
MISSELHORN, CATRIN:

Künstliche Intelligenz und Empathie.

Vom Leben mit Emotionserkennung, Sexrobotern & Co

Stuttgart: Reclam 2021.

ISBN 978-3-15-014047-5; 150 S.; 6,00 €



Rezension von Dietrich PUKAS, Bad Nenndorf

Emotionale Künstliche Intelligenz betrachtet Catrin Misselhorn, Philosophie-Professorin und Spezialistin für Maschinen-Ethik an der Universität Göttingen, als Schlüsseltechnologie der Zukunft, insofern artifizielle Systeme mitfühlend sein und Empathie in uns Menschen auslösen sollen. In dem Band der handlichen Reclam-Reihe vermittelt sie uns einen gut nachvollziehbaren Überblick über die komplexe Thematik der Emotionen und Empathie, neuen Tendenzen der Künstlichen Intelligenz, der sozialen Robotik und Biorobotik. Anhand von praktischen Beispielen erläutert sie die technischen und ethischen Aspekte der Problematik und diskutiert Fragen, wie **synthetische Systeme** menschliche Emotionen erkennen und verarbeiten, ob sie echte Gefühle für uns hervorbringen können und sich die Grenze zwischen Maschinen und biologischen Organismen etwa aufgrund der Entwicklung schmerzempfindlicher Roboter verschiebt oder welche moralischen Konsequenzen sich daraus für unser Mitgefühl ergeben bis hin zu Roboterliebe und Sexrobotern.

Zunächst erörtert die Autorin die Grundbegriffe der künstlichen und emotionalen Intelligenz. Das Forschungsgebiet der **künstlichen Intelligenz (KI)** erstreckt sich darauf, intelligentes Problemlösungsverhalten durch Maschinen nachzubilden oder zu simulieren wie visuelle Wahrnehmung, Spracherkennung und Spracherzeugung, automatisches Schließen und Entscheiden. Durch „schwache KI“ werden konkrete Anwendungsaufgaben vollstreckt, während menschliche Intelligenz höchstens nachgeahmt wird. „Starke KI“ bezog sich anfangs auf Arbeitsvollzüge, wie sie Digitalcomputer und neuronale Netzwerke vollbringen, indem sie große Datenmengen verarbeiten, und strebt heute noch die Übertragung auf unterschiedlichste Aufgabenfelder an. Der Begriff der **emotionalen künstlichen KI** entstand im Rahmen der Debatte um die emotionale Intelligenz in den 1990er Jahren und lässt sich analog zu Ersterer als Disziplin definieren, die sich mit der Nachbildung und Simulation von Emotionen beschäftigt bzw. genauer, wie künstliche Systeme Affekte erkennen und ausdrücken oder gar Gemütsbewegungen selbst aufweisen könnten. Ausführlich setzt sich Misselhorn damit auseinander, was **Emotionen** sind: nämlich „episodische Reaktionen“, körperliche Erregungen, wenn wir von Freude, Schuldgefühlen, Panik ergriffen oder überwältigt werden. Typische Merkmale sind eine subjektive Erlebnisqualität, ein Ausgeliefertsein sowie Anfang und Ende der Gemütsbewegung. Außerdem haben Emotionen einen Objektbezug, insofern sich feststellen lässt, worüber man sich freut, weshalb man wütend ist oder wovor man sich fürchtet, was man indes passiv hinnehmen muss. Emotionen zeigen sich in bestimmten Ausdrucksformen der Sprache,

Haltung, Gestik, Mimik und bewirken oft bestimmte Verhaltensweisen wie z. B. Erstarrung, Flucht oder Angriff bei Furcht.

Die Emotionsforschung hat die Grundlagen in verschiedene Theorien differenziert und spezifiziert, die der **automatisierten Emotionserkennung** als bedeutsamster Zweig der emotionalen KI-Wachstumsbranche dazu dienen, Methoden zu entwickeln und Technologien zur Erkennung menschlicher Emotionen durch künstliche Systeme einzusetzen. Ein Anwendungsverfahren stellt die **gesichtsbasierte Emotionserkennung** durch das Gesichtsbewegungs-Koordinierungssystem dar, das Gesichtsausdrücke aufgrund von Muskelbewegungen klassifiziert und als Code abbildet. Es ist auch möglich, künstliche Systeme mit Hilfe von „Big Data“ darauf zu trainieren, dass sie Gesichtsausdrücke als Muster in Datensätzen erkennen, welche durch Live-Aufnahmen von Kameras, Videoaufzeichnungen, Standbildern aus Videos oder Fotos gewonnen werden. Das erlaubt kommerziellen oder staatlichen Nutzern, die wahren Emotionen von Personen, z. B. potenziellen Kunden oder gar Verbrechern, zu enttarnen, ohne auf deren Selbstinterpretation angewiesen zu sein. Ein weiteres Verfahren ist die **stimm-basierte Emotionserkennung**, deren Analyse sich hauptsächlich auf die Erregungsachse bezieht und Lautstärke, Tonhöhe, Sprechgeschwindigkeit auswertet. Darauf beruhende Apps ermitteln dann Persönlichkeitsmerkmale, um beispielsweise Stellenbewerber oder Mitarbeiter bis hin zum unbewussten Innenleben beurteilen und kontrollieren zu können. Demgegenüber konzentriert sich die **Sentimentanalyse** auf die Sprachbedeutung, indem Ausdrücke, Phrasen, Sätze oder ganze Dokumente untersucht werden, um festzustellen, ob eine Zielgruppe einem Produkt, einer Dienstleistung, Marke, Institution, politischen Partei positiv oder negativ gegenüber steht, was hinsichtlich der Stimmungs-Entwicklung auf Aktien- und Finanzmärkten oder beim Entstehen von Protestbewegungen oder sozialen Unruhen, aber ebenfalls zur Einschätzung der Gefühlslage auf sozialen Medien für eine strategische Einflussnahme betroffener Akteure wichtig sein kann. **Bio-Sensoren** werden in „Smart Watches“ verwendet, damit man durch Messung von Pulswellen, Sauerstoffsättigung im Blut, Atemfrequenz oder Hauttemperatur Gesundheitszustände oder mentale Erkrankungen wie Depressionen feststellen kann zur Erleichterung von Stressmanagement und medizinischen Anwendungen. Nicht zuletzt versucht man durch multimodale Trainings-Datensätze die Zuverlässigkeit der automatischen Emotionserkennung zu steigern.

Neben den technischen Herausforderungen setzt sich Misselhorn schließlich auch mit den **ethischen Fragen** der Problematik auseinander. Da etwa die gesichtsbasierte Emotionserkennung Unsicherheiten hinsichtlich der Diversität aufweist, wird ein Verbot der Technologien gefordert, bis die Risiken vollständig geklärt und angemessene rechtliche Vorschriften geschaffen sind. Die Wahrung der **Privatsphäre** gilt es ferner bei der Erhebung und Übertragung der sensiblen persönlichen Daten zu gewährleisten, denn sie dürfen nicht ohne Zustimmung der Betroffenen weitergeleitet werden. Selbst ihre Anonymisierung könnte als Verstoß gegen die **informationelle Selbstbestimmung** gewertet werden, wenn über die Speicherung und Nutzung der Daten nicht informiert wird. Der freiwillige Verzicht auf informationelle Selbstbestimmung birgt Gefahren, nämlich die Preisgabe der Intimsphäre, der persönlichen Gedanken und Gefühle sowie die unbewusste Beeinflussung unseres Verhaltens. Catrin Misselhorn führt dazu eine umfangreiche Liste bzw. Tabelle von zahlreichen Anwendungsbereichen emotionaler KI, nämlich Manipulations-Möglichkeiten auf.

Schließlich konstatiert sie die **Empathielosigkeit künstlicher Systeme** und geht auf Ansätze ein, menschliche Emotionen nicht nur zu erkennen, sondern maschinelles Mitempfinden, also **artifizielle Empathie** zu erzeugen. Als Terminus technicus wird Empathie allgemein als Fähigkeit bezeichnet, die Gefühle anderer Personen zu verstehen. Dabei wird grundsätzlich zwischen kognitiver und affektiver Empathie unterschieden, was heißt, die Emotionen anderer rational nachzuvollziehen und andererseits mitzuempfinden. Auf Letzteres kommt es an, nämlich Kongruenz, Übereinstimmung der affektiven Zustände verschiedener Personen als Gefühle derselben Art herzustellen. Die Autorin untersucht gründlich Bedingungen und Formen der Empathie als Schlüsselbegriff, indem sie die Möglichkeit, artifizielle Empathie in künstlichen Systemen zu realisieren, an der Bedeutung **zwischenmenschlicher Empathie** misst. Im Mittelpunkt steht die **perzeptuelle Empathie als Resonanz** oder „innere Imitation“, die mit ihrer Intensität die Emotion anzeigt und dadurch mehr oder weniger Empathie hervorruft. Dabei begünstigen oder erschweren Ähnlichkeit, Vertrautheit, die kommunikative Situation und ihre Bewertung das Empfinden von Empathie. Weiterhin wird die **Empathiefähigkeit** von der Stimmung, dem Erregungsgrad, der jeweiligen Persönlichkeit, dem Geschlecht, Alter, emotionalen Repertoire und dem Umgang mit den eigenen Emotionen beeinflusst. Empathie im menschlichen Zusammenleben ist notwendig, damit wir Zugang zu den Gefühlen anderer Personen bekommen, ihre Motive und Verhaltensweisen verstehen oder gar vorhersagen können, z. B. akute Schmerzen zu erkennen und schnell zu helfen. Die Empathie als eine Wurzel pro-sozialen Handelns entsteht mit der Entwicklung des Kindes, hemmt Aggressionen und stellt die Weichen für altruistische und kooperative Einstellungen und Verhaltensweisen. So ist sie mitverantwortlich oder konstitutiv für unser Verständnis von Moral, die Herausbildung und Rechtfertigung moralischer Normen sowie die Beweggründe, sich nach ihnen zu richten. Misselhorn veranschaulicht das überzeugend mit einem „Schema: Empathie und moralisches Urteilen“ am Beispiel der Schmerzempfindung. Empathie als Quelle moralischer Rechtfertigung ist fehleranfällig, wie auch Vernunft nicht unfehlbar ist, aber beide können als gegenseitiges Korrektiv zum Wohle der Menschen zusammenwirken, beispielsweise bei der Rationalisierung von Selbstsucht auf ein verträgliches Maß.

Der perzeptuellen stellt Misselhorn die **artifizielle Empathie** gegenüber. Wenn Menschen mit künstlichen Systemen wie Robotern zusammen arbeiten und leben sollen, ist naheliegend, Letztere mit einem entsprechenden Empathievermögen auszustatten, wozu sich unter dem Begriff **„Artificial Empathy“** ein adäquates Forschungsfeld aufgetan hat und bereits als relevante Disziplin der sozialen Robotik gilt. Anwendungsgebiete, in denen es auf die Qualität der sozialen Interaktion ankommt, ergeben sich im Bereich der Bildung, Pflege, in gesundheitlich-medizinischen Kontexten, bei Computerspielen oder Robotern als dereinst gesellschaftliche Gefährten. Das geht über die künstliche Emotionserkennung hinaus, weil die Maschine dazu über einen „Körper mit Sensoren und Aktoren“ verfügen muss, um mit der Umwelt und anderen Individuen zu interagieren. Ergänzt werden soll die kognitive Architektur eines derartigen Systems durch die Simulation neurophysiologischer Prozesse, die der menschlichen Empathie zugrunde liegen und sie gleichfalls ausdrücken können. Da die Konstruktion und Herstellung von Robotern aufwendig und kostspielig sowie nicht für alle Anwendungsgebiete erforderlich ist, beschränken sich etliche Forschungsansätze auf die Entwicklung **virtueller Agenten**. Das sind autonome, grafisch modellierte und animierte Figuren, die sich in einer simulierten Umgebung

befinden, jedoch mithilfe von Sprache, Mimik und Gestik in Echtzeit mit Menschen kommunizieren können. Außerdem lassen sich mit ihnen die Reaktionen der Nutzer einfacher beobachten und messen, indem meistens über einen Bildschirm interagiert wird. Misselhorn beschreibt als theoretische Grundlagen artifizierlicher Empathie die Entwicklung **computationaler Modelle**, die verschiedene Komponenten von Empathie berechenbar und auf einem Computer implementierbar machen: neurophysiologisch theoriebasiert und mit datengetriebener Methode durch Training von Systemen, gesteuert von Experimentatoren. Dabei treten noch infolge großer Abstraktionen Probleme mit der Detailgenauigkeit, mit Theorie-Verkürzungen und -verzerrungen, Informationsverlust und Implementier-Schwierigkeiten auf, sodass hier hinsichtlich der Ergebnisse verschiedener Forschergruppen Vorsicht geboten erscheint.

Unter der Überschrift „Von empathischen Pflegerobotern und virtuellen Seelenklempnern“ wird an einem Beispiel aus der Altenpflege sowie der klinischen Psychologie aufgezeigt, wie artifizierliche Empathie in der Praxis eingesetzt werden kann. **Empathische Roboter** sollen ältere Menschen im täglichen Leben unterstützen und als Schnittstelle zu weiteren Technologien fungieren und die Beleuchtung und Raumtemperatur regeln, Rollläden öffnen und schließen, den Kühlschrank steuern („Ambient Assisted Living“). Es gibt bereits ein Projekt NICA, bei dem ein kleiner „menschenähnlicher Roboter“ die Schnittstelle zu einer Smart-Home-Umgebung bildet und alten Menschen als sozialer Gefährte dient. Im anderen Beispiel handelt es sich um einen **virtuellen Agenten**, der psychische Störungen diagnostizieren soll. Beide Fälle sind aussagekräftig hinsichtlich der Methoden, Reichweite und Anwendungsbereiche künstlicher Empathie und wurden wissenschaftlich durchgeführt sowie dokumentiert. Ergebnis: Die Nutzer interpretierten die Simulation der Roboter zwar als empathisch, empfanden jedoch nicht wirklich Empathie, sondern nur Sympathie. Andernfalls müssten die künstlichen Systeme selbst Emotionen haben können.

Diesem diffizilen Anliegen widmet sich die Autorin mit der Untersuchung zur **Schmerzempfindung** in den neuen Tendenzen der **Biorobotik**. Damit Maschinen fühlen können, müssten die künstlichen Systeme über entsprechende neurophysiologische Voraussetzungen für Empathie wie etwa Spiegelneuronen verfügen. Demgemäß Forschungsansätze versuchen deshalb, ein Analogon des Nervensystems zu bilden, um zu echter artifizierlicher Empathie zu gelangen. Dabei steht die Nachbildung des Schmerzsystems im Mittelpunkt, weil der Empathie bei der Schmerzempfindung besondere soziale und moralische Bedeutung zukommt, denn aufgrund eigener Schmerzerfahrungen geht es darum, anderen, die Schmerzen haben und äußern, schnell zu helfen und Linderung zu verschaffen. Die Realisierungsansätze begannen bescheiden damit, einem Roboterarm mit einem Fingerspitzen-Sensor beizubringen, dass er sich bei Berührung zurückzieht. Die dazu eingesetzten Mechanismen orientieren sich zwar an menschlichen Schmerzreflexen, erschöpfen sich jedoch in einem einfachen Reiz-Reaktions-Bogen ohne phänomenal bewusste Schmerzempfindung und simulieren lediglich die Funktion von Schmerzen, schädliche Einflüsse durch Ausweichen zu vermeiden. Doch auch ein in Japan entwickelter **hyperrealistischer Roboter** in Kindergestalt, bei dem 116 Punkte im Gesicht mit taktilen Sensoren gekoppelt wurden, um bei Berührung mit schmerzverzerrten Gesichtsbewegungen zu reagieren, förderte lediglich rein behavioristische Verhaltensreaktionen zutage. Auch Misselhorns **Gedankenexperiment** zur großangelegten Simulation des menschlichen Gehirns einschließlich aller Wahrnehmungsprozesse und Bewegungsimpulse läuft darauf hinaus, dass alles

Wissen über den Gehirnaufbau und das Zusammenspiel mit dem Körper **prinzipiell kein phänomenales Schmerzbewusstsein** hervorbringen kann.

So fragt Catrin Misselhorn in dem Zusammenhang rhetorisch, ob ein unauflöslicher **Dualismus zwischen Körper und Geist** bestehe. Die klassische Dualismus-Position in der Philosophie geht auf Descartes zurück und besagt, dass Körper und Geist zwei grundsätzlich voneinander getrennte Seinssphären, zwei grundverschiedene Substanzen oder Eigenschaftsarten, darstellen. Die körperliche Dinge sind im Raum lokalisiert, verfügen über Masse, sind allgemein beobachtbar und weisen keinen Erlebnisaspekt auf, während die Inhalte des geistigen Bereichs nicht räumlich bestimmt sind, keine Masse haben, nicht öffentlich anzuschauen sind. Aber sie besitzen eine Erlebnisperspektive, die sich nur durch Introspektion in der Ich-Dimension unmittelbar subjektiv erfassen lässt. Misselhorn widerlegt sowohl den Dualismus als auch den **Epiphänomenalismus**, wonach das Geistige aus dem Körperlichen emergiere, geistige Zustände oder Ereignisse als Nebenwirkungen körperlicher Prozesse keinerlei Einfluss auf Körper und Umgebung ausüben würden. Denn die Alltagserfahrung mit dem subjektiven Schmerzerlebnis wie beim Anfassen einer heißen Herdplatte zeigt, dass wir – als „biologische Funktion des phänomenalen Bewusstseins“ – sehr wohl flexibel darauf reagieren, um körperliche Schäden zu vermeiden und letztlich zu überleben. Überhaupt erweist sich der Mensch als Paradebeispiel für die Einheit von Körper und Geist oder Implikation von gegenständlicher und transzendentaler Welt.

Die Konsequenz daraus für die Biorobotik bedeutet, Roboter nach dem biologischen **Prinzip der Homöostase oder Selbsterhaltung** zu konstruieren, indem lebende Organismen für ihre Weiterexistenz homöostatische Gleichgewichtszustände herstellen, etwa aufgrund des Hungergefühls Nahrung aufnehmen. So sollen **Soft Robotics und maschinelles Lernen** dem Prozess des Lebens möglichst nahe kommen, insofern Soft Robots nicht aus Metall und hartem Kunststoff bestehen, sondern aus nachgiebigen Materialien, die mit ihren Eigenschaften natürlichen Geweben wie Haut, Muskeln, Knorpeln ähneln. Es gibt schon eine elektronische Gelhaut, die auf Berührung, Druck und Zug, Temperatur (Sinnesmodalitäten) eingeht und sich sogar nach einem Schnitt selbst zu heilen vermag. Mit Hilfe maschinellen Lernens sollen die Informationen aus verschiedenen Sinnesmodalitäten zusammengeführt und ständig abgeglichen werden, damit sich das System in einem homöostatischen Gleichgewicht oder „Wohlbefinden“ halten kann. Auf diese Weise können homöostatische Roboter erweiterte Intelligenz sowie Autonomie erreichen und so tun, als ob sie Gefühle haben, aber aufgrund ihrer synthetischen Beschaffenheit kein echtes Lebewesen sein. Der nächste folgerichtige Schritt besteht daher darin, **biohybride Roboter** zu erschaffen, die nicht nur aus weichen, sondern aus organischen Materialien wie Gewebe- oder Muskelzellen geformt werden. Dazu gab es 2020 den Modellversuch „**Xenobots**“ zur Herstellung programmierbarer Organismen, die mit Computereinsatz aus Froschstammzellen erzeugt wurden, sich selbst organisieren und fortbewegen können, aber es sind Winzlinge, kleiner als 1 mm und von kurzer Lebensdauer. Sie weisen nicht die klassischen Merkmale der Autopoiese von Lebewesen zur Selbsterhaltung auf und können nicht auf Reize der Umwelt reagieren, sodass die Forscher noch weit davon entfernt sind, aus den Zellhaufen der Xenobots empfindungsfähige künstliche Wesen mit phänomenalem Bewusstsein zu erschaffen.

Hier kehrt Misselhorn die Frage um, ob künstliche Systeme Empathie mit Menschen verspüren können, und untersucht, inwiefern **Menschen Empathie mit Robotern** verspüren können und wieweit dies moralisch bedeutsam ist. Dazu setzt sie bei **Science-Fiction-Werken** an, die von künstlichen Systemen ausgehen und extrem menschenähnlich erscheinen wie auch „Alexa“ von Amazon. Sie entwirft Schemata zur Entwicklung von Empathie und Moral im Umgang mit Robotern, zeigt die Gefahr des Missbrauchs auf, dass man zum Opfer von Manipulation werden kann, und sie bringt Beispiele positiver Arbeit mit Robotern für Demenzzranke und die Kinder-Therapie bei Erkrankungen des Autismus-Spektrums. Schließlich widmet sie sich dem Thema „Freundschaft, Liebe und Sex mit Robotern“ und betrachtet **soziale Roboter**, die auf ihre Umwelt reagieren können und lernfähig sind, sie verfügen über eine Palette von Verhaltensweisen und suggerieren, eigenständig zu handeln. Ausführlich setzt sie sich mit dem **Einsatz von Sexrobotern** auseinander. Es handelt sich um anatomisch korrekte „Sexpuppen“, deren Skelett mit fleischartig synthetischer Haut überzogen ist, die Menschen in Gestalt, Bewegung und Verhaltensweisen ähneln. Sie verfügen über ein Mindestmaß an Interaktionsfähigkeit mit integriertem Chatbot und wenigen interaktiven Features für diejenigen, die sich ihrer bedienen. Das sind hauptsächlich Männer, sodass es überwiegend weibliche Sexroboter gibt. Ein großes Akzeptanzproblem sieht die Autorin indes nicht, sie vermutet, dass künftig manche Leute lieber Sex mit Robotern als mit Artgenossen wollen könnten, wenn er genussvoll und perfekt auf ihre Bedürfnisse und Vorlieben abgestimmt ist. Befürworter meinen, Bordelle mit Sexrobotern wie in Barcelona und Turin könnten gar Prostitution und sexuelle Gewalt mindern oder die Geräte können als Segen für Menschen ohne Partner, aber mit Traumata reichen. Catrin Misselhorn konstatiert, dass Sex mit Robotern die Definition von **Pornografie** erfüllt, insofern es sich um subjektsimulierende Maschinen handelt, die den sexuellen Vorgang bis zum Orgasmus vortäuschen und mit denen primär Frauen als sexuelle Gebrauchsobjekte benutzt werden bis hin zur Vergewaltigung von Sexrobotern mit der diesbezüglichen Absenkung der Hemmschwelle, was moralisch problematisch ist. Weil in der Liebe zu einem Roboter das „Subjekt in der Echokammer seines Selbst gefangen bleibt, in der immer nur seine eigenen Gefühle verstärkt zurückhalten“ (**Solipsismus** als Selbstverdinglichung und deformiertes Selbstverhältnis), plädiert sie mit Nachdruck für Freundschaft, Liebe und Sex mit echten, empfindungsfähigen Menschen, die einander anerkennen.

Der Hinweis auf Sexroboter im Buchtitel mag um der Sache willen verstärkt Leseranreiz provozieren und ein großes Leserpublikum, nicht zuletzt Berufs- und Wirtschaftspädagogen sowie Berufsschullehrer/-innen, gewinnen helfen, um so der Gesamtaufklärung über das wichtige Thema „emotionale künstliche Intelligenz“ zu dienen, die noch beachtlichen Einfluss auf die Gestaltung unserer Gesellschaft und künftigen Welt nehmen dürfte.

Zitieren dieser Rezension

Pukas, D. (2021): bwp@-Rezension zu Catrin Misselhorn: Künstliche Intelligenz und Empathie. Vom Leben mit Emotionserkennung, Sexrobotern & Co. Stuttgart, 1-6. Online: http://www.bwpat.de/rezensionen/rezension_11-2021_misselhorn.pdf (19.12.2021).
