

I. Künstliche Intelligenz in der radiologischen Diagnostik: Medizinische Aspekte

1. Einführung und begriffliche Einordnung

Der Begriff »künstliche Intelligenz« (KI) beschreibt datenverarbeitende Systeme, die in der Lage sind, aus großen und komplexen Datenbeständen zu lernen, Muster zu erkennen und Entscheidungen basierend auf diesen Informationen abzuleiten. Hierdurch wird intelligentes Verhalten nachgebildet, wobei die hierfür verwendeten Ansätze historisch vielfältig waren. Auch wenn das Feld der künstlichen Intelligenz weitaus mehr umfasst, werden in Hinsicht auf die heutige und absehbar zukünftige Anwendung in der Radiologie mit dem Begriff künstliche Intelligenz vor allem Methoden des maschinellen Lernens, insbesondere des tiefen maschinellen Lernens mit neuronalen Netzwerken (*deep learning*), in Verbindung gebracht (Moawad, 2022). Letztere zeigen beeindruckende Leistungen insbesondere in der Bildanalyse, aber auch in der Sprachverarbeitung. Eine weitere, momentan viel Beachtung findende Variante der KI sind generative Verfahren, welche nach einer Aufforderung selbstständig Texte oder Bilder basierend auf gelernten Assoziationen zwischen Begriffen erstellen (Koochi-Moghadam, 2023). Gemein ist diesen KI-Verfahren, dass sie Muster in hochdimensionalen Daten gelernt haben und basierend auf diesem »Wissen« (welches in den entsprechenden Modellparametern verankert, jedoch nicht explizit abrufbar ist oder erklärt werden kann) Entscheidungen auf Grundlage neuer Daten treffen können. Das heißt, die zentrale Fähigkeit solcher KI-Modelle ist ein selbstständiger Lernprozess auf Basis vorhandener Trainingsdaten, wobei die hierdurch gelernten Regeln auch auf neue Situationen, Fälle oder Konstellationen generalisieren (Chen, 2023). Dementsprechend wird eine Reihe von Verfahren, auch wenn sie sich für Nutzer*innen oft ähnlich präsentieren, nicht

dem maschinellen Lernen zugerechnet und daher nicht unter dem Begriff KI im engeren Sinne verstanden. Dies gilt für einfache Automatisierungen, regelbasierte Entscheidungsbäume oder deterministische Maschinen, die lediglich vorprogrammierte Abläufe abarbeiten, ohne selbstständig dazuzulernen. Beispiele hierfür wären unter anderem Segmentierung von Bilddaten in der Radiologie anhand von Schwellwerten oder, vielleicht am eindrucklichsten, klassische Ansätze zur Routenplanung per Navigationssystem, welche über automatisierte (Optimierungs-)Prozesse verfügen, jedoch nicht über KI-Qualitäten im Sinne eines Lernens aus vorhandenen Daten mit Generalisierung auf neue Situationen.

Während die öffentliche Diskussion zu »KI in der Medizin« oft von der Idee geprägt ist, dass ein Computer mehr oder weniger selbstständig Entscheidungen über Diagnosen und damit Therapien trifft (Heinrichs, 2020), d. h., analog zu einem selbstfahrenden Auto agiert, muss angemerkt werden, dass die Nutzung von KI-Verfahren im medizinischen Bereich deutlich komplexer ist.

Versteckte KI arbeitet oft im Hintergrund in medizinischen Geräten und verbessert beispielsweise Bildqualität bei Magnetresonanztomographien (MRT), Computertomographien (CT) und Ultraschall. Sie rekonstruiert Aufnahmen, entfernt Störungen oder optimiert Kontraste, ohne dass Anwendende dies explizit bemerken (Mazurowski, 2019). Auch in der Verarbeitung und Aufbereitung von Daten spielt KI schon heute eine relevante Rolle. Wie in der Optimierung von Bilddaten steht sie hier jedoch im Hintergrund, da die entsprechenden Verfahren nur genutzt werden, um Ärzt*innen möglichst optimale Informationen und Daten bereitzustellen – nicht jedoch, um diese selbst zu bewerten. In einer einfachen Analogie können diese KI-Ansätze daher mit dem Autofokus einer Smartphone-Kamera verglichen werden: Sie übernehmen Hintergrundprozesse, vereinfachen damit den Nutzer*innen die Arbeit und verbessern das Ergebnis.

Als ein Sonderfall können »closed-loop KI-Systeme« angesehen werden, welche Geräte wie Beatmungsmaschinen, Infusionssysteme oder Überwachungsmonitore selbstständig steuern und anhand von Patientendaten Parameter in Echtzeit anpassen (Wang, 2025). Die Abgrenzung zu nicht-intelligenten, regelbasierten Systemen ist dabei in der Regel schwierig.

Autonome Systeme zur Diagnose, Therapie und Prognose, welche Ärzt*innen nicht nur unterstützen, sondern sie potenziell ersetzen, erscheinen auch mittelfristig eher unwahrscheinlich (Langlotz, 2019). Große Hoffnungen werden jedoch in unterstützende Systeme gelegt, welche Daten wie Laborergebnisse oder bildgebende Befunde nicht nur verarbeiten, sondern auch auswerten (Hedderich, 2021). Solche Anwendungen haben anhand einer großen Anzahl von vorherigen Fällen Zusammenhänge gelernt, welche es erlauben, auch bei neuen Einzelfällen Hinweise auf mögliche Befunde oder Diagnosen zu geben. Diese (Vorab-)Informationen sollten dann zu einer Entlastung bzw. Effizienzsteigerung bei Ärzt*innen führen. Dies gilt insbesondere für Disziplinen wie die Radiologie, in der sehr ausge-reifte Algorithmen zur Bildverarbeitung (*»computer vision«*) nicht nur eine erste Beurteilung der Bilder erlauben, sondern insbesondere auch unerwartete Nebenergebnisse für die ärztliche Beurteilung hervorheben können.

Eine Variante solcher Ansätze sind in Assistenzsystemen zu sehen, welche Diagnostik und Therapieplanung unterstützen, indem sie klinisches Wissen, Leitlinien und aktuelle Studien systematisch auswerten und Ärzt*innen basierend auf diesen Informationen konkrete Vorschläge machen, etwa für Risikoeinschätzungen oder Therapieoptionen (Bizzo, 2019). Da solche Systeme nur medizinisches Erfahrungswissen auswerten und verknüpfen, jedoch nicht auf Primärdaten trainiert werden, können sie als eine Art indirekte medizinische KI angesehen werden.

Zusammenfassend muss festgehalten werden, dass »KI in der Medizin« ein sehr heterogenes Feld umfasst, was nicht zuletzt dem Umstand geschuldet ist, dass der Begriff KI oft sehr unscharf gebraucht wird. So nehmen ethische und rechtliche Debatten in diesem Zusammenhang oft die momentan noch nicht in greifbarer Nähe erscheinenden autonomen Systeme in den Fokus (Eickhoff, 2021). Für den praktischen Alltag heute und in der absehbaren Zukunft dürften aber vor allem KI-Anwendungen im Vordergrund stehen, welche ärztliche Entscheidungen unterstützen, indem z. B. Daten aufbereitet oder analysiert werden. Aufgrund der Vorreiterrolle, die der Radiologie in Hinblick auf die Einbindung von KI in der Medizin zukommt, fokussieren die folgenden Abschnitte vor allem die Anwendung von KI in diesem Bereich.

2. Historische Einordnung

Auch wenn die aktuelle mediale Präsenz des Themas und öffentliche Debatte etwas anderes vermuten lassen, ist künstliche Intelligenz keine neuartige Entwicklung der letzten Jahre und weder an sich noch für die Anwendung in der Medizin ein gänzlich neues Thema. Tatsächlich reichen die Anfänge und konzeptionellen Grundlagen von künstlicher Intelligenz bis in die 1950er Jahre zurück. Als einer der Wegbereiter der künstlichen Intelligenz gilt der britische Mathematiker Alan Turing, der in seinen Arbeiten die relevanten theoretischen Grundsteine künstlicher Intelligenz legte. In seiner wegweisenden, im Jahr 1950 veröffentlichten Arbeit »Computing Machinery and Intelligence« (Turing, 1950) stellte er die Frage »Can machines think?« und präsentierte das Konzept des sogenannten »imitation game«. Bei diesem später als »Turing-Test« bekannt gewordenen Gedankenspiel handelt es sich um ein theoretisches Experiment, das zur Beantwortung der Frage herangezogen wird, ob eine Maschine ein dem Menschen gleichwertiges Denkvermögen aufweist. Hierbei kommuniziert ein*e menschliche*Fragesteller*in ohne Sicht- und Hörkontakt über schriftliche Nachrichten zum einen mit einer menschlichen Person und zum anderen mit einer Maschine. Ist der oder die Fragesteller*in nach intensiver Befragung nicht dazu in der Lage, zu differenzieren, bei welchem Gesprächspartner es sich um die menschliche Person und bei welchem es sich um die Maschine handelt, hat die Maschine den Turing-Test bestanden und gilt als (künstlich) intelligent. Trotz diverser Kritik findet der Turing-Test bis heute Beachtung und bietet zwar keine umfassende Definition, aber zumindest einen gut verständlichen Zugang, um den abstrakten Begriff der künstlichen Intelligenz besser zu fassen. Obwohl das Bestehen des Turing-Tests lange als unerreicht galt, geht man heute davon aus, dass moderne sprachbasierte KI-Systeme (*large language models*) dazu in der Lage sind, den Turing-Test zumindest bestimmten Interpretationen zufolge zu bestehen (Jones 2025). Nach den theoretischen Vorarbeiten von Turing gilt die im Jahr 1956 veranstaltete »Dartmouth Summer Research Project on Artificial Intelligence«-Konferenz als eigentliche Geburtsstunde der künstlichen Intelligenz, bei der erstmals der Begriff »künstliche Intelligenz« geprägt wurde und von einigen führenden Mathematikern und Wissenschaftlern relevante Konzepte dieses neuen Feldes kon-

stituiert wurden. Die weitere Entwicklung der künstlichen Intelligenz in den kommenden Jahren war vor allem durch frühe Entwicklungen geprägt, wie beispielsweise einfache Entscheidungssysteme auf Basis von Entscheidungsbäumen, oder den Prototyp eines ersten Chat-Computers mit dem Namen ELIZA Mitte der 1960er Jahre (Weizenbaum, 1966). In diese Zeit fallen auch erste Versuche, KI auf medizinische Bilddaten anzuwenden. So nutzte etwa Gwilym S. Lodwick ein auf dem Bayes-Theorem beruhendes probabilistisches System, um Knochentumoren anhand bestimmter Bildeigenschaften in Röntgen-Bildern zu diagnostizieren (1965). Während die theoretischen Grundlagen der heute gebräuchlichen künstlichen neuronalen Netzwerke bereits in den 1940er Jahren durch die Neurowissenschaftler Warren S. McCulloch und Walter Pitts geprägt wurden, wurde mit dem von Frank Rosenblatt 1958 vorgestellten *Perzeptron*, einem vereinfachten künstlichen neuronalen Netzwerk, welches für einfache Klassifikationsaufgaben geeignet ist (1958), zudem eine wichtige Grundlage für die Entwicklung neuronaler Netzwerke in dieser Zeit gelegt. 1969 führten Marvin Minsky und Seymour Papert jedoch den Beweis, dass nicht-lineare Probleme durch das Perzeptron nicht lösbar sind (1969). Diese und weitere Erkenntnisse hinsichtlich der Limitationen dieser frühen KI-Systeme führten Mitte der 1970er Jahre zu einer großen Enttäuschung der bis dahin stark überzogenen Erwartungen an künstliche Intelligenz, weshalb diese Zeit heute gemeinhin als erster »KI-Winter« bezeichnet wird.

Einen erneuten Aufschwung der künstlichen Intelligenz brachten in den 1980er Jahren die sogenannten »Experten-Systeme«, die eine computer-basierte Entscheidungsunterstützung für komplexe Probleme durch eine Menge definierter »Wenn-dann-Regeln« sowie eine zugrundeliegende Wissensdatenbank liefern sollten. Diese Experten-Systeme fanden insbesondere auch im Kontext medizinischer Problemstellungen Anklang. Eines der ersten solcher Systeme war CASNET (Weiss, 1978), ein Experten-System, welches die Diagnose und Therapie des Glaukoms unterstützen sollte. Ein weiteres Beispiel war MYCIN (Shortliffe, 1977), welches mit einem enorm umfangreichen Regelsystem Krankheiten anhand von eingegebenen Patientendaten erkennen und Therapievorschlüsse für Antibiotika-Regime vorschlagen sollte und aus dem später das noch umfangreichere System INTERNIST-1 entstand (Miller, 1985). Letztendlich erreichten die Experten-Systeme insbesondere bei komplexen Pro-

blemstellungen nicht die erhoffte Leistungsfähigkeit und enttäuschten die hochgesteckten Erwartungen an diese Technologie, was Anfang der 1990er den Beginn des zweiten KI-Winters markierte.

Ab den 1990er Jahren traten an die Stelle der zuvor regelbasierten Ansätze der Experten-Systeme vor allem stärker datengetriebene Ansätze. Methoden wie die *support vector machines* (Cortes, 1995), Bayes'sche Netzwerke (Castillo, 2012) oder Ensemble-Methoden wie *random forest* (Breiman, 2001) wurden in dieser Zeit etabliert und fanden zunehmend Anwendung auf medizinischen Daten. Auf dieser Basis entwickelten sich in den 1990er Jahren auch zunehmend kommerziell verfügbare »CAD-Systeme« (*computer aided detection*), die Radiolog*innen bei der Erkennung bestimmter Pathologien unterstützen sollten, beispielsweise bei der Brustkrebs-Erkennung in Mammographien (Freer, 2001) oder der Detektion von Lungenrundherden in Computertomographien der Lunge (Das, 2006). Teilweise konnte mit dem Einsatz dieser Systeme bereits eine Steigerung der Detektionsrate entsprechender Läsionen demonstriert werden; insgesamt blieb die klinische Akzeptanz dieser Systeme jedoch verhalten.

Eine CAD-Systeme beruhten bereits auf künstlichen neuronalen Netzwerken. Bereits in den 1980er Jahren gab es gegenüber dem Perzeptron relevante theoretische Weiterentwicklungen im Bereich künstlicher neuronaler Netzwerke, unter anderem die architekturelle Entwicklung der später im *deep learning* vornehmlich verwendeten *convolutional neural networks* (CNNs) und die Anwendung des Backpropagation Algorithmus auf solche mehrschichtigen neuronalen Netzwerke (Fukushima, 1980; LeCun, 1989). Obwohl es in der Folgezeit kontinuierliche Weiterentwicklungen im Bereich künstlicher neuronaler Netzwerke gab, blieb der große Durchbruch jedoch zunächst aus – vor allem aufgrund der noch begrenzten Rechenleistung damaliger Computersysteme. Erst mit dem deutlichen Anstieg der verfügbaren Rechenkapazitäten und dem Einsatz von Grafikprozessoren ab den späten 2000er und insbesondere ab den 2010er-Jahren konnten tiefe neuronale Netzwerke ihre volle Leistungsfähigkeit entfalten. Als Meilenstein dieses Durchbruchs gilt dabei das Jahr 2012, als das Netzwerk *AlexNet* im *ImageNet Large Scale Visual Recognition Challenge* Wettbewerb einen bis dahin unerreichten Leistungssprung erzielte (Krizhevsky, 2012) und damit den Beginn der modernen *deep learning* Ära markierte. Der losgetretene Trend

übertrag sich unmittelbar auf die Anwendung in der Medizin. Prominente Beispiele für solche Anwendungen waren unter anderem ein *convolutional neural network*, das Hautkrebs in Fotografien und dermatoskopischen Bildern auf fachärztlichem Niveau klassifizieren konnte (Esteva, 2017) oder ein *deep learning* Netzwerk, welches diabetische Retinopathie anhand von Fotografien des Augenfundus erkennt (Gulshan, 2016). Aufgrund der besonderen Eignung von *deep learning* Verfahren für Mustererkennung in Bilddaten rückte jedoch insbesondere die medizinische Bildgebung und Radiologie in den Fokus des wissenschaftlichen Interesses. Die Zahl wissenschaftlicher Publikationen zu künstlicher Intelligenz in der Radiologie stieg in dieser Zeit exponentiell von zuvor ca. 300 Publikationen pro Jahr im Jahr 2010 auf über 2000 Publikationen im Jahr 2020 an (Decuyper, 2021). Nahezu jede Bildgebungsmodalität und Krankheitsentität wurde wissenschaftlich mithilfe von *deep learning* Netzwerken analysiert, vor allem, um Pathologien in Bildern zu erkennen oder Zusammenhänge zwischen medizinischen Bilddaten und Krankheits- oder Therapieverläufen herzustellen. Beispiele hierfür waren Systeme basierend auf *convolutional neural networks* zur Brustkrebserkennung in Mammographien, die hinsichtlich ihrer Leistungsfähigkeit klassische CAD-Systeme übertraf und eine Erkennungsrate auf dem Niveau von Radiolog*innen aufwies (Kooi, 2017), Algorithmen zur zuverlässigen Erkennung von Pathologien wie Tuberkulose oder Pneumonien in Röntgen-Thorax Bildern (Lakhani, 2017; Rajpurkar, 2017) oder Systeme zur Segmentierung von Hirntumoren in Magnetresonanztomographien des Schädels (Havaei, 2017). Während der *deep learning* Hype in der medizinischen Bildgebung noch nicht abgeklungen war/ist, kam es Anfang der 2020er Jahre zu einem darauf aufsetzenden Hype um sogenannte generative künstliche Intelligenz. Dieser Hype wurde durch relevante Weiterentwicklungen sogenannter *Transformer-Modelle* ausgelöst (Vaswani, 2017) und erlangte spätestens durch die zu dieser Zeit verblüffenden Fähigkeiten des auf der *Transformer-Architektur* beruhenden und im November 2022 von der Firma OpenAI veröffentlichten Chat-Systems ChatGPT (OpenAI, San Francisco, USA) eine breite öffentliche Wahrnehmung. Neben großen Sprachmodellen erlangten zu dieser Zeit auch weitere generative KI-Systeme zur Erzeugung von Bildern, gesprochener Sprache, Musik oder sogar Videos aus Texteingaben große Beachtung. Schnell wurden die Potentiale generativer KI

auch für die Medizin erkannt. Neben der Generierung täuschend echter synthetischer medizinischer Bilder wie Röntgen-Thorax-Bilder (Bluethgen, 2025), denen das Potential zugeschrieben wird, die Wahrung des Datenschutzes beim Training von KI Modellen mit großen Datenmengen im medizinischen Kontext zu erleichtern, konnten generative Modelle vor allem auch eingesetzt werden, um aus stark verrauschten medizinischen Bilddaten Bilder von diagnostischer Qualität zu erzeugen und so z. B. die Strahlen-Exposition in der Computertomographie (Sadia, 2024) oder die Scan-Dauer in Magnetresonanztomographien relevant zu senken (Kadavigere, 2024). Auch verschiedene Anwendungsmöglichkeiten großer generativer Sprachmodelle im medizinischen Kontext wurden bereits in zahlreichen Arbeiten erforscht, unter anderem, um Fehler in radiologischen Befunden zu erkennen (Gertz, 2024), strukturierte Daten aus Freitextbefunden zu Qualitätssicherungszwecken zu extrahieren (Lehnen, 2024), Fachwissen aus medizinischen Leitlinien abzuleiten (Vach, 2025) oder Bildgebungsprotokolle für Magnetresonanztomographien anhand der klinischen Fragestellung automatisch zu generieren (Boschenriedter, 2025). Auch die automatische Generierung radiologischer Befunde direkt aus medizinischen Bildern auf dem Niveau von Radiolog*innen konnte erfolgreich demonstriert werden (Zhou, 2025). Zuletzt wurde in wissenschaftlichen Arbeiten auch vermehrt sogenannte agentische künstliche Intelligenz fokussiert – generative KI-Systeme, die durch Zugriff auf Werkzeuge, Wissensdatenbanken und Interaktionsmöglichkeiten umfassende Möglichkeiten haben, Aufgaben durch Planung und Erweiterung ihres Kontextes zu lösen – und ihre Anwendbarkeit im medizinischen Kontext demonstriert (Ferber, 2025).

3. Aktuelle Entwicklungen und Herausforderungen

Nach den Entwicklungen der letzten Jahre zeigt sich aktuell ein zunehmender Einzug von künstlicher Intelligenz in die klinische Anwendung. In diesem Zuge zeigte sich in den letzten zehn Jahren ein enormer Zuwachs von Zulassungen KI-basierter Software-Systeme als Medizinprodukte. Nach der ersten Zulassung eines KI-basierten Software-Produktes durch die für die Regulierung von Medizinprodukten in den USA zuständige Food and Drug Administration

(FDA) im Jahr 1995 kam es in den folgenden Jahren lediglich sporadisch zu weiteren Zulassungen. So wurden im Jahr 2015 lediglich sechs Produkte durch die FDA neu zugelassen. Seither zeigt sich jedoch ein rapider Anstieg neu zugelassener Medizinprodukte mit beispielsweise 130 Neuzulassungen im Jahr 2021 und zuletzt 235 Neuzulassungen im Jahr 2024 (FDA, 2025). Bemerkenswert ist hierbei, dass der Radiologie eine deutliche Vorreiterrolle für die Adaptation von KI in der Medizin zukommt: Von den insgesamt 1247 zugelassenen KI-basierten Software-Medizinprodukten, welche im August 2025 von der FDA geführt werden, entfallen 956 Produkte, also 76,7 %, auf den Bereich Radiologie (FDA, 2025). Bei diesen Software-Produkten handelt es sich überwiegend um Software-basierte Assistenzsysteme, die Radiolog*innen bei der Interpretation von Bildern oder bei der Priorisierung der Befundung unterstützen sollen.

Trotz dieses immensen Angebots an KI-basierten Software-Produkten in der Radiologie bleibt die tatsächliche Anwendung in der klinischen Radiologie bisher vergleichsweise zurückhaltend. In der täglichen Praxis haben sich bislang kaum Systeme flächendeckend etabliert, und ihr Beitrag zur Routineversorgung ist bislang begrenzt. Ein wesentlicher Grund hierfür liegt in der Frage der Finanzierung (van Leeuwen, 2024). Viele der derzeit verfügbaren KI-Anwendungen sind kostenintensiv in Anschaffung und Wartung, während belastbare Nachweise für einen klaren ökonomischen Nutzen fehlen. Solange nicht belegt ist, dass KI-Systeme die Effizienz radiologischer Prozesse messbar steigern oder Kosten im Gesundheitssystem senken, fällt es schwer, entsprechende Investitionen im ohnehin stark budgetierten klinischen Umfeld zu rechtfertigen.

Hinzu kommen erhebliche Hürden bei der Integration in bestehende klinische Workflows. Radiologische Abteilungen arbeiten mit komplexen IT-Landschaften, die Bildarchivierungs- und Kommunikationssysteme (engl. Picture Archiving and Communication System, PACS), Radiologie-Informationssysteme (RIS) und Krankenhaus-Informationssysteme (KIS) umfassen. KI-Anwendungen, die nicht nahtlos in diese Strukturen eingebettet werden können, führen im Arbeitsalltag eher zu Mehraufwand als zu Entlastung. Es ist naheliegend, dass Systeme, die nur isoliert funktionieren oder zusätzliche Arbeitsschritte erfordern, nur begrenzten Nutzen entfalten und häufig nicht akzeptiert und rasch wieder verworfen werden. Daher

ist es relevant, dass sich KI-Anwendungen nahtlos in bestehende Arbeitsabläufe und -infrastrukturen eingliedern und z. B. Arbeitslisten priorisieren, auffällige Befunde markieren oder standardisierte Messungen automatisiert durchführen, ohne den klinischen Workflow zu stören.

Eng mit diesen Aspekten verknüpft ist die noch immer unzureichende Evidenzlage in Bezug auf den tatsächlichen klinischen und ökonomischen Nutzen von KI. Zwar gibt es zahlreiche wissenschaftliche Arbeiten zu KI in der Radiologie, die teils beeindruckende Leistungswerte einzelner Modelle berichten, doch handelt es sich hierbei ganz überwiegend um retrospektive, monozentrische Fallsammlungen, während groß angelegte, randomisiert kontrollierte Studien, die die Effektivität von KI-Systemen unter Alltagsbedingungen belegen, aktuell kaum vorhanden sind (Gorenstein, 2023). Ohne diese Evidenz bleibt unklar, ob der Einsatz von KI tatsächlich die diagnostische Genauigkeit, die Versorgungsqualität oder die Sicherheit der Patient*innen verbessert und ob er einen messbaren Mehrwert für das Gesundheitssystem generiert.

Über die genannten praktischen Hürden hinaus bestehen grundsätzliche methodische und wissenschaftliche Herausforderungen. Ein zentrales Problem stellt die sogenannte Domänen-Verschiebung (*domain shift*) dar. Modelle, die in einer bestimmten Institution, Population oder mit Daten eines bestimmten Herstellers trainiert wurden, verlieren häufig an Genauigkeit, wenn sie auf neue Patientenkollektive, Scanner oder Bildgebungsprotokolle angewandt werden (Guan & Liu, 2021). Diese mangelnde Generalisierbarkeit begrenzt die klinische Nutzbarkeit und erfordert Verfahren zur Domänenanpassung sowie standardisierte externe Validierungen.

Ein weiteres Problemfeld betrifft Verzerrungen (*biases*) in KI-Systemen. Wenn Trainingsdaten nicht repräsentativ für die gesamte Patientenpopulation sind, entstehen Verzerrungen, die zu systematisch ungleichen Ergebnissen führen können. In der Radiologie betrifft dies beispielsweise Unterschiede zwischen verschiedenen Altersgruppen, Geschlechtern oder ethnischen Hintergründen (Banerjee, 2023). Solche Verzerrungen können die diagnostische Qualität gefährden und werfen zudem grundlegende ethische Fragen nach Gerechtigkeit und Fairness auf (Lara, 2022).

Ein bislang nur wenig beachteter Aspekt ist zudem die Mensch-Maschine-Interaktion. Während die meisten wissenschaftlichen Ar-

beiten vor allem die technische Validität von KI-Systemen bei der Anwendung auf Bildungsdaten fokussieren, wird das Zusammenspiel der Anwender*innen mit den KI-Systemen und hieraus möglicherweise resultierender Effekte wenig beachtet. Bei der Verwendung automatisierter Systeme kann es zu relevanten Beeinflussungen durch kognitive Verzerrungen der Anwender*innen durch übermäßiges Vertrauen in die Empfehlungen einer Maschine kommen, selbst wenn diese fehlerhaft sind (Parasuraman & Manzey, 2010). Dieses als *automation bias* bezeichnete Phänomen konnte unter anderem eindrucksvoll in einem Laborexperiment demonstriert werden, bei dem die Genauigkeit von Radiolog*innen bei der KI-unterstützten Befundung von Mammografien abhängig vom Erfahrungsgrad von ca. 80 % auf bis zu ca. 20 % sank, wenn die KI falsche Entscheidungen suggerierte (Dratsch, 2023). Strategien zur Kommunikation von Unsicherheit von KI-Systemen und zur Vermeidung solcher Verzerrungen sowie zur Förderung eines kritischen Umgangs mit KI-Ausgaben sind daher entscheidend.

Ethische und regulatorische Fragen bilden einen weiteren zentralen Bereich. Der Einsatz von KI in der Radiologie betrifft hochsensible Patientendaten und Entscheidungen von erheblicher Tragweite. Entsprechend hoch sind die Anforderungen an Datenschutz, Transparenz und Nachvollziehbarkeit. Gleichzeitig müssen regulatorische Systeme wie der europäische AI Act und Institutionen wie die US-amerikanische FDA sicherstellen, dass KI-Produkte nicht nur bei der Zulassung, sondern auch während ihres gesamten Lebenszyklus überwacht und regelmäßig evaluiert werden. Eine besondere Herausforderung ergibt sich hier bei adaptiven Systemen, die sich nach der Zulassung durch kontinuierliches Lernen verändern können. Hier gilt es, einen regulatorischen Rahmen zu schaffen, der Innovation ermöglicht, aber zugleich Sicherheit und Verlässlichkeit garantiert.

Zusammenfassend lässt sich feststellen, dass die aktuelle Nutzung von KI in der Radiologie noch weit hinter den Erwartungen zurückbleibt, die die technologische Entwicklung in den letzten Jahren geweckt hat. Neben praktischen Hürden wie Finanzierung, Integration und Evidenzlage sind es auch grundlegende wissenschaftliche, ethische und regulatorische Fragen, die eine breite Anwendung hemmen. Zugleich verdeutlichen diese Herausforderungen, dass die Weiterentwicklung der KI in der Radiologie nicht nur eine technische,

sondern auch eine gesellschaftliche Aufgabe ist, die einen interdisziplinären Diskurs und klare Rahmenbedingungen erfordert.

4. Perspektiven für KI in der Radiologie

In Anbetracht der enormen Dynamik, die die Entwicklung der künstlichen Intelligenz in den vergangenen Jahren entfaltet hat, lassen sich belastbare Prognosen über zukünftige Einsatzformen in der Radiologie nur schwer treffen. Gleichwohl zeichnen sich zentrale Handlungsfelder ab, die über die erfolgreiche und verantwortungsvolle Integration dieser Technologien entscheiden werden. Eine wesentliche Aufgabe der kommenden Jahre besteht darin, die bestehenden und oben erläuterten Herausforderungen gezielt zu adressieren, um einen vertrauensvollen Umgang mit dieser potenten, zugleich aber auch hochsensiblen Technologie im Bereich der Medizin zu ermöglichen. Hierzu gehört insbesondere die Schaffung einer ausreichenden wissenschaftlichen Evidenzbasis, die nicht nur technische Leistungskennzahlen, sondern auch den tatsächlichen Nutzen von KI für die Versorgung der Patient*innen und das Gesundheitssystem belegt.

Eng damit verbunden ist die Frage, wie sich ein hinreichendes Maß an Kontrolle über KI-Systeme bewahren lässt, um deren sicheren und gewinnbringenden Einsatz gewährleisten zu können. Von besonderer Bedeutung wird dabei die Gestaltung der Mensch-Maschine-Interaktion sein. Nur wenn klar definiert bleibt, welche Aufgaben von KI übernommen werden können und wo ärztliche Verantwortung unverzichtbar ist, kann das Vertrauen in diese Systeme gestärkt werden. Auch regulatorische Vorgaben wie der europäische AI Act, der für Hochrisiko-KI-Systeme ausdrücklich eine menschliche Überwachungspflicht vorsieht, unterstreichen die Relevanz dieses Aspekts.

Auf technologischer Ebene eröffnet insbesondere die Entwicklung sogenannter *foundation models* weitreichende Perspektiven. Diese Modelle werden auf sehr großen, multimodalen Datensätzen trainiert, die unterschiedliche Datenquellen wie radiologische Bilddaten, Textdaten aus Arztbriefen oder auch medizinisches Fachwissen aus der Literatur umfassen können. Dadurch entsteht ein breites, abstrahiertes medizinisches Wissen, das in vielfältige Anwendungs-

aufgaben – sogenannte *downstream tasks* – überführt werden kann. Denkbar sind auf dieser Basis sowohl Assistenzsysteme zur Erkennung ganz unterschiedlicher Pathologien in Bilddaten als auch Systeme, die Radiolog*innen beim Erstellen von Befunden oder Arztbriefen unterstützen, oder auch Chatbots, die administrative Tätigkeiten wie etwa die Terminorganisation übernehmen (D’Antonoli, 2025). Im Vergleich zu den gegenwärtig vorherrschenden Insellösungen, die jeweils nur eine eng umgrenzte Aufgabe erfüllen – etwa die Detektion von Lungenrundherden in der Computertomographie, ohne in der Lage zu sein, andere Pathologien wie beispielsweise Frakturen zu erkennen oder andere klinische Dokumente zu verarbeiten – erscheint die Integration eines einzigen, generalistischen Systems auf Basis eines *foundation models* für die klinische Anwendung mittelfristig deutlich realistischer.

Eine weitere Entwicklungsperspektive ergibt sich durch sogenannte *KI-Agenten*. Diese Systeme verfügen über die Möglichkeit, auf Wissensbasen wie lokale Datenbanken von Gesundheitseinrichtungen oder externe Quellen aus dem Internet sowie auf Software- und Kommunikationstools zuzugreifen, um hiermit ihren Kontext eigenständig zu erweitern und komplexe Aufgaben weitgehend autonom zu bearbeiten. Das Potenzial solcher Agenten reicht von der medizinischen Assistenz – etwa bei der Befundung oder Diagnoseunterstützung – bis hin zur administrativen Entlastung in radiologischen Abteilungen (Ferber, 2025). Gerade bei dieser Technologie ist es jedoch essenziell, sorgfältig zu regulieren, in welchem Umfang der Zugriff auf Daten und Werkzeuge im medizinischen Kontext erlaubt werden kann, um die notwendige Kontrolle über die Systeme zu gewährleisten und Missbrauch vorzubeugen.

Auch wenn die Leistungsfähigkeit von KI-Systemen in einzelnen Aufgabenfeldern bereits das Niveau erfahrener Radiolog*innen erreichen kann, ist ein vollständiger Ersatz ärztlicher Expertise in absehbarer Zeit nicht zu erwarten (Langlotz, 2019). Wahrscheinlicher ist eine Verschiebung des beruflichen Rollenbildes: Durch die Automatisierung repetitiver Tätigkeiten wie der Mustererkennung in Bilddaten oder bestimmter administrativer Arbeiten werden Radiolog*innen zunehmend in die Rolle von »Orchestrator*innen« und Supervisor*innen treten, die den Einsatz von KI-Systemen überwachen, Ergebnisse kritisch reflektieren und in den klinischen Gesamt-

kontext einordnen. Damit wird ihre Tätigkeit zwar nicht weniger zentral, wohl aber grundlegend anders strukturiert sein.

Vor diesem Hintergrund gewinnt auch die Aus- und Weiterbildung in der Radiologie neue Bedeutung. Es wird darauf ankommen, Konzepte zu entwickeln, die Studierende und Nachwuchsärztinnen und -ärzte gezielt auf die Arbeit an der Schnittstelle zwischen Medizin und Informationstechnologie vorbereiten. Neben der klassischen radiologischen Expertise müssen Kompetenzen im Umgang mit KI, in der Bewertung algorithmischer Ergebnisse und in der verantwortungsvollen Nutzung digitaler Werkzeuge vermittelt werden. Nur so lässt sich sicherstellen, dass die Radiologie von den Chancen der KI profitiert, ohne dass dabei die Sicherheit der Patient*innen oder die ärztliche Verantwortung kompromittiert werden.

Nicht zuletzt bestimmen regulatorische Rahmenbedingungen, ethische Grundsätze und ökonomische Überlegungen die zukünftige Rolle der KI in der Radiologie. Es stellt sich die Frage, wie die Kosten für Entwicklung, Implementierung und Wartung von KI-Systemen in ein Gesundheitssystem integriert werden können, das bereits erheblich unter ökonomischem Druck steht. Akzeptanz seitens der Radiolog*innen wird nur dann entstehen, wenn die Systeme nachweislich die Versorgungsqualität verbessern, gleichzeitig aber auch die Arbeitsbelastung reduzieren.

Literaturverzeichnis

- Banerjee, I., Bhattacharjee, K., Burns, J. L., Trivedi, H., Purkayastha, S., Seyyed-Kalantari, L., Patel, B. N., Shiradkar, R., & Gichoya, J. (2023). »Shortcuts« causing bias in radiology artificial intelligence: Causes, evaluation, and mitigation. *Journal of the American College of Radiology*, 20(9), 842–851. <https://doi.org/10.1016/j.jacr.2023.06.025>
- Bizzo, B. C., Almeida, R. R., Michalski, M. H., & Alkasab, T. K. (2019). Artificial intelligence and clinical decision support for radiologists and referring providers. *Journal of the American College of Radiology*, 16(9), 1351–1356. <https://doi.org/10.1016/j.jacr.2019.06.010>
- Bluethgen, C., Chambon, P., Delbrouck, J. B., van der Sluijs, R., Polacin, M., Zambrano Chaves, J. M., Abraham, T. M., Purohit, S., Langlotz C. P., & Chaudhari, A. S. (2025). A vision–language foundation model for the generation of realistic chest X-ray images. *Nature Biomedical Engineering*, 9(4), 494–506. <https://doi.org/10.1038/s41551-024-01246-y>

- Boschenriedter, C., Rubbert, C., Vach, M., & Caspers, J. (2025). Automated protocol suggestions for cranial MRI examinations using locally fine-tuned BERT models. *Clinical Neuroradiology*, epub ahead of print. <https://doi.org/10.1007/s00062-025-01554-z>
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Castillo, E., Gutierrez, J. M., & Hadi, A. S. (2012). *Expert systems and probabilistic network models*. Springer Science & Business Media.
- Chen, J., Patil, K. R., Yeo, B. T., & Eickhoff, S. B. (2023). Leveraging machine learning for gaining neurobiological and nosological insights in psychiatric research. *Biological Psychiatry*, 93(1), 18–28. <https://doi.org/10.1016/j.biopsych.2022.07.025>
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>
- D’Antonoli, T. A., Bluethgen, C., Cuocolo, R., Klontzas, M. E., Ponsiglione, A., & Kocak, B. (2025). Foundation models for radiology: fundamentals, applications, opportunities, challenges, risks, and prospects. *Diagnostic and Interventional Radiology*. epub ahead of print. <https://doi.org/10.4274/dir.2025.253445>
- Das, M., Mühlenbruch, G., Mahnken, A. H., Flohr, T. G., Gündel, L., Stanzel, S., Kraus, T., Günther, R. W., & Wildberger, J. E. (2006). Small pulmonary nodules: Effect of two computer-aided detection systems on radiologist performance. *Radiology*, 241(2), 564–571. <https://doi.org/10.1148/radiol.2412051139>
- Decuyper, M., Maebe, J., Van Holen, R., & Vandenberghe, S. (2021). Artificial intelligence with deep learning in nuclear medicine and radiology. *EJN-MMI Physics*, 8(1), 81. <https://doi.org/10.1186/s40658-021-00426-y>
- Dratsch, T., Chen, X., Rezazade Mehrizi, M., Kloeckner, R., Mähringer-Kunz, A., Püsken, M., Baeßler, B., Sauer, S., Maintz, D., & Pinto dos Santos, D. (2023). Automation bias in mammography: The impact of artificial intelligence BI-RADS suggestions on reader performance. *Radiology*, 307(4), e222176. <https://doi.org/10.1148/radiol.222176>
- Eickhoff, S. B., & Heinrichs, B. (2021). Der vorhersagbare Mensch. *Der Nervenarzt*, 92(11), 1140–1148. <https://doi.org/10.1007/s00115-021-01197-8>
- 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. <https://doi.org/10.1038/nature21056>

- Ferber, D., El Nahhas, O. S., Wölflein, G., Wiest, I. C., Clusmann, J., Leßmann, M. E., Foersch, S., Lammert, J., Tschochohei, M., Jäger, D., Salto-Tellez, M., Schultz, N., Thrun, D., & Kather, J. N. (2025). Development and validation of an autonomous artificial intelligence agent for clinical decision-making in oncology. *Nature Cancer*, 6, 1337–1349. <https://doi.org/10.1038/s43018-025-00991-6>
- Food and Drug Administration (FDA) (2025). *Artificial intelligence and machine learning (AI/ML)- enabled medical devices*. USA Department of Health and Human Services. <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-enabled-medical-devices> (Zugriff 18.08.2025).
- Freer, T. W., & Ulissey, M. J. (2001). Screening mammography with computer-aided detection: Prospective study of 12,860 patients in a community breast center. *Radiology*, 220(3), 781–786. <https://doi.org/10.1148/radiol.2203001282>
- Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36(4), 193–202. <https://doi.org/10.1007/BF00344251>
- Gertz, R. J., Dratsch, T., Bunck, A. C., Lennartz, S., Iuga, A. I., Hellmich, M. G., Persigehl, T., Pennig, L., Gietzen, C. H., Fervers, P., Maintz, D., Hahnfeldt, R., & Kottlors, J. (2024). Potential of GPT-4 for detecting errors in radiology reports: Implications for reporting accuracy. *Radiology*, 311(1), e232714. <https://doi.org/10.1148/radiol.232714>
- Gorenstein, L., Soffer, S., Apter, S., Konen, E., & Klang, E. (2023). AI in radiology: is it the time for randomized controlled trials? *European Radiology*, 33(6), 4223–4225. <https://doi.org/10.1007/s00330-022-09381-3>
- Guan, H., & Liu, M. (2021). Domain adaptation for medical image analysis: A survey. *IEEE Transactions on Biomedical Engineering*, 69(3), 1173–1185. <https://doi.org/10.1109/TBME.2021.3117407>
- Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., Kim, R., Raman, R., Nelson, P. C., Mega, J. L., & Webster, D. R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*, 316(22), 2402–2410. <https://doi.org/10.1001/jama.2016.17216>
- Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P., & Larochelle, H. (2017). Brain tumor segmentation with deep neural networks. *Medical Image Analysis*, 35, 18–31. <https://doi.org/10.1016/j.media.2016.05.004>
- Hedderich, D. M., & Eickhoff, S. B. (2021). Machine learning for psychiatry: getting doctors at the black box? *Molecular Psychiatry*, 26(1), 23–25. <https://doi.org/10.1038/s41380-020-00931-z>

- Heinrichs, B., & Eickhoff, S. B. (2020). Your evidence? Machine learning algorithms for medical diagnosis and prediction. *Human Brain Mapping*, 41(6), 1435–1444. <https://doi.org/10.1002/hbm.24886>
- Jones, C.R., & Bergen, B.K. (2025). Large language models pass the turing test. *arXiv*, Artikel 2503.23674. <https://doi.org/10.48550/arXiv.2503.23674>
- Priyanka, Kadavigere, R., Nayak S, S., Chandran M, O., Shirlal, A., Pires, T., & Pendem, S. (2024). Impact of artificial intelligence assisted compressed sensing technique on scan time and image quality in musculoskeletal MRI—A systematic review. *Radiography*, 30(6), 1704–1712. <https://doi.org/10.1016/j.radi.2024.08.012>
- Koohi-Moghadam, M., & Bae, K. T. (2023). Generative AI in medical imaging: Applications, challenges, and ethics. *Journal of Medical Systems*, 47(1), Artikel 94. <https://doi.org/10.1007/s10916-023-01987-4>
- Kooi, T., Litjens, G., Van Ginneken, B., Gubern-Mérida, A., Sánchez, C. I., Mann, R., den Heeten, A., & Karssemeijer, N. (2017). Large scale deep learning for computer aided detection of mammographic lesions. *Medical Image Analysis*, 35, 303–312. <https://doi.org/10.1016/j.media.2016.07.007>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In Pereira, F., Burges, C.J., Bottou, L., Weinberger, K.Q. (Hrsg.). *Advances in Neural Information Processing Systems*, 25.
- Lakhani, P., & Sundaram, B. (2017). Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology*, 284(2), 574–582. <https://doi.org/10.1148/radiol.2017162326>
- Langlotz, C. P. (2019). Will artificial intelligence replace radiologists? *Radiology: Artificial Intelligence*, 1(3), Artikel e190058. <https://doi.org/10.1148/ryai.2019190058>
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1(4), 541–551. <https://doi.org/10.1162/neco.1989.1.4.541>
- Lehnen, N. C., Dorn, F., Wiest, I. C., Zimmermann, H., Radbruch, A., Kather, J. N., & Paech, D. (2024). Data extraction from free-text reports on mechanical thrombectomy in acute ischemic stroke using ChatGPT: A retrospective analysis. *Radiology*, 311(1), Artikel e232741. <https://doi.org/10.1148/radiol.232741>
- Lodwick, G. S. (1965). A probabilistic approach to the diagnosis of bone tumors. *Radiologic Clinics of North America*, 3(3), 487–497. [https://doi.org/10.1016/S0033-8389\(22\)02844-5](https://doi.org/10.1016/S0033-8389(22)02844-5)

- Mazurowski, M. A., Buda, M., Saha, A., & Bashir, M. R. (2019). Deep learning in radiology: An overview of the concepts and a survey of the state of the art with focus on MRI. *Journal of Magnetic Resonance Imaging*, 49(4), 939–954. <https://doi.org/10.1002/jmri.26534>
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4), 115–133. <https://doi.org/10.1007/BF02478259>
- Minsky, M., & Papert, S. (1969). *Perceptrons. An introduction to computational geometry*. MIT Press.
- Miller, R. A., Pople Jr, H. E., & Myers, J. D. (1985). INTERNIST-I, An experimental computer-based diagnostic consultant for general internal medicine. In B. I. Blum, J. A. Reggia, & S. Tuhim (Hrsg.), *Computer-Assisted Medical Decision Making* (S. 139–158). Springer New York. https://doi.org/10.1007/978-1-4612-5108-8_8
- Moawad, A. W., Fuentes, D. T., ElBanan, M. G., Shalaby, A. S., Guccione, J., Kamel, S., Jensen, C. T., & Elsayes, K. M. (2022). Artificial intelligence in diagnostic radiology: where do we stand, challenges, and opportunities. *Journal of Computer Assisted Tomography*, 46(1), 78–90. <https://doi.org/10.1097/RCT.0000000000001247>
- Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, 52(3), 381–410. <https://doi.org/10.1177/0018720810376055>
- Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M. P., & Ng, A. Y. (2017). CheXNet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv*, Artikel 1711.05225. <https://doi.org/10.48550/arXiv.1711.05225>
- Ricci Lara, M. A., Echeveste, R., & Ferrante, E. (2022). Addressing fairness in artificial intelligence for medical imaging. *Nature Communications*, 13(1), Artikel 4581. <https://doi.org/10.1038/s41467-022-32186-3>
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386–408. <https://doi.org/10.1037/h0042519>
- Sadia, R. T., Chen, J., & Zhang, J. (2024). CT image denoising methods for image quality improvement and radiation dose reduction. *Journal of Applied Clinical Medical Physics*, 25(2), Artikel e14270. <https://doi.org/10.1002/acm2.14270>
- Shortliffe, E. H. (1977). Mycin: A knowledge-based computer program applied to infectious diseases. In *Proceedings of the Annual Symposium on Computer Application in Medical Care* (S. 66).
- Turing, A.M. (1950). Computing machinery and intelligence. *Mind*, LIX(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>

- Vach, M., Gliem, M., Weiss, D., Ivan, V.L., Hauke, F., Boschenriedter, C., Rubbert, C., & Caspers, J. (2025). Evaluating retrieval augmented generation-enhanced large language models for question answering on german neurovascular guidelines. *Clinical Neuroradiology*, epub ahead of print. <https://doi.org/10.1007/s00062-025-01562-z>
- van Leeuwen, K. G., de Rooij, M., Schalekamp, S., van Ginneken, B., & Rutten, M. J. (2024). Clinical use of artificial intelligence products for radiology in the Netherlands between 2020 and 2022. *European Radiology*, 34(1), 348–354. <https://doi.org/10.1007/s00330-023-09991-5>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In Guyon, I., Von Luxburg, U., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (Hrsg.), *31st Conference on Neural Information Processing Systems (NIPS). Advances in Neural Information Processing Systems*, 30.
- Wang, X., Si, J., Li, Y., Tse, P., Zhang, G., Wang, X., Ren, J., Xu, J., Sun, J., & Yao, X. (2025). Effectiveness and safety of AI-driven closed-loop systems in diabetes management: a systematic review and meta-analysis. *Diabetology & Metabolic Syndrome*, 17(1), Artikel 238. <https://doi.org/10.1186/s13098-025-01819-0>
- Weiss, S., Kulikowski, C. A., & Safir, A. (1978). Glaucoma consultation by computer. *Computers in Biology and Medicine*, 8(1), 25–40. [https://doi.org/10.1016/0010-4825\(78\)90011-2](https://doi.org/10.1016/0010-4825(78)90011-2)
- Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45. <https://doi.org/10.1145/365153.365168>
- Zhou, H. Y., Acosta, J. N., Adithan, S., Datta, S., Topol, E. J., & Rajpurkar, P. (2024). MedVersa: A generalist foundation model for medical image interpretation. *arXiv*, Artikel 2405.07988. <https://doi.org/10.48550/arXiv.2405.07988>

