

# Ethische und soziologische Aspekte der Mensch-Roboter-Interaktion



baua: Bericht kompakt

**Immer häufiger arbeiten Menschen unmittelbar mit Robotern zusammen. Neben technischen wirft diese neue Form der Zusammenarbeit auch ethische und soziologische Fragen auf. Der Bericht der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA) „Ethische und soziologische Aspekte der Mensch-Roboter-Interaktion“ (MRI) geht diesen systematisch nach. Neben der Rolle des Menschen und dem Autonomiegrad des Roboters thematisiert der Bericht auch die äußere Gestaltung von Robotern unter ethischen und soziologischen Gesichtspunkten.**

In vielen Anwendungsbereichen nimmt der Einsatz von Robotern zu, die direkt mit Menschen zusammenarbeiten. Ob bei medizinischen Eingriffen in der Chirurgie, im Krankenhausflur bei der Begegnung mit einem fahrerlosen Transportsystem oder im Haushalt, wo der Staubsauger-Roboter im Einsatz ist – viele Lebensbereiche sind von der Mensch-Roboter-Interaktion beeinflusst. Auch in der Produktion wird diese Technologie immer stärker genutzt. Mit ihrer Hilfe lassen sich sowohl Produktivität und Effizienz verbessern als auch Aufgaben bewältigen, die ohne diese Kombination aus menschlicher Arbeit und Robotereinsatz nicht lösbar wären.

Dabei sollte laut BAuA-Bericht Ziel sein, die Tätigkeit der Mitarbeitenden mit dem robotischen System möglichst optimal im Sinne einer menschengerechten Gestaltung der Arbeit zu kombinieren. Für dieses Gestaltungsziel sind nicht nur technische Aspekte zu berücksichtigen. Ebenso wichtig ist auch die Frage nach der Qualität des neuen Arbeitsumfeldes sowie der sich darin ergebenden Tätigkeiten aus der menschlichen Perspektive. Eben diese ethischen und sozialen Blickwinkel auf MRI nimmt der Bericht ein, der die Ergebnisse eines Forschungsprojekts der BAuA vorstellt.

## Das Projekt im Überblick

„Ethische und soziologische Aspekte der Mensch-Roboter-Interaktion“ – so der Titel des Projekts, das Human-Factors-Consult (HFC) zusammen mit dem Institut für Philosophie, Literatur-, Wissenschafts- und Technikgeschichte der TU Berlin im Auftrag der BAuA durchgeführt hat. Ziel war es, ethische und soziale Aspekte der MRI im industriellen Kontext herauszuarbeiten. Dazu erfolgte in einem ersten Projektschritt eine umfassende Literaturrecherche zur MRI im ethisch-moralischen Diskurs. Im Fokus standen dabei Fragen des ethisch-moralischen Status autono-

mer Entscheidungen von Robotern, der Verantwortung in der MRI und der Technikfolgen der MRI. Darauf aufbauend wurden im zweiten Arbeitsschritt kritische Aspekte der MRI aus ethisch-sozialer Perspektive ermittelt und deren Bedeutsamkeit auf der Grundlage von publizierten experimentellen Studien untersucht. Um herauszufinden, welche Variablen die Zusammenarbeit von Mensch und Roboter beeinflussen, wurden wissenschaftliche Studien ausgewählt, die Relevanz für diese kritischen Aspekte haben, und nach einer etablierten Taxonomie klassifiziert. Dabei konnten drei Einflussgrößen auf ethisch-moralische Aspekte der MRI ermittelt werden, die wesentlich die gefühlsmäßige Einstellung und Bewertung der MRI, die wahrgenommene Verantwortung in der Aufgabenbearbeitung und die Akzeptanz des „Kollegen Roboters“ in unterschiedlichen Formen der Zusammenarbeit beeinflussen: die Rolle des Menschen, der Autonomiegrad und die Morphologie (also äußere Gestaltung) des Roboters. Der Bericht formuliert abschließend auf Basis der Projektergebnisse einige Gestaltungshinweise für die industrielle MRI nach ethisch-sozialen Kriterien.

## MRI im ethisch-moralischen Diskurs

Derzeit besteht das Ziel ethischer Beiträge zur Robotik vorrangig darin, für ethische Problemstellungen zu sensibilisieren und das Feld der Handlungsoptionen abzustecken. Die wertebasierte, normative Bewertung technologischer Entwicklungen sowie die Entscheidung für oder gegen bestimmte technische Anwendungen liegt letztlich bei den Verantwortlichen in Technik, Politik und Gesellschaft. Die Aufgabe von Ethikfachleuten besteht damit darin, die konzeptionellen Grundlagen zu erarbeiten, auf denen möglichst sachgerechte und ethisch verantwortungsvolle Entscheidungen getroffen werden können. Zudem liefern sie wichtige Diskussionsanregungen.

Diskutiert werden beispielsweise ethisch-moralische Fragen und Probleme, die sich infolge der Autonomie von Robotern stellen. Dabei können drei Problembereiche ausgemacht werden:

1. die Substitution oder Verbesserung ethisch relevanter Entscheidungen durch Roboter
2. die Eröffnung ethisch relevanter Entscheidungsspielräume durch Roboter
3. die Idee einer ethischen Reflexionsfähigkeit von Robotern.

#### Ethisch-moralische Entscheidungen von Robotern

Beispielhafte Einsatzbereiche der Robotik, für die entsprechende Fragen diskutiert werden, sind etwa Militär, Medizin und Robotik in gesellschaftlichen Zusammenhängen. In diesen Bereichen müssen Entscheidungen mit z. T. dramatischer ethischer Reichweite getroffen werden. Während Menschen solche Entscheidungen mit großer Selbstverständlichkeit treffen, müssen Robotern entsprechende Entscheidungsfähigkeiten erst „beigebracht“ werden. Damit verbunden sind grundsätzliche Fragen: Welche ethischen Prinzipien autonomer Entscheidungen können technisch realisiert werden? Sollten Roboter überhaupt genauso entscheiden, wie Menschen an ihrer Stelle entscheiden würden, oder müssten sie nicht „besser“ entscheiden? Wie können solche Entscheidungen unsere moralischen und ethischen Ansprüche erfüllen? Und welcher der in Kontrast stehenden ethischen Theorien (z.B. Utilitarismus) sollte ein Roboter folgen? Bislang gibt es keine allgemein anerkannte Lösung einer „programmierten Ethik“; sie müsste jedoch in jedem Fall auf einer gesellschaftlichen Konsensbildung basieren.

Auch die Öffnung von neuen Entscheidungsräumen ist mit Fragen verbunden: Wie geht man mit ethischen Dilemmata um, wenn durch die im Voraus geschehende Programmierung von Robotern genügend Zeit und Steuerungsmöglichkeiten bestehen, um bewusste, reflektierte ethische Entscheidungen zu treffen. Das viel zitierte Trolley-Dilemma – ein Fahrer, der die Wahl hat zwischen dem Unfalltod mehrerer Fußgänger oder der Selbsttötung durch Ausweichen – bekommt eine neue Qualität, wenn eine solche Entscheidung bei autonom fahrenden Fahrzeugen im Vorhinein programmiert werden könnte. Wenngleich in diesem Fall eine präventive, sicherheitstechnische Lösung, die das Entstehen einer solchen Situation unmöglich macht, die Prämisse sein muss, zeigt er doch auf, wie neue technische Möglichkeiten wie die Vorbereitung von Entscheidungen oder die Verfügbarkeit großer Datenmengen neue ethische Fragestellungen mit praktischer Relevanz aufwerfen können. Damit erzeugen Technologien autonomer Steuerungen neue Ansprüche an eine ethisch verantwortungsvolle Gestaltung.

Der dritte Bereich schließlich – die Idee einer ethischen Reflexionsfähigkeit von Robotern – spielt praktisch (noch) keine Rolle. Sie geht über die reine Programmierung moralischer Regeln hinaus, setzt allerdings damit starke Künstliche Intelligenz (KI) voraus, die zeitnah nicht zur Verfügung steht. Es bleibt die Frage, ob ein eigenes Reflexionsvermögen umsetzbar, benötigt oder gewünscht wäre. Der Einsatz technischer Systeme mit autonomen Entscheidungsfähigkeiten ist nach Ansicht der Autoren in solchen Bereichen ethisch problematisch, die auch sonst von menschlichen, ethisch-moralischen Konflikten geprägt sind. Technisch realisierte Entscheidungen mit ethisch-moralischer Dimension sind danach nur dort geboten, wo ein Konsens über die ethische Richtigkeit der Entscheidung existiert. Der Entscheidungsspielraum für autonome Systeme im Bereich der Produktion muss daher die etablierten Anforderungen an einen sicheren und gesunden Betrieb erfüllen, so dass sich alle autonomen Entscheidungen an solchen unumstrittenen ethischen Standards des Menschen orientieren können und müssen.

#### Die Zuordnung von Verantwortung für Roboter-Handlungen

Die im menschlichen Falle enge und geregelte Verknüpfung einer Handlung mit der entsprechenden Verantwortung wird bei Roboter-Handlungen gelöst. Damit stellt sich sowohl für den rechtlichen wie auch den ethisch-moralischen Bereich das identische Problem: Roboter handeln zwar, sind aber keine Handlungssubjekte im üblichen Sinne – weder bei Haftungsfragen noch bei Fragen von Schuld. Wenn Roboter also Schäden verursachen, dann sind sie – wie andere technische Gegenstände – keine Teilnehmenden der menschlichen ethisch-moralischen Praxis. Roboter tragen insofern keine moralische Schuld. Während rechtliche Verantwortung im Sinne von Haftung den entsprechenden beteiligten Personen zuzuordnen ist, könnte unter bestimmten Umständen die Zuschreibung von Verantwortung schwierig werden. Autonome Roboter könnten in dieser Hinsicht einen ähnlichen Status wie Tiere haben.

#### Kritische Aspekte der MRI aus ethisch-sozialer Perspektive

Mensch-Roboter-Interaktionen erscheinen in vielerlei Hinsicht als vorteilhaftes Szenario für die Zukunft der Arbeit. Vom robotischen System erwartet man dabei keine Ersetzung des Menschen, sondern vielmehr hilfreiche Mitarbeit. Damit werden Produktivitäts- und Effizienzsteigerungen möglich oder auch die belastungsoptimierte Gestaltung von Arbeit. Die MRI eröffnet durch Kooperation oder Kollaboration neue Möglichkeiten für den Einsatz von Robotern und auch neue Gestaltungsoptionen für die Arbeit des Menschen. Mit der MRI werden folglich neue Arbeitsfelder für die Teilautomatisierung erschlossen. Das kann auch dazu führen, dass sich neue Arbeitsbeziehungen entwickeln: Der Mensch nutzt den Roboter unter

Umständen nicht wie andere Maschinen, sondern tritt in eine komplexere Arbeitsbeziehung mit ihm ein, spätestens dann, wenn der Roboter Entscheidungsfunktionen übernehmen kann. Wie diese ethisch und sozial zu bewerten ist, ist eine zentrale Frage für die Gestaltung von MRI, da eingespielte Praktiken und Werte der Teamarbeit auf die enge Zusammenarbeit von Mensch und Roboter zumindest übertragen oder dafür sogar neu gedacht werden müssen.

Für die Persönlichkeitsentwicklung des Menschen bei der Arbeit hat die Qualität sozialer Beziehungen entscheidende Bedeutung. Auch im Kontext der Einführung neuer Technologien wie kollaborierender Roboter muss dieser Aspekt betrachtet werden: Hier stellen sich ganz neue Fragen, wenn gezielt nicht mehr (nur) Personen, sondern Roboter als Interaktionspartner des Menschen in Betracht kommen. Angesprochen sind dabei z. B. Aspekte wie die Selbstbestimmung des Menschen, der Ausdruck der Wertschätzung seiner Arbeitskraft und damit das sinnstiftende Potenzial der Arbeit. Probleme entstehen z. B. dann, wenn eine konkrete Ausgestaltung der MRI die Persönlichkeit des Menschen einschränkt oder nicht angemessen fördert.

Daneben sind solche Aspekte der MRI bedeutsam, die bestimmte Einstellungen des Menschen in Bezug auf den Roboter beeinflussen. So nehmen z. B. morphologische Gestaltungen und Metaphern Einfluss auf den Umgang mit Robotern. Hierbei wird ein ethischer Grundkonflikt von Mensch-Roboter-Interaktionen sichtbar: MRI kann mit Zuschreibungen typisch menschlicher Eigenschaften zu Robotern einher gehen, wie z. B. Absichten oder Entscheidungen, während Roboter rational betrachtet „nur“ Maschinen bzw. komplexe Werkzeuge sind, die unter Umständen mit Autonomie und einer gewissen „Lernfähigkeit“ ausgestattet sind, aber keine Prozesse der menschlichen Willensbildung und absichtsvollen Handlung vollziehen. Probleme entstehen also vorrangig dann, wenn die Grenze zwischen Mensch und Roboter im Rahmen der Zusammenarbeit auf eine Weise verschwimmt, die Auswirkungen auf menschliche Handlungen und die Qualität des Zusammenlebens hat. In einem weiteren Teil des Projektes wurde dazu die Fachliteratur gesichtet und systematisch analysiert.

#### Analyse von MRI-Studien in Bezug auf ethische und soziale Aspekte

Ziel der Analyse war es, konkrete Gestaltungsmerkmale der Zusammenarbeit von Mensch und Roboter zu identifizieren, die ethisch-soziale Aspekte der MRI sowohl positiv als auch negativ beeinflussen können. Hierzu wurden 14 empirische Studien zum Thema MRI identifiziert (siehe vollständiger Bericht unter weiterführende Informationen), die ethisch relevante Aspekte beleuchten. Die darin

untersuchten Wirkgrößen wurden anhand der Taxonomie von Onnasch und Kollegen (2016, siehe weiterführende Informationen) klassifiziert. Hieraus ergaben sich drei relevante Einflussgrößen auf ethisch-soziale Aspekte der MRI: die Interaktionsrolle des Menschen, der Autonomiegrad des Roboters sowie die Morphologie des Roboters in Bezug auf Kontext, Gestalt, Kommunikation und Bewegung. Für die Wahl der Interaktionsrolle gilt: Je stärker der Mensch in einer gemeinsamen Aufgabe die Interaktion führt, desto höher ist auch das Verantwortungsgefühl für die Gesamtaufgabe von Roboter und Mensch. Gleichzeitig wird festgestellt, dass Menschen Roboter im Arbeitskontext sowohl in gleichgestellter als auch in führender Funktion akzeptieren können.

Zudem weist auch der Autonomiegrad des Roboters einen wesentlichen Einfluss auf die wahrgenommene Verantwortung des Menschen für die Gesamtaufgabe auf. Positiv zu bewerten ist, dass eine hohe Roboterautonomie zu einer erhöhten Akzeptanz führt. Allerdings zeigen sich auch Hinweise auf zunehmende Manipulationsmöglichkeiten, d. h. auf eine unmerklichen Beeinflussung des Menschen durch einen hohen Autonomiegrad.

Hinsichtlich der Gestaltmorphologie weisen die Studien darauf hin, dass der Grad des Anthropomorphismus des Roboters, also der Menschenähnlichkeit, im Arbeitskontext möglichst gering ausgeprägt sein sollte. Andernfalls erhöht sich die Wahrscheinlichkeit einer nichtintendierten affektiven Bindung und somit auch das Manipulationspotenzial stark. Studien zeigen, dass mit einem steigenden Ähnlichkeitsgrad des Roboters zum Menschen sowohl die Empathie als auch die Kooperationsbereitschaft zunehmen. Diese Bereitschaft geht sogar so weit, dass Menschen unter Umständen zu Gunsten des Roboters auf ihren eigenen Maximalnutzen verzichten. Deshalb sollte mit einer anthropomorphen Gestaltung der äußeren Form des Roboters im Arbeitskontext ausgesprochen vorsichtig umgegangen werden und sie sollte nach den Erfordernissen der gemeinsam auszuführenden Aufgabe so gering wie möglich dosiert werden. Das ist auch mit Blick auf die wahrgenommene Aufgabenverantwortung des Menschen in der MRI ratsam, da mit steigendem Anthropomorphismus die wahrgenommene Verantwortung des Menschen für die Gesamtaufgabe abnimmt.

Ähnliches gilt auch für die Kontextmorphologie. Hierzu fand sich in der Studienrecherche eine Studie, deren Befunde jedoch eine eindeutige Sprache sprechen. Es wurde der Einfluss von framing („in einen Kontext setzen“) auf die Bereitschaft von Probanden untersucht, einen kleinen Roboter mit einem Hammer zu zerstören. Vor der Aufforderung, den Roboter mit dem Hammer zu zerschlagen, bekamen die Probanden entweder keine weiteren Informationen zum Roboter, eine personifizierte Geschichte zum

Roboter (Name und Eigenschaften), oder eine Geschichte über Erfahrungen des Roboters, die dieser bereits gemacht hatte. Es zeigte sich, dass Probanden sich erheblich länger weigerten bzw. den Roboter gar nicht zerstörten, wenn dieser eine Hintergrundgeschichte besaß. Dies führt zu der Annahme, dass je reicher der Kontext ausgestaltet wird, desto stärker wird die (affektive) Bindung von Menschen gegenüber Robotern. Dies kann in bestimmten Kontexten wünschenswert sein, z. B. beim Einsatz von Robotern in der Therapie. Im konventionellen Arbeitskontext ist Kontextmorphologie aber mit Vorsicht einzusetzen. Einerseits bietet framing ein hohes Potenzial um z. B. Berührungängste abzubauen, andererseits ist eine affektive Bindung an Roboter aus ethischer Perspektive u. U. kritisch zu bewerten, da sich daraus ein gewisses unerwünschtes Manipulationspotenzial beim Menschen ergibt.

Ein positives Bild zeichnen die Befunde zur anthropomorphen Bewegungsgestaltung des Roboters. Diese führt zu einer verbesserten Koordinationsleistung zwischen Mensch und Roboter, während sich nach der aktuellen Studienlage keine Hinweise auf mögliche negative Konsequenzen zeigen.

#### Empfehlungen für eine ethisch-sozial angemessenen MRI

Die Gestaltung, Auswahl und der Einsatz eines kollaborativen Roboters erfordert folglich wesentlich mehr als die weit verbreitete Betrachtung der rein technischen Funktionen und Leistungen. Die nachfolgenden aufgeführten Empfehlungen sollten bei der Planung des Einsatzes frühzeitig berücksichtigt werden.

Nach derzeitigem Kenntnisstand – der bislang noch auf einer sehr begrenzten Anzahl von Studien basiert – lassen sich die folgenden Empfehlungen ableiten:

**Autonomie und Transparenz:** Mit steigender Autonomie des Roboters sollte die Transparenz bzgl. seiner Funktionsweise und Entscheidungsgrundlagen ein wesentliches Gestaltungsmerkmal darstellen. Transparenz wirkt möglicher Verantwortungsdiffusion entgegen und kann eine Beeinflussbarkeit des Menschen abmindern.

**Die Interaktionsrolle in Abhängigkeit der Aufgabenallokation:** Die Aufgabenteilung zwischen Mensch und Roboter sollte für die jeweiligen Fähigkeiten und Kompetenzen der Beschäftigten angemessen sein. Einflussspielräume für den Menschen über die technische Steuerung von Arbeitsprozessen erlaubt es, den menschlichen Überblick

über Prozesse einzubeziehen. In Bezug auf den Roboter sollten Begrifflichkeiten wie „Führungsrolle“ vermieden werden, um keine falschen Erwartungen hinsichtlich sozialer Kompetenzen und Verantwortungsfähigkeit zu wecken.

**Gestaltmorphologie:** Um Verantwortungsdiffusion und unangemessene emotionale Bindungen zu vermeiden, sollte eine anthropo- oder zoomorphe Robotergestaltung im Arbeitskontext entsprechend der Erfordernisse möglichst minimiert werden.

**Bewegungsmorphologie:** Eine anthropomorphe Gestaltung von Bewegungsmustern ist hingegen grundsätzlich zu empfehlen, da dies die wahrgenommene Intentionalität und resultierend die Koordinationsleistung von Mensch und Roboter erhöht.

**Kommunikationsmorphologie:** Um eine Übertragung menschlicher Stereotype zu vermeiden, ist eine nonverbale Kommunikation zu bevorzugen, die keine unerfüllbaren Erwartungen an Roboterfähigkeiten weckt. Implizite nonverbale Kommunikation könnte sich zusätzlich günstig auswirken, da sie keinen kognitiven Mehraufwand erfordert.

**Kontextmorphologie:** Künstlich kontextbildende Maßnahmen sind in industriellen Anwendungsbereichen nur mit Vorsicht einzusetzen. Es sollte darauf geachtet werden, keine ungewünschte affektive Bindung zu erzeugen, gerade in Kombination mit humanoider oder zoomorpher Ausgestaltung des Roboters.

#### Weiterführende Informationen

1 Onnasch L., Jürgensohn T., Remmers P., Asmuth, C.: Ethische und soziologische Aspekte der Mensch-Roboter-Interaktion. 1. Auflage. Dortmund: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin 2019. Seiten 73, Projektnummer: F 2369, PDF-Datei, DOI: 10.21934/baur:bericht20190128.

2 Onnasch L., Maier X., Jürgensohn, T.: Mensch-Roboter-Interaktion - Eine Taxonomie für alle Anwendungsfälle. 1. Auflage. Dortmund: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin 2016. (baur: Fokus) Seiten 12, Projektnummer: F 2369, PDF-Datei, DOI: 10.21934/baur:fokus20160630.