

Zusammenfassung

Künstliche Intelligenz (KI) und Robotik sind bahnbrechende Technologien, die Wirtschaft und Gesellschaft grundlegend verändern. Um diese Transformation zu verstehen und zu antizipieren, müssen Politikverantwortliche zunächst einmal wissen, was diese Technologien leisten können und was nicht. Die OECD hat das Projekt *Artificial Intelligence and the Future of Skills* eingeleitet, um ein Programm aufzulegen, in dem die Fähigkeiten von KI und Robotik sowie ihre Auswirkungen auf das Bildungswesen und die Arbeitswelt beurteilt werden können. Dieser Bericht ist der erste Schritt zur Erarbeitung der Methodik dieses Projekts. Er überprüft Taxonomien und Tests, die in der Psychologie und Informatik entwickelt wurden, und erörtert ihre Stärken und Schwächen sowie ihre Anwendbarkeit zur Beurteilung der Fähigkeiten von Maschinen.

Die Beurteilung der Fähigkeiten von KI und Robotik ist unerlässlich, um ihre Auswirkungen auf das Bildungswesen, die Arbeitswelt und die Gesellschaft insgesamt zu verstehen.

Ein kontinuierliches Programm zur Beurteilung von KI und Robotik wird die internationalen Vergleichsgrößen der OECD, die den Politikverantwortlichen helfen, menschliche Kompetenzen zu verstehen, entscheidend erweitern. Die Internationale Schulleistungsstudie (PISA) beschreibt den Zusammenhang zwischen dem Bildungssystem und der Entwicklung menschlicher Kompetenzen, und die Internationale Vergleichsstudie der Kompetenzen Erwachsener (PIAAC) untersucht, wie sich diese Kompetenzen auf die Berufstätigkeit und andere wichtige Aufgaben von Erwachsenen auswirken. Mit dem neuen Programm kommt nun eine Untersuchung der Bedeutung menschlicher Kompetenzen im Kontext von KI und Robotik hinzu. Dadurch wird eine Brücke zwischen KI und Robotik, ihren Auswirkungen auf Bildung und Arbeit und den daraus resultierenden sozialen Umwälzungen der kommenden Jahrzehnte geschlagen.

Es gibt eine Reihe von Taxonomien und Tests für menschliche Kompetenzen. Sie bieten unterschiedliche Perspektiven und Möglichkeiten für das Verständnis von KI-Fähigkeiten.

Die Taxonomien der kognitiven Psychologie sind hierarchische Modelle breit gefasster kognitiver Fähigkeiten wie fluide Intelligenz, Erinnern/Lernen sowie visuelle und auditive Wahrnehmung, die durch Faktorenanalysen kognitiver Fähigkeitstests beurteilt werden. Diese Tests werden gerne zur Beurteilung menschlicher Kompetenzen eingesetzt und wurden vielfach validiert.

Das Forschungsinteresse an sozialen und emotionalen Kompetenzen nimmt zu und die entsprechenden Tests machen Fortschritte. Während es bei diesen Kompetenzen zunächst einmal um Persönlichkeit, Temperament, Einstellungen, Integrität und Interaktion von Einzelpersonen geht, beschäftigen sich neuere Forschungsarbeiten auch mit kollektiven Fähigkeiten. Im Rahmen dieser Studien werden die Faktoren der „kollektiven Intelligenz“ untersucht und Tests entwickelt, um sie zu messen.

Die Bildungsforschung hat ebenfalls zu einem besseren Verständnis menschlicher Kompetenzen beigetragen. Dabei geht es um fachspezifisches Wissen (z. B. im Bereich der Mathematik, Biologie oder Geschichte), Grundkompetenzen wie Lesen und Rechnen sowie komplexere übergreifende Fähigkeiten, von Problemlösen, Zusammenarbeit und Kreativität bis hin zu digitaler und globaler Kompetenz. In groß angelegten internationalen und nationalen Bildungserhebungen sind verschiedene Testformate zu finden.

Kompetenzen können auf Arbeitstätigkeiten und Berufe bezogen und durch komplexe Eignungstests gemessen werden.

Ein weiterer wichtiger Bereich – Arbeits- und Organisationspsychologie – untersucht den Zusammenhang zwischen Fähigkeiten und berufsspezifischen Tätigkeiten. Die daraus resultierenden umfassenden Berufstaxonomien klassifizieren Berufe nach Arbeitsaufgaben und den dafür erforderlichen Kompetenzen, Kenntnissen und Fähigkeiten. Die am häufigsten genutzten Klassifikationen sind die O*NET-Datenbank des US-Arbeitsministeriums und die Europäische Klassifizierung für Fähigkeiten, Kompetenzen, Qualifikationen und Berufe (ESCO). Die für Beurteilungen herangezogenen beruflichen Eignungsprüfungen haben unterschiedliche Formate.

Gesunde Erwachsene besitzen einige Grundkompetenzen, die KI-Systeme nicht haben.

Viele Taxonomien zur Beurteilung von Kompetenzen übersehen wichtige einfache oder grundlegende kognitive Kompetenzen. Diese werden bei erwachsenen Menschen nur in Ausnahmefällen beurteilt, weil sich außer bei schweren Behinderungen nur selten signifikante Unterschiede zwischen den Proband*innen ergeben. Bei KI-Systemen sind sie aber nicht zwangsläufig vorhanden (z. B. die Navigation in einem komplexen physischen Umfeld, grundlegende Sprachkenntnisse oder allgemeines Weltwissen). In den Bereichen tierische Kognition, kindliche Entwicklung und Neuropsychologie sind diese Kompetenzen bereits Gegenstand von Taxonomien und Beurteilungen. Die Erkenntnisse dieser Psychologiefelder werden in einem neuen Forschungszweig genutzt, um die grundlegenden (einfachen) Kompetenzen von KI-Systemen zu beurteilen.

Die Evaluierung von KI- und Robotiksystemen ist komplex und die für Menschen entwickelten Tests sind nicht immer übertragbar.

Die KI-Beurteilung konzentriert sich auf Fähigkeiten intelligenter Mechanismen, wie Wissensrepräsentation, logisches Schlussfolgern, Wahrnehmung, Navigation und Verarbeitung natürlicher Sprache. Diese hängen eng mit dem technischen Verfahren zusammen, das dem System zugrunde liegt. Viele Fähigkeiten überschneiden sich zwar mit den Fähigkeitskategorien, die in der Psychologie für Menschen entwickelt wurden, aber die Übereinstimmung ist nicht exakt. Außerdem werden zahlreiche Fähigkeiten, die im Bereich der KI derzeit entwickelt werden – wie beispielsweise Sprachidentifikation und die Erzeugung realistischer Bilder – von den Taxonomien oder Tests für menschliche Kompetenzen nicht gut erfasst.

Des Weiteren wird bei der Konzeption von Tests für Menschen unterstellt, dass alle Proband*innen bestimmte grundlegende Merkmale menschlicher Intelligenz besitzen, was bei KI völlig anders sein kann. Menschen fällt es beispielsweise leicht, grundlegende Kompetenzen wie das Verstehen natürlicher Sprache und Objekterkennung in verschiedene Aufgaben zu integrieren. Demgegenüber werden die meisten KI-Systeme trainiert, eine bestimmte eng abgegrenzte Aufgabe auszuführen, und sind nicht (oder nur selten) in der Lage, diese Kompetenzen bei verschiedenen Tätigkeiten einzusetzen. Deshalb ist es schwierig,

von der Leistung eines KI-Systems bei einem bestimmten für Menschen konzipierten Fähigkeitstest auf eine zugrunde liegende KI-Kompetenz oder allgemeine Intelligenz zu schließen.

Es gibt verschiedene empirische Ansätze zur Beurteilung von KI-Fähigkeiten, die Studien sind aber punktuell und nicht systematisch.

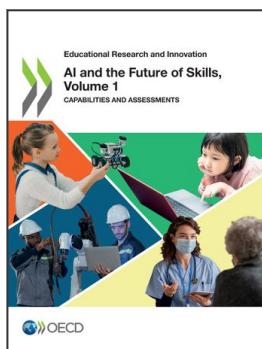
Es gibt eine Vielzahl von Benchmarks und Wettbewerben, um KI-Systeme empirisch zu beurteilen und zu vergleichen. Die Ergebnisse wurden bisher jedoch noch nicht systematisch klassifiziert. Immer mehr Institutionen führen strenge Evaluierungen durch, um die Fähigkeiten von KI und Robotiksystemen zu beurteilen. Einerseits werden einzelne Funktionen, d. h. separate Fähigkeiten wie Selbstlokalisierung, erfasst. Andererseits werden aber auch vollständige Aufgaben evaluiert, die eine sinnvolle Tätigkeit darstellen, wie autonomes Fahren oder Textzusammenfassung. In bestimmten Bereichen wie Sprachverstehen ist die Evaluierung von KI-Systemen besonders gut entwickelt. Insbesondere der Bereich maschinelles Übersetzen bietet viele richtungweisende Ansätze für die Beurteilung von KI.

Eine systematische Beurteilung von KI erfordert einen umfassenden Rahmen, der alle für die Arbeitswelt und das tägliche Leben notwendigen menschlichen Kompetenzen erfasst.

Belastbare, verlässliche und aussagekräftige Messgrößen für die Fähigkeiten von KI und Robotik setzen einen umfassenden Ansatz voraus, der verschiedene Forschungstraditionen und komplementäre Methoden zusammenführt. Dabei sollten die folgenden Aspekte berücksichtigt werden: das gesamte Spektrum relevanter menschlicher Fähigkeiten; die zusätzlichen Fähigkeiten, die speziell für die Prüfung von KI erforderlich sind (weil sie für KI schwierig sind und in den Listen menschlicher Kompetenzen häufig vernachlässigt werden) sowie das gesamte Spektrum von Aufgaben, die im Bildungsbereich, am Arbeitsplatz und im täglichen Leben wichtig sind.

Bei der Beurteilung von KI und Robotiksystemen ist es wichtig, mit den Grundprinzipien der Evaluierung vertraut zu sein und verschiedene Herangehensweisen einzubeziehen.

Ein multidisziplinärer Ansatz erfordert einen theoretischen Unterbau, der die Herausforderungen einer Beurteilung von KI und Robotik in Bezug auf menschliche Kompetenzen berücksichtigt. Die verschiedenen Herangehensweisen können in zwei Kategorien eingeordnet werden. Bei der ersten Kategorie geht es darum, ob Kompetenztaxonomien und -tests in erster Linie *menschliche oder KI-Fähigkeiten* messen. Die zweite Kategorie betrifft die Frage, ob sie *einzelne (isolierte) Fähigkeiten oder komplexe Aufgaben* messen, die eine Vielzahl von Fähigkeiten erfordern. Die künftige systematische Beurteilung von KI-Fähigkeiten sollte verschiedene Ansätze auf Basis dieser beiden Kategorien zusammenführen und ihr Potenzial zielführend integrieren, um belastbare Schlussfolgerungen über die Zukunft von Arbeit und Bildung ziehen zu können.



From:
AI and the Future of Skills, Volume 1
Capabilities and Assessments

Access the complete publication at:

<https://doi.org/10.1787/5ee71f34-en>

Please cite this chapter as:

OECD (2022), "Zusammenfassung", in *AI and the Future of Skills, Volume 1: Capabilities and Assessments*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/cf9db84c-de>

Das vorliegende Dokument wird unter der Verantwortung des Generalsekretärs der OECD veröffentlicht. Die darin zum Ausdruck gebrachten Meinungen und Argumente spiegeln nicht zwangsläufig die offizielle Einstellung der OECD-Mitgliedstaaten wider.

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area. Extracts from publications may be subject to additional disclaimers, which are set out in the complete version of the publication, available at the link provided.

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at <http://www.oecd.org/termsandconditions>.