

Kriterien für die Mensch-Maschine-Interaktion bei KI

Whitepaper von Norbert Huchler et al.
AG Arbeit/Qualifikation,
Mensch-Maschine-Interaktion



Kurzfassung

Der Einsatz von Künstlicher Intelligenz (KI) eröffnet vielfältige Potentiale für ein sicheres, eigenverantwortliches und selbstbestimmtes Arbeiten sowie attraktive und wettbewerbsfähige Arbeitsplätze. So können KI-basierte Assistenzsysteme Beschäftigte beispielsweise von anstrengenden oder gefährlichen Tätigkeiten entlasten und bei komplexen Prozessen und Entscheidungen unterstützen. Gleichzeitig verändern KI-Systeme die Interaktion von Mensch und Technik in der Arbeitswelt. Mensch und Maschine werden künftig noch stärker – und anders – als bisher zusammenwirken, da Maschinelles Lernen (ML) und ähnliche Technologien Maschinen befähigen, bestimmte Aufgaben selbstständig auszuführen und dabei laufend dazuzulernen.

Die zunehmende Kollaboration zwischen Mensch und Technik macht beim Einsatz von Künstlicher Intelligenz eine Neujustierung der Aufgabenverteilung notwendig. Um diese im Sinne des Menschen zu gestalten, muss die Technologie an die Vorteile und Potentiale menschlichen Denkens und Handelns anknüpfen und die wechselseitige Ergänzung – und nicht Ersatz oder Konflikt – in den Mittelpunkt der Interaktion stellen. Eine koordinierte Balance, die sowohl den Beschäftigten als auch den technologischen und wirtschaftlichen Potentialen der Künstlichen Intelligenz gerecht wird, erhöht die Chancen für einen individuell und gesellschaftlich akzeptierten Einsatz von Künstlicher Intelligenz in der Arbeitswelt.

Nur durch klar definierte Orientierungsmaßstäbe für die neue Aufgabenverteilung können sichere Arbeitsplätze geschaffen, qualifizierte Arbeitskräfte ausgebildet sowie die Gestaltung guter und menschengerechter Arbeit umgesetzt werden. An dieser Stelle setzt das Whitepaper an und legt einen Kriterienkatalog für die Mensch-Maschine-Interaktion im Arbeitskontext vor. Die Kriterien, die auf eine nachhaltig zukunftsorientierte und menschenzentrierte Gestaltung der Mensch-Maschine-Interaktion zielen, lassen sich zu vier Clustern zusammenfassen:

Tabelle 1: Kriterien für die Gestaltung der Mensch-Maschine-Interaktion

Schutz des Einzelnen	
Sicherheit und Gesundheitsschutz	<ul style="list-style-type: none"> • Vermeidung von Risiken für die physische und psychische Gesundheit von Beschäftigten • Schutz vor Unfall- und Schadensereignissen (Personen- und Sachschäden) • Vorbeugung gegen negative physische oder psychische Beanspruchungsfolgen
Datenschutz und verantwortungsbewusste Leistungserfassung	<ul style="list-style-type: none"> • Schutz der Persönlichkeitsrechte, Datensparsamkeit und Zweckbindung der Datennutzung • Vermeidung von Datenanalysen zur ungerechtfertigten Leistungsmessung • Entwicklung einer positiven Kultur des Leistungsfeedbacks • Transparenz über und Befähigung zu Datenanalysen und deren Nutzung
Vielfaltssensibilität und Diskriminierungsfreiheit	<ul style="list-style-type: none"> • Schutz vor Diskriminierung von Individuen oder Gruppen • Bestehende Rechtsordnung als Grundlage für die Vielfaltsensibilität und Diskriminierungsfreiheit
Vertrauenswürdigkeit	
Qualität der verfügbaren Daten	<ul style="list-style-type: none"> • Vermeidung von qualitativ unzureichendem Datenmaterial • Vorbeugung von verzerrten Datensätzen, Fehlern/ Fehlinterpretationen und Diskriminierung • Verbesserung der Mensch-Maschine-Interaktion durch verlässliche Datengrundlagen
Transparenz, Erklärbarkeit und Widerspruchsfreiheit	<ul style="list-style-type: none"> • Umsetzung von Ansätzen einer erklärbaren Künstlichen Intelligenz (Explainable AI) • Entwicklung von Methoden zur Nachvollziehbarkeit Lernender Systeme • Herstellung (abgestufter) Transparenz über Entscheidungsprozesse Lernender Systeme • Verhinderung von Demotivation durch widerspruchsfreie Gestaltung der Mensch-Maschine-Interaktion
Verantwortung, Haftung und Systemvertrauen	<ul style="list-style-type: none"> • Transparenz über und Zuschreibbarkeit von Verantwortung • Kompetenz und Kontrolle über das System als Voraussetzung für die Übernahme von Verantwortung • Ausmaß der Systemkontrolle als Gradmesser für Art und Umfang der Verantwortung • Ansatzpunkte entlang der Konzepte einer vertrauenswürdigen Künstlichen Intelligenz (Trustworthy AI)
Sinnvolle Arbeitsteilung	
Angemessenheit, Entlastung und Unterstützung	<ul style="list-style-type: none"> • Angemessene Arbeitsinhalte und Anforderungen • Komplementäre Ergänzung der Fähigkeiten von Menschen und Maschinen für eine sinnvolle Arbeitsteilung • Nachhaltige Entlastung und Unterstützung von Beschäftigten durch KI-basierte Assistenzsysteme • Befähigung zur Arbeit mit KI-Systemen
Handlungsträgerschaft und Situationskontrolle	<ul style="list-style-type: none"> • Gezielte und transparente Gestaltung der Handlungsträgerschaft und Situationskontrolle • Minimierung und Vermeidung von Risiken und negativen Beanspruchungsfolgen
Adaptivität, Fehlertoleranz und Individualisierbarkeit	<ul style="list-style-type: none"> • Befähigung Lernender Systeme, sich flexibel und situationspezifisch an den Bedarfen und Bedürfnissen sowie an der Arbeitspraxis der Nutzerinnen und Nutzer ausrichten zu können

Förderliche Arbeitsbedingungen

Handlungsräume und reichhaltige Arbeit	<ul style="list-style-type: none"> • Sicherung und gegebenenfalls Ausweitung von Handlungsspielräumen von Beschäftigten (insbesondere Autonomie und Entscheidungsfreiheit sowie Vielfalt an Handlungsmöglichkeiten) • Einbeziehung von menschlichen Grundbedürfnissen für eine sinnvolle, motivierende und gesundheits- sowie persönlichkeitsförderliche Arbeit
Lern- und Erfahrungsförderlichkeit	<ul style="list-style-type: none"> • Ermöglichung des wechselseitigen Lernens des Menschen von der Maschine und umgekehrt • Nachvollziehbare, adaptive Gestaltung der Systeme zur Integration von Wissen und Erfahrung • Gewährleistung der Überführung von Daten in Information beziehungsweise von Information in Wissen
Kommunikation, Kooperation und soziale Einbindung	<ul style="list-style-type: none"> • Doppelte Sensibilisierung von Künstlicher Intelligenz für die sozialen Kontexte und Strukturen • Unterstützung zwischenmenschlicher Kommunikation, Kooperation und Verbundenheit • Künstliche Intelligenz als Kooperationspartner

Diese Kriterien richten sich an Akteure aus der Planung und Entwicklung Lernender Systeme sowie an Akteure, die an der Implementierung von KI-Systemen in Unternehmen beteiligt sind. Der Kriterienkatalog soll Handreichungen für die Gestaltung der Aufgabenteilung zwischen Mensch und Technik bei der Anwendung Lernender Systeme anbieten. Außerdem sollen die Kriterien zur Weiterentwicklung bestehender Regelungen inspirieren – etwa in der Normung, Gesetzgebung oder Sozialpartnerschaft – und künftig ein flexibleres, selbstbestimmteres und eigenverantwortlicheres Arbeiten ermöglichen.

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: Juni 2020 | Bildnachweis: Westend61/gettyimages

Diese Kurzfassung entstand auf Grundlage des Whitepapers Kriterien für die Mensch-Maschine-Interaktion bei KI – Ansätze für die menschengerechte Gestaltung in der Arbeitswelt, München, 2020. Es wurde erstellt von der Arbeitsgruppe Arbeit/Qualifikation, Mensch-Maschine-Interaktion der Plattform Lernende Systeme. Die Originalfassung der Publikation ist online verfügbar unter: <https://www.plattform-lernende-systeme.de/publikationen.html>



GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

acatech
DEUTSCHE AKADEMIE DER
TECHNIKWISSENSCHAFTEN