



Alkohol



Cannabis



Illegal Drogen



Medikamente



Nikotin



Glücksspiel



Online / Media

Landesstelle für Suchtfragen Schleswig-Holstein e.V.

TERMINI

ÜBER UNS



SUCHTHILFE

Unterstützung und
Vernetzung
stationärer und
ambulanter
Suchthilfe-
Einrichtungen.



PRÄVENTION

Konzeptionierung
von Maßnahmen
sowie Schulung
und Unterstützung
der Prävention in
Schleswig-
Holstein.



BETRIEBLICHE SUCHTHILFE

Schulungen, Fort-
und Weiterbildung
im Bereich der
Betrieblichen
Suchtarbeit.



GLÜCKSSPIEL UND MEDIEN

Koordination der
Fachstellen für
Glücksspiel- und
Medien-
abhängigkeit.



SUCHT- SELBSTHILFE

Vernetzung der
Sucht-Selbsthilfe-
verbände und
Gremienarbeit in
Schleswig-
Holstein.



DOKU

Konzept und
Entwicklung von
Dokumentations-
strategien in der
Suchthilfe und
Suchtprävention.

KI und Sucht

Prof. Dr. Kerstin Prechel

Professorin für Digitalisierung und Ethik

Verantwortungsbewusste KI



“Und ein herzliches
Willkommen an unsere
Freunde aus Italien.”



Verantwortungsbewusste KI



„Natürlich kann ich
italienisch – das hörst du
nur. Ich könnte den
heutigen Vortrag über Ethik
der KI auf Italienisch
halten. Gar kein Problem.“



Verantwortungsbewusste KI



“Nein, das ist nicht betrügen. Jeder weiß
heute, dass man KI nutzen kann, um
eine Sprache zu übersetzen – und dann sprichst
du Spanisch, Holländisch, Chinesisch. Ich
möchte Ihnen sagen, dies ist ein Service für
Sie, so dass Sie Inhalte in ihrer
gewohnten Sprache erhalten. Ihr findet es
bestimmt gut, dass ich eure Sprachen sprechen
kann oder? Hmm?”

Warum reden wir über KI?

Und warum heute?

- Aktuelle schnelle technologische Entwicklung
- Ratsuchende nutzen KI bereits aktiv („ChatGPT sagt, …“)
- Suchtberatung trifft auf neue Kommunikationsformen
- Corona-Zeit: Schub für Online-Beratungen



Wer hat diese Woche schon KI benutzt?

Wer hat heute schon KI benutzt?

◆ Übersicht mit KI

Die DIAKO Nordfriesland gGmbH ist ein Anbieter von Leistungen für die psychische und seelische Gesundheit in Schleswig-Holstein, der verschiedene Fachbereiche wie Allgemeinpsychiatrie, Psychosomatik, Psychotherapie und Suchtmedizin abdeckt. Sie betreibt unter anderem Fachkliniken, Tageskliniken, Ambulanzen, Beratungsstellen und Einrichtungen für Eingliederungshilfe und Rehabilitation an zahlreichen Standorten, beispielsweise in Flensburg, Husum, Niebüll, Kiel und Bredstedt. Das Angebot reicht von Prävention und Beratung über Behandlung und medizinische/berufliche Rehabilitation bis hin zur Nachsorge. ⓘ

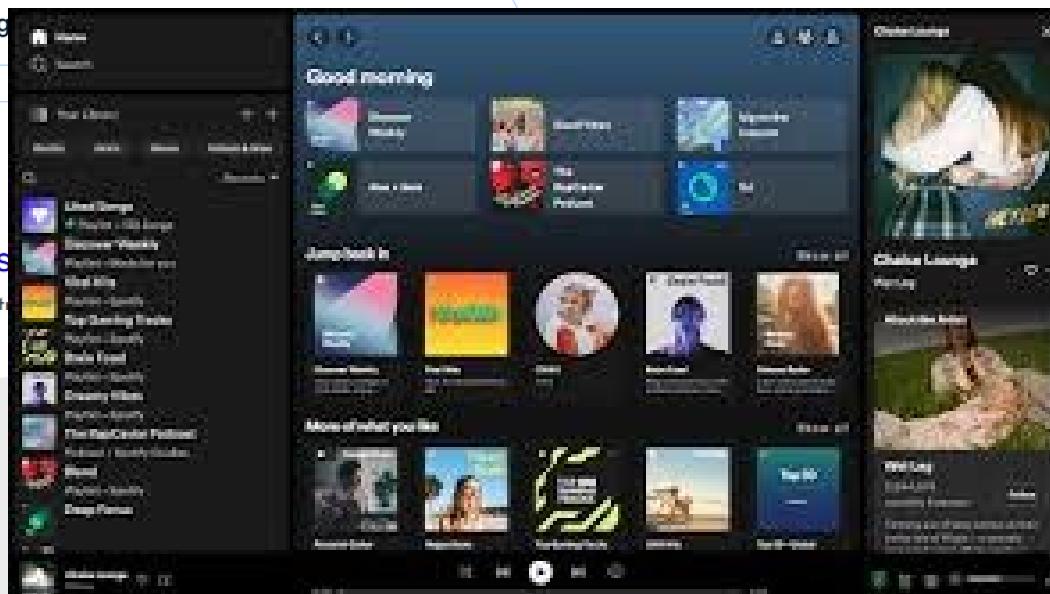
Angebote und Tätigkeitsbereiche

- Fachkliniken: Spezialisierte Kliniken für Psychiatrie, Psychosomatik und

Mehr anzeigen

 DIAKO Nordfriesland
<https://www.diako-nf.de> ::[DIAKO Nordfriesland - Psychiatrie, Ps](#)

Zu unseren Fachgebieten gehören Allgemeinpsychiatrie,



Fachklinik für Rehabilitation -
Abhängigkeitserkrankungen und ...



Fachklinik für Rehabilitation -
Abhängigkeitserkrankungen und...

↓ DIAKO Nordfriesland ::

DIAKO Husumer Insel - Nachsorge und
Eingliederungshilfe



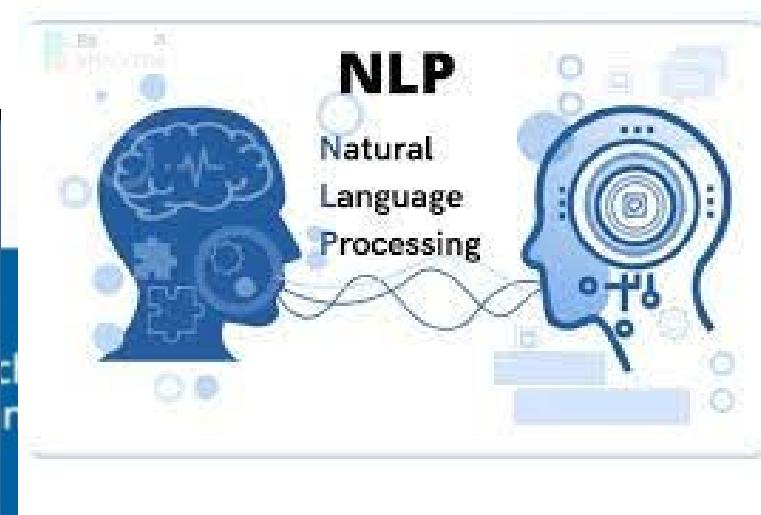
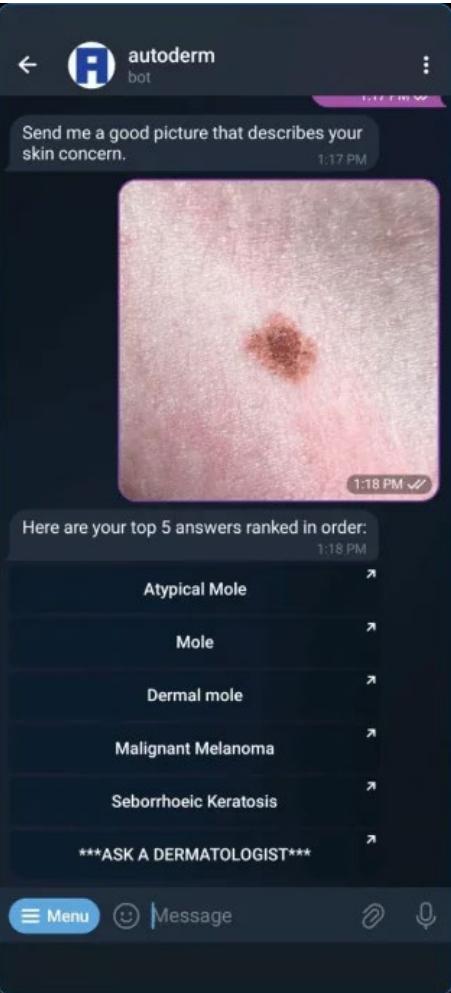
Unterstützung für Menschen mit psychischen
Erkrankungen. Die Husumer Insel ist ein Ange...



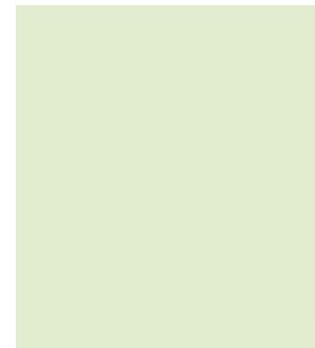
Eure Erfahrungswerte

Wo entlastet KI eure Arbeit?

Wo belastet sie euch?



KI im 1st und 2nd Level Kundensupport



Sozial Roboter Navel



Augmented Reality Brillen

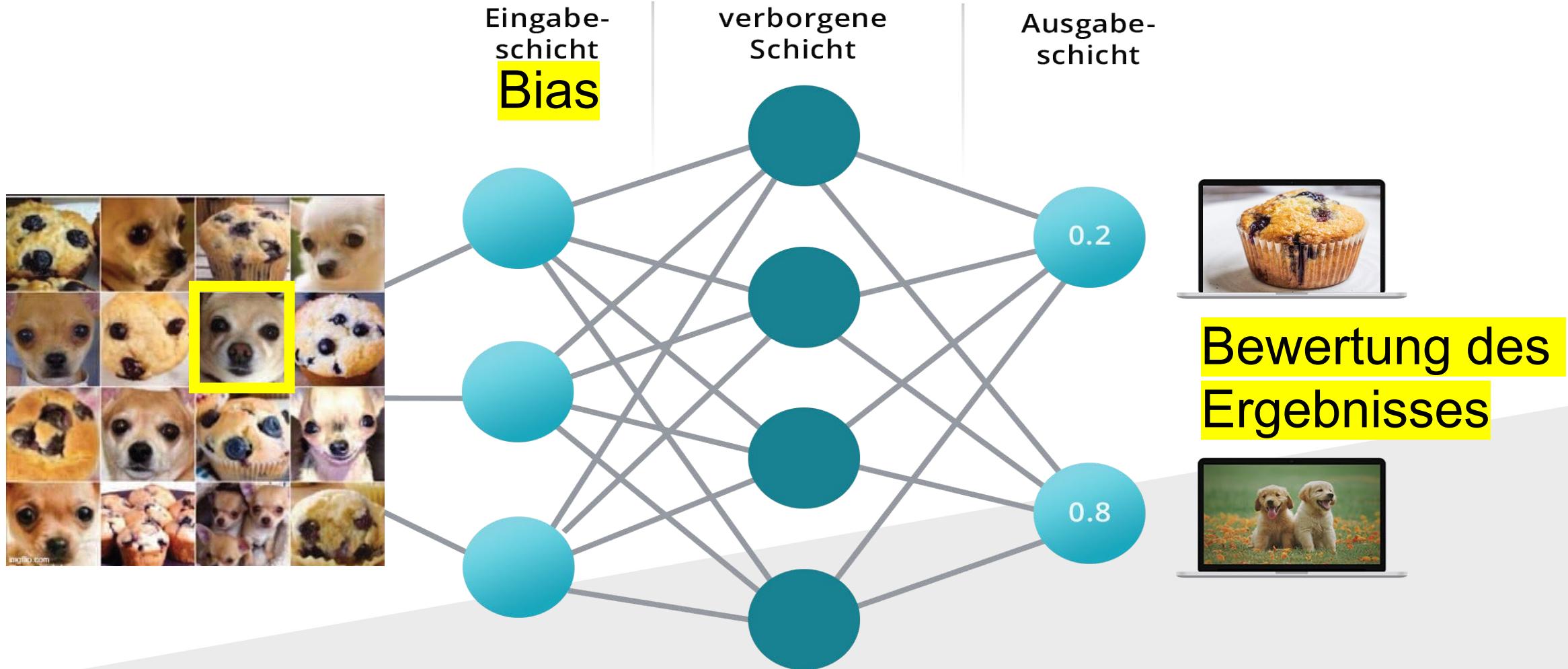


Kurzstudium Informatik – Schwerpunkt Künstliche Intelligenz

Digitale Souveränität als Präventionskonzept

Neuronale Netzwerke eine Form der Künstlichen Intelligenz

Black Box Problem



Katze oder Croissant?



Halluzinationen und statistische Papageien

Bendner et al 2021

ARTICLE | OPEN ACCESS

X in ↗ f ↗

On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?



Authors: [Emily M. Bender](#), [Timnit Gebru](#), [Angelina McMillan-Major](#), [Shmargaret Shmitchell](#) | [Authors Info & Claims](#)

FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency • Pages 610 - 623
<https://doi.org/10.1145/3442188.3445922>

PDF eReader

ent and
s variants,
of the
sing these
isks





KI als Spiegel der Gesellschaft

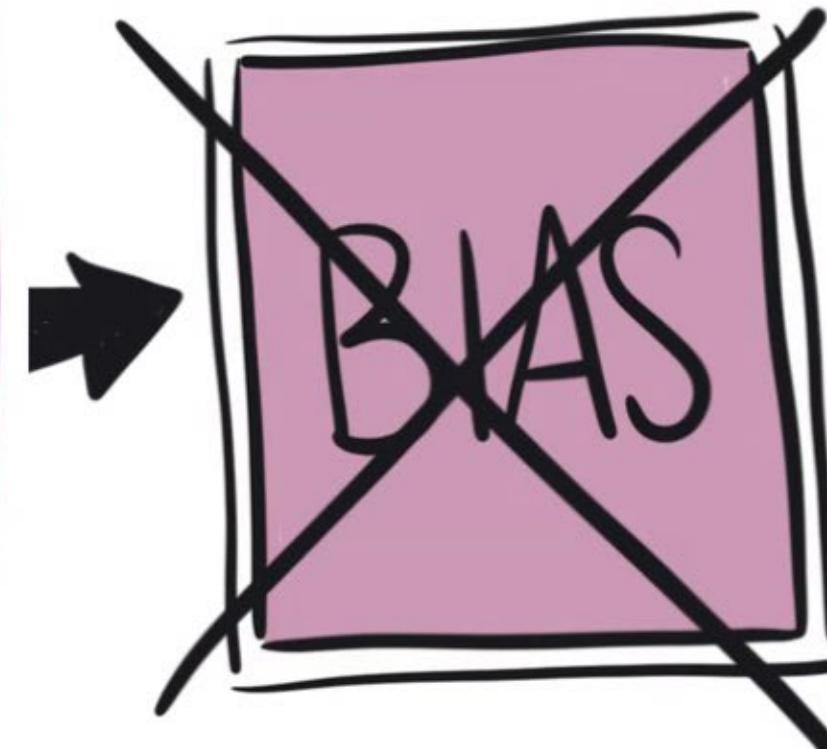
- Existierende Verzerrungen werden in Daten wiedergespiegelt
- Biases in LLMs verstärken Ungleichheit (z. B. Geschlecht, Ethnizität)
- Lösungen sind oft primär entwicklungsorientiert



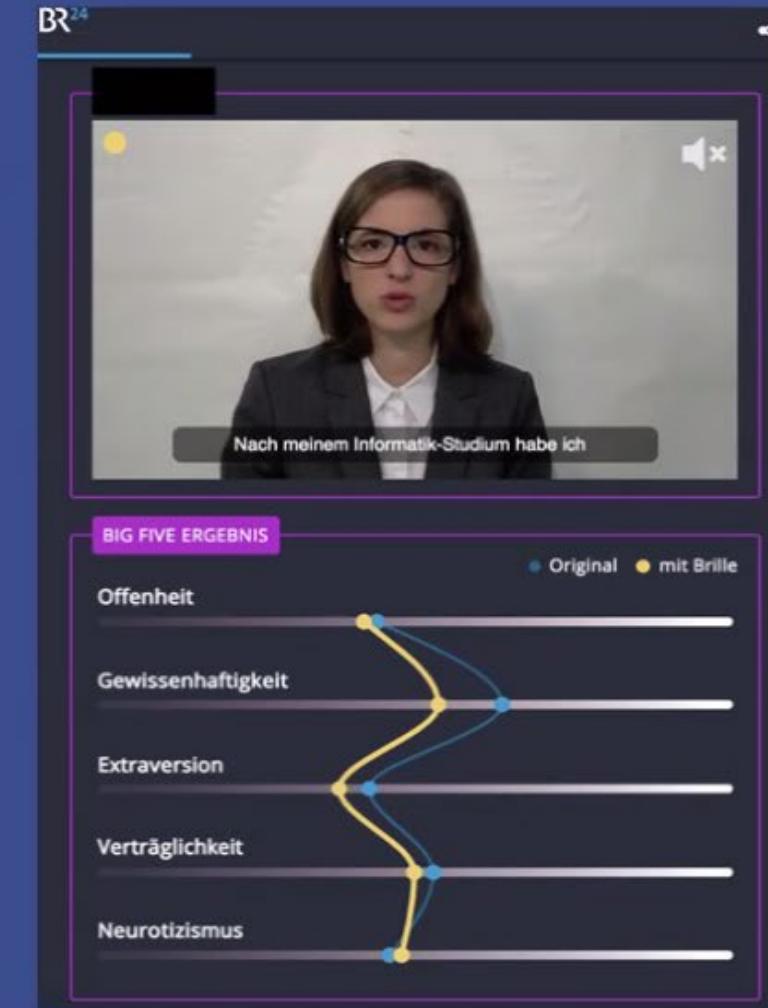
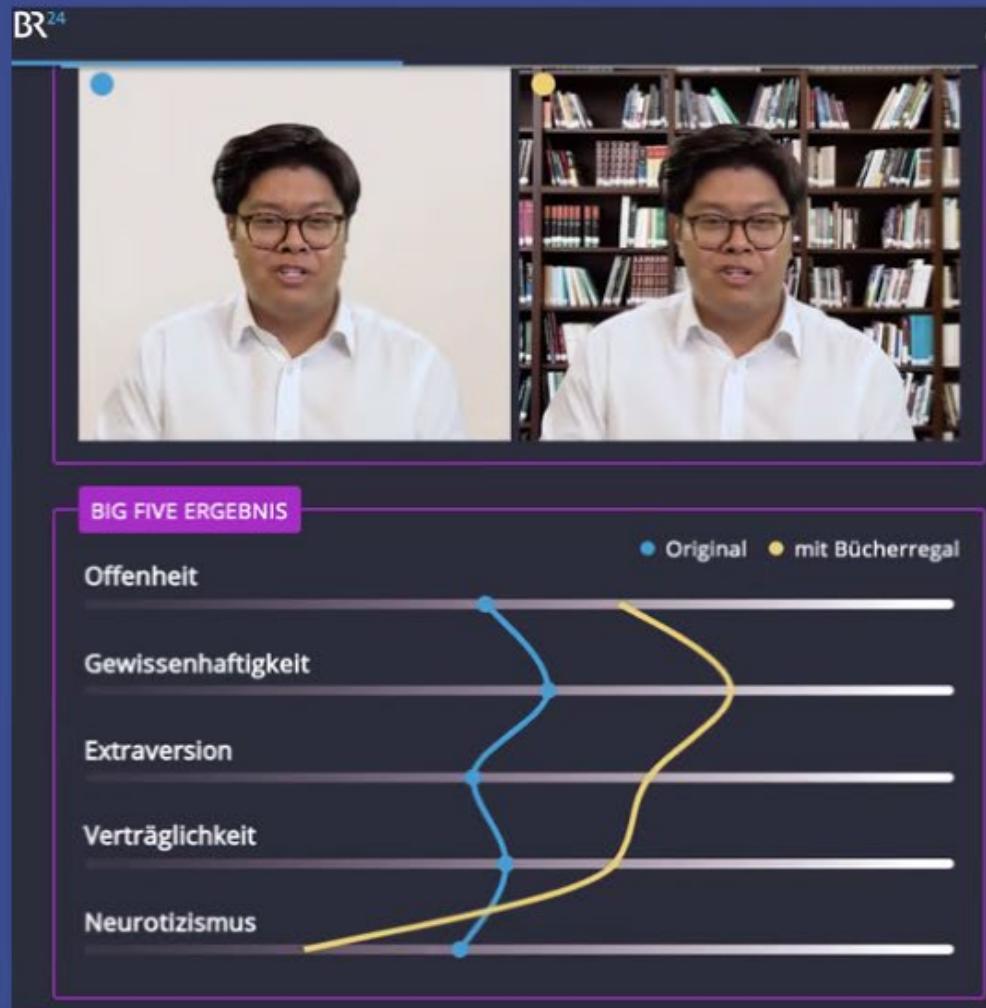


Die Herausforderungen von KI-Bias

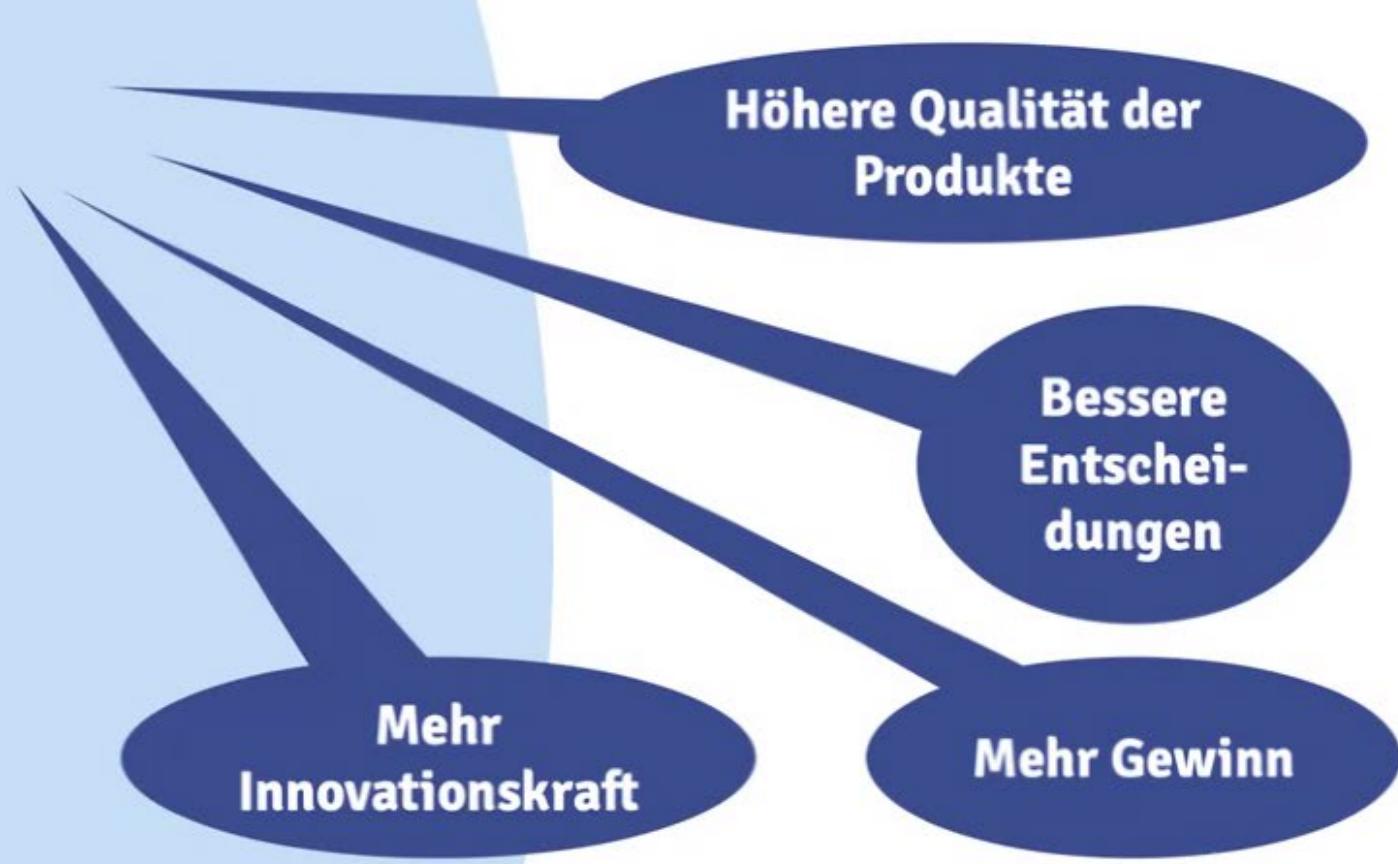
- **KI prägt kritische Entscheidungen**
Einstellungen, Kredite, Justiz
- **Bias verstärkt soziale Ungleichheit**
Geschlecht, Ethnizität, sozioökonomischer Status
- **Lösungen fokussieren auf Entwickler*innen**
für Endnutzende oft unzugänglich
- **Moralisches Disengagement**
„Bias ist unvermeidbar“ – verbreitete Haltung



KI zur Beurteilung von Bewerber*innen

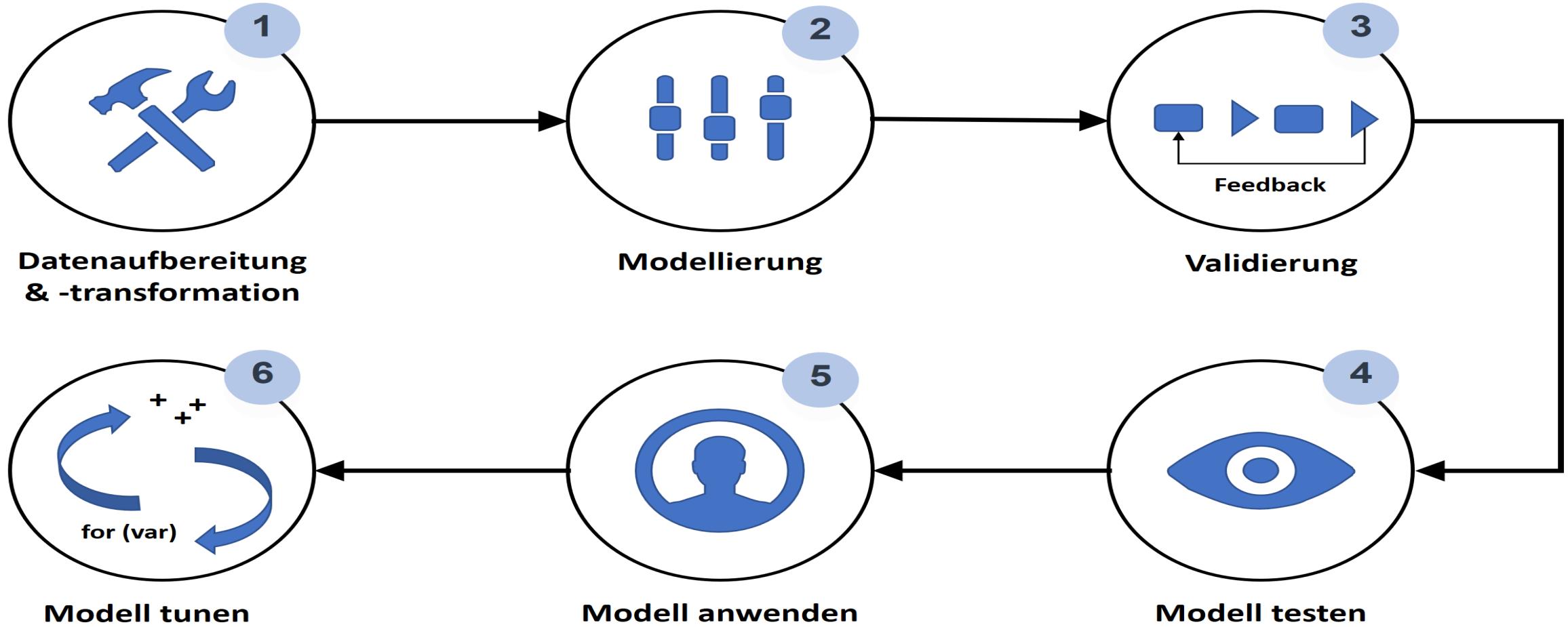


Das Ziel: Potenzial diverser Teams nutzen



Hunt, Vivian, Prince, Sara, Dixon-Fyle, Sundiatu, & Yee, Lareina. (2018). Delivering through Diversity. Retrieved from <https://www.mckinsey.com/business-functions/organization/our-insights/delivering-through-diversity>

Phasen bei der Einführung von KI-Lösungen



Kurzstudium Soziale Berufe – Schwerpunkt Künstliche Intelligenz

Digitale Souveränität als Präventionskonzept



Zwischen Entlastung und Entfremdung: KI im sozialen Bereich 2025

Eine kritische Betrachtung der
der Transformation der
Profession im digitalen Zeitalter
Zeitalter

Prof. Dr. Michael Garkisch: Mit dem Konzept Konzept der *Artificial Social Work* etabliert sich etabliert sich ein völlig neues Verständnis der Verständnis der Disziplin.

Laut IWF werden 60 % der Arbeitsplätze durch KI beeinflusst – von Automatisierung bestehender Prozesse bis zur Entstehung völlig neuer Tätigkeitsfelder und Anforderungsprofile

Neue Frühwarnsysteme können Krisen wie drohende Obdachlosigkeit oder Gewaltsituationen bereits erkennen, bevor sie eskalieren – ein präventives Potenzial, das uns bisher verschlossen blieb.

Bis zu 30 % der Arbeitszeit entfallen auf Administration – KI befreit wertvolle Ressourcen für das Wesentliche: die Beziehungsarbeit mit Klientinnen und Klienten.





TARGET BOARD 1 board +

Palantir RADIANT FALCON ...

Search targets...

Stage Target Status Battlespace

Recommend Taskings

Vessel OP RADIANT FALCON FIX MTS//HCS/SI/TK//OC/NF

VESSEL - ▲ No Target Location OP RADIANT FALCON FIX MTS//HCS/SI/TK//OC/NF

#0a196b / SA-5 IADS OP RADIANT FALCON ENGAGE MTS//HCS/SI/TK//OC/NF

Vessel OP RADIANT FALCON ENGAGE MTS//HCS/SI/TK//OC/NF

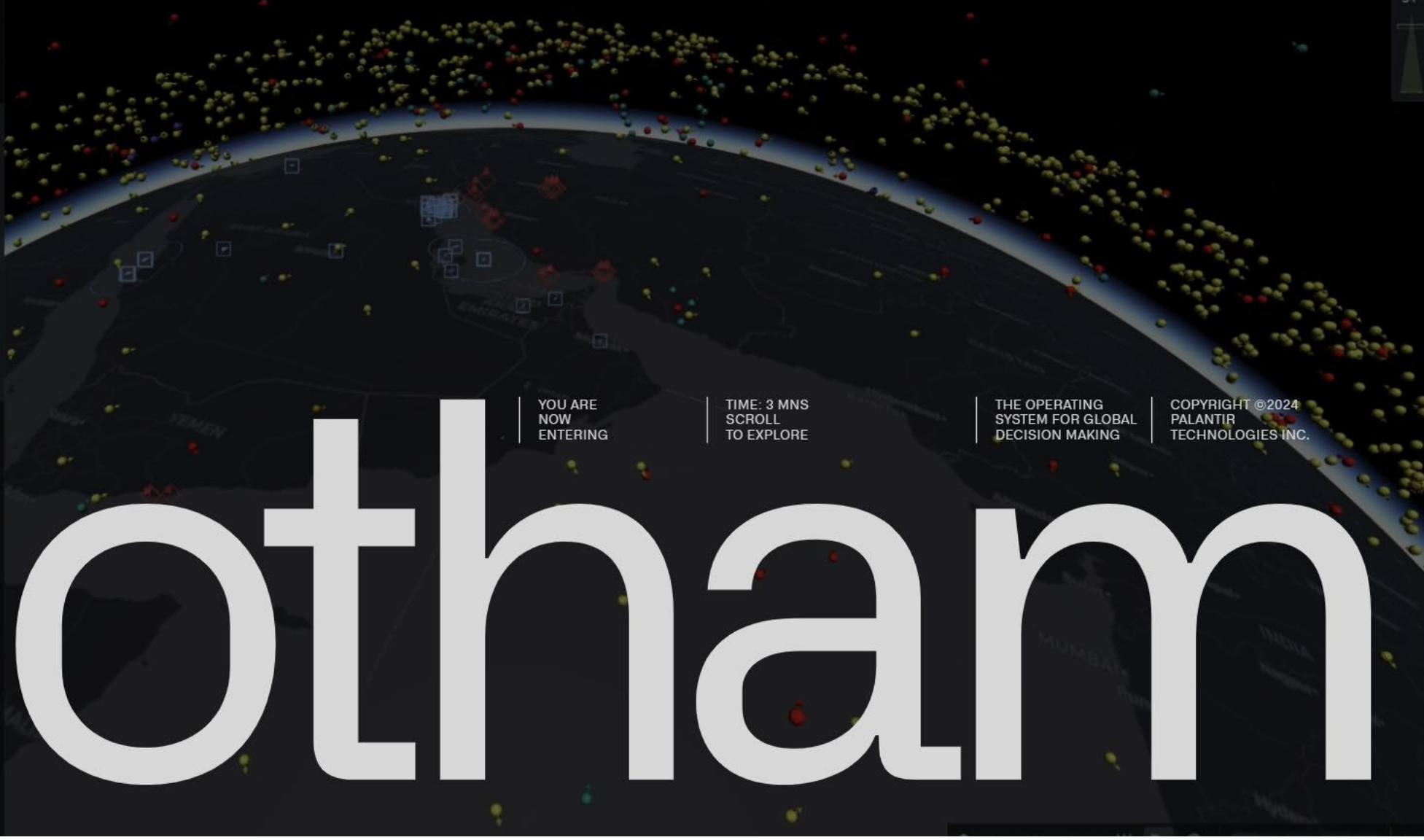
OP RADIANT FALCON FIX MTS//HCS/SI/TK//OC/NF

HQ-2 HQ-2 ▲ No Linked Intelligence OP RADIANT FALCON TRACK MTS//HCS/SI/TK//OC/NF

818d0 / SA-5 DS ▲ No Linked Intelligence OP RADIANT FALCON FIND MTS//HCS/SI/TK//OC/NF

M4241 / HQ-2 DS ▲ No Linked Intelligence OP RADIANT FALCON FIX MTS//HCS/SI/TK//OC/NF

#56c2dd / HQ-2 DS ▲ No Linked Intelligence OP RADIANT FALCON FIX MTS//HCS/SI/TK//OC/NF





Beschleunigt –
– aber zu
welchem Preis?
Preis?



Always-On-Mentalität

Zwanghaftes Checken von Systemen und E-Mails trotz Erschöpfung – ein klares Warnsignal für digitale Überlastung

Digitale Medien und permanente Benachrichtigungen führen zu weniger Aufmerksamkeit und verminderter Konzentrationsfähigkeit bei komplexen Aufgaben.

48 % der Beschäftigten berichten von hohem Leistungs- und Zeitdruck im digitalen Arbeitsumfeld. Die Erwartung ständiger Verfügbarkeit wird zur Belastung.



Technologie trifft Fürsorge

KI-gestützte Früherkennung psychischer Risiken



Datenanalyse

Erfassung relevanter Belastungsfaktoren



Früherkennung

Identifikation von Risikosignalen



Intervention

Gezielte Unterstützungsmaßnahmen

Psychische Gefährdungsbeurteilung mit KI-Unterstützung (Wirtschaftspsychologie Heute, 2024)

Kann KI süchtig machen?“

Was bedeutet „Sucht“ im Kontext KI?

KI kann verstärkende Mechanismen enthalten:

- sofortiges Feedback
- scheinbare Empathie
- personalisierte Antworten
- unendliche Verfügbarkeit

Risiko einer Medienabhängigkeit durch KI

- KI ersetzt soziale Interaktionen
- „KI versteht mich besser als Menschen“ (gefährliche Dynamik)
- Gefahr emotionaler Abhängigkeit
- Potenzial für „KI-bedingte Einsamkeit“
- Risiko: Verlängerte Nutzung von Substanzen durch entlastende Aussagen („Das ist nicht so schlimm...“)

Spezielle Herausforderungen für Jugendliche

- Vapes, Cannabis, Nikotin
- KI normalisiert Trends (Gefahr der Verharmlosung)
- Schnell verfügbare „Informationen“, ungeprüft und nicht altersgerecht

Kurzstudium Ethik und Recht – Schwerpunkt Künstliche Intelligenz

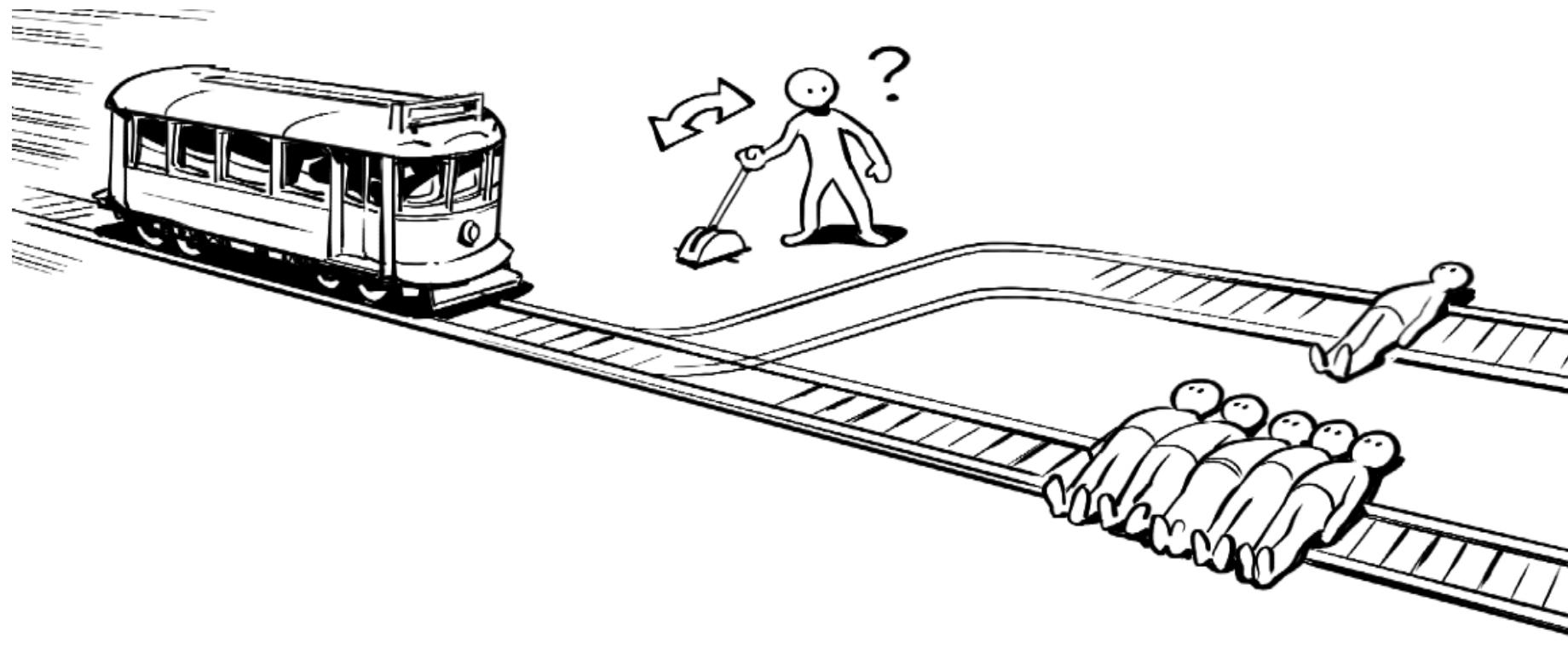
Digitale Souveränität als Präventionskonzept

Ethische Anforderungen an KI



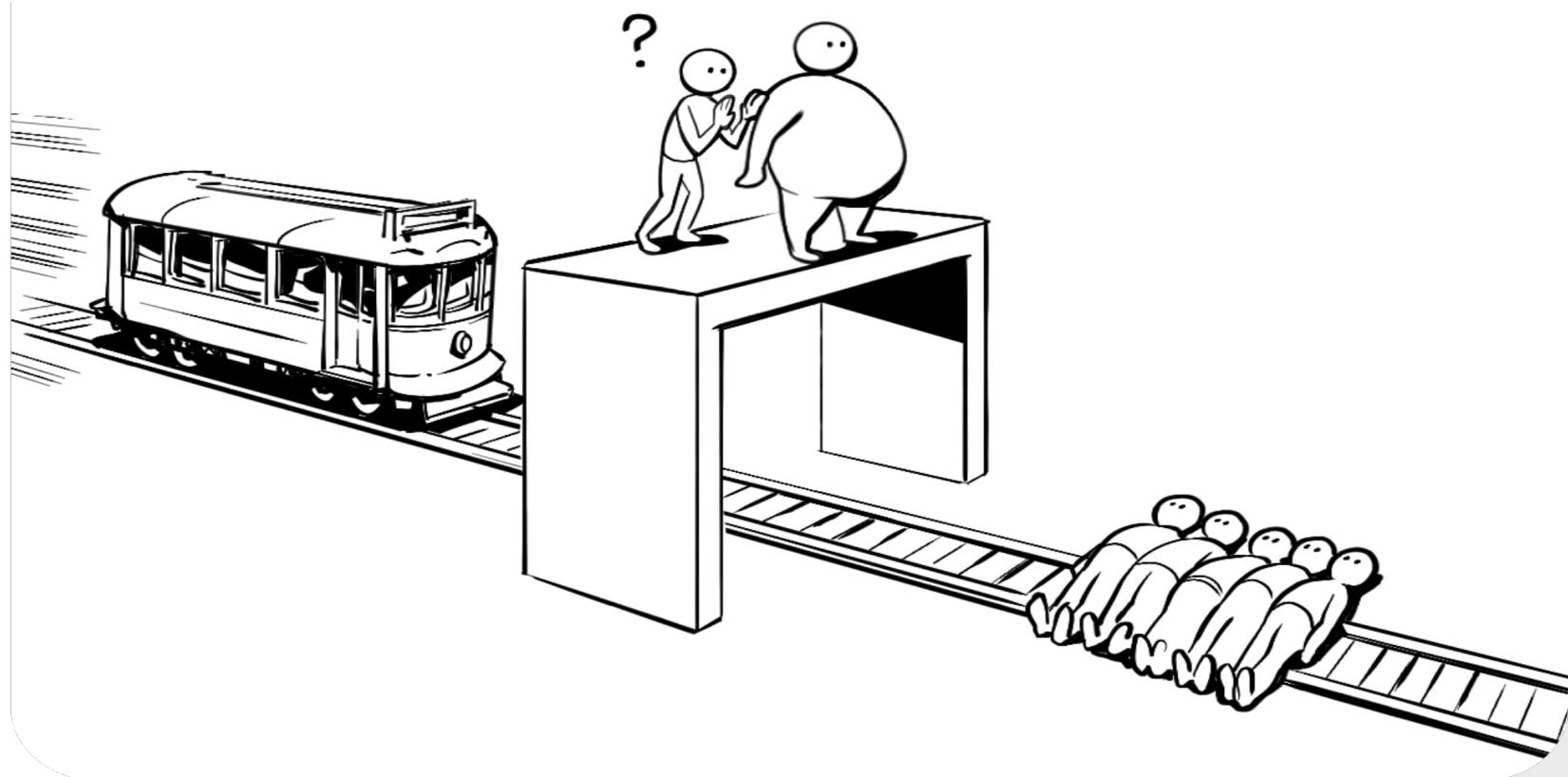
Trolley-Problem

Straßenbahndilemma



Fat Man

Brücke



Results

Most Saved Character



Most Killed Character



Saving More Lives



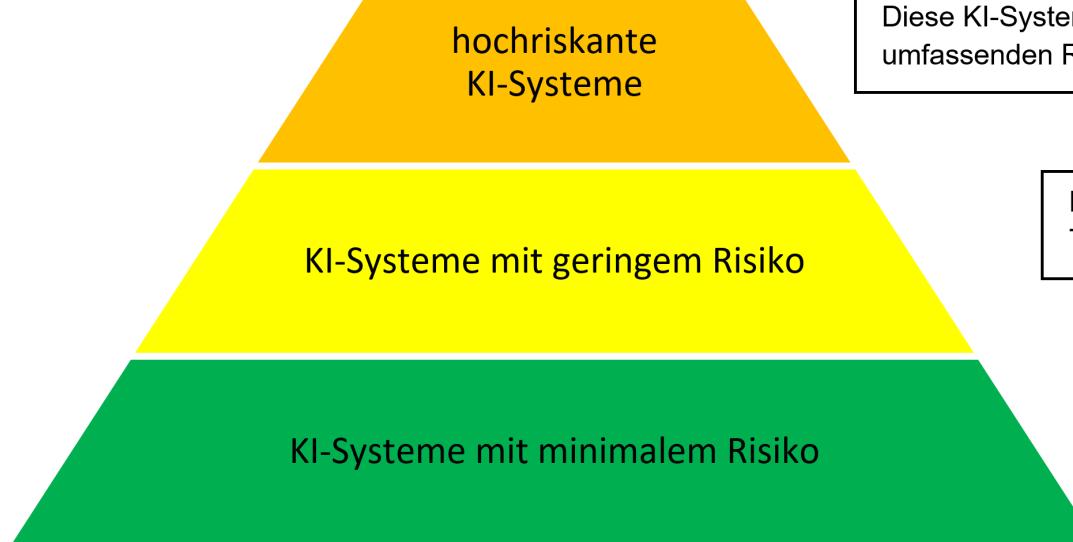
Protecting Passengers





**"WITH GREAT
POWER COMES
GREAT
RESPONSIBILITY."**

SPIDER-MAN



EU KI Verordnung: Ein Kompass

Diese KI-Systeme sind gem. Art. 5 der KI-VO verboten.

Diese KI-Systeme unterliegen gem. Art. 6 bis 51 KI-VO einer umfassenden Regulierung und weitreichenden Verpflichtungen.

Diese KI-Systeme unterliegen den Transparenzpflichten nach Art. 52 KI-VO.

Diese KI-Systeme unterfallen dem Anwendungsbereich der KI-VO nicht und unterliegen daher keinen diesbezüglichen Beschränkungen.

Schlaraffenland oder Dystopie?

Welche Veränderungen kommen auf die Suchthilfe zu?

- KI wird zu einem „unsichtbaren Dritten“ im Beratungsgespräch
- Ratsuchende kommen mit KI-Vorwissen – oft fehlerhaft
- Bedarf an **digitaler Gesundheitskompetenz** bei Fachkräften
- Beratungsstellen müssen strategischer mit KI umgehen



Grenzen & Gefahren

- Datenschutz & Vertraulichkeit
- Fehlentscheidungen / Falschinformationen
- Keine Diagnosen, keine empathische Beziehung
- Gefahr, vulnerable Menschen falsch zu beraten
- Bias & Diskriminierungen

Chancen für die Suchthilfe

- Entlastung im Hintergrund
- Bessere Zugänge für schwer erreichbare Gruppen
- Unterstützung bei Dokumentation & Anamnese
- niedrigschwellige Erstkontakte (Infos, Screening)



Bedeutung von Künstlicher Intelligenz in der Sozialen Arbeit; G. Linnemann, Julian Löhe, Beate Rottkemper. Soziale Passagen, 2023.

KI und Soziale Arbeit – Was ist heute möglich?; Aleksandra Poltermann, Eric Rudolph, P. Steigerwald, R. Lehmann, Sozialwirtschaft, 2024.

Wie kann Künstliche Intelligenz im Sozialmanagement sinnvoll eingesetzt werden?; Ulrich Gartzke, Michaela Preis, Blätter der Wohlfahrtspflege, 2024.

Mein neuer Teamkollege ist ein Roboter! Wie soziale Roboter die Zukunft der Arbeit verändern können; Sabine T. Koeszegi, A. Weiss, Management-Reihe Corporate Social Responsibility, 2021.

Künstliche Intelligenz; Christian Montag, Künstliche Intelligenzen als moralisch verantwortliche Akteure? 2020.

Künstliche Intelligenz in der Sozialen Arbeit – Zwischen Bedenken und Optionen; Reinhold Gravelmann, TUP - Theorie und Praxis der Sozialen Arbeit, 2024.

Künstliche Intelligenz in der Sozialen Arbeit; Olivier Steiner, D. Tschopp, Sozial Extra, 2022.

KI als Schlüsseltechnologie für Pflege, Teilhabe und Bildung; Laura Goretzka, David Große Dütting, Sozialwirtschaft, 2025.

Von Potenzialen zu spürbaren Mehrwerten; Julia Berner, Sandra Frings, Stefan Strunck, Sozialwirtschaft, 2025.

KI-Chatbots und digitale Suchtberatung: Eine Interviewstudie zu wahrgenommenen Chancen und Herausforderungen von KI-Expert*innen; Nima Montaseri, Matthias Morfeld, Mark Helle, Suchttherapie, 2025.

Wenn Algorithmen für uns entscheiden: Chancen und Risiken der künstlichen Intelligenz; M. Christen, Christen Clemens, Johann Čas, Tarik Abou-Chadi, A. Bernstein, et al. 2020.

Die Einstellung der Konsumenten gegenüber der Nutzung von neuen Technologien und künstlicher Intelligenz; Corina Pelau, I. Ene, Ruxandra Badescu, 2021.

[algorithmwach.org/de/suche/Soziale+Arbeit](https://algorithmwatch.org/de/suche/Soziale+Arbeit)

The image shows the homepage of algorithmwach.org. At the top, there's a navigation bar with icons for back, forward, search, and a star. Below it is a header with the website's URL. The main content area is divided into six sections, each containing a thumbnail image, a date, and a title. The first three sections are on the top row, and the last three are on the bottom row.

Date	Title	Image Description
4. SEPTEMBER 2025	Diskriminierende KI bei Job-Bewerbungen: AlgorithmWatch CH zeigt mit Forschungsprojekt neue Lösungen	Abstract geometric shapes in purple, blue, and white.
29. AUGUST 2025	Jobs	People sitting on the grass outdoors.
28. AUGUST 2025	Just Hiring! So vermeiden wir Diskriminierung durch Algorithmen in Einstellungsverfahren	Abstract geometric shapes in purple, blue, and white.
9. AUGUST 2025	Die KI-Revolution frisst ihre Gigworker	Illustration of people working at desks.
8. AUGUST 2025	Zivilgesellschaft kritisiert Unsicherheitspaket 2.0	Graphic with text: 'Dobrindt sagt Danke! Dein Selfie als Spende für staatliche Überwachung'.
7. AUGUST 2025	Biometrische Überwachung: Minister Dobrindt spielt mit Rechtsbrüchen	People standing on a sidewalk with white boxes around them.



Last Refresh: 14-Nov-2025 07...

Dashboard created by Nidhal Jegham

Sustainability Metrics

i Average Carbon Emissions (gCO2e)

Average Energy Consumption (Wh)

Average Water Consumption (Site & ...)

Average Water Consumption (Site, mL)

Average Water Consumption (Source,...)

Company

Alle

Model Size

Alle

Query Length

Medium

Number of Queries

100B

Reset Settings

Download Data (CSV)

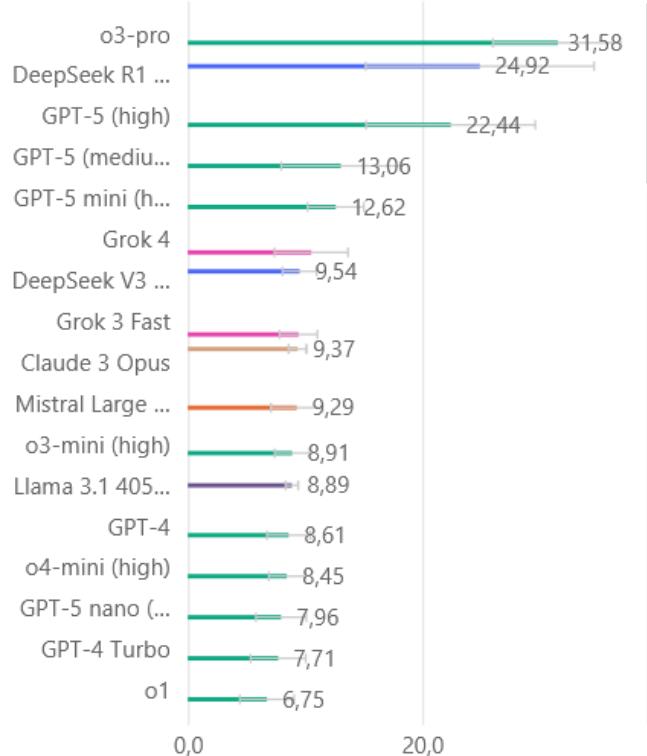
How Hungry is AI? | Dashboard Overview



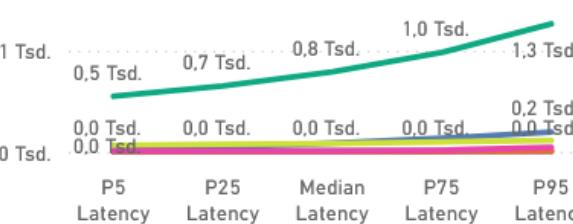
Last Refresh: 14-Nov-2025 07...

① Company
Select a ModelModel
Select a ModelHost
Select a ModelHardware
Select a ModelModel Size
Select...PUE
Select...WUE (Site)
(L/kWh)
Select...WUE (Source)
(L/kWh)
Select...CIF
(KgCO2/kWh)
Select...

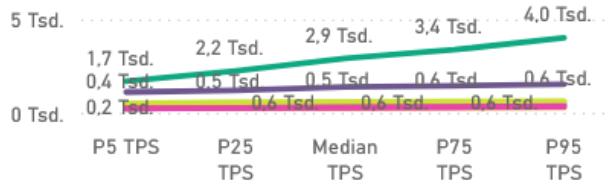
Average Energy Consumption (Wh) | Multiple Companies Models | Multiple Sizes | Medium Prompts



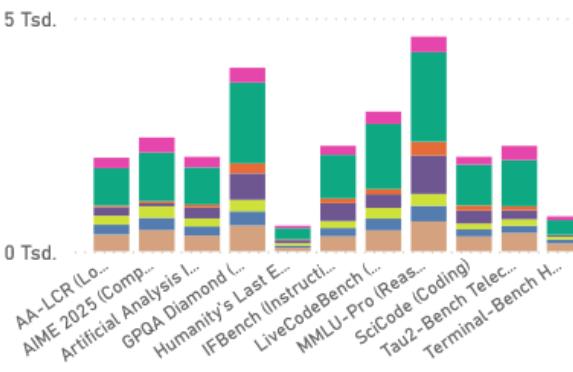
Latency (s) for Multiple Models



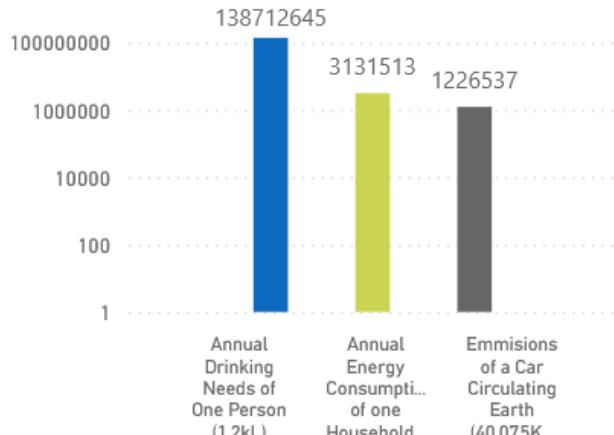
Tokens-per-Second (TPS) for Multiple Models



Performance Overview of Multiple Models



Environmental Footprint Equivalent of Multiple Models for 100B Queries



Auswahl Relevante Studien zum Thema KI

- (Buchempfehlung nicht wissenschaftlich: Marc-Uwe Kling "Views")
Abdul-Kader, S. A., & Woods, J. (2015). Survey on chatbot design techniques in speech conversation systems. *International Journal of Advanced Computer Science and Applications*, 6(7), 72-80.
- Alvarado, R. (2022). What kind of trust does AI deserve, if any? *AI and Ethics*. DOI: 10.1007/s43681-022-00224-x.
- Amann, J., Vetter, D., Blomberg, S.N., Christensen, H.C., Coffee, M., Gerke, S., Gilbert, T.K., Hagendorff, T., Holm, S., Livne, M., Spezzatti, A., Strümke, I., Zicari, R.V., & Madai, V.I. (2022). To explain or not to explain? Artificial intelligence explainability in clinical decision support systems. *PLOS Digital Health* 1(2), e0000016. DOI: 10.1371/journal.pdig.0000016 [Open Access].
- Arbelaez Ossa, L., Starke, G., Lorenzini, G., Vogt, J.E., Shaw, D.M., & Elger, B.S. (2022). Re-focusing explainability in medicine. *Digital Health*, 8. DOI: 10.1177/20552076221074488 [Open Access].
- Babushkina, D. (2022). Are we justified attributing a mistake in diagnosis to an AI diagnostic system? *AI and Ethics*. DOI: 10.1007/s43681-022-00189-9 [Open Access].
- Baile, W. F., Buckman, R., Lenzi, R., Globor, G., Beale, E. A., & Kudelka, A. P. (2000). SPIKES—A six-step protocol for delivering bad news: Application to the patient with cancer. *The Oncologist*, 5(4), 302-311.
- Benrimoh, D., Hawco, C., & Fratila, R. (2020). Using artificial intelligence to support patients facing cancer: From chatbot to clinical decision-making tools. *Current Oncology Reports*, 22(11), 1-8.
- Bickmore, T. W., & Schulman, D. (2011). Health dialog systems for patients and consumers. *Journal of Biomedical Informatics*, 39(5), 556-571.
- Bickmore, T. W., & Schulman, D. (2022). Practical approaches to comforting patients with relational agents. *Interacting with Computers*, 23(3), 279-288.
- Bleher, H., & Braun, M. (2022). Diffused responsibility: Attributions of responsibility in the use of AI-driven clinical decision support systems. *AI and Ethics*, 2(4), 747-761. DOI: 10.1007/s43681-022-00135-x [Open Access].
- Chen, H., Gomez, C., Huang, C.-M., & Unberath, M. (2022). Explainable medical imaging AI needs human-centered design: Guidelines and evidence from a systematic review. *npj Digital Medicine*, 5, 156. DOI: 10.1038/s41746-022-00699-2 [Open Access].
- Combi, C., Amico, B., Bellazzi, R., Holzinger, A., Moore, J.H., Zitnik, M., & Holmes, J.H. (2022). A manifesto on explainability for artificial intelligence in medicine. *Artificial Intelligence in Medicine*, 133, 102423. DOI: 10.1016/j.artmed.2022.102423 [Open Access].
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115-118.
- Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1), 1-15.
- Friedrich, A.B., Mason, J., & Malone, J.R. (2022). Rethinking explainability: Toward a postphenomenology of black-box artificial intelligence in medicine. *Ethics and Information Technology*, 24, 8. DOI: 10.1007/s10676-022-09631-4.
- Funer, F. (2022). Accuracy and Interpretability: Struggling with the Epistemic Foundations of Machine Learning-Generated Medical Information and Their Practical Implications for the Doctor-Patient Relationship. *Philosophy & Technology* 35(5). DOI: 10.1007/s13347-022-00505-7 [Open Access].
- Gardner, A., Smith, A.L., Steventon, A., Coughlan, E., & Oldfield, M. (2022). The Deception of Certainty: how Non-Interpretable Machine Learning Outcomes Challenge the Epistemic Authority of Physicians. *Medicine, Health Care and Philosophy*, 25, 167-178. DOI: 10.1007/s11019-022-10076-1 [Open Access].
- Hasani, N., Morris, M.A., Rhamim, A., Summers, R.M., Jones, E., Siegel, E., & Saboury, B. (2022). Ethical funding for trustworthy AI: Proposals to address the responsibility of funders to ensure that projects adhere to trustworthy AI practice. *AI and Ethics*, 2, 277-291.
- Grote, T., & Berens, P. (2020). On the ethics of algorithmic decision-making in healthcare. *Journal of Medical Ethics*, 46(3), 205-211.
- Hallowell, N., Badger, S., Sauerbrei, A., Nelläker, C., & Kerasidou, A. (2022). "I don't think people are ready to trust these algorithms at face value": Trust and the use of machine learning algorithms in the diagnosis of rare disease. *BMC Medical Ethics*, 23, 112. DOI: 10.1186/s12910-022-00842-4 [Open Access].
- Herzog, C. (2022). Trustworthy Artificial Intelligence in Medical Imaging. *PET Clin*, 17(1), 1-12. DOI: 10.1016/j.cpet.2021.09.007.
- Herzog, C. (2022). On the risk of confusing interpretability with explicability. *AI and Ethics*, 2, 219-225.
- Hetherley, J., Sparrow, R., & Howard, M. (2022). On the ethical and epistemological utility of explicable AI in medicine. *Philosophy & Technology*, 35, 50. DOI: 10.1007/s13347-022-00546-y [Open Access].
- Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., ... & Wang, Y. (2017). The virtues of interpretable medical artificial intelligence. *Cambridge Quarterly of Healthcare Ethics*. DOI: 10.1017/S0963180122000305 [Open Access].
- Jobin, A., lenca, M., & Vayena, E. (2019). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, 2(4), 230-243.
- Kawamleh, S. (2022). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389-399.
- Kawamleh, S. (2022). Against explainability requirements for ethical artificial intelligence in health care. *AI and Ethics*. DOI: 10.1007/s43681-022-00212-1.
- Kawamleh, S. (2022). Against explainability requirements for ethical artificial intelligence in health care. *AI and Ethics*. DOI: 10.1007/s43681-022-00212-1.
- Kemp, H., Freyer, N., & Nagel, S.K. (2022). Justice and the normative standards of explainability in healthcare. *Philosophy & Technology* 35, 100. DOI: 10.1007/s13347-022-00598-0 [Open Access].
- Kerasidou, C., Kerasidou, A., Buscher, M., & Wilkinson, S. (2022). Before and beyond trust: Reliance in medical AI. *Journal of Medical Ethics*, 48(11), 852-856. DOI: 10.1136/kmemb-2022-008522.
- Lütge, C. (2020). Transparency of AI in healthcare as a multilayered system of accountabilities: Between legal requirements and technical limitations. *Frontiers in Artificial Intelligence*, 5, 879603. DOI: 10.3389/frai.2022.879603.
- McDougall, R. J. (2019). The Ethics of AI and Robotics: A German Perspective. In *Ethics of Artificial Intelligence and Robotics: Fundamentals and Applications* (pp. 33-51). Springer.
- McTear, M. F., Callejas, Z., & Grisol, D. (2016). Computer knows best? The need for value-flexibility in medical AI. *Journal of Medical Ethics*, 45(3), 156-160.
- Milne-Ives, M., de Cock, C., Lim, E., Shehadeh, M. H., de Pennington, N., Mole, G., & Meinert, E. (2020). The role of conversational agents in healthcare: A literature review. *Journal of Medical Systems*, 40(7), 1-12.
- Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). The effectiveness of artificial intelligence conversational agents in health care: Systematic review. *Journal of Medical Internet Research*, 22(10), e20346.
- Ott, T., & Dabrock, P. (2022). From what to how: An initial review of publicly available AI ethics tools, methods, and research to translate principles into practices. *Science and Engineering Ethics*, 26(4), 2141-2168.
- Patch, J., Di, S., & Nelson, W. (2022). Transparent human – (non-)transparent technology? The Janus-faced call for transparency in AI-based health care technologies. *Frontiers in Genetics* 13, 902960. DOI: 10.3389/fgene.2022.902960 [Open Access].
- Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., ... & Ng, A. Y. (2017). Opening the black box: The promise and limitations of explainable machine learning in cardiology. *Canadian Journal of Cardiology* 38(2), 204-213. DOI: 10.1016/j.cjca.2021.09.004 [Open Access].
- Salahuddin, Z., Woodruff, H.C., Chatterjee, A., & Lambin, P. (2022). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. *arXiv preprint arXiv:1711.05225*.
- Sand, M., Durán, J.M., & Jongsma, K.R. (2022). Transparency of deep neural networks for medical image analysis. A review of interpretability methods. *Computers in Biology and Medicine* 140, 105111. DOI: 10.1016/j.combiomed.2021.105111 [Open Access].
- Schmitz, R., Werner, R., Repici, A., Bisschops, R., Meining, A., Zornow, M., Messmann, H., Hassan, C., Sharma, P., & Rösch, T. (2022). Responsibility beyond design: Physicians' requirements for ethical medical AI. *Bioethics* 36(2), 162-169. DOI: 10.1111/bioe.12887 [Open Access].
- Shickel, B., Tighe, P. J., Bihorac, A., & Rashidi, P. (2018). Artificial intelligence in GI endoscopy: Stumbling blocks, gold standards and the role of endoscopy societies. *Gut* 71(3), 451-454. DOI: 10.1136/gutjnl-2020-323115.
- Starke, G., & van den Brule, R., Elger, B.S., & Haselager, P. (2022). Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. *IEEE Journal of Biomedical and Health Informatics*, 22(5), 1589-1604.
- Stärke, G., & lenca, M. (2022). Intentional machines: A defence of trust in medical artificial intelligence. *Bioethics* 36, 154-161.
- Topol, E. J. (2019). Misplaced trust and distrust: How not to engage with medical artificial intelligence. *Cambridge Quarterly of Healthcare Ethics*. DOI: 10.1017/S0963180122000445 [Open Access].
- Timmermann, C., & Steger, F. (2022). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44-56.
- Verdicchio, M., & Perin, A. (2022). Explicability of artificial intelligence in radiology: Is a fifth bioethical principle conceptually necessary? *Bioethics* 36(2), 143-153. DOI: 10.1111/bioe.12918 [Open Access].
- Wadden, J.J. (2022). When doctors and AI interact: On human responsibility for artificial risks. *Philosophy & Technology* 35, 11. DOI: 10.1007/s13347-022-00506-6 [Open Access].
- Winter, P.D., & Carusi, A. (2022). Defining the undefinable: The black box problem in healthcare artificial intelligence. *Journal of Medical Ethics* 48(10), 764-768. DOI: 10.1136/medethics-2021-107529.
- Winter, P.D., & Carusi, A. (2022). If you're going to trust the machine, then that trust has got to be based on something: Validation and the co-constitution of trust in developing artificial intelligence (AI) for the early diagnosis of pulmonary hypertension (PH). *Science & Technology Studies* 35(4), 58-77. DOI: 10.23987/sts.102198 [Open Access].
- Yoon, C.H., Torrance, R., & Scheinerman, N. (2022). (De)troubling transparency: Artificial intelligence (AI) for clinical applications. *Medical Humanities*. DOI: 10.1136/medhum-2021-012318.
- Yu, K. H., Beam, A. L., & Kohane, I. S. (2018). Machine learning in medicine: Should the pursuit of enhanced interpretability be abandoned? *Journal of Medical Ethics* 48(9), 581-585. DOI: 10.1136/medethics-2020-107102 [Open Access].
- Zimmermann, C., & Steger, F. (2022). Artificial intelligence in healthcare. *Nature Biomedical Engineering*, 2(10), 719-731.