**Aktuelle philosophische Texte zu künstlicher Intelligenz**

# Philosoph:innen

* Sally Haslanger
* Joanna J. Bryson
* Nick Bostrom
* Daniel Kahneman
* Marcus Hutter
* Jörg Noller
* Kai Denker
* Jonathan D. Geige
* Dr. Manuela Lenzen
* Daniel Martin Feige

# Schwerpunkte

* Überblick/ Einstieg/ „Leitfaden“
  + Klaus Mainzer: Philosophisches Handbuch Künstliche Intelligenz
  + Jens Kipper: Künstliche Intelligenz: Fluch oder Segen?
  + Bert Heinrichs, Jan-Hendrik Heinrichs und Markus Rüther: Künstliche Intelligenz
  + Manuela Lenz: Künstliche Intelligenz. Was sie kann & was uns erwartet
  + Gerhard Paaß/Dirk Hecker: Künstliche Intelligenz. Was steckt hinter der Technologie der Zukunft?
  + Alan Turin: Kann eine Maschine denken? Kursbuch 8, 1967.
* Allgemein
  + Margaret A. Boden: Artificial Intelligence. A very short Introduction, Oxford: Oxford University Press 2018.
  + Nick Bostrom: Superintelligenz: Szenarien einer kommenden Revolution (UB)
  + Nick Bostrom: Die Zukunft der Menschheit. Aufsätze, Berlin: Suhrkamp 2018.
  + Brian Cantwell Smith: The Promise of Atifcial Intelligence. Reckoning and Judgement, Cambridge/Mass: MTI Press 2019.
  + Noël Carroll, The Philosophy of Motion Pictrues, Malden/Ma.: Blackwell 2008.
  + Gabriele Gramelsberger: Phiilosophie des Digitalen zur EInführung, Hamburg: Junius 2023.
  + John Haugeland, Artificial Intelligence. The very Ideal, Cambridge/Mass.: MIT Press 1989.
  + Stefan Heidenreich: Flip Flop. Digitale Datenströme und die Kultur des 21,. Jahrhunderts. München: Hanser 2004.
  + Christoph Hubig: Die Kunst ders Möglichen I. Technikphilosophie als Reflexion der Medialität. Bielefeld: Transcript 2006.
  + Justin Joque,: Revolutionary Mathematics. Artifivial Intelligence, Statics and the Logic of Capitalism, London: Verso 2022.
  + Klaus Kornwachs: Philosophie der Technik. Eine Einführung, München: beck 2013.
  + Sybille Krämer: „Medienphilosophie des digitalen. Warum und wie die Philosophie über das Digitale reflektieren sollte, aber dies so wenig tut, in:Sybille Krämer und Jörg Nöller (Hg.), Was ist digitale Philosophie? Phäänomene, Formen und Methoden, Paderborn: Brill Mentis 2022, S. 3-30.
  + Erik J. Larson: The Myth of Artificial Intelligence. Why computers can´t think the way we o, Cambridge/Mass.: Belknap 2021.
  + John Pollock, „What am I? Virtual Maschines and the Mind/Body Problem, in: Philosophy and Phenomenological Research 72/2 (2008), S. 237-309.
  + Arno Schubbach:“Judging Maschines. Philosophical Aspects of Deep Learning, in: Synthese 198/2 (2021), S. 1807-1827.
  + Joseph Weizenbaum, Die Macht der Computer und die Ohnmacht der Vernunft, Frankfurt a.M.: Suhrkamp 1978.
  + Gold, Peter: Philosophische Aspekte Künstlicher Intelligenz, in: Der Mensch in der Perspektive der Kognitionswissenschaften, herausgegeben von Peter Gold und Andreas K. Engel
  + Kai Denker: „Gefährliche Texte: Lems Waffensysteme des 21. Jahrhunderts“, in: Friedrich, Alexander et al. (Hrsg.): Kosmos Stanisław Lem. Zivilisationspoetik, Wissenschaftsanalytik und Kulturphilosophie, Wiesbaden: Harrassowitz Verlag 2021 (=Veröffentlichungen des Deutschen Polen-Instituts Darmstadt, Band 36), S. 127–144.
  + Ray Kurzweil: KI : das Zeitalter der künstlichen Intelligenz (darin einige interessante Beiträge)
  + Herbert Simon: Die Wissenschaft vom Künstlichen, Hamburg o. J.
  + Margret Boden: Min das Machine. A History of Cognitive Science.
  + Margret Boden: The Philosophy of Artificial Life. 1996.
  + Ray Kurzweil: Menschheit 2.0. Die Singularität naht. 2013.Jaron Lanier: Wenn Träume erwachsen werden. Ein Blick auf das digitale Zeitalter. München 2015.
* Ethik/Moral
  + Krischke, Wolfgang (2020): „Der Algorithmus kennt keine Moral“, Frankfurter Allgemeine Zeitung (2020/4), S. N. 4.
  + Lauer, Christopher (2024): „Deutschland braucht die Staats-KI“, Frankfurter Allgemeine Zeitung (2024/337), S. 11.
  + Priddat, Birger P./Sven Wilhelm (2022): Digitale Welten, Marburg: Metropolis.
  + Jenna Burell: How the Machine 'Thinks:' Understanding Opacity in Machine Learning Algorithms
  + Sebastian Knell: Essay: Künstliche Intelligenz und menschliche Würde – ein aporetisches Verhältnis?, in: Zeitschrift für Ethik und Moralphilosophie, 203-229
  + Karsten Wendland: Dürfen Maschinen denken (können)? Warum KI eine Ethik braucht ([hier](https://arxiv.org/pdf/2208.07402v1).)
  + Andreas Bischof: Soziale Maschinen bauen. Epistemische Praktiken der Sozialrobotik. Bielefeld 2017.
  + Christisn Heller: Post-Privacy. Prima leben ohne Privatsphäre. München 2011.
  + Nowotny, Helga (2023): Die KI sei mit euch. Macht, Illusion und Kontrolle algorithmischer Vorhersage, Berlin: Matthes & Seitz.
  + Mannino Adriano et al. (2015): Künstliche Intelligenz: Chancen und Risiken, in: Diskussionspapiere der Stiftung Für Effektiven Altruismus 2:1-17.
  + Yvonne Hofstetter: Das Ende der Demokratie. Wie künstliche Intelligenz die Demokratie übernimmt und uns entmündigt. München 2016.
* Mensch-Maschine-Interaktion
  + Müßgens, Jean (2022): „Eine Untersuchung algorithmischer Kommunikationsverhältnisse unter Rekonstruktion einer Mensch/Maschine-Interaktion im Licht der Philosophie von Emmanuel Levinas und Edmund Husserl“, in Birger P. Priddat/Sven Wilhelm (Hg.), Digitale Welten, Marburg: Metropolis 2022, S.121–148.
  + Peter Seele: Künstliche Intelligenz und Maschinisierung des Menschen
  + Elena Esposito: Artificial Communication: How Algorithms Produce Social Intelligence
  + (zusammen mit Alisa Kronberger): „When I touch the person on the screen. Der digitale Körper zwischen Materialität und Virtualität“ (2024). In: Sibylle Krämer, Jörg Noller (Hrsg.): Was ist digitale Philosophie? Phänomene, Formen und Methoden. Paderborn: Brill | Mentis, S. 156–179.
  + Florian Rössler: Smart Cities im Cyberwar. Frankfurt a.M. 2016.
  + Ipke Wachsmuth: Menschen, Tiere und Max. Natürliche Kommunikation und künstliche Intelligenz. Berlin 2013.
* Ästhetik/Umgang mit digitalen Räumen/ …
  + Arne Manzeschke und Bruno Gransche: Bilder machen Menschen. Zur Bildermacht der Künstlichen
  + Intelligenz
* Neuronale Netze/„Schachcomputer“
  + Ray Kurzweil: Das Geheimnis des menschlichen Denkens. Einblicke in das Reverse Engineering des Gehirns. Berlin 2014.
  + Thomas Metzinger: Der Ego-Tunnel. Vom Mythos des Selbst zur Ethik des Bewusstseins. Berlin 2009.
  + Thomas Rid: Maschinendämmerung. Eine kurze Geschichte der Kybernetik. Berlin 2016.
  + John Searl
  + Ray Kurzweil: Wann Computer die Rechenkapazität des Menschen erreichen, in: ders.: Menschheit 2.0 : die Singularität naht.
  + John von Neumann: Die Rechenmaschine und das Gehirn. München 1965.
  + John Searle: Geist, Hirn und Wissenschaft. Frankfurt a.M. 1986.
  + Achim Stephan/Sven Walter: Handbuch Kognitionswissenschaft, Stuttgart 2013.
  + Alan Turin: Turing, Alan M.: Computing Machinery and Intelligence/Können Maschinen denken? (gibt es auch als Reclam-Ausgabe)

# Sonstiges

* Stanford Encyclopedia of Philosophy - Artificial Intelligence (insb. Teile 1-5)
* Eingesprochenes Interview mit Sybille Anders (Astrophysikerin und Philosophin): <https://www.faz.net/podcasts/f-a-z-wissen-der-podcast/kuenstliche-intelligenz-droht-der-kontrollverlust-bei-ki-19303468.html>
* Webseite von Geoffrey E. Hinton; <https://www.cs.toronto.edu/~hinton/> (für eine tiefere Auseinandersetzung nicht unbedingt phil. Art)
* Kate Crawford: Atlas der KI. DIE MATERIELLE WAHRHEIT HINTER DEN NEUEN DATENIMPERIEN. à Für Zwischendurch und nebenbei lesen, um zu entdecken, welch eine Ökonomisierung auch dahinter steckt und welche Bereiche für die Datenspeicherungen usw. benötigt werden à „der Körper hinter der transzendentalen KI“
* Grace, Katja/John Salvatier/Allan Dafoe/Baobao Zhang/Owain Evans (2018): „When Will AI Exceed Human Performance? Evidence from AI Experts“, Journal of Artificial Intelligence Research (2018/62), https://doi.org/10.48550/arXiv.1705.08807, S. 729–754.
* Rouvroy, Antoinette (2019): „Die Herrschaft der Algorithmen ist ziellos“, philomag.de 22.5.2019, https://www.philomag.de/artikel/antoinette-rouvroy-die-herrschaft-der-algorithmen-ist-ziellos?utm\_source=philomag.de&utm\_campaign=43342ed610-EMAIL\_CAMPAIGN\_2021\_02\_09\_04\_38\_COPY\_01&utm\_medium=email&utm\_term=0\_8d7052c3cd-43342ed610-1204767306 (Zugriff: 12/8/2024).
* Zachary C. Lipton: The Mythos of Model Interpretability
* Bohlmann, Markus; Berger, Annika. 2024. ‘ChatGPT and the Writing of Philosophical Essays. An Experimental Study with Prospective Teachers on How the Turing Test Inverted.’ Teaching Philosophy 47, Nr. 2: 233–253. doi: 10.5840/teachphil202451200.
* Vortrag Kai Denker: „Was ist differenzphilosophische Algorithmenkritik?“, 13. Juni 2018, Sommerkolloquium des Instituts für Philosophie, TU Darmstadt.
* Michael Feiten: Hannah Arendt und das Digitale
* Gewinnen oder verlieren wir Macht durch das Digitale?, in: ders. Digitalisierung und Digitalität.
* Költzsch, T. (2018) Eine künstliche Intelligenz als Bürgermeister. ([hier](https://www.golem.de/news/ki-kandidat-%20eine-kuenstliche-intelligenz-als-buergermeister-1804-133830.html))
* Lamarr (Das Lamarr-Institut für Maschinelles Lernen und Künstliche Intelligenz): <https://lamarr-institute.org/de/lamarr/>
* MIT (Massachusetts Institute of Technology): <https://web.mit.edu/about/>
* DeepmindBlog: <https://deepmind.google/discover/blog/>
* Jaron Lanier: Gadget. Warum die Zukunft uns noch braucht. Berlin 2010.
* Ders.: Wem gehört die Zukunft? Du bist nicht Kunde der Internet-Konzerne, du bist ihr Produkt. Hamburg 2014.
* Jaron Lanier: Wenn Träume erwachsen werden. Ein Blick auf das digitale Zeitalter. München 2015.
* Stefan Lorenz Sorgner: Transhumanismus. „Die gefährlichste Idee der Welt“ !?, Freiburg i.Br. 2016.
* Rory Cellan-Jones: Stephen Hawking warns artificial intelligence could end mankind, ([hier](https://www.bbc.com/news/technology-30290540))
* Envisioning: Human-Machine interactions ethics ([hier](https://www.envisioning.io/signals/human-machine-interaction-ethics))

**Empfehlungen Shirin:**

KI Texte:

- Brian Cantwell Smith: The Promise of Artificial Intelligence

- Artikel von Rainer Mühlhoff (die meisten auch auf Englisch... Empfehlung v.a.: "Human-aided artificial intelligence")

- Matteo Pasquinelli: Das Auge des Meisters (historisch-materialistische Studie von KI) (Pasquinelli hat auch kürzere Paper zu KI/Algorithmic thinking geschrieben)

Daniel Martin Feiges "Kritik der Digitalität"

drei KI-Diskurse, die aktuell in der Philosophie virulent sind, ist m.E. der rund um LLMs (—> inwiefern handelt es sich um "Denken" oder "Sprechen" oder gar "Bewusstsein"?!), um KI-generierte Kunst und um ethische Fragen (Privatheitsdebatten, Ausbeutung, Diskriminierung, etc.).

Je nachdem, wo du da einsteigen möchtest, könnte ich noch weiter überlegen, welche Texte empfehlenswert sind (es gibt leider auch sehr viel Schrott...).

Katias kritische KI-Philosophie finde ich super.

Z.B.: Schwerzmann: “Abolish! Against the Use of Risk Assessment Algorithms at Sentencing in the US Criminal Justice System.” Philosophy & Technology 34, no. 4 (2021): 1883–1904. https://doi.org/10.1007/s13347-021-00491-2