

Handlungsoptionen für Nutzende von KI-Systemen:

Selbstbestimmung

- Verwendung von Produkten, die den Einsatz von KI-Systemen transparent kennzeichnen
- Wahl von Produkten, die die Marktpluralität fördern

Gerechtigkeit

- Vermeidung des Einsatzes von KI in Bereichen mit hohem Diskriminierungspotential
- Stichprobenartige Überprüfung der Empfehlungen von KI-Systemen auf Plausibilität

Privatsphäre

- Einholen von Einwilligungserklärungen zur Erhebung und Verarbeitung der Daten
- Verwendung anonymisierter Datensätze

Handlungsmöglichkeiten für Betroffene von KI-Systemen:

Selbstbestimmung

- Bereitschaft, bestehende Informationsangebote zu KI-Systemen anzunehmen
- Aktives Einfordern von Informationsangeboten

Gerechtigkeit

- Verzicht auf die Angabe von Daten, die zu Diskriminierung führen könnten (sofern möglich)

Privatsphäre

- Bereitstellung von Daten nur gegenüber seriösen Anbietern

Interviews mit Vertreterinnen und Vertretern ausgewählter Unternehmen haben ergeben, dass die Umsetzung dieser Vorschläge in der Praxis möglich ist. Weiterführende Informationen zu den einzelnen Anforderungen und Maßnahmen enthält das Whitepaper *Ethik-Briefing – Leitfaden für eine verantwortungsvolle Entwicklung und Anwendung von KI-Systemen*.

Impressum

Herausgeber: Lernende Systeme –
Die Plattform für Künstliche Intelligenz

Geschäftsstelle: c/o acatech | Karolinenplatz 4 | D-80333 München

kontakt@plattform-lernende-systeme.de
www.plattform-lernende-systeme.de
Folgen Sie uns auf Twitter: @LernendeSysteme

Stand: Oktober 2020

Druck: Druck- und Verlagshaus Zarbock GmbH & Co. KG

Bildnachweis: tortoon/AdobeStock/Titel

Diese Kurzfassung entstand auf Grundlage des Whitepapers *Ethik-Briefing – Leitfaden für eine verantwortungsvolle Entwicklung und Anwendung von KI-Systemen*, München, 2020.

Es wurde erstellt von der Arbeitsgruppe IT-Sicherheit, Privacy, Recht und Ethik der Plattform Lernende Systeme. Die Originalfassung der Publikation ist online verfügbar unter:

<https://www.plattform-lernende-systeme.de/publikationen.html>



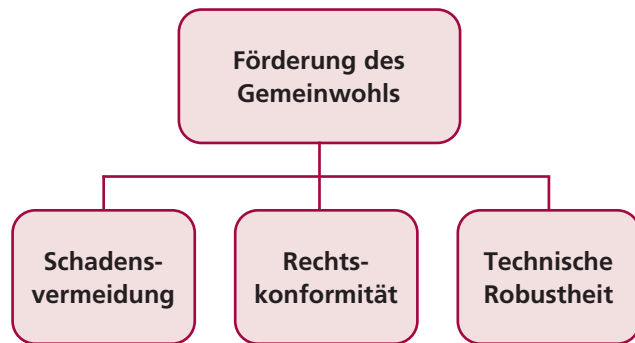
Ethik- Briefings

Leitfaden für eine verantwortungsvolle Entwicklung und Anwendung von KI-Systemen

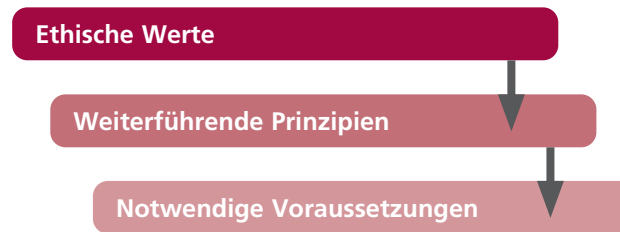
GEFÖRDERT VOM

Bei der verantwortungsvollen Entwicklung und Anwendung von KI-Systemen stellt sich die Frage, welche ethischen Werte beachtet werden müssen und wie deren Umsetzung konkret aussehen kann. Antwortmöglichkeiten gibt die Plattform Lernende Systeme in ihrem Ethik-Briefing. Der Leitfaden für eine verantwortungsvolle Entwicklung und Anwendung von KI-Systemen dient als Orientierung für Entwickelnde, Anbietende, Nutzende und Betroffene dieser Systeme. Dieser Flyer stellt die wichtigsten ethischen Herausforderungen vor und zeigt erste Handlungsempfehlungen auf.

KI-Systeme sollten immer mit dem Ziel entwickelt und eingesetzt werden, zur Förderung des Gemeinwohls beizutragen. KI-Systeme sollen also so eingesetzt werden, dass **Schaden** für Einzelpersonen, Gemeinschaften und die Umwelt **vermieden wird**, die **Rechtskonformität** von KI-Systemen und Akteuren gewährleistet ist und das eingesetzte KI-System die notwendige **technische Robustheit** erfüllt, sodass es zu keinem Zeitpunkt ein unannehmbares Sicherheitsrisiko darstellt.



Darauf aufbauend lassen sich Kriterien für eine verantwortungsvolle Entwicklung und Anwendung von KI-Systemen definieren. Dabei können drei übergeordnete Kategorien ethischer Werte unterschieden werden: **Selbstbestimmung**, **Gerechtigkeit** und **Schutz der Privatheit**. An diese Werte schließen weiterführende ethische Prinzipien sowie notwendige Voraussetzungen zur Realisierung an.



Selbstbestimmung begründet 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 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 Anwendungen. Voraussetzung für die Realisierung sind Rechenschaftspflichten.

Privatsphäre bezeichnet einen Raum, der sich von der Öffentlichkeit abgrenzt und der geschützt ist vor äußeren Eingriffen. Mit zunehmendem Einsatz von KI wird der Schutz der Privatsphäre eine große Herausforderung. **Datenminimierung** und **Transparenz** bei Datenerhebung und -verarbeitung sind Voraussetzungen zum Schutz der Privatsphäre.

Diese ethischen Werte können durch konkrete Maßnahmen von unterschiedlichen Akteuren realisiert werden.

Handlungsoptionen für Entwickelnde eines KI-Systems:

Selbstbestimmung

- Entwicklung von KI-Systemen als Erklärbare KI
- Transparente Kommunikation des Einsatzes und Autonomiegrades der KI
- Entwicklung von KI-Systemen, die den Menschen als Letztentscheider vorsehen

Gerechtigkeit

- Vermeidung von diskriminierenden Variablen in Datensätzen
- Etablierung von Kontrollmechanismen

Privatsphäre

- Verwendung anonymisierter Datensätze
- Etablierung hoher Sicherheitsstandards

Handlungsoptionen für Anbietende von KI-gestützten Produkten und Dienstleistungen:

Selbstbestimmung

- Offene Kommunikation der Verwendung von KI-Systemen
- Anbieten von Schulungen für Beschäftigte

Gerechtigkeit

- Überprüfung von Verletzungen von Gerechtigkeitsprinzipien durch Formen der Selbstregulierung
- Etablierung von Erkennungs- und Reaktionsmuster für Schäden

Privatsphäre

- Einholen von Einwilligungserklärungen zur Erhebung und Verarbeitung der Daten und Anbieten von leicht verständlichen Widerrufsmöglichkeiten zu diesen
- Verwendung anonymisierter Datensätze