

Stand der Gestaltungsempfehlungen zu KI

Dezember 2020

1. Einleitung

Im Projekt „Ethische und sozial verträgliche KI in Unternehmen“ der Universitäten Hohenheim und Stuttgart werden anhand ausgewählter Fallbeispiele Kriterien für eine erfolgreiche, ethische, sozial verträgliche und menschenzentrierte Gestaltung und Einführung von KI-Systemen im Mittelstand entwickelt. Die gewonnenen Erkenntnisse werden für KMU sowie weitere Interessierte zugänglich gemacht. Die ersten Handreichungen, veröffentlicht im September 2020, bieten einen Überblick über den aktuellen Stand des Begriffs KI, Hilfe bei der Standortbestimmung im eigenen Unternehmen zu diesem Thema und zum Stand der Empfehlungen zu Gestaltung von Systemen mit KI-Anteilen.

2. Welche Aspekte sind bei der Gestaltung von Systemen mit KI-Anteilen relevant?

Um dieser Frage näher zu kommen, hilft die Betrachtung des eigenen Unternehmens anhand des Modells MTO – Mensch, Technik, Organisation. Dies gilt sowohl für eine übergeordnete Betrachtung auf all diesen Ebenen und hinsichtlich ihrer Wechselwirkungen (vgl. hierzu ausführlicher Handreichung 03) als auch für

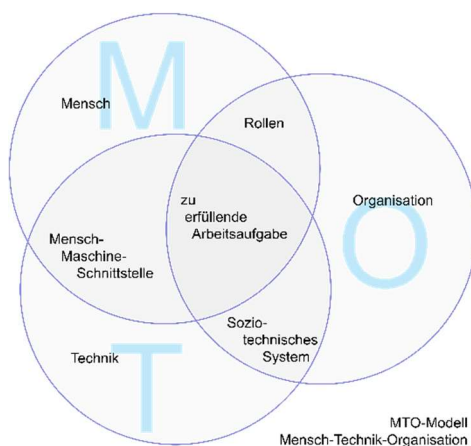


Abbildung 2: MTO Modell

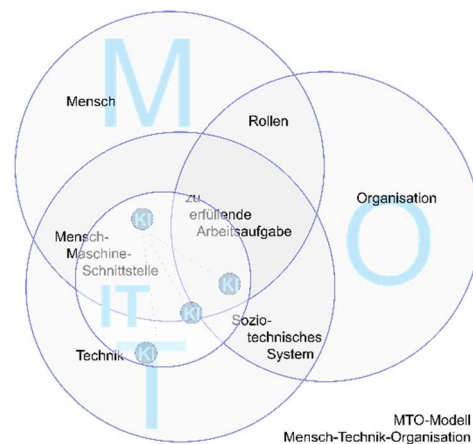


Abbildung 2: Positionierung der IT im MTO

einen Blick speziell auf die Technikgestaltung sowie unter der Bedingung, aus der Perspektive der Technologie auf alle Dimensionen zu blicken.

Denn wird das Modell um die Rolle, die KI und damit auch IT und Digitalisierung spielen, erweitert, so wird deutlich, dass der KI-Anteil am Gesamtsystem unterschiedlich ausgeprägt sein kann. Nicht nur der Umfang dessen, wie komplex und verschiedenartig die eingesetzten Algorithmen und Verfahren sein können, variiert, sondern auch die Orte im Unternehmen, an denen sie eingesetzt werden. Darüber hinaus unterscheiden sich Systeme mit KI-Anteilen darin, inwiefern sie miteinander verbunden sind und auch hinsichtlich des Anteils, der KI für die Anwender und Anwenderinnen tatsächlich sichtbar wird.

So kann eine KI-Anwendung für die Nutzenden als solche überhaupt nicht bzw. kaum erkennbar sein: In einem solchen Fall könnte die Rolle der enthaltenen KI als reiner Technikfortschritt zu sehen sein, der ein bestehendes Verfahren verbessert. Dieses Verfahren ist rein äußerlich nicht von einem regelbasierten System zu unterscheiden. Und trotzdem können Effekte des KI-Einsatzes an verschiedenen Stellen auftreten, weshalb eine umfassende Betrachtung und Technikfolgenabschätzung immer ratsam ist, um etwa auch nicht-intendierte Nebenfolgen nicht aus dem Blick zu verlieren.

Anders sieht es aus, wenn die KI-Anteile für die Nutzenden sichtbar werden, und sich womöglich auch das Verhalten des Systems mit der Zeit verändert. In einem solchen Fall wird auch von der Rolle des „human in the loop“ gesprochen, wo sich das menschliche Verhalten und das Vorhandensein von Datenreihen auf das Verhalten des Systems auswirkt und das System sich selbst optimiert, sei es in automatisch angestoßenen Lernzyklen oder in von Menschen begleiteten Prozessen.

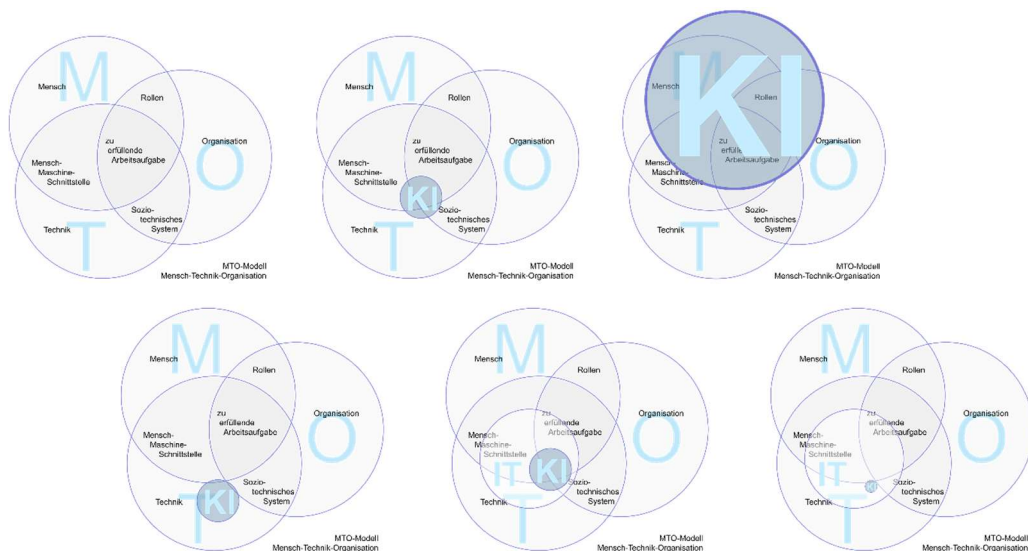


Abbildung 3: Unterschiedliche Rollen von KI in Unternehmen

Abbildung 3 zeigt unterschiedliche Positionierungen von KI im MTO-Modell. Während das hierin enthaltene zweite Szenario oben mittig eine langfristig realistische und gut eingebettet erscheinende Positionierung von KI-Anwendungen zeigt, dreht sich die breite gesellschaftliche Debatte im Augenblick meistens um deutlich weniger realistische Szenarien, wie oben rechts mit einer starken KI-Überlagerung

visualisiert. Dort wird KI als alles, vor allem auch den Menschen bestimmend diskutiert, ohne darauf zu achten, dass solche Anwendungen hochgradig unterschiedlich sind und damit auch ihre Auswirkungen auf Unternehmen, Menschen und Arbeit stark variieren. Links unten zeigt ein realistisches Bild für die nahe Zukunft, d. h. etwa für die nächsten zehn Jahre, in dem KI-Anwendungen zunächst auf Seiten der Technik die größte Rolle spielen. Im fünften Szenario, Mitte unten, sind die aktuelle Fachdiskussion rund um das Thema KI und dessen Einbettung in die vorhandene IT der Unternehmen zu sehen und im sechsten, unten rechts, ist schließlich der aktuelle Stand von KI-Anwendungen in Unternehmen veranschaulicht.

3. Warum ein menschenzentrierter Entwicklungsprozess auch bei KI-Projekten wichtig ist

In der Realität von KI-Projekten liegen bei verschiedenen Stakeholdern unterschiedliche Motivationen vor, sich mit dem Thema zu beschäftigen: Von wirtschaftlichen Abwägungen über Technikbegeisterung bis Neugier gibt es eher technologie- oder datengetriebene Ansätze, bis hin zu aufgaben- oder prozessorientierten Ansätzen und schließlich noch den menschenzentrierten Entwicklungsprozess. Klassischerweise ist der menschenzentrierte Entwicklungsprozess hinsichtlich der Wahl einer Technologie ergebnisoffen und stellt die Bedürfnisse und das Erleben der Nutzer und Nutzerinnen in den Mittelpunkt. Kennzeichnend für dieses Vorgehen ist, dass so lange weitere Prototypen und Iterationen erstellt werden, bis das Ergebnis den Anforderungen der Nutzer und Nutzerinnen entspricht. Doch verschiedene Motivationen müssen sich überhaupt nicht widersprechen und ein Entweder-Oder darstellen. Vielmehr sind sie als Sowohl-als-Auch zu denken und miteinander zu verschränken, um eine bestmögliche Entwicklung, Anpassung der (Arbeits-)Organisation, Akzeptanz und Nutzung zu erreichen, die im Sinne aller Stakeholder ist. Hiermit befasst sich das vorliegende Forschungsprojekt.

Eine Besonderheit des Einsatzes von Anwendungen mit KI-Komponenten ist, dass sich durch die dafür oft notwendige Verbesserung der Datengrundlage ganz neue Anwendungsmöglichkeiten ergeben. Welche Daten oder Auswertungen relevant werden, ergibt sich manchmal erst mit der Erfahrung. D. h. die neuen Möglichkeiten wecken einen weiteren Bedarf.

4. Empfehlungen zur Gestaltung von KI und Prozessen rund um KI

Bislang vorliegende Empfehlungen sind primär zu unterscheiden hinsichtlich Empfehlungen unternehmensorganisatorischer Art und auf ethischer Ebene (a) sowie hinsichtlich zur konkreten technischen Ausgestaltung des Funktionsumfangs und der Schnittstelle zu den Nutzenden des Systems (b). Beide betreffen sowohl die die Arbeits- als auch die Technikgestaltung und sind wichtig für das Gesamtergebnis.

Allerdings ist diese Unterscheidung nicht trennscharf: dass zum Beispiel beim Maschinellen Lernen die Datenbasis eine Rolle spielt – beispielsweise in Bezug auf die Berücksichtigung aller Menschen, Fairness, Erklärbarkeit –, ist sowohl eine ethische Fragestellung als auch eine funktionale. Denn häufig wird trotz zahlreicher Lehrstunden immer noch unterschätzt, dass KI-Software Entscheidungen vorschlägt oder Schlüsse zieht, die „unfair“ oder „biased“ sind. Doch solche Systeme sind entgegen des Mythos vermeintlich objektiver Technik eben nicht ohne Weiteres frei von höchst subjektiven Einschlägen bzw. eben einem bias (vgl. hierzu ausführlicher Handreichung 03).

(a) Hierzu hat etwa die Arbeitsgruppe Arbeit/Qualifikation, Mensch-Maschine-Interaktion der Plattform Lernende Systeme in einem Whitepaper zwölf Kriterien für eine gute Zusammenarbeit von KI und Beschäftigten formuliert. Diese sind in vier Cluster zusammengefasst: „Schutz des Einzelnen vor negativen Folgen des KI-Einsatzes“, „Vertrauenswürdigkeit der KI-Systeme“, „sinnvolle Arbeitsteilung zwischen Beschäftigten und KI-System“ und „förderliche Arbeitsbedingungen für die Beschäftigten“ (Huchler et al. 2020). Eine darüber hinaus gehende, konkreter auf Arbeitsgestaltungsbedingungen hin ausgerichtete Systematisierung bieten Zirnic et al. (2021) an, die im Rahmen des vorliegenden Forschungsprojektes entwickelt und empirisch untersucht wird. Hierbei werden insbesondere auch Aspekte der Arbeitsorganisation und der spezifisch notwendigen Arbeitsgestaltung (b) mitberücksichtigt, da nur ein ganzheitlicher Ansatz zielführend ist.

Auch von namhaften Technologiekonzernen (z. B. Google, Microsoft) existieren Leitfäden, die sich jedoch primär an Entwicklungsteams von KI richten und Empfehlungen geben zur Gestaltung der Rolle der „humans in the loop“, also der Rolle, die die Nutzenden in einem sich selbst regelnden System spielen. Es wird dabei als Herausforderung angesehen, zu erklären, wie ein KI-System genau funktioniert, also Transparenz dahingehend zu schaffen. Dabei ist es jedoch sehr wichtig, dass Nutzende Wissen hierüber haben, also eine explainable AI entwickelt wird.

(b) In anderen Publikation wie bei Schröder/Franz (2019) wird beschrieben, wie wichtig ein zielorientierter und ehrlicher Aushandlungsprozess auf der betrieblichen Ebene ist, was den Zweck des Einsatzes von KI anbelangt und wie bei solchen Systemen, die große zeitliche Vorteile bringen oder neue Funktionen möglich machen, die Aufgabenteilung zwischen Mitarbeitenden und System ist – d. h. es müssen zum Beispiel Fragen geklärt werden wie: „Haben Mitarbeitende dadurch mehr Zeit für einzelne Kunden und Kundinnen oder müssen sie sich in Zukunft mehr Fälle kümmern?“ Profund betrachtet werden solche Aspekte ebenfalls bei Zirnic et al. (2021), mit einem Fokus auf Dienstleistungsarbeit in Anbetracht von KI.

Die Europäische Union hat eine High Level Expert Group on AI (HLEG) berufen, die neben einer Definition von KI Handreichungen zu Mensch-zentrierter KI und Ethischen Richtlinien (2019) ausgibt und anhand deren Leitlinien sich auch das vorliegende Projekt orientiert. Wichtigster Faktor aus Sicht der HLEG ist es, einen Rah-

men für vertrauenswürdige KI zu schaffen. Dazu müssen die Faktoren „Rechtmäßige KI“, „Ethische KI“ und „Robuste KI“ detailliert betrachtet werden. Der Aspekt der Ethischen KI wird noch einmal detaillierter beschrieben und enthält als Fundament vier ethische Grundsätze (Achtung der menschlichen Autonomie, Schadensverhütung, Fairness, Erklärbarkeit). An die Umsetzung von KI werden sieben Kernanforderungen gestellt und zur Überprüfung eine Checkliste angeboten. Die Überführung dieser eher noch etwas abstrakten Leitlinien in die Praxis verbunden mit arbeitsorganisatorischen Betrachtungen und Empfehlungen ist Gegenstand des vorliegenden Projekts.

Forschende der AI Ethics Impact Group um Hallensleben/Hustedt (2020) schlagen zudem etwa ein Kennzeichnungssystem für KI-Anwendungen vor, das sich an den Energie-Effizienz-Labels orientiert. Als Grundwerte legen sie Gerechtigkeit, Umweltfreundlichkeit (ökologische Nachhaltigkeit), Verantwortlichkeit (Verantwortung), Transparenz, Privatsphäre und Zuverlässigkeit fest. Ihr Ziel ist es, ein praxistaugliches Framework zur Verfügung zu stellen, das es verschiedenen Stakeholder und Stakeholderinnen ermöglicht, ethische Werte in ein KI-System zu integrieren. Um zur Bewertung eines KI-Systems zu kommen, werden auf der einen Seite beobachtbare Faktoren und Indikatoren so aggregiert, dass Charakteristiken gebildet werden, die aufzeigen, in wie weit das System einen Wert (z. B. Transparenz) erfüllt. Auf der anderen Seite wird der Applikationskontext klassifiziert, um so die Minimalanforderungen zu definieren, die das KI-System im spezifischen Kontext erfüllen muss. Dieses Modell kann bei der Entwicklung einer KI-Anwendung ebenfalls hilfreich sein, um die manchmal aus der konkreten Anwendungssituation schwer zu erfassenden ethischen Fragestellungen zu adressieren.

5. Besondere Situation in KMU

In unserem Projekt wollen wir die spezifischen Herausforderungen von KMU bei der Einführung von KI kennenlernen und analysieren, um branchenübergreifende Handlungsempfehlungen ableiten zu können. Die Herausforderungen für kleine und mittelständische Unternehmen, KI-Systeme einzuführen, sind heterogener und größer aufgrund der unterschiedlichen Gegebenheiten der Unternehmen: Von Gestaltungskompetenz bei Mitarbeitenden direkt im Unternehmen bis zur Kooperation mit externen Anbietern kann alles vorkommen. Je nach Größe des Unternehmens kann es im Extremfall vorkommen, dass verschiedene Rollen von Stakeholdern so auch in einer Person zusammenkommen, zum Beispiel Geschäftsführung und End-Anwendende oder dass Funktionen wie Betriebsrat und Datenschutz gar nicht vorhanden sind. Neben einem Bewusstsein für die Notwendigkeit der Gestaltung ist dabei auch die Verfügbarkeit von passenden Daten eine wichtige Voraussetzung für KI-Anwendungen in KMU.

Die Notwendigkeit der Gestaltung speziell in Bezug auf KI hängt wie oben beschrieben von der Sichtbarkeit der KI-Anteile, deren Komplexität, und der Art der Interaktion mit den End-Anwendenden ab (für weitere Details zur Situation in KMU siehe auch Handreichung 02).

6. Resümee

Es gibt bereits diverse Publikationen zu Themen rund um KI, deren Gestaltung sowie z. T. auch zu (arbeits-)organisatorischen Flankierungen. Sie sind häufig jedoch noch sehr abstrakt und kaum auf Besonderheiten der Unternehmenslandschaft hierzulande zugeschnitten. Kleine und mittelständisch Unternehmen, die die Kapazität haben, eigene Expertise aufbauen zu können, setzen derzeit allerdings bereits erste KI-Funktionen um. Komplexe, zusammenhängende Anwendungen, deren User Interface aufwendig gestaltet wird, sind das jedoch meist noch nicht. Und auch wird die Perspektive der Nutzer und Nutzerinnen meist nicht hinreichend betrachtet, die jedoch erfolgskritisch ist. Deshalb ist es wichtig, Beschäftigte und ihre Vorstellung von KI und damit verbundenen Einführungs- und Arbeitsprozessen bereits in die Entwicklung sowie letztlich Darstellung und Beschreibung der Systeme und deren erwünschtem Output mit ein zu binden und die Systeme so zu gestalten, dass eine realistische Vorstellung der Funktionsweise der Komponenten mit KI-Anteil entsteht. Die richtigen Erwartungen zu wecken, ist wichtig, damit aus den vorhandenen Informationen keine falschen Schlüsse gezogen oder Systeme an den Erfordernissen der Praxis vorbei entwickelt werden.

7. Literatur

- Hallensleben, Sebastian; Hustedt, Carla (2020): From Principles to Practice: An interdisciplinary framework to operationalise AI ethics. Bertelsmann Stiftung. URL: [Internetseite Bertelsmann Stiftung](#) (Stand: 21.07.2020).
- Huchler, Norbert; Adolph, Lars; André, Elisabeth; Bauer, Wilhelm; Bender, Nadine; Müller, Nadine; Neuburger, Rahild; Peissner, Matthias; Steil, Jochen und Stowasser, Sascha (2020): Kriterien für die menschengerechte Gestaltung der Mensch-Maschine-Interaktion bei Lernenden Systemen – Whitepaper der Plattform Lernende Systeme. München: Lernende Systeme – Die Plattform für Künstliche Intelligenz.
- Schröder, Lothar; Franz, Markus (2019): Eine warme Stimme schleicht sich in Dein Ohr. Fluch und Segen von Künstlicher Intelligenz. Gewerkschaftliche Antworten. Hamburg: VSA.
- Smuha, Nathalie (2018): Ethik-Leitlinien für eine vertrauenswürdige KI, Brüssel: Hocharangige Expertengruppe für künstliche Intelligenz, Europäische Kommission. Link: [Internetseite Europäische Kommission](#).
- Zirnic, Christopher; Jungtäubl, Marc; Ruiner, Caroline (2021 i.E.): Menschengerechte Gestaltung von KI bei Dienstleistungsarbeit. In: Bruhn, Manfred; Hadwich, Karsten (Hrsg.): Künstliche Intelligenz im Dienstleistungsmanagement. Forum Dienstleistungsmanagement.

Ihre Ansprechpartnerinnen und -ansprechpartner

Jj Link

Jasmin.Link@iao.fraunhofer.de

Patrick Stern

Patrick.Stern@iat.uni-stuttgart.de

Projekt Homepage:

[Internetseite Ministerium für Wirtschaft, Arbeit und Wohnungsbau Baden-Württemberg](#)