

Künstliche Intelligenz und menschliche Emotionen

Christa Schmid-Meier

Zusammenfassung

In den letzten Jahren hat die Entwicklung von Anwendungen und Technologien, die auf Künstlicher Intelligenz (KI) basieren, grosse Fortschritte gemacht. Während KI viele Möglichkeiten und Vorteile bietet, gibt es auch Bedenken hinsichtlich ihres missbräuchlichen Einsatzes. Immer mehr KI-Programme sind so programmiert, dass sie die Emotionen der Anwender:innen lesen und verwenden können. Daher ist es unabdingbar, dass sich Menschen beim Gebrauch solcher Anwendungen bewusst sind, dass sie auf emotionaler Ebene manipuliert werden können. Welche Bedeutung diese Entwicklungen für die Heilpädagogik haben können, wird in diesem Beitrag aufgezeigt.

Résumé

Ces dernières années, le développement d'applications et de technologies basées sur l'intelligence artificielle (IA) a fait de grands progrès. Si l'IA offre de nombreux avantages et ouvre de nouvelles possibilités, son utilisation abusive suscite également des inquiétudes. De plus en plus de programmes d'IA sont conçus pour lire et utiliser les émotions des utilisatrices et des utilisateurs. Il est donc essentiel que les personnes qui utilisent ces applications soient conscientes qu'elles peuvent être manipulées sur le plan émotionnel. Cet article démontre l'importance de ces développements pour la pédagogie spécialisée.

Keywords: digitale Transformation, Künstliche Intelligenz, Emotion, Sucht, Medienkompetenz / transformation numérique, intelligence artificielle, émotion, dépendance, compétences médiatiques

DOI: <https://doi.org/10.57161/z2023-09-05>

Schweizerische Zeitschrift für Heilpädagogik, Jg. 29, 09/2023



Künstliche Intelligenz

Künstliche Intelligenz (KI) hat das Potenzial, viele Bereiche unseres Lebens zu optimieren. Allerdings kann Künstliche Intelligenz auch erheblichen Schaden anrichten. Deswegen erfordert die zunehmende Verwendung von Emotionserkennung durch Künstliche Intelligenz ein besseres Verständnis der Funktionsweise und des Einsatzes dieser Technologie.

Der Begriff «Künstliche Intelligenz» wird kontrovers diskutiert und es existieren verschiedene Definitionen. Für diesen Text wird die folgende Definition verwendet (Europäisches Parlament, 2023):

Künstliche Intelligenz ist die Fähigkeit einer Maschine, menschliche Fähigkeiten wie logisches Denken, Lernen, Planen und Kreativität zu imitieren. KI ermöglicht es technischen Systemen, ihre Umwelt wahrzunehmen, mit dem Wahrgenommenen umzugehen und Probleme zu lösen, um ein bestimmtes Ziel zu erreichen. Der Computer empfängt Daten (die bereits über eigene Sensoren, zum Beispiel eine Kamera, vorbereitet oder gesammelt wurden), verarbeitet sie und reagiert. KI-Systeme sind in der Lage, ihr Handeln anzupassen, indem sie die Folgen früherer Aktionen analysieren und autonom arbeiten.

Künstliche Intelligenz ist momentan eine der am schnellsten wachsenden Technologien. Sie hat das Potenzial, die Art und Weise zu verändern, wie wir leben, arbeiten und miteinander interagieren. Während der Begriff noch vor einem Jahr nur wenigen Menschen geläufig war, ist er ins öffentliche Bewusstsein gerückt, seit im November 2022 *ChatGPT* erschienen ist. Doch Künstliche Intelligenz umfasst viel mehr als *ChatGPT*. Es gibt bereits viele KI-gestützte Anwendungen, die zunehmend in allen Lebensbereichen eingesetzt werden, darunter dem Gesundheitswesen, der Landwirtschaft und dem öffentlichen Sektor. Wichtig zu wissen ist, dass Künstliche Intelligenz autonom arbeitet. Während

frühere Algorithmen Regeln der Form «Wenn ..., dann ...» folgten, werden KI-gestützte Algorithmen mit einem grossen Datensatz trainiert. Sie suchen programmierten Anweisungen folgend selbstständig nach auffälligen Mustern, die in Verbindung mit einem bestimmten Ergebnis auftreten. KI-Algorithmen «lernen» also an Beispielen. Die gefundenen Muster werden dann als Regeln auf neue Daten angewendet. Welche Faktoren in welcher Kombination und in welcher Gewichtung dabei im Hintergrund den Ausschlag geben, ist selbst für Programmierer:innen solcher Algorithmen grösstenteils unklar. Das macht es schwierig, die Entscheidungen eines Modells nachzuvollziehen und zu überprüfen, ob sie fair, ethisch und korrekt sind. Hinzu kommt, dass KI-Systeme verzerrte Weltansichten transportieren können aufgrund der Trainingsdaten, mit denen sie gefüttert wurden (Sinha et al., 2023). Dies wird dadurch verstärkt, dass KI-Algorithmen auf statistischen Annahmen und Werten beruhen und die Berechnungen entsprechend zur Mitte tendieren. Bisher sind deshalb Diversität, alternative Herangehensweisen oder Menschen mit Behinderung in KI-Systemen noch kaum mitgedacht.

Die Entwicklung steht jedoch erst am Anfang. Viele KI-Systeme integrieren neue Daten laufend und optimieren sich so selbst (Hamisch & Kruschel, 2022). Es zeichnet sich ab, dass sie sich künftig immer individueller an die Benutzer:innen anpassen und so neue Anwendungsfelder erschlossen und Systeme kombiniert werden. Dies bietet auch in der Heilpädagogik erhebliche Chancen, insbesondere in Bezug auf Lernen und Assistenz. So kann KI beispielsweise dazu beitragen, Technologien auf die spezifischen Bedürfnisse und Fähigkeiten jedes einzelnen Menschen zuzuschneiden. Sie können die individuelle Aussprache einer Einzelperson verstehen lernen, persönliche Bewegungsmuster analysieren und deuten oder eine auf ein Individuum abgestimmte Applikation zur Unterstützten Kommunikation (UK) bereitstellen. Künstliche Intelligenz ist per se also weder gut noch schlecht. Ihre Wirkung und ihr Nutzen hängen vielmehr stark davon ab, wie sie angewendet wird.

Bisher sind Diversität, alternative Herangehensweisen oder Menschen mit Behinderung in KI-Systemen noch kaum mitgedacht.

Menschliche Emotionen treffen auf Künstliche Intelligenz

Der Mensch ist ein soziales Wesen. Seit Jahrtausenden (über-)lebt er in Gemeinschaften. Emotionen spielen eine entscheidende Rolle in sozialen Interaktionen und der Art und Weise, wie Menschen auf soziale Signale reagieren. Emotionen helfen uns nicht nur dabei, positive soziale Beziehungen aufzubauen und zu pflegen, sondern auch dabei, unsere soziale Position im Vergleich zu anderen zu etablieren oder zu erhalten (Fischer & Manstead, 2016). Menschliche Emotionen können durch *interpersonale Emotionsregulation* beeinflusst werden, also den ausdrücklichen oder impliziten Versuch, die Gefühle einer anderen Person zu verändern (ebd.). Dies kann auf verschiedene Weisen geschehen, beispielsweise durch Kommunikation, Mitgefühl oder Unterstützung. Dabei wenden sich Menschen oft aktiv an andere, um Hilfe bei der Gestaltung ihres affektiven Lebens zu erhalten (Zaki & Williams, 2013). Ein verwandtes Konzept ist die *Co-Regulation*. Damit handelt es sich um einen Prozess, bei dem Menschen gegenseitig ihre Emotionen beeinflussen und regulieren (Fischer & Manstead, 2016). Dieser Prozess der gegenseitigen emotionalen Beeinflussung kann sowohl konstruktiv als auch destruktiv sein. Beide Konzepte spielen eine zentrale Rolle in unseren sozialen Interaktionen und Beziehungen und beeinflussen unsere emotionale Gesundheit und unser Wohlbefinden.

Während sozialer Austausch in früheren Zeiten fast ausschliesslich von Angesicht zu Angesicht stattfand, verlagert sich dieser heute verstärkt in den digitalen Raum. Viele Menschen verbringen erhebliche Zeit mit digitalen Medien und sind daher oft den emotionalen Ausdrücken anderer Menschen ausgesetzt – sei es im direkten Austausch mit ihnen oder beim Konsum von digitalen Inhalten und Programmen. Dies kann dazu führen, dass ihre eigenen Gefühlsäusserungen den emotionalen Ausdrücken, die sie online wahrnehmen, ähnlicher werden. Dieses Phänomen wird als *digital emotion contagion* (digitale Emotionsansteckung) bezeichnet (Goldenberg & Gross, 2020). Während bei persönlichen Interaktionen die emotionalen Reaktionen des Gegenübers auf direktem Weg wahrgenommen werden, werden

Interaktionen über digitale Medien häufig durch Unternehmen gesteuert und auch manipuliert. Oftmals zielen sie darauf ab, die positiven Emotionen der Benutzer:innen zu verstärken (ebd.).

Der Teilbereich der Künstlichen Intelligenz, der sich mit Emotionen befasst, wird *Emotion-AI* oder auch *affective computing* genannt¹. *Emotion-AI* hat zum Ziel, die Interaktion zwischen Menschen und Maschine zu humanisieren und intuitiver zu gestalten. Sie soll verschiedene Emotionen identifizieren und eine Wahrscheinlichkeit für jede Emotion berechnen in Datenmaterial wie Bild- und Audiodaten, aber auch in physiologischen Daten wie Herzschlag oder Atemfrequenz (Monteith et al., 2022). Diese Technologien könnten beispielsweise in der Strafverfolgung, Krankenpflege oder Bildung eingesetzt werden, um Emotionen zu erkennen und entsprechend zu reagieren. Die weltweite Nachfrage nach solchen Systemen steigt seit Jahren steil an (Peters, 2021). Sie sind jedoch noch weit davon entfernt, perfekt zu sein. So hat eine Überprüfung von über 1000 Studien ergeben, dass der emotionale Zustand nicht zuverlässig aus Gesichtsbewegungen abgeleitet werden kann (Feldman Barrett et al., 2019). Dies wirft Fragen zur Genauigkeit und Ethik (z. B. Verwechslungsgefahr von Personen, Datenschutz, Fehlzuschreibungen) von solchen Anwendungen auf. Zudem unterstreicht es die Komplexität menschlicher Emotionen sowie die Schwierigkeiten, diese präzise zu messen und zu interpretieren. Dies ist besonders relevant, wenn man an Menschen mit atypischen Ausdrucks- oder Verhaltensweisen denkt, bei denen herkömmliche Methoden der Emotionserkennung möglicherweise nicht anwendbar oder ungenau sind. Ausserdem ist zu befürchten, dass die Ergebnisse solcher *Emotion-AI* als Fakten verwendet werden und dadurch die Stigmatisierung und Diskriminierung von Menschen mit psychischen Störungen verstärken könnten (Monteith et al., 2022). Die Möglichkeiten, wie Künstliche Intelligenz die Emotionen von Menschen beeinflussen kann, sind vielfältig. Im Folgenden werden daher einige exemplarische Beispiele angeführt.

Es gibt vielfältige Möglichkeiten, wie Künstliche Intelligenz die Emotionen von Menschen beeinflussen kann.

Manipulation

KI-Systeme, die darauf ausgelegt sind, menschliche Emotionen zu verstehen und darauf zu reagieren, können zur Manipulation von Gefühlen und Verhalten verwendet werden. Diese Manipulation kann in vielen Kontexten stattfinden: von politischer Propaganda, die unsere Meinungen und Wahlverhalten formen, bis hin zu Anwendungen im Rahmen von therapeutischen Kontexten. Bekannt ist die Möglichkeit zur Manipulation durch personalisierte Inhalte bereits seit einiger Zeit, beispielsweise unter dem Begriff *Filterblase*. Dies beschreibt ein Phänomen, das in Sozialen Medien und Online-Plattformen auftritt. Die Benutzer:innen sehen hauptsächlich Inhalte, die ihren bestehenden Überzeugungen und Vorlieben entsprechen (Messingschlager & Holtz, 2020). Diese Personalisierung wird durch Algorithmen ermöglicht, die das Verhalten und die Präferenzen der Nutzer:innen analysieren und dann entsprechende Inhalte liefern. Die Emotionserkennung verschärft diese Problematik. Indem soziale Medienplattformen die Emotionen der Benutzer:innen erkennen und analysieren, können sie weitaus spezifischere und zielgerichtete Inhalte liefern. Dies kann zu einer Reihe von negativen Auswirkungen führen, von verstärkter Polarisierung und Desinformation bis hin zu negativen Auswirkungen auf die psychische Gesundheit der Benutzer:innen.

Abhängigkeit

Es besteht die Gefahr, dass Benutzer:innen von Programmen mit KI ein Suchtverhalten entwickeln, insbesondere, wenn diese Systeme positive Emotionen hervorrufen. Ein aktuelles Beispiel hierfür ist *Tiktok*. Diese App verfügt über eines der fortschrittlichsten KI-Algorithmus-Systeme und gilt momentan im Vergleich zu anderen sozialen Medienplattformen als diejenige, die am ehesten süchtig macht (Qin et al., 2022). Die App erreicht dies, indem sie ihren Benutzer:innen Momente des *Flow*-Erlebnisses bietet (Zhao & Wagner, 2023). *Flow* ist ein mentaler Zustand, in dem man vollständig in eine

¹ Der Begriff *Emotion-AI* wird auch im Deutschen so verwendet.

Tätigkeit vertieft ist und ein Gefühl von müheloser Konzentration und Freude erlebt. Werden Gefühle durch KI auf Knopfdruck abrufbar, besteht die Gefahr, dass Menschen eine Abhängigkeit von diesen – sofort und mit minimalem Aufwand erzeugten – Gefühlen entwickeln, anstatt diese Gefühle durch aufwendigere, langwierigere Erfahrungen und Interaktionen im realen Leben zu suchen.

Übermässiges Vertrauen

KI-gesteuerte Chatbots und virtuelle Assistenzen können so programmiert werden, dass sie menschenähnliche Konversationen führen und empathisch auf die Emotionen der Benutzer:innen reagieren. Es konnte gezeigt werden, dass (vermeintliche) Nähe entsteht, wenn diese Systeme «persönliche» Informationen preisgeben (Tsumura & Yamada, 2023). Dies beruht wahrscheinlich auf dem Phänomen, dass Menschen einander Vertrauen und Offenheit zeigen durch den Austausch von persönlichen Geheimnissen und dadurch Beziehungen stärken. Offenbar findet dieser Effekt auch statt, wenn sie mit einer virtuellen Assistenz kommunizieren. Weiter können die Bequemlichkeit und Effizienz, die eine Technologie bietet, dazu führen, dass Menschen sich zunehmend auf sie verlassen. Es gibt beispielsweise immer wieder Menschen, die beim Autofahren gedankenlos den Anweisungen der Navigations-App folgen und sich damit in Gefahr begeben. Auch kann das Nichtverstehen von modernen Technologien und insbesondere der scheinbare Intellekt von KI dazu führen, dass Menschen ihnen mehr Vertrauen schenken, als dies angemessen wäre. So hat eine Studie die Reaktion von Menschen auf Roboter in Evakuierungsszenarien untersucht. Trotz vorheriger schlechter Leistung des Roboters bei einer Navigationsaufgabe folgten die Teilnehmenden ihm in der Notfallsituation (Robinette et al., 2016). Der Einsatz von KI kann auch die Risikobereitschaft erhöhen (Villa et al., 2023). Es wird angenommen, dass dieses Phänomen auf ähnliche Weise wirkt wie der Placebo-Effekt, indem Menschen sich durch die Künstliche Intelligenz unterstützt fühlen und dadurch bereit werden, mehr Risiken einzugehen. Auffallend ist, dass viele Menschen respektvoll mit virtuellen Assistenzen umgehen und sich sogar bei ihnen bedanken. Dies legt die Schlussfolgerung nahe, dass Menschen virtuelle Assistenzen vermenschlichen, was ebenfalls zum übermässigen Vertrauen gegenüber KI-gestützten Systemen beitragen könnte.

*Indem Menschen virtuelle Assistenzen vermenschlichen,
wächst das Vertrauen in sie übermässig.*

Einsichten, Ansichten und Aussichten

Künstliche Intelligenz hat also das Potenzial, menschliche Emotionen zu erkennen und zu beeinflussen. Daraus ergibt sich eine neue Facette der Medienkompetenz: sich seiner Menschlichkeit bewusst zu sein. Dies beinhaltet, seine eigenen menschlichen Fähigkeiten, Emotionen und Reaktionen zu erkennen und zu verstehen, während man digitale Anwendungen nutzt. In einer zunehmend von KI geprägten Welt ist diese Kompetenz entscheidend, um eine bewusste und ethisch sensibilisierte Interaktion mit digitalen Systemen zu gewährleisten, die Kontrolle über unsere Technologieerfahrungen zu behalten und unsere Menschlichkeit zu respektieren. Die Heilpädagogik kann hierfür einen wertvollen Beitrag leisten. Obwohl sie auf anderen Grundvoraussetzungen als die Maschinenlogik basiert, sind sich die beiden Gebiete näher, als dies auf den ersten Blick erscheinen mag: Viele der heute in Laptops, Tablets und Smartphones eingesetzten Anwendungen und Technologien haben ihren Ursprung im Gebiet der assistiven Technologien. Beispiele hierfür sind *Text-to-Speech* sowie individuell einstellbare Textgrösse und Kontrast auf Geräten, die ursprünglich für Menschen mit Sehbehinderung entwickelt wurden. Mit ihrem Blick auf die Menschen in ihrer Vielfalt, mit ihren individuellen Ressourcen und Herausforderungen ist die Heilpädagogik ein idealer Ausgangspunkt dafür, die Inklusion aller Menschen in die digital geprägte Zukunft zu denken.



Christa Schmid-Meier
MEd, MA
Fachstelle ICT for Inclusion
Interkantonale Hochschule für Heilpädagogik, Zürich
christa.schmid-meier@hfh.ch

Literatur

- Europäisches Parlament. (2023). *Was ist künstliche Intelligenz und wie wird sie genutzt?* www.europarl.europa.eu/news/de/headlines/society/20200827STO85804/was-ist-kunstliche-intelligenz-und-wie-wird-sie-genutzt [Zugriff: 12.08.2023]
- Feldman Barrett, L., Adolphs, R., Marsella, S., Martinez, A. M. & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological science in the public interest*, 20 (1), 1–68.
- Fischer, A. H. & Manstead, A. S. (2016). Social functions of emotion and emotion regulation. *Handbook of emotions*, 4, 424–439.
- Goldenberg, A. & Gross, J. J. (2020). Digital Emotion Contagion. *Trends in Cognitive Sciences*, 24 (4), 316–328. <https://doi.org/10.1016/j.tics.2020.01.009>
- Hamisch, K. & Kruschel, R. (2022). Zwischen Individualisierungsversprechen und Vermessungsgefahr – Die Rolle der Schlüsseltechnologie Künstliche Intelligenz in der inklusiven Schule. *Grenzen. Gänge. Zwischen. Welten*, 108.
- Messingschlager, T. & Holtz, P. (2020). Filter Bubbles und Echo Chambers. In M. Appel (Hrsg.), *Die Psychologie des Postfaktischen: Über Fake News, «Lügenpresse», Clickbait & Co.* (S. 91–102). Springer. https://doi.org/10.1007/978-3-662-58695-2_9
- Monteith, S., Glenn, T., Geddes, J., Whybrow, P. C. & Bauer, M. (2022). Commercial use of emotion artificial intelligence (AI): Implications for psychiatry. *Current Psychiatry Reports*, 24 (3), 203–211.
- Peters, R. (2021). *Emotionserkennung mittels künstlicher Intelligenz – Perspektiven und Grenzen von Technologien zur Analyse von Gesichtsbewegungen*. <https://publikationen.bibliothek.kit.edu/1000134317>
- Qin, Y., Omar, B. & Musetti, A. (2022). The addiction behavior of short-form video app TikTok: The information quality and system quality perspective. *Frontiers in Psychology*, 13, 932805. <https://doi.org/10.3389/fpsyg.2022.932805>
- Robinette, P., Li, W., Allen, R., Howard, A. M. & Wagner, A. R. (2016). *Overtrust of robots in emergency evacuation scenarios*. 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 101–108. <https://doi.org/10.1109/HRI.2016.7451740>
- Sinha, A., Sapra, D., Sinwar, D., Singh, V. & Raghuvanshi, G. (2023). Assessing and Mitigating Bias in Artificial Intelligence: A review. *Recent Advances in Computer Science and Communications*, 16. <https://doi.org/10.2174/2666255816666230523114425>
- Tsumura, T. & Yamada, S. (2023). Influence of agent's self-disclosure on human empathy. *PLOS ONE*, 18 (5), e0283955. <https://doi.org/10.1371/journal.pone.0283955>
- Villa, S., Kosch, T., Grelka, F., Schmidt, A. & Welsch, R. (2023). The placebo effect of human augmentation: Anticipating cognitive augmentation increases risk-taking behavior. *Computers in Human Behavior*, 146, 107787. <https://doi.org/10.1016/j.chb.2023.107787>

Zaki, J. & Williams, W. C. (2013). Interpersonal emotion regulation. *Emotion*, 13 (5), 803–810.

<https://doi.org/10.1037/a0033839>

Zhao, H. & Wagner, C. (2023). How TikTok leads users to flow experience: Investigating the effects of technology affordances with user experience level and video length as moderators. *Internet Research*, 33 (2), 820–849.

<https://doi.org/10.1108/INTR-08-2021-0595>