

Alles unter einem D-A-CH

Wie KI, Robotics und Bionic Accessibility und Inklusion beeinflussen (könn(t)en)?

Und: Brauchen wir Zertifizierung für KI ?

#AllesUnterEinemDach

Definition von KI lt. HLG zur AI

„Künstliche Intelligenz (KI) bezeichnet Systeme mit einem „intelligenten“ Verhalten, die ihre Umgebung analysieren und mit einem gewissen Grad an Autonomie handeln, um bestimmte Ziele zu erreichen.

KI-basierte Systeme können rein softwaregestützt in einer virtuellen Umgebung arbeiten (z. B. Sprachassistenten, Bildanalysesoftware, Suchmaschinen, Sprach- und Gesichtserkennungssysteme), aber auch in Hardware-Systeme eingebettet sein (z. B. moderne Roboter, autonome Pkw, Drohnen oder Anwendungen des ‚Internet der Dinge‘).

Definition von KI lt. HLG zur AI

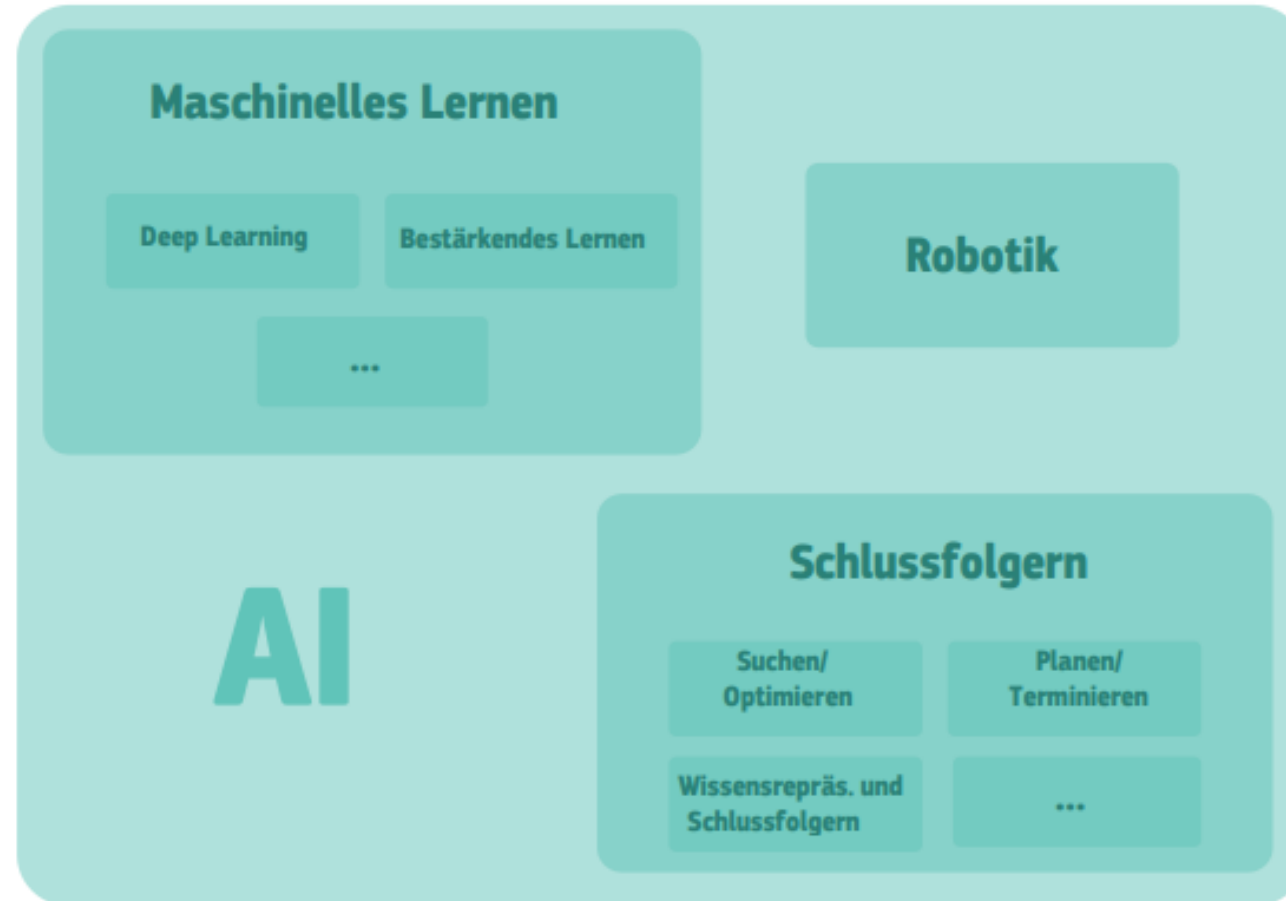


Abbildung 2: Vereinfachter Überblick über die Teilidisziplinen der KI und ihre Zusammenhänge. Maschinelles Lernen und Schlussfolgern umfassen jeweils eine Reihe anderer Techniken; die Robotik wiederum beinhaltet Techniken, die nicht zur KI gehören. Die KI als Ganzes ist eine Teilidisziplin der Informatik.

- **Enge oder angewandte (schwache) KI und allgemeine (starke) KI.** Ein allgemeines KI-System soll die meisten Aufgaben erfüllen können, zu denen auch der Mensch in der Lage ist. Enge (oder angewandte) KI-Systeme hingegen sollen eine oder einige wenige spezifische Aufgaben erledigen. Bei den heute verwendeten KI-Systemen handelt es sich um Beispiele der engen KI.
- **Datenqualität und Verzerrungen.** Da viele KI-Systeme – zum Beispiel solche mit Komponenten des überwachten maschinellen Lernens – enorme Datenmengen brauchen, um gute Ergebnisse zu erzielen, ist es wichtig, zu verstehen, wie Daten das Verhalten des KI-Systems beeinflussen.
- **Black-Box-KI und Erklärbarkeit.** Einige Verfahren des maschinellen Lernens sind zwar im Hinblick auf die Genauigkeit sehr erfolgreich, aber gleichzeitig auch sehr undurchsichtig, was die Art und Weise ihrer Entscheidungsfindung betrifft.
- **Zielgerichtete KI.** Die heutigen KI-Systeme arbeiten zielgerichtet. Das bedeutet, dass ihnen vom Menschen ein bestimmtes Ziel vorgegeben wird, welches sie mithilfe bestimmter Verfahren umsetzen. Sie bestimmen ihre Ziele nicht selbst

Angepasste Definition von KI lt. HLG zur AI

„Systeme der künstlichen Intelligenz (KI-Systeme) sind vom Menschen entwickelte Softwaresysteme (und gegebenenfalls auch Hardwaresysteme), die in Bezug auf ein komplexes Ziel auf physischer oder digitaler Ebene handeln, indem sie ihre Umgebung durch Datenerfassung wahrnehmen, die gesammelten strukturierten oder unstrukturierten Daten interpretieren, Schlussfolgerungen daraus ziehen oder die aus diesen Daten abgeleiteten Informationen verarbeiten, und über das bestmögliche Handeln zur Erreichung des vorgegebenen Ziels entscheiden.

KI-Systeme können entweder symbolische Regeln verwenden oder ein numerisches Modell erlernen, und sind auch in der Lage, die Auswirkungen ihrer früheren Handlungen auf die Umgebung zu analysieren und ihr Verhalten entsprechend anzupassen.

Als wissenschaftliche Disziplin umfasst die KI mehrere Ansätze und Techniken wie z. B. maschinelles Lernen (Beispiele dafür sind „Deep Learning“ und bestärkendes Lernen), maschinelles Denken (es umfasst Planung, Terminierung, Wissensrepräsentation und Schlussfolgerung, Suche und Optimierung) und die Robotik (sie umfasst Steuerung, Wahrnehmung, Sensoren und Aktoren sowie die Einbeziehung aller anderen Techniken in cyber-physische Systeme).“

Rahmen für eine vertrauenswürdige KI

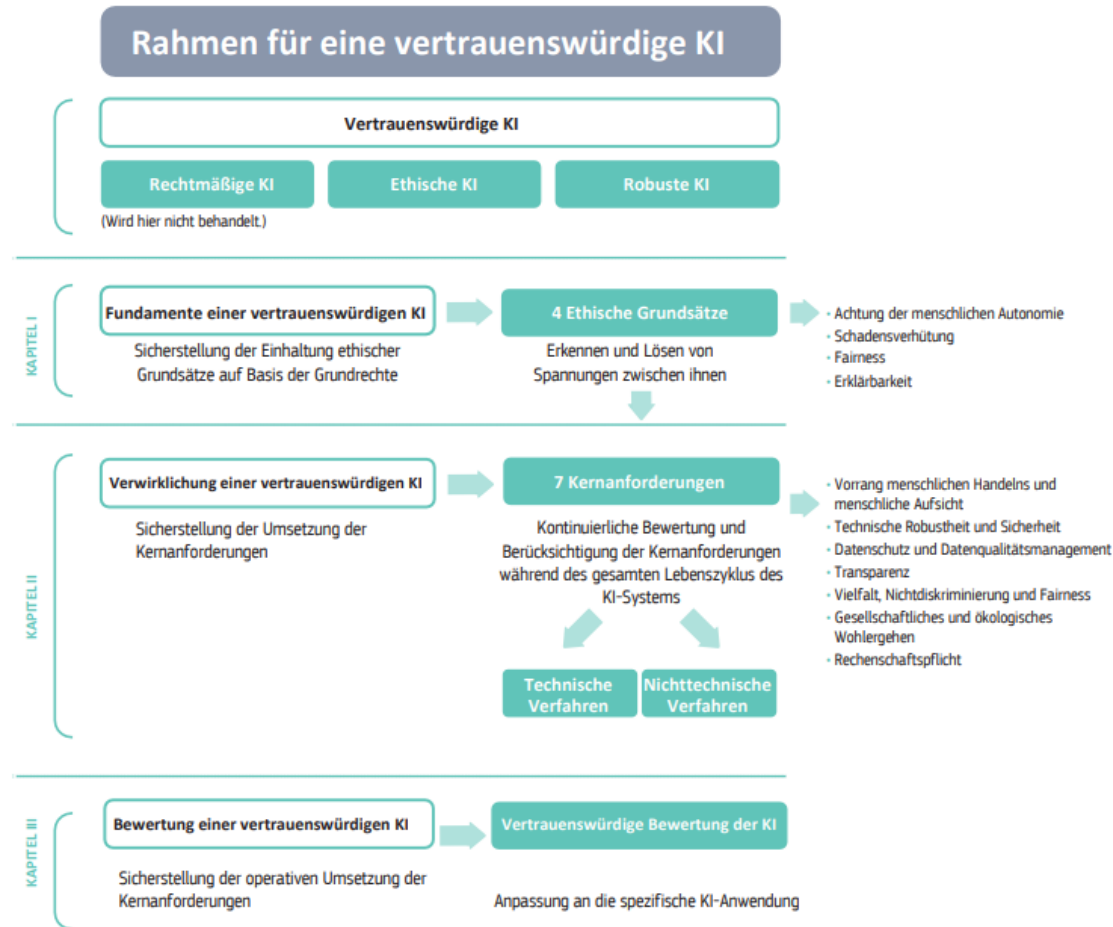


Abbildung 1: Die Leitlinien als Rahmen für eine vertrauenswürdige KI

7 Kernanforderungen

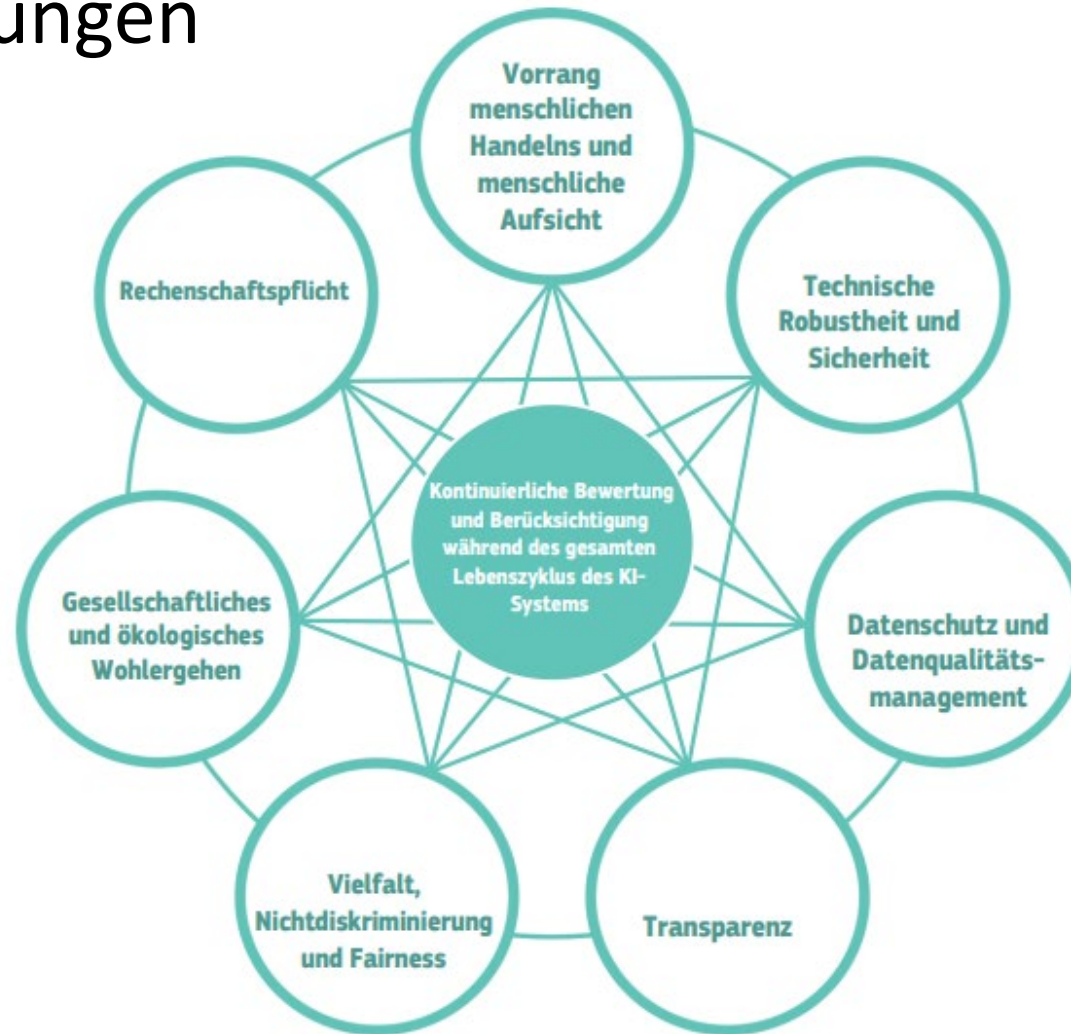


Abbildung 2: Beziehung zwischen den sieben Anforderungen: Alle sind in Bezug auf ihre Bedeutung gleichrangig, unterstützen sich gegenseitig und sollten während des gesamten Lebenszyklus eines KI-Systems umgesetzt und bewertet werden.

ALTAI – Assessment List for Trustworthy AI

- Im Anschluss an eine Pilotphase, an der 350 Stakeholder teilnahmen, wurde die Assessment Liste in ein Onlinetool angepasst und modifiziert (aufgrund der Einlassungen der Teilnehmer). Sie steht nur online – in EN – zur Verfügung.
- Die 7 Kernprinzipien wurden anhand eines Fragekatalogs ausformuliert, und jeder Teilnehmer erhält im Anschluss eine Bewertung, ob seine Organisation diesen Kriterien entspricht.
- Sie fußt auf dem Schutz der Grundrechte aller Menschen, welche in den Gründungsverträgen der EU festgehalten sind sowie auf internationalen Rechtsgrundlagen, wie der UNCRPD

<https://futurium.ec.europa.eu/en/european-ai-alliance/pages/altai-assessment-list-trustworthy-artificial-intelligence>

- 1. Diskriminiert das KI-System Menschen möglicherweise negativ aus einem der folgenden Gründe (nicht taxativ): Geschlecht, Rasse, Hautfarbe, ethnische oder soziale Herkunft, genetische Merkmale, Sprache, Religion oder Weltanschauung, politisch anderweitig begründeten Meinung, Zugehörigkeit zu einer nationalen Minderheit, Geburt, Behinderung, Alter oder sexuelle Orientierung?
 - Haben Sie Prozesse eingerichtet, um mögliche negative Diskriminierungen (Bias) während der Entwicklungs-, Bereitstellungs- und Nutzungsphasen des KI-Systems zu testen und zu überwachen?
 - Haben Sie Prozesse eingerichtet, um mögliche negative Diskriminierungen (Verzerrungen) im KI-System anzugehen und zu korrigieren?
- 2. Respektiert das KI-System die Rechte des Kindes, beispielsweise in Bezug auf den Kinderschutz und unter Berücksichtigung des Kindeswohls?

- Schützt das KI-System personenbezogene Daten von Personen im Einklang mit der DSGVO?
 - Haben Sie Verfahren eingerichtet, um die Notwendigkeit einer Datenschutz-Folgenabschätzung im Detail zu bewerten, einschließlich einer Bewertung der Notwendigkeit und Verhältnismäßigkeit der Verarbeitungsvorgänge in Beziehung zu ihrem Zweck in Bezug auf die Entwicklungs-, Bereitstellungs- und Nutzungsphasen des KI-Systems?
 - Haben Sie Maßnahmen ergriffen, um die Risiken zu berücksichtigen, einschließlich Schutzmaßnahmen, Sicherheitsmaßnahmen und Mechanismen, um den Schutz personenbezogener Daten in Bezug auf die Entwicklungs-, Bereitstellungs- und Nutzungsphasen des KI-Systems zu gewährleisten?

- Respektiert das KI-System die Meinungs- und Informationsfreiheit und / oder die Versammlungs- und Vereinigungsfreiheit?
 - Haben Sie Prozesse eingerichtet, um mögliche Verstöße gegen die Meinungs- und Informationsfreiheit und / oder die Versammlungs- und Vereinigungsfreiheit während der Entwicklungs-, Bereitstellungs- und Nutzungsphasen des KI-Systems zu testen und zu überwachen?
 - Haben Sie Prozesse eingerichtet, um mögliche Verstöße gegen die Meinungs- und Informationsfreiheit und / oder die Versammlungs- und Vereinigungsfreiheit im KI-System zu beheben und zu korrigieren?

Fragen in Bezug auf Vielfalt, Nichtdiskriminierung und Fairness

- Vermeidung von unfairem Bias
 - Mechanismen und Strategien, um Bias zu verhindern bzw. aufzudecken. Testen von Use Cases durch die entsprechenden Zielgruppen, Diversität und Abdeckung aller End-User berücksichtigen
- Accessibility und Universal Design
 - Wurde sichergestellt, dass die Präferenzen und Möglichkeiten der Gesellschaft abgedeckt wurden – no one behind. Kann das System von Menschen mit Behinderung bedient werden, ist es barrierefrei. Wurde Zielgruppen/User von Assistierenden Technologien einbezogen während der Planungs- und Implementierungsphase
- Stakeholder Participation
 - Beteiligung eines möglichst breiten Spektrums möglicher Akteure am Entwurf und der Entwicklung des KI-Systems

Wo könnten KI – Technologien für Menschen mit Behinderungen hilfreich sein – eine Aufzählung

- Sprachsteuerung von Computern und Smartphones
- Automatische Übersetzung und Untertitelung
- Bilderkennung und automatisierter ALT-Text, automatisierte Audiobeschreibung
- Zusammenfassen / Löschen von Texten für Personen mit Leseschwierigkeiten
- Gesichtserkennung
- Pflegeroboter, Medizinroboter

Wo könnten KI – Technologien für Menschen mit Behinderungen hilfreich sein – eine Aufzählung

- Automatisiertes Testen von Websites, Reengineering, Anpassen von Websites, automatisierte Korrektur
- Avatare für die on the fly Übersetzung von Gebärdensprache
- Autonome Autos - blind fahren? Oder wenn Sie eine Mobilitätseinschränkung haben.
- Virtuelle Realität und Augmented Reality (zum Lernen, Erkunden der Umwelt ..)
- Smart Cities / Smart Homes / Intelligente Umgebung

Bedenken?

- 88% waren besorgt über mangelnde Barrierefreiheit
- 60% waren besorgt über mangelnde Standardisierung
- 56% waren besorgt über die Interoperabilität mit der von ihnen bereits verwendeten assistierenden Technologien
- 50% waren besorgt über neue Technologien, die zu Diskriminierung führen
- 40% waren besorgt über neue Technologien, die sich auf ihre Privatsphäre auswirken
- 42% waren besorgt darüber, dass ET ihre Sicherheit beeinträchtigen könnte
- 20% hatten andere Bedenken, einschließlich Bedenken hinsichtlich Benutzerfreundlichkeit, Erschwinglichkeit und Mangel an digitale Fähigkeiten

Lt. eine Umfrage des EDF: PLUG AND PRAY? A disability perspective on artificial intelligence, automated decision-making and emerging technologies

Quellen:

- <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>
- <http://www.edf-feph.org/content/uploads/2020/12/edf-emerging-tech-report-accessible.pdf>
- 17th International Conference, ICCHP 2020, Lecco, Italy, September 9–11, 2020, Proceedings, Part I

