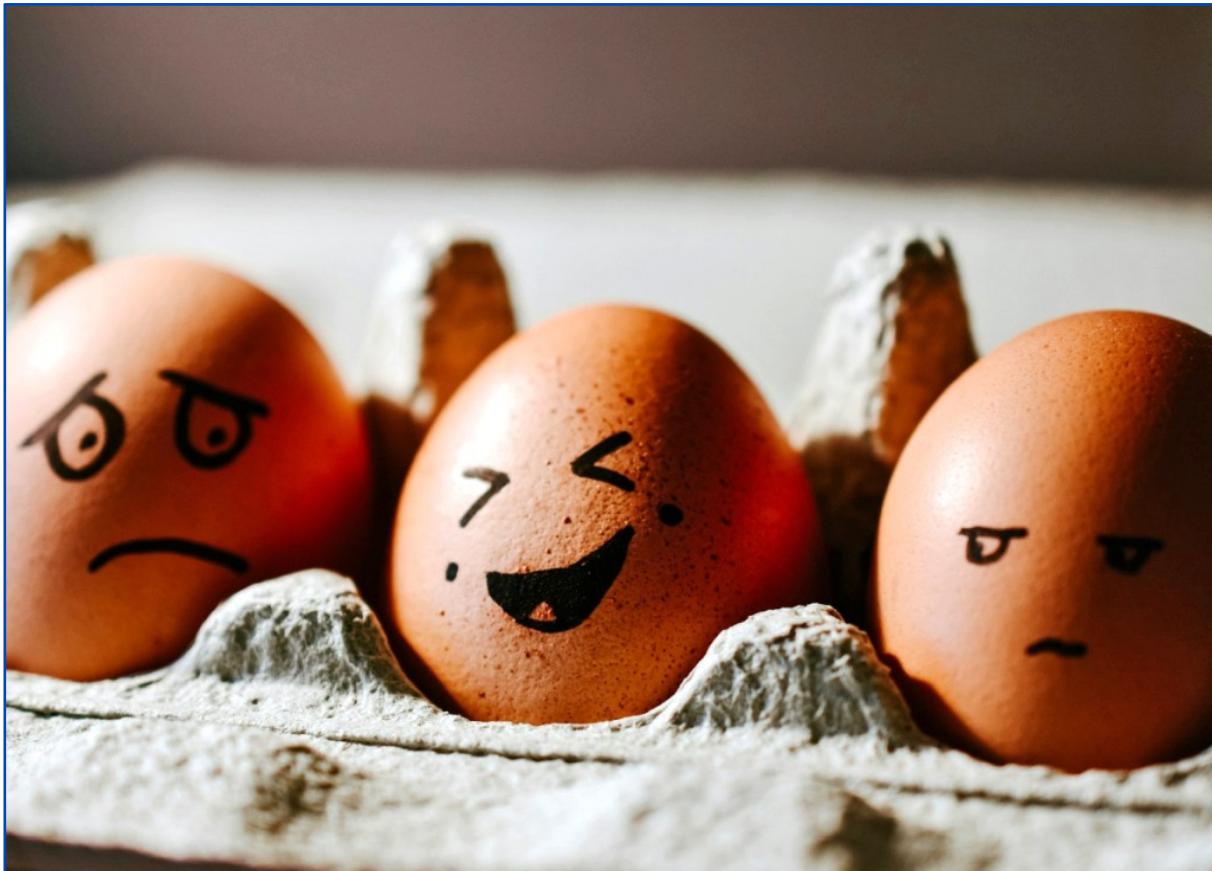


EMOTIONSERKENNUNG ALS HOFFNUNGSGEBIET FÜR KI?



© CC0 (Nik/unsplash)

ZUSAMMENFASSUNG

KI wird zunehmend dazu eingesetzt, Muster in Gesichtern, Sprache, Texten, Körpersignalen und Verhaltensweisen zu erkennen und daraus Analysen und Vorhersagen über emotionale Zustände, Vorlieben und Persönlichkeitsmerkmale von Menschen zu erstellen. Die Systeme können in vielfältigen Bereichen eingesetzt werden – etwa durch Sicherheitsbehörden an Grenzen und im öffentlichen Raum, bei der therapeutischen Betreuung in der Medizin, bei der Unterstützung von Schüler:innen in der Bildung, im personalisierten Marketing oder der Personalentwicklung in der Arbeitswelt. Die Verlässlichkeit der Systeme muss jedoch aufgrund der dahinterliegenden Modelle und psychologischen Annahmen kritisch eingestuft werden. Zudem ergeben sich vielfältige Herausforderungen bei Fragen der Diskriminierung, des Datenschutzes, der ethischen Verantwortung für besonders schutzwürdige Gruppen, des Strukturwandels politischer Öffentlichkeiten und der Verschärfung sozialer Ungleichheit in der Arbeitswelt.

ÜBERBLICK ZUM THEMA

Mit dem Fortschritt im Bereich der künstlichen Intelligenz eröffnen sich neue Möglichkeiten, menschliches Verhalten einer bestimmten *Emotionskategorie zuzuordnen* und daraus weiter Schlüsse über emotionale Zustände, charakterliche Eigenschaften oder Verhaltensweisen von Menschen zu ziehen.¹ 1995 führte Rosalind Picard dafür den Begriff „Affective Computing“ ein (Picard 2000). Entsprechende KI-Systeme können dabei auf verschiedene Datensorten wie biometrische Signale, Bilder, Texte, Interaktions- und Audio- und Videodaten zurückgreifen. Eine bereits lang erprobte Variante ist die Analyse und Voraussage von Charaktereigenschaften und Vorlieben auf Social-Media-Plattformen. Genutzt werden dabei Interaktionsmuster, um Profile von Nutzer:innen zu erstellen (Kosinski, Stillwell & Graepel 2013). Eine weitere Möglichkeit besteht in der Erfassung visueller Daten durch Systeme der Gesichtserkennung, die über die Mimik Gefühlszustände von Personen erfassen. Die heutigen Systeme zur Gesichtserkennung basieren vor allem auf sogenannten holistischen Ansätzen. Im Gegensatz zu merkmalsbasierten Ansätzen, in denen tatsächlich vorhandene, biometrische Daten wie der Abstand der Augen vermessen werden, erfassen holistische Systeme ausschließlich Helligkeitsverteilungen in einem Bild (Meyer 2019, S. 321-359). Darüber hinaus können auch biometrische Daten wie die Herzfrequenz, der Pupillendurchmesser oder die Muskelaktivität zur emotionalen Erkennung verwendet werden. Die benötigten Sensoren können inzwischen auch mobil eingesetzt werden – etwa in Form von Wearables wie Smart Watches (Khare et al. 2024). Insbesondere mit der Verbreitung von Large-Language-Models (LLM) ist zudem das Erkennen von Emotionen aus Texten immer effektiver möglich. Teilweise werden Emotionen direkt benannt, teilweise werden sie aus Beschreibungen, Ausrufen oder anderen rhetorischen Wendungen ermittelt (Yang 2023).

KI zur Erkennung von Emotionen kann in vielfältigen Bereichen zum Einsatz kommen. So sind Kfz-Assistenzsysteme, die Müdigkeit aufgrund des Lenkverhaltens, der Pedalnutzung oder über ein Kamerasystem die Kopfhaltung oder das Blickverhalten kontrollieren, bereits auf dem Markt erhältlich.² Das Fraunhofer-Projekt EMOBIO will beispielsweise Systeme auf KI-Basis weiterentwickeln, die physiologische Daten für die Erkennung von (Mikro-)Emotionen verwenden, um so verschiedene Fahrerzustände zu ermitteln.³ Die EU ließ im Projekt iBorderCtrl einen Lügendetektor entwickeln, der Einreisende aus Drittstaaten automatisch aufgrund von Mikroexpressionen im Gesicht analysieren soll.⁴ Ebenso könnten KI-Systeme zur Emotionserkennung in vielen Bereichen der öffentlichen Daseinsvorsorge angewendet werden. In der psychologischen Betreuung werden vor allem von privaten Anbietern in den letzten Jahren Chatbots wie Woebot⁵ oder Wysa⁶ basierend auf KI-Sprachmodellen angeboten. Die Anwendungen sollen den

*Emotionserkennung:
neue Möglichkeiten
mit KI*

*Vielfältige
Anwendungsgebiete:
Sicherheit, öffentliche
Daseinsvorsorge,
Marketing ...*

¹ t3n.de/news/emotion-ai-maschinen-lernen-1149606/.

² bussgeldkatalog.org/muedigkeitserkennung.

³ hci.iao.fraunhofer.de/de/Human-Centered-AI/feinfuehlige-technik/KI-gestuetzte-Emotionserkennung.html.

⁴ mdr.de/wissen/naturwissenschaften-technik/eu-grenzschutz-kuenstliche-intelligenz-einreise-lugendetektor-100.html.

⁵ woebotehealth.com.

⁶ wysa.com.

Zugang zu *therapeutischen Angeboten* vor dem Hintergrund knapper Therapieplätze erleichtern und mittels strukturierter Übungen und Selbsthilfeinhalten kurzfristige Hilfe bei Depressionen oder Angststörungen bieten (Swartz 2023). In der Pflege wird ein Einsatz im Zusammenhang mit Pflege-Robotern und deren Probleme diskutiert (Manzeschke & Assadi 2023). Die Systeme könnten dort die Emotionen von Pflegebedürftigen erkennen, um damit das Interaktionsverhalten von Robotern zu steuern. Auch im Online-Marketing auf Social-Media- und Marktplatzplattformen liegt ein Einsatzgebiet. Hier können Produktempfehlungen an die emotionalen Zustände von Personen angepasst werden, um Kaufentscheidungen zu stimulieren (Zeng et al. 2016). Visuelle Emotionserkennung soll darüber hinaus die Lücke zwischen digitaler und analoger Welt im Bereich personalisierter Werbung schließen. Über Videoüberwachungen sollen etwa die Gesichter von Kund:innen im Einzelhandel erfasst werden, um ihre Reaktionen auf bestimmte Produkte zu analysieren und daraufhin maßgeschneiderte Werbung abspielen zu können (Meyer 2021, S. 42). Ho et al 2021 stellen eine Fokus-Verschiebung des affektiven Computings fest, weg von Diagnose und Erkennung psychischer Krankheiten hin zu mehr kommerziellen Anwendungen (**KI-Tutoren**).

Ein weites Einsatzfeld bieten zudem prinzipiell die Arbeitswelt und der Bildungsbereich. Im Bildungswesen könnte die Technologie zur Weiterentwicklung von Lernumgebungen genutzt werden – von der Unterstützung lernbeeinträchtigter Menschen über die flexible Anpassung von Lerninhalten an emotionale Reaktionen bis hin zur Konzentrationsüberwachung von Schüler:innen oder Studierenden (Büchling et al. 2019, Yuvarai 2025). In der Arbeitswelt lassen sich KI-Systeme zur Emotionserkennung sowohl für Bewerbungsprozesse als auch zur internen Personalentwicklung und Leistungskontrolle verwenden. Per Video aufgezeichnete Bewerbungsgespräche könnten automatisch nach Mimik, Gestik und Sprache ausgewertet werden und so Auskunft über die Persönlichkeiten der Bewerber:innen geben.⁷ Ebenso könnten sie betrieblich bei der internen Bewertung von Mitarbeiter:innen eingesetzt werden, um auf Grundlage einer Persönlichkeits- und Leistungsanalyse Personalentscheidungen zu treffen. Solche Systeme werden beispielsweise in den USA bereits eingesetzt. In der EU ist Emotionserkennung, die mit dem Arbeitsplatz oder dem Bildungsbereich in Zusammenhang steht, gemäß KI-Gesetz (AI Act) künftig verboten. Sie bergen das Risiko der Diskriminierung und können als Überwachungstechnologien eingesetzt werden, die auf intime zwischenmenschliche oder körperliche Äußerungen von Beschäftigten bzw. Schüler:innen abzielen. Dadurch könnten gesetzlich verankerte Rechte, etwa die Streikfreiheit, die gewerkschaftliche Vereinigungsfreiheit oder das Recht auf Gründung eines Betriebsrates, unterminiert werden.

Für alle KI-Systeme zur Emotionserkennung gilt, dass sie mit bestehendem Material wie Bildern oder Texten trainiert werden, um daraus Muster zu erkennen, denen Emotionen zugeordnet werden. Die Ableitung von Emotionen und Persönlichkeitsmerkmalen aus beobachtbarem Verhalten und Aussehen von Personen ist jedoch umstritten. Die EU-Kommission weist darauf hin, dass „im Hinblick auf die wissenschaftliche Grundlage von KI-Systemen, die darauf abzielen, Emotionen zu erkennen oder abzuleiten, [...] ernsthafte Bedenken“ bestehen. Emotionen sind nicht unmittelbar empirisch überprüfbar, weshalb ihr Ersatz durch messbare Stellvertretermerkmale (proxy features) nicht überprüft werden kann. Dadurch

**Hochrisiko-Anwendungen:
Arbeitswelt und
Bildungsbereich**

Herausforderungen:

- Zuverlässigkeit der Modelle
- Bias
- Diskriminierung
- Datenschutz

⁷ algorithmwatch.org/de/sprachanalyse-hr/.

steht die Verlässlichkeit der Modelle grundsätzlich in Frage. Zum einen müssen Trainingsdaten oftmals von Menschen mit bestimmten Emotionen klassifiziert werden, um eine KI zu trainieren. Dazu werden meist Klassifikationsansätze verwendet, in denen zwischen sechs sogenannten Basisemotionen unterschieden wird – Trauer, Freude, Angst, Ekel, Überraschung und Ärger –, denen eine bestimmte typisierte Ausdrucksweise – etwa eine Mimik – zugewiesen wird (Ekman 1971). Die Systeme abstrahieren damit weitgehend von der Vielfalt emotionaler Zustände, ihrer Repräsentation in Mimik, Interaktionsweisen oder Texten und ihrem Zusammenhang mit stabilen Persönlichkeitsmerkmalen zu Gunsten einer schematischen Zuordnung stereotyper Ausdrücke. Insbesondere für Modelle zur Gesichtserkennung konnte festgestellt werden, dass die zugeschriebenen Emotionen oftmals der Vielfalt emotionaler Ausdrücke in realen Alltagssituationen und in verschiedenen kulturellen Räumen nicht gerecht werden (Feldman Barrett et al. 2019). Daran anschließend eröffnet sich das Risiko der Diskriminierung. Denn weil die dominanten Modelle hinter emotionaler Mustererkennung eher stereotyp sind und sie meist mit historischem Datenmaterial trainiert werden, laufen sie Gefahr, Bevölkerungsgruppen und emotionale Ausdrucksweisen zu diskriminieren, die nicht als normal und durchschnittlich gelten oder bisher in bestimmten Datensätzen nicht repräsentiert waren. Eingesetzt zur Überwachung des öffentlichen Raums oder bei Grenzkontrollen bergen die Systeme die Gefahr, dass sie vor allem den Gesichtsausdrücken nicht-weißer Menschen negative Emotionen zuschreiben und sie damit als Gefährder:innen markieren (Rhue 2018). Eine Herausforderung ist auch der Datenschutz, da in vielen Anwendungsfällen die Verarbeitung (inkl. Speicherung und Analyse) sensibler personenbezogener Daten erfolgt, und zwar in einer Situation des Machtungleichgewichts zwischen Überwachenden und Überwachtem. Durch die zunehmende Nutzung generativer KI eröffnen sich zudem Risiken der emotionalen Manipulation durch KI-Chatbots. Affektive Systeme können beispielsweise verstärkend auf psychisch labile Jugendliche einwirken, die im Extremfall in den Suizid getrieben werden.⁸

Unabhängig von der Verlässlichkeit der Modelle und Trainingsdaten stellen sich insbesondere in wirtschaftlichen Anwendungsbereichen grundlegende gesellschaftspolitische Fragen. Im Online-Marketing kann die Personalisierung von Produkten und Dienstleistungen durch KI-basierte Emotionserkennungssysteme zur Manipulation von sozialem Verhalten führen, weil Kund:innen durch die Analyse unbewusster emotionaler Zustände zu Konsumententscheidungen bewegt werden. Oftmals werden die Daten zur emotionalen Analyse dabei in digitalen Diensten gesammelt, die Nutzer:innen wegen ganz anderer Zwecke – etwa der sozialen und politischen Teilhabe oder der Informationsbeschaffung – verwenden. Dies birgt die Gefahr, dass kulturelle und politische Öffentlichkeiten durch ökonomische Zwecke strukturiert werden, was Folgen auf die Funktionsfähigkeiten des demokratischen Austauschs und den sozialen Zusammenhalt haben kann. Beispielsweise ist nachgewiesen, dass die an Werbung orientierte, algorithmische Sortierung von Inhalten auf Social-Media-Plattformen den emotionalen Zustand von Nutzer:innen beeinflusst (Kramer et al. 2014) und damit auch politische Mobilisierung lenken kann.

Emotionalisierung politischer Öffentlichkeiten

⁸ the-decoder.de/wie-chatgpt-die-suizidalen-gedanken-eines-16-jaehrigen-bestraert.

RELEVANZ DES THEMAS FÜR DAS PARLAMENT UND FÜR ÖSTERREICH

In der KI-Verordnung (AI-Act) der Europäischen Union wurde 2023 ein Rechtsrahmen vorgelegt, der den Einsatz KI-basierter Emotionserkennungssysteme auch für Österreich regelt.⁹ Die Regulierung erfolgt entsprechend dem risikobasierten Ansatz des AI-Act mit verschiedenen Regulierungsinstrumenten, gestuft je nach Risikokategorie. In die Kategorie der nicht-akzeptablen Risiken fallen Emotionserkennungssysteme, die im Bereich des Arbeitsplatzes und in Bildungseinrichtungen eingesetzt werden sollen. Sie sind verboten, es sei denn, sie werden aus medizinischen und Sicherheitsgründen eingesetzt. Als Hochrisiko-KI-Systeme werden unter anderem solche Modelle bezeichnet, die im Bereich der Beschäftigung, des Personalmanagements, des Zugangs zu grundlegenden öffentlichen Diensten, in bestimmten Bereichen der Strafverfolgung und Grenzkontrollen eingesetzt werden. Derartige Systeme sind erlaubt, allerdings bestehen zahlreiche regulatorische Anforderungen, wie z. B. der kontinuierliche Betrieb eines Risikomanagementsystems.

Daraus ergibt sich, dass der Einsatz von KI-Systemen zur Emotionserkennung in Österreich in einigen oben beschriebenen, sensiblen Bereichen zwar bereits verboten oder unter strenge Auflagen gestellt ist. Allerdings ist ihr Einsatz in allen anderen Bereichen grundsätzlich erlaubt. Zudem ergeben sich gesetzgeberische Lücken und Grauzonen. Das Verbot betrifft nur KI-Systeme zur Emotionserkennung aus biometrischen Daten. Textbasierte Analysen bleiben erlaubt (EU-Kommission 2025).¹⁰ Die Verordnung erlaubt zudem die Erfassung physischer Zustände wie Schmerz oder Ermüdung, etwa für Fahrerassistenzsysteme. Die Abgrenzung zu Emotionen bleibt jedoch unscharf. Außerdem ist der Einsatz im Bereich des Arbeitsplatzes verboten, aber die Anwendung im Personalmanagement erlaubt. Welche Daten dabei beispielsweise von Personalabteilungen für das Personalmanagement erhoben werden dürfen und welche in den Bereich des Arbeitsplatzes fallen, dürfte in den folgenden Jahren Gegenstand von Rechtsprechungen und sozialen Aushandlungen werden. Die Klärung der Auslegung der Regulierung soll zwar durch delegierte Rechtsakte, gemeinsame Spezifikationen, Standardisierungen, Leitlinien etc. von verschiedenen Akteur:innen auf verschiedenen Ebenen innerhalb der EU erfolgen. Es kann aber auch vermutet werden, dass Gültigkeits- und Auslegungsfragen erst langfristige durch Gerichtsentscheidungen geklärt werden oder Regelungslücken bestehen bleiben.

*Grauzonen rechtlicher
Regulierung*

VORSCHLAG WEITERES VORGEHEN

In einer TA-Studie wären die Grenzen KI-basierter Emotionserkennungssysteme kritisch zu hinterfragen. Aufgrund der vielfältigen Einsatzmöglichkeiten solcher Systeme und der einsetzenden Regulierung durch die EU wäre es zudem lohnenswert, zu überprüfen, in welchen rechtlichen Grauzonen sie angewendet werden

⁹ Siehe Version vom 13. März 2024, [europarl.europa.eu/RegistreWeb/search/
simpleSearchHome.htm?references=P9_TA\(2024\)0138&language=de](http://europarl.europa.eu/RegistreWeb/search/simpleSearchHome.htm?references=P9_TA(2024)0138&language=de).

¹⁰ [digital-strategy.ec.europa.eu/de/library/commission-publishes-guidelines-prohibited-
artificial-intelligence-ai-practices-defined-ai-act](http://digital-strategy.ec.europa.eu/de/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act).

können, obwohl dadurch ethische bzw. soziale Probleme entstehen könnten. Beispielsweise bieten die teils unklaren Regulierungen in der Arbeitswelt Streitpotential, in welchen Bereichen solche KI-Systeme eingesetzt werden dürfen und wo nicht. Daran anschließend sollte untersucht werden, wo Betroffene aufgrund besonderer Abhängigkeitsverhältnisse und Vulnerabilität wenig Möglichkeiten haben, geltendes Recht einzuklagen. Dies betrifft etwa den Einsatz in den Grenzbehörden, aber auch in sozial wenig abgesicherten, prekären Arbeitsverhältnissen.

ZITIERTE LITERATUR

- Ekman, P. (1971). Universal and cultural differences in facial expression of emotion. In: Cole J. (Hg.): Nebraska Symposium on Motivation. Lincoln, 207–282.
- Feldman Barret, L. et al. (2019). Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. Psychological science in the public interest 20(1), 1–68, doi.org/10.1177/1529100619832930.
- Ho, M.-T., Mantello, P., Nguyen, H.-K. T., & Vuong, Q.-H. (2021). Affective computing scholarship and the rise of China: A view from 25 years of bibliometric data. *Humanities and Social Sciences Communications*, 8(1), 282. doi.org/10.1057/s41599-021-00959-8.
- Khare, S. K. et al. (2024). Emotion recognition and artificial intelligence: A systematic review (2014–2023) and research recommendations. *Information Fusion* 102, 102019. doi.org/10.1016/j.inffus.2023.102019.
- Kosinski, M. et all (2013). Private traits and attributes are predictable from digital records of human behavior. *PNAS* 110 (15), 5802-5805, doi.org/10.1073/pnas.1218772110.
- Kramer, Adam D. I. et al. (2014). Experimental Evidence of Massive-Scale Emotional Contagion through Social Network. *Proceedings of the National Academy of Sciences of the United States of America* 111 (29), 8788–8790, doi.org/10.1073/pnas.1320040111.
- Manzeschke, A., & Assadi, G. (2023). Künstliche Emotion – Zum ethischen Umgang mit Gefühlen zwischen Mensch und Technik. *Ethik in der Medizin*, 35(2), 201–219. doi.org/10.1007/s00481-023-00766-6.
- Meyer, R. (2019). Operative Porträts. Eine Bildgeschichte der Identifizierbarkeit von Lavater bis Facebook. Konstanz University Press: Konstanz.
- Meyer, R. (2021). Gesichtserkennung. Digitale Bildkulturen. Verlag Klaus Wagenbach: Berlin.
- Picard, R. (2000). Affective Computing. MIT Media Laboratory Perceptual Computing Section Technical Report No. 321.
- Rhue, L. (2018). Racial Influence on Automated Perceptions of Emotions, dx.doi.org/10.2139/ssrn.3281765.
- Swartz, H. A. (2023). Artificial Intelligence (AI) Psychotherapy: Coming Soon to a Consultation Room Near You? *American Journal for Psychotherapy*, (76) 2., 55-56, doi.org/10.1176/appi.psychotherapy.20230018.
- Yang, Tao et al. (2023): PsyCoT: Psychological Questionnaire as Powerful Chain-of-Thought for Personality Detection. In: Bouamor, H. et al. (Hg.): Findings 2023. Singapore. doi.org/10.18653/v1/2023.findings-emnlp.216.
- Yuvraj, R., Mittal, R., Prince, A. A., & Huang, J. S. (2025). Affective Computing for Learning in Education: A Systematic Review and Bibliometric Analysis. *Education Sciences*, 15(1), 65. doi.org/10.3390/educsci15010065.