Einordnung vor allem menschlicher Akteure in eine der vier Gruppen situationsabhängig: So kann eine Mitarbeiterin des Arbeitsamtes Repräsentantin eines Nutzenden von KI-Systemen sein. Zeitgleich kann sie beispielsweise als Kundin in Interaktion mit einem Kreditinstitut, das KI-Systeme zur Bewertung der Bonität einsetzt, Betroffene von KI-Systemen sein.

Ergänzt werden diese Akteure durch die Kategorie **politische Entscheidungsträgerinnen und Entscheidungsträger**. Unter diesen werden sowohl nationale als auch europäische politische Entscheidungsträgerinnen und Entscheidungsträger verstanden.

Die Handlungsoptionen für die verschiedenen Akteure stellen Anstöße dar, welche Maßnahmen ergriffen werden können, um einen verantwortungsvollen KI-Entwicklungs- und KI-Anwendungsprozess zu erreichen. Im konkreten Einzelfall ist dann im Sinne einer Abwägung zwischen diesen Optionen zu wählen.

2.2.1 Selbstbestimmung

Wert

Selbstbestimmung ist das grundlegende Prinzip demokratischer Gesellschaften und garantiert jeder und jedem die freie Entfaltung seiner Persönlichkeit, soweit die Rechte anderer nicht verletzt werden. Die Selbstbestimmung des Einzelnen muss im Kontakt mit anderen immer abgewogen werden gegen die Prinzipien der Gleichheit und Gerechtigkeit.

Weiterführende Prinzipien

Förderung der Autonomie: KI-Systeme können menschliche Akteure bei der Entscheidungsfindung unterstützen und entlasten. KI-Systeme können deshalb zu einem höheren Maß an Selbstbestimmung beitragen. Gleichzeitig können automatisierte Entscheidungen, die über Menschen getroffen werden, auch eine Einschränkung der Selbstbestimmung bedeuten, insbesondere wenn nicht klar ist, dass und unter welchen Bedingungen das passiert. Mit dem Einsatz von KI-Systemen sollte immer die Stärkung menschlicher Selbstbestimmungsinteressen verbunden sein.

Verantwortungswahrnehmung: Aus der Selbstbestimmung des handelnden Individuums ergibt sich Verantwortung für die eigenen Handlungen. Gleichwohl steht die Verantwortung des Einzelnen nicht für sich, sondern ist eingebettet in einen staatlichen Bezugsrahmen. Die Achtung vor der Selbstbestimmung verbietet es, eigene Pflichten unter Verweis auf die freie Wahl des Gegenübers abzuwälzen (z. B. Übertragung der Pflicht zur Wahrung der Privatsphäre in (unverständlichen) AGBs auf die freie Entscheidung der Kundin oder des Kunden).

Notwendige Voraussetzungen

Erklärbarkeit/Transparenz als Voraussetzung für Selbstbestimmung: Ein wesentliches Element, um auch im Umgang mit KI-Systemen Selbstbestimmung zu erhalten, ist Transparenz. Nur, wenn die Endnutzerin oder der Endnutzer weiß, dass sie oder er mit einem KI-System interagiert (und wie dieses funktioniert), kann sie oder er selbstbestimmt und eigenverantwortlich Entscheidungen fällen (AI High Level Expert Group 2019). Welche Anforderungen Transparenz an den Einsatz von KI konkret richtet, ist abhängig von der oder dem jeweiligen Betroffenen und dem Einsatzgebiet beziehungsweise der Aufgabe des KI-Systems. Für Betroffene einer KI-basierten Entscheidung muss beispielsweise keine Transparenz auf technischer Ebene bestehen, für den für die Überwachung des Systems Zuständigen ist diese Transparenz hingegen notwendig.

Nachvollziehbarkeit: Es muss für jede und jeden am Einsatz von KI-Systemen Beteiligten nachvollziehbar sein, welche Daten das KI-System verarbeitet und warum es zu bestimmten Ergebnissen kommt (Floridi et al. 2018). Nachvollziehbarkeit geht über die Forderung nach einer „informationstechnischen“ Transparenz im Sinne von offenen Quellcodes, Optimierungs- und Zielfunktionen oder Daten hinaus. Diese werden nachvollziehbar, wenn zwischen der informationstechnischen Erkenntnis und der Lebenswelt der Betroffenen eine gelingende Übersetzungsarbeit geleistet wird (Beining 2019).

Vielfalt des Marktangebots: Wenn zwischen Entwickelnden/Anbietenden und Nutzenenden/Betroffenen von KI-Systemen einseitige Abhängigkeitsverhältnisse bestehen, können viele ethische Werte wie Selbstbestimmung oder Gerechtigkeit, aber beispielsweise auch die Bereitschaft von kleinen Unternehmen, sich digitalen Angeboten zu öffnen, darunter leiden (Mittelstand Digital 2019). Gerade wenn Unternehmen monopolartige Strukturen entwickeln, werden solche Abhängigkeitsverhältnisse befördert. Im Zuge einer ethisch reflektierten Etablierung von KI-Systemen muss daher darauf geachtet werden, mit einer Vielfalt des Angebots die freie Wahl der Nutzenden zwischen mehreren Produkten und somit Konsumentensouveränität sicherzustellen.

Offene Schnittstellen und Systemoperabilität: Offene Schnittstellen ermöglichen es, Softwaredienste externer Anbietender in eigene Applikationen zu integrieren. Damit steigt die Verwendbarkeit der Software für verschiedene Anbietende und Nutzende und das mögliche Angebot an Diensten. Systemoperabilität meint die Kompatibilität von zwei oder mehr Systemen. Ist ein erfolgreiches KI-System nicht auf Interoperabilität mit anderen Anwendungen ausgelegt, kann es diese langfristig vom Markt verdrängen. Offene Schnittstellen und Systemoperabilität beugen monopolartigen Strukturen vor und tragen so zu selbstbestimmten Wahlmöglichkeiten bei. Durch die Auswahlmöglichkeit verschiedener Dienste wird Selbstbestimmung gestärkt.

In Zukunft werden beispielsweise Ärztinnen und Ärzte zur Diagnose und Behandlung einer Krebserkrankung durch KI-Assistenzsysteme unterstützt werden: Die KI-Systeme können bei der Diagnose absichern, Empfehlungen für eine Behandlung geben und auch bei einer möglichen Operation assistieren (Plattform Lernende Systeme 2019b). Sowohl die beteiligten Ärztinnen und Ärzte als auch die Patientin oder der Patient wissen, dass ein KI-System am Diagnoseprozess beteiligt ist und Empfehlungen für eine Behandlung gibt. Entscheidend für die Behandlungsempfehlung ist aber eine Gruppe menschlicher Spezialistinnen und Spezialisten; letztlich relevante Entscheidungen werden von Ärztinnen und Ärzten in Absprache mit den Patientinnen und Patienten getroffen.

Die **Wahrung der Selbstbestimmung** ist für einen ethisch reflektierten Technikentwicklungs- und Technikanwendungsprozess unerlässlich. Hierfür stehen folgende Handlungsoptionen für die unterschiedlichen Akteure zur Verfügung:

■ **Entwickelnde**

- gestalten ihre Systeme als erklärbare KI-Systeme (Explainable Artificial Intelligence, XAI) in einer Weise, dass Ausgaben und Entscheidungen nachvollziehbar werden, und unterstützen so die Selbstbestimmung der oder des Nutzenden;
- konzipieren KI-Systeme so, dass immer ein menschlicher Akteur die Letztentscheidung trifft (Ausnahmen könnten getroffen werden, wenn automatisierte Entscheidungen ausschließlich und garantiert zum Vorteil der oder des Nutzenden erfolgen), oder konzipieren KI-Systeme (in Fällen, in denen keine Interaktionen mit oder Auswirkungen auf Menschen vorliegen, z. B. in der Industrie) so, dass die Verantwortung für die Entscheidung zugerechnet werden kann.²

■ **Anbietende** von KI-gestützten Produkten

- legen eigenständig offen, ob und in welchem Umfang KI-Systeme in ihren Produkten verwendet werden;
- implementieren Hinweise in ihr KI-System, welchen Autonomiegrad die in ihren Produkten eingesetzte Software hat, und, dass die Verantwortlichkeit für Entscheidungen, die mit auf Algorithmen basierenden Empfehlungen zu Stande kamen, immer bei einem Menschen liegt;
- richten in Kooperation mit anderen Unternehmen und der öffentlichen Hand ein zentrales KI-Register ein, über das sich Verbraucher über den Einsatz von KI in verschiedenen Produkten transparent informieren können, und tragen dort ihre KI-Anwendungen beziehungsweise KI-basierten Produkte selbstständig ein;
- bieten für Mitarbeiterinnen und Mitarbeiter, die regelmäßig mit KI-Systemen arbeiten, spezielle Schulungen an.

■ **Nutzende und Betroffene** achten im Umgang mit und beim Konsum von digitalen Technologien mit besonderer Aufmerksamkeit auf den Einsatz von KI-Systemen und treffen ihre Entscheidung zur Benutzung dieser Technologien im Sinne der Konsumentensouveränität und Wahlfreiheit wohlüberlegt und abwägend.

■ Die **politischen Entscheidungsträgerinnen und Entscheidungsträger** können den Selbstbestimmungs-Grundsatz in Gesetzen und Verordnungen (wie beispielsweise in Art. 22 DSGVO) konkretisieren.

² Zur Rechtsgrundlage vgl. hierzu Art. 22 EU DSGVO.

KI kann Menschen unterstützen und sie in ihrer Selbstbestimmung stärken. Dazu ist es aber notwendig, dass die oder der Nutzende eines KI-Systems weiß, dass sie oder er mit einem KI-System interagiert, mit Informationen eines KI-Systems arbeitet oder (sie oder er) von Entscheidungen eines KI-Systems unterstützt wird. Um einen vertrauenswürdigen Umgang mit KI-Systemen zu gewährleisten, muss **deren Erklärbarkeit** sichergestellt sein:

■ **Entwickelnde**

- gewährleisten die lückenlose Begleitung und Überwachung des Entwicklungsvorgangs von KI-Systemen;
- benennen eine Ombudsperson/-stelle, die die Einhaltung dieser Verpflichtungen neutral überprüft und gegebenenfalls sanktioniert, und stellen gleichzeitig unternehmensintern das nötige Bewusstsein für ethische Problemstellungen sicher (AI High Level Expert Group 2019).

■ **Anbietende** von KI-gestützten Produkten und Dienstleistungen

- gehen eine Selbstverpflichtung (KI Bundesverband e.V. 2019) ein, Transparenz bezüglich der Verwendung von KI-Systemen zu schaffen;
- sorgen für Transparenz, indem sie ihren Kunden den Einsatz von KI-Systemen in ihren Produkten/Anwendungen/Software offen und auf die Interessen der Kunden zugeschnitten kommunizieren (Europäische Kommission 2020). Anschauliche Beispiele können den Prozess erleichtern (Beining 2019).

■ **Entwickelnde** und **Anbietende** stellen Weiterbildungsangebote und -materialien bereit, um ein grundlegendes Verständnis in der Bevölkerung zu Arbeitsweise und Einfluss von KI-Systemen zu fördern. Öffentliche Einrichtungen unterstützen hierbei.

■ **Nutzende**

- achten darauf, soweit möglich nur Produkte mit KI-Systemen einzusetzen, die hinsichtlich der Verwendung der KI transparent sind;
- kommunizieren den Betroffenen den Einsatz von KI-Systemen transparent.

■ **Betroffene**

- bringen die Bereitschaft mit, bestehende Informationsangebote zu KI-Systemen anzunehmen, und regen durch aktives Nachfragen Anbietende und Entwickelnde der KI-Systeme zur Bereitstellung dieser Informationen an;
- üben, wenn nötig, durch die Äußerung von Kritik oder durch individuelle oder gemeinschaftliche Beschwerde Druck auf die Entwickelnden und Anbietenden aus, diese Informationsangebote bereitzustellen.

■ Den **politischen Entscheidungsträgerinnen und Entscheidungsträgern** stehen folgende mögliche Ansatzpunkte zur Verfügung:

- Sie können, wo die oben aufgeführten Aspekte und die geltenden Rechtsvorschriften nicht ausreichen (bspw. bei KI-Anwendungen in Anwendungskontexten mit

einer hohen Kritikalität), durch rechtliche Rahmenbedingungen die nötige Transparenz vorschreiben (Datenethikkommission 2019, Handlungsempfehlungen 38, 40 und 41 sowie S. 25).³

- Sie können die bestehenden Aufsichtsinstitutionen im Rahmen ihrer Zuständigkeit stärken, sie neu ausrichten und, wo erforderlich, auch neue Institutionen schaffen (Datenethikkommission 2019, Handlungsempfehlungen 55 und 56).
- Sie können institutionelle Formen der Auskunftspflicht und Beschwerdestellen schaffen (entsprechend zur Datenschutzgrundverordnung und analog zum Netzwerkdurchsetzungsgesetz).

Gerade im immer stärker wachsenden Bereich der Plattformökonomie, also einer Marktstruktur, in der digitale Plattformen Anbietende und Konsumierende über eine unkomplizierte Schnittstelle zusammenbringen, zeichnet sich die Problematik von The-winner-takes-it-all-Strukturen ab. Große Betreiber von digitalen Verkaufsplattformen können damit potenziell ihre Macht ausnutzen, um Verkäufer unter Druck zu setzen und damit für den Verkäufer schlechte Vertragsbedingungen durchzusetzen. Da die Programmierung von leistungsstarken KI-Systemen ressourcenaufwendig und kostenintensiv ist, ist zu erwarten, dass von großen Anbietern eingesetzte, erfolgreiche KI-Systeme diese Macht weiter vergrößern.

Um die **Vielfalt auch im Bereich der digitalen Ökonomie zu gewährleisten**, Wahlmöglichkeiten zu erhalten und damit Selbstbestimmung zu fördern, kommen folgende Optionen in Betracht:

- **Entwickelnde** richten einen Fokus darauf, bei der Entwicklung von KI-Systemen
 - offene Schnittstellen in ihre Systeme zu integrieren;
 - ihre Systeme interoperabel zu konzipieren.
- **Anbietende** und **Nutzende** von KI-Systemen achten im Rahmen ihrer Möglichkeiten darauf, nur KI-Systeme bzw. Produkte mit KI-Systemen zu verwenden, die Marktpluralität fördern.
- Den **politischen Entscheidungsträgerinnen und Entscheidungsträgern** stehen folgende mögliche Ansatzpunkte zur Verfügung:
 - Sie können Start-ups in den Bereichen KI und Plattformökonomie durch den Abbau von bürokratischen Hürden und besondere finanzielle Förderung unterstützen, um die Pluralität der Angebote zu erhöhen (Plattform Lernende Systeme 2019a).
 - Sie unterstützen bestehende KMU bei der Implementierung von Plattformen und beim Aufbau eigener digitaler Plattformen.

³ Zum Themenkomplex Innovation und Regulierung siehe u. a. Heesen et al. 2020.

- Sie können rechtliche Rahmenbedingungen für Kooperationen für den Betrieb von digitalen Plattformen schaffen und damit Wettbewerbern aus der Realwirtschaft die Zusammenarbeit in neuen digitalen Märkten ermöglichen (Mittelstand Digital 2019).
- Sie fördern F&E-Projekte zur Entwicklung von plattformbasierten Geschäftsmodellen (Mittelstand Digital 2019).
- Sie können das Wettbewerbs- und Kartellrecht auf die Herausforderungen der Digitalisierung und der Plattformökonomie vorbereiten und es auf die nationalen, europäischen und internationalen Herausforderungen anpassen, um Konsumentensouveränität und Wahlmöglichkeiten in der Nutzung zu ermöglichen.⁴
- Sie bauen die Portabilität von Daten praktisch nutzbar aus und fordern offene Schnittstellen rechtlich verbindlich ein.⁵
- Sie können die Voraussetzungen schaffen, dass das Wettbewerbs- und Kartellrecht tatsächlich eingehalten werden müssen, damit Konsumentensouveränität und Wahlmöglichkeiten in der Nutzung überhaupt möglich werden.

2.2.2 Gerechtigkeit

Wert

Gerechtigkeit ist das Prinzip eines individuellen oder gesellschaftlichen Handelns, das jeder und jedem gleichermaßen sein Recht gewährt und dabei eine kontextbezogene Beurteilung von Ausgleich und Gleichbehandlung im Blick hat.

Weiterführende Prinzipien

Gleichheit: Das Prinzip der Gleichheit umfasst die Gleichheit vor dem Gesetz, die Uneingeschränkbarkeit der Menschenwürde sowie Diskriminierungsfreiheit und Chancengerechtigkeit.⁶

Diskriminierungsfreiheit: Diskriminierung (von lat. *discriminare*, unterscheiden) bedeutet in normativer Perspektive Ungleichbehandlung von Gleichem oder Gleichbehandlung von Ungleichem ohne sachlichen Grund. Dies ist besonders im Hinblick auf diskriminierungsfreie Lernprozesse und Datengrundlagen relevant.

⁴ Diese Aspekte werden auch bei der Überarbeitung durch das Bundesministerium für Wirtschaft und Energie im Rahmen der 10. Novelle des Gesetzes gegen Wettbewerbsbeschränkungen (GWB-Digitalisierungsgesetz) diskutiert.

⁵ Hierzu gibt es bereits in Art. 20 DSGVO Vorschriften, diese könnten aber angepasst werden.

⁶ Siehe hierzu Allgemeine Erklärung über die Menschenrechte Art.1 und Art.7, Art. 20 – 23 EU-GRC und Art. 3 GG.

Diversität und Vielfalt im Bereich der KI: Diversität und Vielfalt im Bereich der KI beziehen sich sowohl auf diverse Teams, bestehend aus Entwicklerinnen und Entwicklern mit unterschiedlichen Perspektiven und Erfahrungshintergründen als auch auf die Betrachtung der Zielgruppe als heterogene Gruppe (Hagendorff 2019). Auch die Einbindung von Perspektiven, die möglichst breit gesellschaftliche Strukturen abbilden, ist aus Perspektive der Diversität wichtig. Für das Training von KI-Systemen muss auf eine Repräsentierung der vielfältigen gesellschaftlichen Gruppen in den Trainingsdaten geachtet werden.

Fairer Zugang zu den Vorteilen von KI: Das Prinzip der Fairness fordert im Zusammenhang mit KI die prinzipiell gleiche Möglichkeit des Zugangs aller zu KI-Systemen und zu von KI gesammelten und generierten Daten sowie die faire Verteilung des durch den Einsatz von KI-Systemen erwirtschafteten Nutzens (AI High Level Expert Group 2019). Wie jede Technik werden auch KI-Systeme für einige Menschen Nachteile haben. Unter der Perspektive der Fairness muss deshalb auch gefragt werden, wie gegebenenfalls entstehende Nachteile ausgeglichen werden können.

Nachhaltigkeit: Intergenerationelle Gerechtigkeit beschreibt die Verpflichtung gegenüber künftigen Generationen, für einen Erhalt der Lebensumstände zu sorgen. KI-Systeme sollen zur Entwicklung einer nachhaltigen Gesellschaft beitragen und nachhaltig entwickelt und angewandt werden. Nachhaltigkeit hat dabei eine ökonomische, ökologische und soziale Dimension (Datenethikkommission 2019).

Notwendige Voraussetzungen

Auskunftspflichten im Entwicklungs- und Anwendungszyklus: Auskunftspflichten im Entwicklungs- und Anwendungszyklus beschreiben die Pflicht, Auskunft über etwas zu geben, wofür man verantwortlich ist. KI-Anwendungen durchlaufen als Lernende Systeme unterschiedliche Entwicklungen (KI Bundesverband e.V. 2019). Für diese Entwicklungen ist die oder der Entwickelnde verantwortlich, ihre oder seine Rechenschaftspflicht bezieht sich also auf den gesamten Entwicklungszyklus der KI. Dies trifft besonders auf KI-Systeme zu, deren Einsatz potenziell Fragen der Gerechtigkeit berührt.

Zurechenbarkeit von Verantwortung und Rechenschaft: Je wichtiger KI-Systeme im Alltag werden, desto wichtiger wird auch die Frage, wer für die Auswirkungen von KI-Systemen verantwortlich ist, denn die Zuordnung von Rechenschaftspflichten ist für einen Rechtsstaat unerlässlich. Ob und in welchen Fällen Entwickelnde oder Anbietende des Systems für mögliche Schäden haften, muss klar und transparent geklärt sein. Denn auch wenn sich Entwickelnde und Anbietende bemühen, einen ethisch reflektierten Technikentwicklungsprozess durchzuführen und somit ein vertrauenswürdigen KI-System zu entwickeln und anzubieten, kann nicht ausgeschlossen werden, dass KI-Systeme von Zeit zu Zeit Schaden verursachen. Für diese Fälle muss Rechenschaft bestehen und der entstehende Schaden bestmöglich ausgeglichen werden.

Ein mögliches Beispiel für die Verletzung von Gerechtigkeit (hier: für Diskriminierung durch KI-Systeme) liefert die Bewerbersoftware einer großen Firma: Die Personalverantwortlichen sortieren mithilfe eines Algorithmus die eingegangenen Bewerbungen vor. Was sie nicht wissen: Bewerbungen von Frauen bewertet der Algorithmus systematisch schlechter als jene von Männern. Denn die Datenbasis, mit der der Algorithmus trainiert wurde, beinhaltet die Daten der in den letzten zehn Jahren erfolgreich eingestellten Beschäftigten. Da diese überwiegend männlich waren, hat der Algorithmus gelernt, diese Eigenschaft als positiv zu bewerten (Beck et al. 2019).

Um dem Risiko einer mangelnden **Gerechtigkeit** zu begegnen und ihm vorzubeugen, werden den beteiligten Akteuren folgende Möglichkeiten empfohlen:

■ **Entwickelnde** von KI-Systemen

- filtern als diskriminierend bewertete Variablen (z. B. Hautfarbe) in Trainingsdatensätzen heraus (Beck et al. 2019) und legen den dokumentierten Prozess offen (Datenschutzkonferenz 2019) oder verwenden Datensätze, die hinsichtlich der Diskriminierungsmerkmale ausreichend repräsentativ sind (Europäische Kommission 2020);
- legen die Herkunft der Rohdaten, die die Lernbasis von KI-Systemen bilden, offen (Datenschutzkonferenz 2019);
- achten auf neue Diskriminierungsmerkmale (Datenethikkommission 2019), die durch die Eigenheit maschinellen Lernens entstehen;
- stellen die Qualität von Trainingsdaten hinsichtlich Datenauswahl und Datenintegrität sicher (Heesen/Reinhardt/Schelenz 2020);
- achten auf eine vielfältige Zusammensetzung der Entwicklerteams;
- wählen Fairness in Ergänzung zu Effizienz als Entwicklungsziel Maschinellen Lernens (Beck 2019).

- **Anbietende** von KI-Systemen und KI-gestützten Produkten
 - achten beim Einsatz von KI-Systemen darauf, ob diese in Kontakt mit besonders schützenswerten Zielgruppen kommen, unter anderem wenn hohes Risiko für Diskriminierung besteht oder besondere Machtgefälle existieren;
 - bieten außerdem auf die Praxis ausgerichtete Schulungen für Beschäftigte, die KI nutzen, an (Datenethikkommission 2019);
 - schaffen Formen der Selbstregulierung zur Überprüfung der KI-Systeme auf Verletzungen von Gerechtigkeitsprinzipien.

- **Nutzende**
 - vermeiden den Einsatz von KI-Systemen in Bereichen mit hohem Risikopotenzial für Diskriminierung oder andere Aspekte der Gerechtigkeit;
 - lassen Empfehlungen von KI-Systemen (stichprobenartig) von einem menschlichen Akteur auf Plausibilität unter Berücksichtigung von Gerechtigkeitsprinzipien überprüfen.

- **Anbietende** und **Nutzende** sehen Mechanismen vor, die es Betroffenen ermöglichen, Probleme in Zusammenhang mit Diskriminierungen oder anderen Verletzungen von Gerechtigkeitsprinzipien durch KI-Systeme zu melden (Beschwerdemanagement) (AI High Level Expert Group 2019).

- **Betroffene verzichten**, wann immer möglich, im Umgang mit KI-Systemen auf die Angabe von Daten, die zu Diskriminierung führen könnten, oder auf die Verwendung von diskriminierenden KI-Systemen. Da der Verzicht auf Technologie Nachteile für die Nutzenden mit sich bringt, muss gewährleistet sein, dass die freie Auswahl zwischen Angeboten durch eine Vielfalt an Produkten gewährleistet ist (siehe Vielfalt).

- Den **politischen Entscheidungsträgerinnen und Entscheidungsträgern** stehen folgende mögliche Ansatzpunkte zur Verfügung:
 - Sie sehen Transparenz über die Datenbasis von als kritisch bewerteten KI-Systemen und Einspruchsmöglichkeiten für betroffene Personen vor.
 - Sie können den rechtlichen Rahmen zum Schutz der Betroffenen anpassen und Kriterien für Nicht-Diskriminierung festlegen (Datenethikkommission 2019, Handlungsempfehlungen 7, 49 und 53).
 - Sie stärken Fördermaßnahmen zu Forschung für faire KI (Heesen 2020).
 - Sie setzen Standards zu Diskriminierungsfreiheit fest, die es ermöglichen, KI-Systeme zu zertifizieren und zu auditieren (ebd.).
 - Sie stellen die Antidiskriminierungsstelle des Bundes für die Anforderungen durch KI neu auf (ebd.).

Auch wenn Entwickelnde und Anbietende von KI-Systemen den Einsatz von KI offenlegen und Systeme unter Einhaltung aller rechtlichen und ethischen Standards entwickelt und betrieben werden, ist es denkbar, dass einem KI-System Fehler unterlaufen: eine Patien-

tin oder ein Patient bekommt von ihrer oder seiner Versicherung einen Erstattungsantrag für eine Behandlung fälschlicherweise nicht bewilligt, eine Bewerberin oder ein Bewerber wird aufgrund ihres oder seines Geschlechtes abgelehnt. Beide Vorkommnisse sind Risiken beim Einsatz von KI-Systemen. Die Prinzipien der Gerechtigkeit erfordern es, dass potenziell entstehende Schäden in angemessener Weise ausgeglichen werden.

Um solchen Fällen bestmöglich vorzubeugen und bei Eintreten eines Schadens diesen auszugleichen, können **Vorkehrungen** getroffen werden:

■ **Entwickelnde**

- implementieren Verfahren zum sicheren Abbrechen eines KI-gesteuerten Vorgangs in ihre Systeme (AI High Level Expert Group 2019);
- lassen KI-Systeme bei einem externen Kontrollgremium (siehe Gesetzgeber) klassifizieren und zertifizieren;
- übernehmen mit lückenloser Überwachung des Entwicklungszyklus Verantwortung für Veränderungen des KI-Systems (Plattform Lernende Systeme 2019a).

■ **Anbietende und Nutzende** von KI-gestützten Produkten

- bieten Schulungen für Mitarbeiterinnen und Mitarbeiter im Umgang mit Beschwerden von Betroffenen, die Schaden durch nicht richtig funktionierende KI-Assistenzsysteme erlitten haben, an;
- etablieren Erkennungs- und Reaktionsmuster, um zu beurteilen, ob beim Einsatz von KI-Systemen Schäden entstehen (AI High Level Expert Group 2019).

■ **Entwickelnde und Anbietende**

- bewahren die Aufzeichnung über die Entwicklung des KI-Systems und die verwendeten Trainingsdaten auf, um Entscheidungen des KI-Systems zurückverfolgen und überprüfen zu können (Europäische Kommission 2020);
- binden die von dem KI-System betroffenen Stakeholder in den Entwicklungs- und Anwendungsprozess ein, um möglichst viele Perspektiven auf mögliche Problematiken einzufangen.

■ **Nutzende** beugen Schäden durch unsachgemäßen Gebrauch von KI-Systemen vor, indem sie durch geeignete Vorkehrungen im Rahmen der Arbeitsprozesse einen beständigen Kontrolldruck aufrechterhalten und übermäßiges Vertrauen in das KI-System verhindern (AI High Level Expert Group 2019).

■ Den **politischen Entscheidungsträgerinnen und Entscheidungsträgern** stehen folgende mögliche Ansatzpunkte zur Verfügung:

- Sie überprüfen, ob der Einsatz von KI-Systemen eine Anpassung des Zivil- und Strafrechts erfordert.
- Sie tragen dafür Sorge, dass die Aus- und Weiterbildung von juristischem Fachpersonal, das Fälle mit Entscheidungen durch KI-Systeme verhandelt, angepasst wird.

- Sie regeln den Marktzugang für sichere KI-Systeme. Dies kann beispielsweise durch Beauftragung eines zentralen Kontrollgremiums oder mithilfe von Zertifizierungen geschehen. Dies schließt die Frage ein, für welche KI-Systeme in welchen Anwendungskontexten beispielsweise eine Zertifizierung notwendig ist (z.B. Europäische Kommission 2020, Bundesregierung 2020).
- Sie schaffen Standards der Beteiligung von durch KI-Systeme Betroffene bzw. entwickeln bestehende Standards weiter.
- Sie schaffen die rechtlichen Voraussetzungen, um Versicherungen für durch KI-Systeme verursachte Schäden zu ermöglichen.

2.2.3 Schutz der Privatheit und der Persönlichkeit

Wert

Privatheit meint die Abgrenzung eines Bereiches von der Öffentlichkeit, der geschützt ist vor äußeren Eingriffen. In diesem Raum kann der Mensch seine Persönlichkeit und Identität frei entfalten und entwickeln. Der Schutz der Privatheit umfasst den Schutz der Privatsphäre, Anonymität als Privatheit im öffentlichen Raum, das Recht auf informationelle Selbstbestimmung sowie die Integrität der Persönlichkeit.

Weiterführende Prinzipien

Privatsphäre als Rückzug aus der Öffentlichkeit: Das Recht auf Privatheit beinhaltet das Recht, sich aus der Öffentlichkeit zurückzuziehen. Privatheit hat unterschiedliche Dimensionen, etwa informationelle, körperliche oder räumliche Privatheit. Die Anforderungen, die durch den Schutz des Privaten und das Recht auf informationelle Selbstbestimmung entstehen, sind von den jeweiligen Kontextbedingungen abhängig: Beispielsweise kann eine Speicherung von Gesundheitsdaten erwartbar und rechters sein, wenn sie durch einen behandelnden Arzt vorgenommen wird, aber eine Verletzung der Privatsphäre darstellen, wenn die Daten von einem Unternehmen zu Werbezwecken erhoben werden (Matzner/Ochs 2019). Der Schutz des Privaten als elementarem Baustein einer freien Gesellschaft erfordert deshalb in einer datafizierten Umgebung angepasste und flexible Instrumente auf individueller und kollektiver Ebene.

Anonymität als Privatheit in der Öffentlichkeit: Anonymität meint die Freiheit, im Öffentlichen nicht identifiziert und systematisch beobachtet werden zu können. Im Öffentlichen kann das Interesse der oder des Einzelnen auf Privatsphäre unter Umständen mit Interessen der Gemeinschaft kollidieren.

Informationelle Selbstbestimmung: Informationelle Selbstbestimmung beschreibt das Recht, grundsätzlich selbst über die Preisgabe und Verwendung von personenbezogenen Daten zu bestimmen.

Integrität der persönlichen Identität: Die Integrität der persönlichen Identität umfasst das Recht auf Selbstdarstellung als Recht am eigenen Wort und Bild sowie den Schutz vor verfälschenden Darstellungen. Darunter fallen Beleidigungen und Verleumdungen ebenso wie die Erstellung von (falsch) prognostizierenden Persönlichkeitsprofilen (etwa in Bezug auf Straffälligkeit).

Notwendige Voraussetzungen

Minimierung der Datenerhebung und -verarbeitung: Mit zunehmendem Einsatz von KI-Systemen wird der Schutz der Privatsphäre eine große Herausforderung (OECD 2019). Ohne die Auswertung immenser Mengen an Daten (Big Data) (Deutscher Ethikrat 2017) lässt sich ein selbstlernendes KI-System oft nicht aufbauen. Die Erhebung von personenbezogenen Daten ist gleichermaßen von Interesse wie die Weiterverarbeitung dieser Daten, genauer die Auswertung der Daten sowie die mögliche Weitergabe an Dritte. Datenschutz betrifft dabei auch Daten, die das System selbst durch die Interaktion mit der oder dem Nutzenden generiert (AI High Level Expert Group 2019). Aus Sicht des Datenschutzes ist auch die Analyse von Verhaltensmustern aus Big Data relevant (Deutscher Ethikrat 2017).

Transparenz als zentrale Forderung von Datenschutz: Damit Betroffene souverän der Weitergabe von Daten zustimmen können, muss über die Art der Erhebung, Auswertung und Verwendung der Daten Transparenz bestehen.

Privacy by Design: Die Umsetzung des Schutzes der Privatheit kann nicht allein durch gesetzliche Regelungen (wie etwa Vorschriften zur Datenminimierung oder Zweckbindung) normiert werden, sondern muss durch das Design der Technologie realisiert werden.

Eine konkrete Verletzung der Privatheit stellt die flächendeckende Überwachung durch den Einsatz von Gesichtserkennungs-Software dar. Auch wenn die Überwachung sich hier auf den öffentlichen Raum beschränkt, kann besonders dann eine Verletzung der Privatheit vorliegen, wenn die gesammelten Bilder analysiert werden und daraus Verhaltensmuster gebildet werden, durch die Rückschlüsse auf die Persönlichkeit möglich sind. Auch

die Verbindung von Überwachung durch Gesichtserkennungs-Software mit einem sogenannten Social Scoring System mit Belohnungs- bzw. Bestrafungsmechanismen für sozial wünschenswertes bzw. abzulehnendes Verhalten ist nicht mit einem Recht auf Privatheit und informationelle Selbstbestimmung vereinbar.

Eine Verletzung liegt ebenfalls vor, wenn Sprachassistenzsysteme Gespräche aufzeichnen und auswerten, ohne dass der Nutzende/Betroffene sie dazu explizit und bewusst willentlich aufgefordert hat. Für die Verletzung der informationellen Selbstbestimmung spielt es dabei grundsätzlich keine Rolle, ob das Mithören absichtlich geschieht oder auf einem technischen Fehler beruht. Probleme für den Schutz des Privaten ergeben sich auch dann, wenn Menschen wissentlich aufgezeichnete Gespräche auswerten, ohne dass das dem Nutzenden/Betroffenen bekannt ist.

Zum Datenschutz und zum Schutz der Privatsphäre können folgende Maßnahmen ergriffen werden:

- **Entwickelnde** achten bei der Programmierung und im Lauf der entwicklungsbegleitenden Überwachung von KI-Systemen darauf
 - das KI-System weitestmöglich mit anonymisierten oder pseudonymisierten Datensätzen zu trainieren, um möglichem Missbrauch vorzubeugen, und dabei technische Konzepte wie Differential Privacy zu berücksichtigen;
 - ausreichend hohe Sicherheitsstandards, die unterschiedliche Kritikalitätsstufen unter Berücksichtigung unterschiedlicher Autonomielevel geeignet abbilden, für die Systeme zu definieren und umzusetzen, um das Risiko von Angriffen zu minimieren;
 - Standards von Privatheit in die Datenverarbeitung zu integrieren (z. B. Helen Nissenbaums „Transmission Principles“) (Nissenbaum 2019).
- **Anbietende und Nutzende** von KI-Systemen
 - holen Einwilligungserklärungen zur Erhebung und Verarbeitung von Daten mit dem Hinweis auf das verwendete KI-System ein;
 - bieten leicht verständliche Widerrufsmöglichkeiten zu diesen Einwilligungserklärungen an und halten diese ein;
 - sehen das Lernen mit anonymen Daten oder pseudonymen Daten vor und setzen zuverlässige Anonymisierungs- und Pseudonymisierungs-Verfahren ein.
- **Betroffene** achten mit besonderer Sorgfalt darauf, ihre Daten nur seriösen Anbietenden zur Verfügung zu stellen, die deren datenschutzgerechte Anonymisierung oder Pseudonymisierung sicherstellen.
- Den **politischen Entscheidungsträgerinnen und Entscheidungsträgern** stehen folgende Ansatzpunkte zur Verfügung:
 - Sie passen bestehende Datenschutzregelungen (EU-DSGVO und BDSG), die entsprechend den Risiken der KI-Anwendungen differenzierte Datenschutzregelungen fordern, die unter anderem ausreichende Transparenz, Vermeidung der Erhebung

personenbezogener Daten, Zweckbindung und den Schutz besonderer Kategorien personenbezogener Daten oder von Gruppen betroffener Personen (z. B. Kinder) vorsehen, im Hinblick auf Herausforderungen durch den Einsatz von KI-Systemen an und entwickeln sie weiter (Plattform Lernende Systeme 2019a)⁷. Dies schließt die Förderung der Ausarbeitung von in Art. 41 EU-DSGVO vorgesehenen branchenbezogenen Verhaltensregeln („Code of Conducts“) für rechtssichere Pseudonymisierung und Anonymisierung mit ein.

- Sie können neue Gesetze schaffen, die Nutzungen von KI-Systemen im Sinne des Privatheit-Schutzes aktiv regeln (statt wie EU-DSGVO und BDSG den Anbietenden die Möglichkeit zu geben, diese im Rahmen von Datenschutzerklärungen und Allgemeinen Geschäftsbedingungen zu regeln).
- Sie fordern für riskante Anwendungen die Zertifizierung der Datenschutzstandards verpflichtend ein (Bundesregierung 2018).

7 Siehe hierzu ebenfalls: Datenethikkommission (2019); Deutscher Ethikrat (2017).

3. Anwendbarkeit in Unternehmen

3.1 Bedeutung ethischer Kriterien für ausgewählte Unternehmen

Bei der Entwicklung des Leitfadens für eine verantwortungsvolle Entwicklung und Anwendung von KI-Systemen stand von Anfang an die Überzeugung, dass die erarbeiteten Werte, Kriterien und Handlungsoptionen nicht losgelöst von der unternehmerischen Praxis in einem theoretischen Stadium verharren können, wenn sie einen Beitrag zur Realisierung eines ethisch fundierten Technologieentwicklungs- und Anwendungsprozesses leisten sollen.

Ein wichtiges Ziel dieses Leitfadens soll demnach die Rückkopplung der Werte und Handlungsoptionen an die Praxis von Unternehmen in Deutschland sein. Im Zentrum steht dabei die Frage, ob die theoretischen Werte in der Praxis realisierbar sind. Dazu wurden ausgewählte Unternehmen befragt, die Mitglieder in der Plattform Lernende Systeme sind und selbst KI-Systeme entwickeln und diese auch operativ einsetzen. Sie decken also die Bereiche ab, die in diesem Papier „Entwickelnde“ und „Anbietende“ genannt werden. Mit SAP und der Deutschen Telekom werden zwei große Unternehmen betrachtet, die in ihrem Bereich zu den führenden Akteuren auf dem Markt zählen, die datanizing GmbH erweitert als vergleichsweise junges Unternehmen diese Perspektive um die Sicht eines Start-ups.

Bereits heute legen viele Unternehmen Wert auf einen ethisch basierten Entwicklungs- und Anwendungsprozess von KI-Systemen. Besonders **Transparenz, Diskriminierungsfreiheit und Vielfalt** sowie **Verantwortung beziehungsweise Sicherheit** sind hier als Kategorien hervorzuheben, die die Vertreterinnen und Vertreter der befragten Unternehmen häufig als wichtige Werte nennen.

Die Einhaltung von ethischen Werten und Prinzipien, die nicht vorrangig der Logik des Marktes folgen, kann dazu führen, im freien Wettbewerb Nachteile zu erleiden und damit wirtschaftlich abgehängt zu werden, so lautet eine gängige Befürchtung. Diese Sorge teilen die befragten Unternehmen jedoch nicht. Im Gegenteil: „**Die Umsetzung ethischer Werte ist für viele Unternehmen kein Wettbewerbshindernis, sondern kann vielmehr zum Verkaufsargument werden**“, so Stephanie Fischer, Gründerin der datanizing GmbH.

Gleichzeitig steht für die befragten Unternehmen außer Frage, dass für die Umsetzung ethischer Werte eine echte Notwendigkeit besteht. Ein sehr aktuelles Beispiel dafür ist eine Applikation, die die Deutsche Telekom zusammen mit SAP zur Bekämpfung der Corona-Krise 2020 zügig entwickelt hat. Bevor die Telekom und SAP das Projekt übernommen haben, schlichen sich beim Projektkonsortium auch aufgrund der gebotenen Eile

zunächst Fehler ein, Intransparenz und Datenschutz- sowie Privacy-Bedenken waren die Folge. Die Privatsphäre der Nutzenden wäre damit gefährdet worden. Die Entwickelnden von SAP und Telekom haben, als sie das Projekt übernommen haben, in besonderer Weise Transparenz und Nachvollziehbarkeit vorgelebt, um auch die Akzeptanz der Anwendung zu steigern. „**Dieses Beispiel zeigt, dass die Einhaltung notwendiger Kriterien, wie Transparenz oder auch der Fokus auf Sicherheit und Privacy, in der Entwicklung und Anwendung auch von KI-Systemen höchste Priorität haben muss**“, so Thomas Schauf, Senior Expert Public & Regulatory Affairs bei der Telekom.

3.2 Praxisbeispiele

Viele Unternehmen sind sich der Bedeutung ethischer Werte zur Realisierung eines verantwortungsvollen Technologieentwicklungsprozesses nicht nur bewusst, sie haben auch bereits Maßnahmen ergriffen, wie die Einhaltung dieser Werte unter anderem gewährleistet werden kann. Die nachfolgenden Praxisbeispiele sollen exemplarisch aufzeigen, welche unterschiedlichen Maßnahmen möglich sind. Sie haben nicht den Anspruch, alle vorgestellten Werte, Prinzipien und Voraussetzungen zur Realisierung abzubilden, vielmehr sollen sie einen ersten Eindruck vermitteln, wie die Umsetzung einiger dieser Werte, Prinzipien und Voraussetzungen in die Praxis möglicherweise gelingen kann. Die vorgestellten Maßnahmen und Programme sind erste Schritte hin zu einem ethisch reflektierten Technikentwicklungs- und Technikanwendungsprozess. Trotz dieser Ansätze muss im einzelnen Fall immer noch die Frage gestellt werden, ob eine staatliche Regulierung notwendig ist.

Mit seinem Textanalyse-Tool steht datanizing vor der Herausforderung, wie Kundinnen und Kunden der Einsatz von KI vermittelt werden kann. Denn die Kundin oder der Kunde ist in der Regel kein KI-Experte und kann dies auch gar nicht sein. Dennoch legt das Unternehmen Wert auf Transparenz: Der Kundin oder dem Kunden soll die Mitwirkung von KI-Systemen an effizienter Text-Analyse von Anfang an klar sein. Um weder die komplexen Systeme unrechtmäßig vereinfacht zu erklären, noch die Erwartung zu haben, dass sie als komplexe Systeme in ihrer Gesamtheit verstanden werden, verwendet datanizing **Bildersprache („Schweizer Taschenmesser für Textanalyse“)**, um die nötige Transparenz zu erreichen.

Für die Gründerin Stephanie Fischer steht fest, dass die Entwicklung von KI-Systemen nicht in den Händen einiger weniger Menschen liegen sollte, die so spezialisiert sind, dass sie nicht in der Lage sein können, die Anforderungen zu erkennen, die sich aus der **Vielfältigkeit** der gesamten Bevölkerung ergeben. Der Quereinstieg in die KI ist deshalb auch aus gesellschaftlicher Sicht wichtig. Datanizing hat deshalb einen **Leitfaden für KI-Quereinsteiger** entwickelt, der insbesondere fachfremden Personen den Einstieg in die KI erleichtern und somit dazu beitragen soll, breitere gesellschaftliche Perspektiven abzubilden.

Als eines der ersten Unternehmen weltweit hat sich die **Deutsche Telekom** im Frühjahr 2018 in Form einer Selbstverpflichtung selbstbindende Leitlinien zur Digitalen Ethik im Umgang mit der Künstlichen Intelligenz gegeben. Die Leitlinien resultieren unter anderem aus dem genuinen unternehmerischen Interesse der Telekom, da ihr Markterfolg essenziell auf dem Vertrauen ihrer Kunden in die Sicherheit und Vertrauenswürdigkeit ihrer Produkte und Services beruht. Die KI-Leitlinien der Telekom enthalten beispielsweise die Auflage, dass **Verantwortungen** für KI-Systeme klar zugeordnet werden können müssen. Ebenso legen die Leitlinien fest, dass der Kundin oder dem Kunden **transparent** sein muss, wann sie oder er mit einem KI-System interagiert und welche Daten von ihr oder ihm durch KI-Systeme verarbeitet werden. Die KI-Leitlinien werden von der Deutschen Telekom unter anderem durch die **Integration der Leitlinien als Prüfschritt in das Privacy and Security Assessment** praktisch umgesetzt. Das Assessment-Verfahren soll gewährleisten, dass alle Projekte zur Entwicklung und Einführung neuer Techniken und Produkte hohe Anforderungen an Sicherheit und Datenschutz erfüllen, und verfolgt dabei unter anderem folgende Ziele: ein einheitliches und adäquat hohes Sicherheits- und Datenschutzniveau und ein standardisiertes Verfahren für technische Sicherheit und Datenschutz als integraler Bestandteil der Entwicklungsprozesse zu gewährleisten.

Für die eigenen Beschäftigten, die mit KI-Systemen arbeiten oder diese entwickeln, hat **SAP** ein internes Handbuch zur Ethik der Künstlichen Intelligenz veröffentlicht. Die Ethik-Leitlinien folgen dabei sieben definierten Werten, unter anderem **vorurteilsfreies Handeln, Transparenz und Integrität** sowie **Sicherheitsstandards**. Das Handbuch deckt verschiedene Phasen des Entwicklungs- und Anwendungsprozesses ab: Beginnend bei strategischer Ausrichtung der Produktentwicklung über Datenauswahl und Datenspeicherung bis hin zu Konformitätsprüfungen im Vertrieb. Außerdem adressiert das Handbuch verschiedene Berufsgruppen, die an den Entwicklungszyklen von KI-Systemen beteiligt sind. Die KI-Leitlinien wurden bei SAP sowohl in einem unternehmensinternen Lenkungskreis für KI-Ethik als auch einem externen, interdisziplinären KI-Ethik-Beirat entwickelt, den SAP Ende 2018 ins Leben gerufen hatte.

4. Fazit und Ausblick

Entwicklung und Anwendung von Technologie sind nie wertneutrale Prozesse. Vielmehr sind sie eingebettet in einen Bezugsrahmen ethischer und gesellschaftlicher Werte und Prinzipien, die in die Technologie einfließen, welche aber umgekehrt auch durch die Technologie beeinflusst werden. Auch die Entwicklung von Lernenden Systemen, die noch in einem frühen Stadium steht, macht hier keine Ausnahme. Im Gegenteil: KI-Systeme stellen aufgrund ihrer besonderen Eigenschaften eine große Herausforderung aus ethischer Perspektive für Prinzipien der Selbstbestimmung, Gerechtigkeit und Privatheit dar.

Diese Prinzipien sind tragende Elemente einer freien, demokratischen Gesellschaft. Damit Lernende Systeme zu einer wertorientierten Gesellschaftsentwicklung beitragen können, müssen die geeigneten Rahmenbedingungen gesetzt werden. Bei der Entwicklung und Anwendung von KI-Systemen ist die Integration ethischer Prinzipien deshalb von zentraler Bedeutung.

Bei dieser Aufgabe sind die Akteure, die KI-Systeme entwickeln und anbieten, besonders gefragt: Sie sind dafür verantwortlich, dass ethische Werte (Gerechtigkeit, Selbstbestimmung sowie Schutz der Privatheit und der Persönlichkeit) und Prinzipien beachtet werden. Aber auch die Nutzenden und Betroffenen von KI-Systemen können mit ihren Entscheidungen sowohl beim Kauf als auch bei der Nutzung von KI-Systemen einen wesentlichen Beitrag leisten.

Bereits heute haben viele Unternehmen die Notwendigkeit der Umsetzung ethischer Prinzipien erkannt und leisten dazu einen konkreten Beitrag, indem sie beispielsweise durch Selbstverpflichtungen Werte verankern, die in die KI-Entwicklung einfließen können. Gleichwohl ist es für einen wertorientierten Einsatz von Lernenden Systemen in manchen Anwendungskontexten notwendig, dass einheitliche Standards durch die politischen Entscheidungsträgerinnen und Entscheidungsträger geschaffen werden.

Literatur

Beck, Susanne et al. (2019): Künstliche Intelligenz und Diskriminierung. Herausforderungen und Lösungsansätze – Whitepaper aus der Plattform Lernende Systeme, München. <https://www.plattform-lernende-systeme.de/publikationen-details/kuenstliche-intelligenz-und-diskriminierung-herausforderungen-und-loesungsansaetze.html> (abgerufen am: 20.09.2020).

Beining, Leonie (2019): Wie Algorithmen verständlich werden. Ideen für Nachvollziehbarkeit von algorithmischen Entscheidungsprozessen für Betroffene. https://www.stiftung-nv.de/sites/default/files/wie_algorithmen_verstandlich_werden.pdf (abgerufen am 26.06.2020).

Bundesregierung – KI – Nationale Strategie für Künstliche Intelligenz (2018): Strategie Künstliche Intelligenz der Bundesregierung. https://www.bmbf.de/files/Nationale_KI-Strategie.pdf (abgerufen am 26.06.2020).

Bundesregierung (2020): Stellungnahme der Bundesregierung der Bundesrepublik Deutschland zum Weißbuch zur Künstlichen Intelligenz – ein europäisches Konzept für Exzellenz und Vertrauen. https://www.ki-strategie-deutschland.de/files/downloads/Stellungnahme_BReg_Weissbuch_KI.pdf (abgerufen am 13.08.2020).

BVDW (2019): Künstliche Intelligenz, Smart Home, vernetzte Gesundheit – Ethik in der Digitalen Wirtschaft (Studie). <https://www.bvdw.org/themen/publikationen/detail/artikel/kuenstliche-intelligenz-smart-home-vernetzte-gesundheit-ethik-in-der-digitalen-wirtschaft/> (abgerufen am 26.06.2020).

Datenethikkommission (DEK) (2019): Gutachten der Datenethikkommission der Bundesregierung. https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf;jsessionid=1A60B9612A0459BD9245B91BC042CB8E.2_cid364?blob=publicationFile&v=4 (abgerufen am 26.06.2020).

Datenschutzkonferenz (DSK) (2019): Positionspapier der DSK zu empfohlenen technischen und organisatorischen Maßnahmen bei der Entwicklung und dem Betrieb von KI-Systemen. https://www.datenschutzkonferenz-online.de/media/en/20191106_positionspapier_kuenstliche_intelligenz.pdf (abgerufen am 26.06.2020).

Deutscher Ethikrat (2017): Big Data und Gesundheit – Datensouveränität als informationelle Freiheitsgestaltung (Stellungnahme). <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-big-data-und-gesundheit.pdf> (abgerufen am 26.06.2020).

Europäische Kommission (2020): Weißbuch zur Künstlichen Intelligenz – ein europäisches Konzept für Exzellenz und Vertrauen. https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_de.pdf (abgerufen am 26.06.2020).

Floridi, Luciano et al. (2018): AI4People – An Ethical Framework for a Good AI Society. Opportunities, Risks, Principles, and Recommendations, in: Minds and Machines, 2018, Vol. 28., No. 4, S. 689 – 707.

Hagendorff, Thilo (2019): The Ethics of AI Ethics – An Evaluation of Guidelines.

Heesen, Jessica/Reinhardt, Karoline/Schelenz, Laura (2020): Diskriminierung durch Algorithmen vermeiden. Analysen und Instrumente für eine digitale demokratische Gesellschaft, in: Gero Bauer et al. (Hrsg.): Diskriminierung/Antidiskriminierung. Transcript (in Vorbereitung).

Heesen, Jessica et al. (2020): Zertifizierung von KI-Systemen – Impulspapier aus der Plattform Lernende Systeme, München. https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG3_Impulspapier_290420.pdf (abgerufen am 22.09.2020)

Jobin, Anna et al. (2019): Artificial Intelligence. The Global Landscape of Ethics Guidelines, in: Nature Machine Intelligence, 7/2019, S. 389 – 39.

KI Bundesverband e.V. (2019): KI Gütesiegel des KI Bundesverbandes. https://ki-verband.de/wp-content/uploads/2019/02/KIBV_Guetesiegel.pdf (abgerufen am 26.06.2020).

Matzner, Tobias/Ochs, Carsten (2019): Privacy, in: Internet Policy Review, 2019, Vol. 8, No. 4. <https://policyreview.info/concepts/privacy> (abgerufen am 26.06.2020).

Mittelstand Digital (2019): Digitale Plattformen als Chance für den Mittelstand. Relevanz, Anwendungen, Transfer. https://www.mittelstand-digital.de/MD/Redaktion/DE/Publikationen/digitale-plattformen-als-chance.pdf?__blob=publicationFile&v=3 (abgerufen am 26.06.2020).

Nissenbaum, Helen (2019): „Contextual Integrity Up and Down the Data Food Chain“. <https://nissenbaum.tech.cornell.edu/papers/Contextual%20Integrity%20Up%20and%20Down.pdf> (abgerufen am 26.06.2020).

OECD (2019): Empfehlung des Rats zu Künstlicher Intelligenz. <http://www.oecd.org/berlin/presse/Empfehlung-des-Rats-zu-kuenstlicher-Intelligenz.pdf> (abgerufen am 26.06.2020).

Plattform Lernende Systeme (2019a): Innovation nutzen, Werte schaffen. Neue Geschäftsmodelle mit Künstlicher Intelligenz Bericht der Arbeitsgruppe Geschäftsmodellinnovationen, München. https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG4_Bericht_231019.pdf (abgerufen am 21.09.2020).

Plattform Lernende Systeme (2019b): Prävention, Diagnose, Therapie. Lernende Systeme im Gesundheitswesen. Bericht der Arbeitsgruppe Gesundheit, Medizintechnik, Pflege, München. https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG6_Bericht_23062019.pdf (abgerufen am 21.09.2020)

The European Commission's High Level Expert Group on Artificial Intelligence (AI HLEG) (04/2019): Ethics Guidelines for Trustworthy AI. <https://ec.europa.eu/futurium/en/ai-alliance-consultation> (abgerufen am 26.06.2020).

VdTÜV (2020): „Sicherheit und Künstliche Intelligenz Erwartungen, Hoffnungen, Emotionen“. https://www.vdtuev.de/dok_view?oid=777991 (abgerufen am 26.06.2020).

Über dieses Whitepaper

Autorinnen und Autoren

PD Dr. Jessica Heesen, Universität Tübingen

Prof. Dr. Armin Grunwald, Karlsruher Institut für Technologie (KIT)

Prof. Dr. Tobias Matzner, Universität Paderborn

Prof. Dr. Alexander Roßnagel, Universität Kassel

Die Autorinnen und Autoren sind Mitglieder der Arbeitsgruppe IT-Sicherheit, Privacy, Recht und Ethik der Plattform Lernende Systeme. Als eine von insgesamt sieben Arbeitsgruppen thematisiert sie Fragen zur Sicherheit (Security), Zuverlässigkeit (Safety) und zum Umgang mit Privatheit (Privacy) bei der Entwicklung und Anwendung von Lernenden Systemen. Sie analysiert zudem damit verbundene rechtliche sowie ethische Anforderungen und steht in engem Austausch mit allen weiteren Arbeitsgruppen der Plattform Lernende Systeme.

Befragte Expertinnen und Experten

Fabian Biegel, SAP SE

Stephanie Fischer, datanizing GmbH

Mary Carol Madigan, SAP SE

Thomas Schauf, Deutsche Telekom AG

Redaktion

Stephanie Dachsberger, Geschäftsstelle der Plattform Lernende Systeme

Alexander Mihatsch, Geschäftsstelle der Plattform Lernende Systeme

Dr. Ursula Ohliger, Geschäftsstelle der Plattform Lernende Systeme

Die Plattform Lernende Systeme

Lernende Systeme im Sinne der Gesellschaft zu gestalten – mit diesem Anspruch wurde die Plattform Lernende Systeme im Jahr 2017 vom Bundesministerium für Bildung und Forschung (BMBF) auf Anregung des Fachforums Autonome Systeme des Hightech-Forums und acatech – Deutsche Akademie der Technikwissenschaften initiiert. Die Plattform bündelt die vorhandene Expertise im Bereich Künstliche Intelligenz und unterstützt den weiteren Weg Deutschlands zu einem international führenden Technologieanbieter. Die rund 200 Mitglieder der Plattform sind in Arbeitsgruppen und einem Lenkungskreis organisiert. Sie zeigen den persönlichen, gesellschaftlichen und wirtschaftlichen Nutzen von Lernenden Systemen auf und benennen Herausforderungen und Gestaltungsoptionen.

Impressum

Herausgeber

Lernende Systeme –
Die Plattform für Künstliche Intelligenz
Geschäftsstelle | c/o acatech
Karolinenplatz 4 | 80333 München
www.plattform-lernende-systeme.de

Gestaltung und Produktion

PRpetuum GmbH, München

Stand

Oktober 2020

Bildnachweis

tortoon/AdobeStock/Titel

Bei Fragen oder Anmerkungen zu dieser
Publikation kontaktieren Sie bitte Johannes Winter
(Leiter der Geschäftsstelle):
kontakt@plattform-lernende-systeme.de

Folgen Sie uns auf Twitter: @LernendeSysteme

Empfohlene Zitierweise

Jessica Heesen et al. (Hrsg.): Ethik-Briefing. Leitfaden
für eine verantwortungsvolle Entwicklung und Anwen-
dung von KI-Systemen – Whitepaper aus der Plattform
Lernende Systeme, München 2020.

Dieses Werk ist urheberrechtlich geschützt.
Die dadurch begründeten Rechte, insbesondere die
der Übersetzung, des Nachdrucks, der Entnahme von
Abbildungen, der Wiedergabe auf fotomechanischem
oder ähnlichem Wege und der Speicherung in Daten-
verarbeitungsanlagen, bleiben – auch bei nur auszugs-
weiser Verwendung – vorbehalten.