Originalarbeit
Themenheft: Perspektiven der DGPs-Fachgruppen auf KI – Teil 1

Von Menschen und Maschinen: Psychologiehistorische Reflexionen über Künstliche Intelligenz

Eine Perspektive der Fachgruppe Geschichte der Psychologie

Fabian Hutmacher[1,2] und Alexander Nicolai Wendt[2,3,4]

[1]Julius-Maximilians-Universität Würzburg, Deutschland
[2]Jungmitgliedervertretung in der DGPs-Fachgruppe Geschichte der Psychologie, Berlin, Deutschland
[3]Sigmund Freud PrivatUniversität Wien, Österreich
[4]Ruprecht-Karls-Universität Heidelberg, Deutschland

Zusammenfassung: Künstliche Intelligenz ist eines der bestimmenden Themen unserer Gegenwart. Es ist aber auch ein Thema mit einer langen Vergangenheit, in der vieles von dem, was heute neu und aktuell erscheint, schon einmal in ähnlicher Weise verhandelt wurde. Im Anschluss an eine knappe Technikgeschichte der künstlichen Intelligenz erörtert der Beitrag die Verschränkung von Mensch und Maschine aus problemgeschichtlicher Perspektive und beschreibt in diskursgeschichtlichen Schlaglichtern typische Muster gesellschaftlicher Debatten beim Auftreten technologischer Innovationen. Die psychologiehistorische Reflexion schafft auf diese Weise Möglichkeiten des durchdachten Hineinwirkens in gegenwärtige Diskurse.

Schlüsselwörter: Künstliche Intelligenz, Geschichte der Psychologie, Problemlösen, Large Language Models

On Humans and Machines: Psychological-Historical Reflections on Artificial Intelligence. A Perspective from the History of Psychology Section

Abstract: Artificial intelligence is one of the defining topics of our time, while also a topic with a long history, in which much of what appears new and current today has already been discussed. Following a brief technical history of artificial intelligence, this article discusses the historical intertwining of humans and machines and the typical patterns of social debate surrounding technological innovations. Reflections on the history of psychology offer opportunities for thoughtful engagement in contemporary discourse.

Keywords: artificial intelligence, history of psychology, problem-solving, large language models

Der Blick in die (Psychologie-)Geschichte verspricht Aufklärung in mehrfacher Hinsicht (Hutmacher & Mayrhofer, 2023): Erstens vermag er uns für die Genese des Gegenwärtigen zu sensibilisieren. Die Vergangenheit kann damit – zweitens – immer auch als expliziter Vergleichspunkt herangezogen werden. Damit erlangen wir – drittens – ein besseres Verständnis des psychologischen Möglichkeitenraumes, also ein besseres Verständnis dessen, was jenseits der aktuellen Trends denkbar und vorstellbar ist. Dies wiederum verspricht viertens einen Zugewinn an Handlungsoptionen und schafft im Idealfall Möglichkeiten des durchdachten Hineinwirkens in gegenwärtige Diskurse. Wie wir im Folgenden zeigen, lassen sich diese allgemeinen Überlegungen auf den Fall der Künstlichen Intelligenz (KI) übertragen.

In seiner schillernden Vielfältigkeit kann KI als summarischer Oberbegriff für all jene maschinellen Systeme aufgefasst werden, die Aufgaben ausführen können, die normalerweise (menschliche) Intelligenz erfordern (Stein et al., 2023). Was wiederum unter *Intelligenz* zu verstehen ist, führt die psychologische Begriffsarbeit bekanntermaßen regelmäßig an ihre Grenzen (Serrano et al., 2014).

Wir entfalten unsere Analyse der Verschränkung von Mensch und Maschine (bzw. KI) im Folgenden in drei Schritten. Im ersten Abschnitt geben wir einen Überblick der (frühen) *Technikgeschichte* der KI. Der zweite Abschnitt widmet sich der *Problemgeschichte* der KI. Zuletzt werfen wir einige Schlaglichter auf die *Diskursgeschichte* der KI, oder anders: auf typische Muster gesellschaftlicher Debatten beim Auftreten technologischer Innovationen.

Technikgeschichte: Von der Analytical Engine zu Large Language Models

Die *Idee*, Automaten zu konstruieren, die selbstständig Probleme lösen oder Menschen nachahmen, reicht so weit in die Vergangenheit zurück, dass sich nicht klar angeben lässt, wer sie das erste Mal geäußert hat (Cave et al., 2020). Ein möglicher Anfangspunkt moderner KI-Entwicklungen ist die von Charles Babbage (1791–1871) unter anderem in Zusammenarbeit mit Ada Lovelace (1815–1852) entworfene *Analytical Engine*. Obwohl diese programmierbare mechanische Rechenmaschine unter anderem mangels finanzieller Mittel nie gebaut wurde, enthielt sie bereits die wichtigsten Grundprinzipien moderner Computer (z. B. Trennung von Speicher- und Recheneinheit, sequenzielle Programmausführung). Interessanterweise scheint es Babbage nicht darum gegangen zu sein, eine *denkende Maschine* zu bauen oder Analogien zu menschlichem Denken zu ziehen: Seine Interessen waren primär ökonomischer und praktischer Natur (Green, 2005).

Ein weiterer Meilenstein sind die Arbeiten Alan Turings (1912–1954), allen voran die nach ihm benannte *Turing-Maschine* (Turing, 1937) sowie der *Turing-Test* (Turing, 1950). Handelt es sich bei der Turing-Maschine um ein einfaches mathematisches Modell eines Computers, das zeigt, wie sich jede berechenbare Aufgabe durch klare Regeln und Schritt-für-Schritt-Verarbeitung von Symbolen lösen lässt, beschreibt der Turing-Test ein (Gedanken-)Experiment, im Rahmen dessen geprüft wird, ob eine Maschine menschliche Intelligenz beziehungsweise menschliches Kommunikationsverhalten so gut nachahmen kann, dass ein Mensch im Gespräch nicht sicher zu sagen vermag, ob er mit einer Maschine oder einer Person spricht. Wie diese Bemerkungen bereits andeuten, waren Turings Beiträge nicht nur aus der Perspektive der Mathematik und Informatik bahnbrechend, sondern stets mit philosophischen – und im weiteren Sinne auch psychologischen – Überlegungen verzahnt – weshalb er uns im zweiten Abschnitt wiederbegegnen wird.

Spätestens nach Turing wird die Geschichte der KI schnell so verzweigt und vielschichtig, dass sich ein komprimierter Abriss auf notdürftige Wegmarken beschränken muss (für einen Überblick siehe He et al., 2025; Shao et al., 2022). Da ist zunächst die eng mit dem Namen Norbert Wiener (1894–1964) verbundene *Kybernetik*, also der Versuch, die Steuerung und Regelung von Maschinen und lebenden Organismen über Rückkopplungsmechanismen zu beschreiben (Wiener, 1948). Das klassische Lehrbuchbeispiel zur Illustration des Grundgedankens ist dasjenige des Thermostats, das auf einen bestimmten Temperatur-Sollwert eingestellt wird und diesen Sollwert dann nicht nur mit dem Istwert abgleicht, sondern basierend auf diesem Vergleich die Wärmezufuhr reguliert. Die kybernetischen Prinzipien der Steuerung und Rückkopplung wurden in der Folge sowohl zur Gestaltung technisch-ingenieurwissenschaftlicher Anwendungen als auch zur Analyse biologischer Systeme und sozialer Prozesse genutzt, schlagen also eine gedankliche Brücke zwischen Mensch und Maschine.

Auf Konrad Zuse (1910–1995) und John von Neumann (1903–1957) geht die Standard-Architektur des Computers zurück, dessen Herzstück der Prozessor mit arithmetisch-logischem Rechen- und Steuerwerk ist. Es handelt sich um ein Speichermodell, das Befehle wie Signale weiterleitet, um arithmetische Operationen durchzuführen. Diese Signale lassen sich nach dem Sender-Empfänger-Modell von Claude E. Shannon (1916–2001) und Warren Weaver (1894–1978) auch als Informationen beschreiben. Eine maßgebliche Eigenschaft dieses Ansatzes ist der sog. von-Neumann-Flaschenhals, der sich dadurch ergibt, dass die gesamten Transformationsleistungen des Systems in einem Prozessor mit beschränktem Arbeitsspeicher durchgeführt werden. Diese Serialität steht im paradigmatischen Gegensatz zu dezentralen und parallelen Konzeptionen der Computerarchitektur.

Dem von-Neumann-Paradigma waren auch Allen Newell (1927–1992) und Herbert A. Simon (1916–2011) verschrieben, die den *General Problem Solver* entwickelten – ein Computerprogramm, das als universelle Problemlösungsmaschine gedacht war (Newell et al., 1959). Etwas vereinfacht gesprochen sollte der General Problem Solver die Differenz zwischen einem Anfangs- und einem Endzustand berechnen und diesen dann mittels geeigneter Operatoren Schritt für Schritt verringern. In etwa denselben Zeitraum fallen auch Frank Rosenblatts (1928–1971) *Perceptron* als Beispiel eines frühen neuronalen Netzes (Rosenblatt, 1958) oder das von Joseph Weizenbaum (1923–2008) Mitte der 1960er Jahre entwickelte Computerprogramm *ELIZA* als Modell natürlicher Sprache (Weizenbaum, 1976). General Problem Solver, Perceptron und ELIZA teilen aber nicht nur den Entstehungszeitraum, sondern auch ihre Begrenztheit: In der Erprobung

aller drei Systeme stellte sich schnell heraus, dass sich weder menschliches Problemlösen noch Lernen und Sprechen so leicht algorithmisieren ließen wie erhofft.

Eine Alternative zur seriellen Computerarchitektur mit ihrer regelgeleiteten Verarbeitung symbolischer Strukturen bedeuteten konnektionistische Modelle wie der maßgeblich von David Rumelhart (1942–2011) und James McClelland (geb. 1948) geprägte *Parallel-Distributed-Processing*-Ansatz (Rumelhart et al., 1986). Grob gesprochen beschreibt Parallel Distributed Processing Modelle, bei denen Informationen nicht an einem Ort gespeichert und nach klar definierten Regeln verarbeitet werden, sondern als Aktivitätsmuster über viele einfache, miteinander verbundene Einheiten verteilt sind, ganz ähnlich – so jedenfalls der Gedanke – wie bei Neuronen im menschlichen Gehirn. Mit diesen und ähnlichen Entwicklungen stehen wir dann bereits an der Schwelle zur aktuellen KI-Forschung – also an jenem Punkt, an dem die Vergangenheit langsam in die Gegenwart übergeht und an denen das Heraufziehen von Large Language Models (LLMs) am Horizont erahnbar wird (für eine detaillierte Beschreibung des aktuellen Sachstandes, siehe He et al., 2025).

Man mag geneigt sein, die bisherigen Beschreibungen als lineare Fortschrittsgeschichte zu lesen, bei der das Erkennen der Beschränkungen eines Ansatzes nahtlos in die Entwicklung leistungsfähigerer Ansätze übergeht. Gleichwohl gehört zur Fortschrittsgeschichte auch immer eine *Scheiternsgeschichte*, die auszuleuchten mindestens genauso aufschlussreich sein kann und die uns unter Umständen davor zu bewahren vermag, allzu unkritisch auf allzu verheißungsvolle Versprechungen einer goldenen Zukunft hereinzufallen. So weist die KI-Geschichte beispielsweise mehrere sog. *KI-Winter* auf (für eine differenzierte Analyse, siehe Haigh, 2023), also Phasen, in denen Finanzierungsquellen versiegten, weil die in die Technologie gesetzten Erwartungen nicht erfüllt wurden und sich herausstellte, dass einiges von dem, was manche zunächst für Kinderkrankheiten einer neuen Technologie zu halten geneigt waren, auf grundsätzliche und nicht zu überwindende Probleme hinwies.

Problemgeschichte: Verschränkung von Mensch und Maschine

Wie weiter oben ausgeführt lässt sich KI als summarischen Oberbegriff für all jene Systeme verstehen, die Aufgaben ausführen können, die normalerweise (menschliche) Intelligenz erfordern. Aus dieser Definition ergibt sich eine – auch historisch nachweisbare (siehe Heßler, 2025) – enge diskursive Verschränkung des Sprechens und Nachdenkens über Menschen und Maschinen (bzw. KI): Wer etwas über die Maschine weiß, lernt demnach etwas über den Menschen; und wer etwas über den Menschen weiß, lernt etwas über die Maschine. Nehmen wir das Beispiel der Computermetapher des menschlichen Geistes: Die Idee, dass sich menschliche Denkprozesse als Informationsverarbeitungsprozesse modellieren lassen, inspirierte zahlreiche kognitionspsychologische Theorien und Modelle, etwa das Arbeitsgedächtnismodell von Baddeley und Hitch (1974). Umgekehrt diente das fortschreitende Verständnis für die Funktionsweise des menschlichen Gehirns beispielsweise als Grundlage für die Entwicklung neuronaler Netze. Selbstverständlich ist dieses Inspirationsverhältnis nicht frei von Spannungen: Zum einen zieht der Versuch, menschliches Erleben und Verhalten mechanistisch zu verstehen, immer wieder Reduktionismusvorwürfe auf sich; und zum anderen zwingt uns die Erkenntnis, dass Maschinen Fähigkeiten entwickeln, die man bis dahin für genuin menschlich gehalten hatte, zum Nachdenken darüber, was den Menschen ausmacht. Diese Spannungen wollen wir an einigen Beispielen etwas genauer erörtern.

Im Idealfall ist KI ein Vehikel zur Problemlösung: Das gilt nicht nur für die gegenwärtigen LLMs, denen Nutzerinnen und Nutzer ihre Fragen und Aufgaben zur Bearbeitung vorlegen, sondern in vergleichbarer Weise auch schon für frühere Modelle wie den General Problem Solver. Was aber bedeutet es, ein Problem zu lösen? Die Antwort auf diese Frage hängt davon ab, welches KI-Modell man betrachtet. Bei LLMs funktioniert Problemlösung über Token Prediction: Basierend auf zuvor gegebenen Tokens (d. h. Wörtern, Satzteilen oder Zeichen) wird das wahrscheinlich nächste Token vorhergesagt. Für den General Problem Solver dagegen ist Problemlösung eine Means-End-Analyse, also der Versuch, die Differenz zwischen Ist- und Soll-Zustand mittels geeigneter Operatoren zu minimieren. Damit ist klar, dass sich KI-Problemlösungsversuche fundamental von menschlichem Problemlösen unterscheiden, das zwar vielleicht mitunter erratischer, vor allem aber vielfältiger und damit schlussendlich oft situationsadäquater ist (Ohlsson, 2012). Daraus folgt auch, dass KI-Problemlösungsversuche nur sehr bedingt als Vorbild dienen können, um menschliches Problemlösen zu modellieren.

Der Kontrast zwischen Mensch und Maschine wird an einem zweiten Punkt noch deutlicher (siehe Wendt, 2024): KI-Systeme reagieren auf den Input der Nutzerinnen und Nutzer. Natürlich reagieren auch Menschen auf den „Input“ ihrer Umwelt. Menschen suchen aber außerdem aktiv verschiedene Umwelten auf und *wählen ihre Probleme*. Das heißt: Bevor sie sich in Problemlösungsversuche stürzen, entscheiden sie zuerst einmal, welchen Problemen sie sich in welcher Form widmen wollen.

Ähnliche Überlegungen werden in der KI-Forschung schon länger diskutiert. In seinem weiter oben bereits zitierten Aufsatz, in dem Turing (1950) den heute so bezeichneten Turing-Test beschreibt, setzt er sich mit *Lady Lovelace's Objection* auseinander, also einem Argument, das sich in Ada Lovelaces Schriften über Babbages Analytical Engine finden lässt. Im Kern lautet Lovelaces These, Maschinen könnten nichts wirklich *Neues* oder *Überraschendes* hervorbringen. Turing entgegnet darauf sinngemäß, es sei vielleicht unklar, ob Maschinen etwas wirklich Neues hervorbringen können – ebenso aber sei unklar, ob Menschen dies können. Im Hinblick auf zeitgenössische KI kann man diese Debatte so zu aktualisieren versuchen: KI handelt immer auf Basis eines *System-Prompts*. Auch menschliches Verhalten aber ist an evolutionär-biologische Rahmenbedingungen gebunden. Bedeutet das, dass Menschen ebenfalls bestimmten „System-Prompts" folgen – oder bezeichnet die Fähigkeit, aus sich heraus Ziele zu setzen, etwas prinzipiell Anderes? Die Antwort auf diese Frage hängt von anthropologischen und ontologischen Überzeugungen ab – und ist eng mit einem anderen Problem verbunden: dem Problem des Verstehens.

Über LLMs wird mitunter gesagt, sie seien nichts anderes als ein *stochastischer Papagei* (Bender et al., 2021). Mit anderen Worten: Mustererkennung und Musterproduktion haben nichts mit echtem Verstehen gemeinsam. Eine der bekanntesten Illustrationen dieses Unterschieds findet sich in John Searles (1980) Gedankenexperiment des *Chinesischen Zimmers*, das sich folgendermaßen zusammenfassen lässt: Man stelle sich eine Person vor, die in einem Zimmer sitzt und die von außen Zettel mit chinesischen Schriftzeichen erhält. Die Person selbst spricht kein Chinesisch. In dem Raum befindet sich jedoch ein Regelbuch, in dem zu finden ist, wie auf die Zeichen mit anderen Schriftzeichen zu reagieren ist. Die Person folgt diesen Anweisungen und gibt die entsprechenden Antworten. Für eine Person außerhalb des Chinesischen Zimmers, so Searles Gedanke, kann der Eindruck entstehen, als würde die Person innerhalb des Zimmers fließend Chinesisch sprechen. Über die Person innerhalb des Zimmers zu sagen, sie verstehe Chinesisch, sei aber im Grunde nichts anderes, als über einen Computer (oder eine KI) zu sagen, sie könne wirklich eine Sprache sprechen und verstehen, nur weil sie in der Lage ist, einen entsprechenden Output zu produzieren (für den aktuellen Diskussionsstand siehe Cole, 2024). Auch dieses Beispiel zeigt noch einmal, wie sehr das Nachdenken über den Menschen und das Nachdenken über die KI beziehungsweise das Nachdenken über Psychologie und Technologie aufeinander bezogen sind: Wer Vorstellungen von künstlicher Intelligenz nachspürt, spürt in der Regel zugleich Vorstellungen darüber nach, was das Denk- und Empfindungsvermögen sowie das Bewusstsein des Menschen ausmacht.

Diskursgeschichte: Zwischen Utopie und Dystopie

Die Einführung neuer Technologien hat immer wieder erhitzte Debatten – will sagen: diametral entgegengesetzte *psychologische* Reaktionen – ausgelöst, die nicht selten zwischen utopischen Versprechungen und dystopischer Panikmache schwanken. Auch hier kann eine (psychologie-)historische Perspektive helfen, vorschnelle Antworten zu vermeiden. Ein besonders häufig anzutreffendes Phänomen ist die Sorge, dass die Verbreitung einer Technologie zum *Verlust bestimmter Fähigkeiten* führen würde. Als illustratives Beispiel mag die Debatte um die Einführung von Taschenrechnern in Schulen dienen. Insbesondere in den westlichen Industrienationen wurden Taschenrechner in den 1970er Jahren zunehmend erschwinglich und verbreiteten sich massenhaft. Vor diesem Hintergrund stellte sich die Frage, ob Taschenrechner auch Einzug in den Mathematikunterricht halten sollten. Die dabei von den Befürwortern ebenso wie den Gegnern ins Feld geführten Argumente ähneln dabei in frappierender Weise jenen im Hinblick auf die Nutzung von KI-Systemen (für einen historischen Rückblick siehe Weigand, 2003). Während die einen betonten, dass Schülerinnen und Schüler in ihrem heimischen Umfeld binnen Kürze ohnehin Zugriff auf Taschenrechner haben würden und es daher besser sei, ihnen den Umgang mit der Technologie beizubringen als deren Existenz zu ignorieren, befürchteten die anderen, dass Schülerinnen und Schüler durch die Nutzung von Taschenrechnern grundlegende Rechenfähigkeiten nicht mehr in angemessenem Maße ausbilden würden. Aus der Parallelität der Argumente in den damaligen und in den gegenwärtigen Debatten folgt selbstverständlich nicht, dass auch die betreffenden Technologien – also Taschenrechner und KI-Systeme – automatisch gleich zu bewerten wären. Argumente, die in einem Kontext überzogen sind, können in einem anderen Kontext durchaus Gültigkeit besitzen. Dennoch ist erhellend, wie die Kompromissformel im Falle der Taschenrechner an den Schulen bis heute üblicherweise aussieht: Bis ein ausreichendes mathematisches Vorverständnis vorhanden ist, bleiben Taschenrechner verboten; ab einer gewissen Klassenstufe werden sie dann aber erlaubt. Ein ähnlich pragmatisch-nüchternes Vorgehen scheint auch im Hinblick auf KI-Systeme wünschenswert – obgleich vielleicht weniger leicht auszutarieren als im Falle des Taschenrechners.

Ein zweites mit technologischen Veränderungen einhergehendes Spannungsfeld entfaltet sich zwischen der *Hoffnung auf Effizienzsteigerung* und der *Furcht vor ökonomischer Verdrängung.* Einerseits liegt in KI-Technologien das Versprechen, dass sie neue Spielräume für Autonomie und Kreativität eröffnen werden. Auch diese Hoffnung ist nicht neu: So war mit der Einführung moderner Haushaltsgeräte wie Geschirrspüler und Waschmaschine die Erwartung verbunden, dass sich damit die Zeit verkürzen lassen würde, die Menschen mit Hausarbeit verbringen. De facto aber veränderten sich mit der Verbreitung der genannten Geräte auch die Maßstäbe für Hygiene und Sauberkeit und stiegen die Erwartungen an Kindererziehung, so dass nicht die erhofften Freizeitgewinne zu verzeichnen waren (Vanek, 1974; differenzierend Gershuny & Harms, 2016). Andererseits ist mit der zunehmenden Verbreitung von KI-Systemen auch die Befürchtung des Wegfalls von Arbeitsplätzen verbunden. Prognosen sind in dieser Frage allein schon deshalb schwierig, weil sich die Wirkung von Technologie nicht einer naturgesetzlichen Logik folgend entfaltet, sondern maßgeblich von politisch-gesellschaftlichen Leitlinien beeinflusst wird. Gewissermaßen als Mahnung, dass technologische Umwälzungen durchaus soziale Sprengkraft entfalten können, sei an dieser Stelle auf die Maschinenstürmer in der Phase der industriellen Revolution verwiesen, etwa im Rahmen der Swing Riots in England oder des Schlesischen Weberaufstandes in Deutschland.

Ein drittes Spannungsfeld entwickelt sich zwischen der *Hoffnung auf Demokratisierung* und der *Sorge um Wahrheit*. Auf der einen Seite steht der Gedanke, dass KI – zum Beispiel als Übersetzer und Konversationspartner – dazu beitragen könnte, Beschränkungen zu überwinden. Das erinnert an die Frühphase des Internets, in der viele Menschen an dessen Potenzial glaubten, politische Teilhabe zu verbessern und die öffentliche Sphäre zu erweitern; gleichzeitig wurde damals schnell klar, dass es so einfach nicht ist und dass sich jede Technologie missbrauchen lässt (z. B. Papacharissi, 2002). Das wiederum lässt sich auch über KI sagen, der oft genug vorgeworfen wird, in den Trainingsdatensätzen enthaltene Verzerrungen zu reproduzieren, regelmäßig falsche oder irreführende Angaben zu machen beziehungsweise – etwa im Falle von Deep Fakes – etwas als Wirklichkeit auszuweisen, das dieser nicht entspricht. Auch diese Debatten sind nicht neu. So ist beispielsweise im Falle der Fotografie die Objektivität der Abbildung von Anfang an fragwürdig: Das liegt schon daran, dass bereits die Auswahl eines Motivs ebenso wie dessen kontextuelle Einbettung interpretative Akte sind. Darüber hinaus begleitet das Fälschen und Manipulieren die Fotografie seit ihrer Erfindung (siehe Fineman, 2012), man denke etwa an die Bildretuschen in der stalinistischen Sowjetunion, bei denen in Ungnade gefallener Mitstreiter später aus Aufnahmen entfernt wurden. Dass KI-Tools derartige Manipulationen in viel größerem Umfang und mit viel geringerem Aufwand erlauben, ändert dabei nichts an der Tatsache, dass das grundsätzliche Problem kein gegenwartsgebundenes Phänomen ist.

Fazit

Künstliche Intelligenz ist eines der bestimmenden Themen unserer Gegenwart. Wir haben auf Basis technologiegeschichtlicher, problemgeschichtlicher und diskursgeschichtlicher Überlegungen deutlich zu machen versucht, dass es sich lohnt, den Blick nicht nur nach vorne, sondern auch nach hinten zu richten – gerade, um ein Gegengewicht zu vereinseitigenden Narrativen zu schaffen. Die gegenwärtigen Debatten sind nicht aus dem Nichts auf uns gekommen: Ihre Kontextualisierung ist daher ein wichtiger Schritt zu ihrer Entdramatisierung.

Literatur

Baddeley, A. D. & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 8, pp. 47 – 89). Cambridge, MA: Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Bender, E. M., Gebru, T., McMillan-Major, A. & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610 – 623). https://doi.org/10.1145/3442188.3445922

Cave, S., Dihal, K. & Dillon, S. (Eds.). (2020). *AI narratives: A history of imaginative thinking about intelligent machines.* Oxford: Oxford University Press.

Cole, D. (2024). The Chinese room argument. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2024 Edition). Redwood City, CA: Stanford University. https://plato.stanford.edu/archives/win2024/entries/chinese-room/

Fineman, M. (2012). *Faking it: Manipulated photography before Photoshop.* New Haven, CT: Yale University Press.

Gershuny, J. & Harms, T. A. (2016). Housework now takes much less time: 85 years of us rural women's time use. *Social Forces*, *95*(2), 503 – 524. https://doi.org/10.1093/sf/sow073

Green, C. D. (2005). Was Babbage's Analytical Engine intended to be a mechanical model of the mind? *History of Psychology, 8*(1), 35 – 45. https://doi.org/10.1037/1093-4510.8.1.35

Haigh, T. (2023). There was no 'first AI winter'. *Communications of the ACM*, *66*(12), 35 – 39. https://doi.org/10.1145/3625833

He, R., Cao, J. & Tan, T. (2025). Generative artificial intelligence: a historical perspective. *National Science Review*, *12*(5), Article nwaf050. https://doi.org/10.1093/nsr/nwaf050

Heßler, M. (2025). *Sisyphos im Maschinenraum. Eine Geschichte der Fehlbarkeit von Mensch und Technologie.* München: C.H. Beck

Hutmacher, F. & Mayrhofer, R. (2023). Psychology as a historical science? Theoretical assumptions, methodological considerations, and potential pitfalls. *Current Psychology, 42*, 18507 – 18514. https://doi.org/10.1007/s12144-022-03030-0

Newell, A., Shaw, J. C. & Simon, H. A. (1959). *Report on a general problem-solving program.* Santa Monica, CA: Rand Corporation.

Ohlsson, S. (2012). The problems with problem solving: Reflections on the rise, current status, and possible future of a cognitive research paradigm. *The Journal of Problem Solving*, *5*(1), Article 7. https://doi.org/10.7771/1932-6246.1144

Papacharissi, Z. (2002). The virtual sphere: The internet as a public sphere. *New Media & Society*, *4*(1), 9 – 27. https://doi.org/10.1177/14614440222226244

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

Rumelhart, D.E., McClelland, J. L. & The PDP Research Group (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations.* Cambridge, MA: MIT Press.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*(3), 417 – 424. https://doi.org/10.1017/S0140525X00005756

Serrano, J. I., del Castillo, M. D. & Carretero, M. (2014). Cognitive? Science? *Foundations of Science*, *19*(2), 115 – 131. https://doi.org/10.1007/s10699-013-9323-1

Shao, Z., Zhao, R., Yuan, S., Ding, M. & Wang, Y. (2022). Tracing the evolution of AI in the past decade and forecasting the emerging trends. *Expert Systems with Applications*, *209*, Article 118221. https://doi.org/10.1016/j.eswa.2022.118221

Stein, J.-P., Messingschlager, T. & Hutmacher, F. (2023). Künstliche Intelligenz. In M. Appel, F. Hutmacher, C. Mengelkamp, J.-P. Stein & S. Weber (Hrsg.), *Digital ist besser?! Psychologie der Online- und Mobilkommunikation* (S. 247 – 260). Springer. https://doi.org/10.1007/978-3-662-66608-1_17

Turing, A. M. (1937). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, *s2 – 42*(1), 230 – 265. https://doi.org/10.1112/plms/s2-42.1.230

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *59*(236), 433 – 460. https://doi.org/10.1093/mind/LIX.236.433

Vanek, J. (1974). Time spent in housework. *Scientific American*, *231*(5), 116 – 121. https://doi.org/10.1038/scientificamerican1174-116

Weigand, H.-G. (2003). Taschenrechner im Mathematikunterricht: Ein retrospektiver Vergleich der Diskussion und Vorgehensweise in der BRD und in der DDR. In H. Henning & P. Bender (Hrsg.), *Didaktik der Mathematik in den alten Bundesländern* (S. 205 – 216). Universität Magdeburg, Universität Paderborn.

Weizenbaum, J. (1976). *Computer power and human reason: From judgment to calculation.* New York, NY: W. H. Freeman and Company.

Wendt, A. N. (2024). Teleogenesis. A Bergsonian view on problem-finding in human and artificial intelligence. *Thaumàzein: Rivista di Filosofia*, *12*(1), 134 – 158. https://doi.org/10.13136/thau.v12i1.283

Wiener, N. (1948). *Cybernetics: or control and communication in the animal and the machine.* Hoboken, NJ: John Wiley.

Onlineveröffentlichung: 21. 07. 2026

Förderung

Open Access-Veröffentlichung ermöglicht durch die Julius Maximilians-Universität Würzburg.

Dr. Fabian Hutmacher
Institut Mensch-Computer-Medien
Julius-Maximilians-Universität Würzburg
Oswald-Külpe-Weg 82
97074 Würzburg
Deutschland
fabian.hutmacher@uni-wuerzburg.de