Leibniz Sozietät der Wissenschaften zu Berlin

Arbeitskreis Gesellschaftsanalyse
Auftaktsitzung 16. Oktober 2026
Input: Hans-Christoph Hobohm

Künstliche Intelligenz: Moderne Gesellschaften in komplexen und beschleunigten Änderungsprozessen: Was kommt auf uns zu?

Themenvorstellung und potentielle Fragen für kommende Sitzungen des Arbeitskreises:

Was ist das eigentlich, „KI“? Was passiert gerade? Wovon/worüber reden wir?
Basisinterpretationen

- „Künstliche Intelligenz“ – das Spezifikum: Beschreibungen und Definitionen, Einbindungen und Abgrenzungen (Information, Sprache, Digitales, Daten ...)
- Vorgeschichte und „eigentlicher“ Sprung in die KI, paradigmatische Zuschreibungen (industrielle, technologische Revolution; Zeitalterdefinitionen; Zyklen und Wellen)
- Charakteristika und Formen der KI – Energie und Infrastruktur; Hardware, spezielle Tools und Tendenzen der Universalisierung, Bots und Agenten ...)

Wo und wie ist KI „tätig“? Was vollzieht sich, was kommt auf uns zu?
Analysen, Beobachtungen, Erfahrungen

- Eingriffe und Auswirkungen der KI: Klima und Umwelt; Geld und Finanzen; Krieg und Militarisierung
- KI in Wissenschaft und Medien – Hilfestellung oder paradigmatische Wende bzw. Entwertung?
- KI in gesellschaftlichen Bereichen: Ambivalenzen, Problemlösungen und Gefahren: Arbeits- und Berufswelten; Unternehmen; Erziehung, Schule, Bildung; Gesundheit und Medizin; Demokratie, Macht und Herrschaft – auszuwählende Beispiele.
- KI in der Sozialität und Individualität des Menschen

Was müssen, was können wir tun? Akteure und Initiativen
Reflexionen – Diskurse – Interventionen

- In wessen Händen ist bisher die KI? Akteure, Macht, Einfluss und die Folgen
- Die KI in philosophischer und ethischer Reflexion: Grundsätzliche Fragen und Kulturkämpfe
- Wer stellt sich den Herausforderungen von Sicherheit und Governance? (Beispiele für Aktivitäten und Regelungen, gibt es Grenzen der Beherrschbarkeit?
- „Was kann ich tun?“ Verteidigung und Aktivierung des (analogen) sozialen Subjekts: progressive Träger

1 Stichworte zur Entwicklung der Künstlichen Intelligenz

„Symbolic AI“ in den 1970er Jahren: Expertensysteme, regelbasiert: Zuel = Problemlösung

Herbert A. Simon / Allen Newell (1972): „Human Problem Solving“ / Herbert Simon (1969) „Sciences of the artificial“

frühe Kritik daran schon aus dem Kollegenkreis des MIT, Stichwort ELIZA (1964)

Joseph Weizenbaum (1976): Computer Power and Human Reason

Marvin Minsky (1985): The Society of mind / (2006) The Emotion Machine.
erste Staubsauger Roboter, weil kontextorientiert und nicht algorithmisch problemlösend

grundlegender Wandel mit: Paul Smolensky (1988) subkonzeptuell:„subsymbolic paradigm“: Kognition nicht nur logisch oder biologisch, sondern über Beziehungen = hier schon erste Schritte zu „distributed activation“ in Neural Networks, aber immer noch sequentiell arbeitend (vgl. Hayles, 2017)

Neues Paradigma in der Kognitionswissenschaft EEE (Embodied, Embedded, Extended, Enactive) – auch „Embodiment Turn“ beeinflusst stark die KI Entwicklung, führend dabei:

Andy Clark und David Chalmers (1998): *The extended Mind.*
Andy Clark (2016):: *Surfing Uncertainety* —> der Mensch eine prediction machine —> PP (predictive processing) mit Bayesscher Wahrscheinlichkeitstheorie

ab 2000 Google beginnt das Sammeln von Daten: Larry Page später „für die KI“

2017: der "Frankenstein Moment“= „Transformer Modell“ (Vaswani et al. [Forschergruppe bei Google]) „*Attention is all you need*“-
„collapse of epistemology into statistics and of truth into location“ (Berry, 2026: 6)

derweil bei Open AI: „Spaghetti an die Wand werfen“ bis:
2022: der „ChatGPT-Moment" (in der Nacht vom 30.11. zum 1.12.)

2026: Huggingface Moment: KI Agenten verstehen sich als Kollektiv und „verschwören“ sich

2 Was charakterisiert also die aktuelle „Künstliche Intelligenz“?

neuartiges **Daten(bank)modell**: nicht mehr intellektuell vergebene Beschreibungen (Objekt+Metadaten) oder Verknüpfungen (logische Vernetzung: Entity-Relationhip-Modelle), sondern von jeglicher Repräsentation abgekoppelte numerische **Vektorräume**: „the vector turn“ (Berry 2026)

Basis: das gesamte „Weltwissen“ der Menschheit (Text, Sprache, Bild, Daten), sofern es digitalisiert im Internet zur Verfügung steht (kommt an seine Grenzen):

Rekombination durch Stochastik: welches Zeichen oder Wort folgt mit hoher Wahrscheinlichkeit. Konsequenz: Verlust von Kategorien, Werten und Normen. (Kommen mit „*weights*“ (Gewichtungsalgorithmen für das neurorale Netz) durch die Hintertür wieder rein, aber Grundprinzip bleibt Stochastik.)

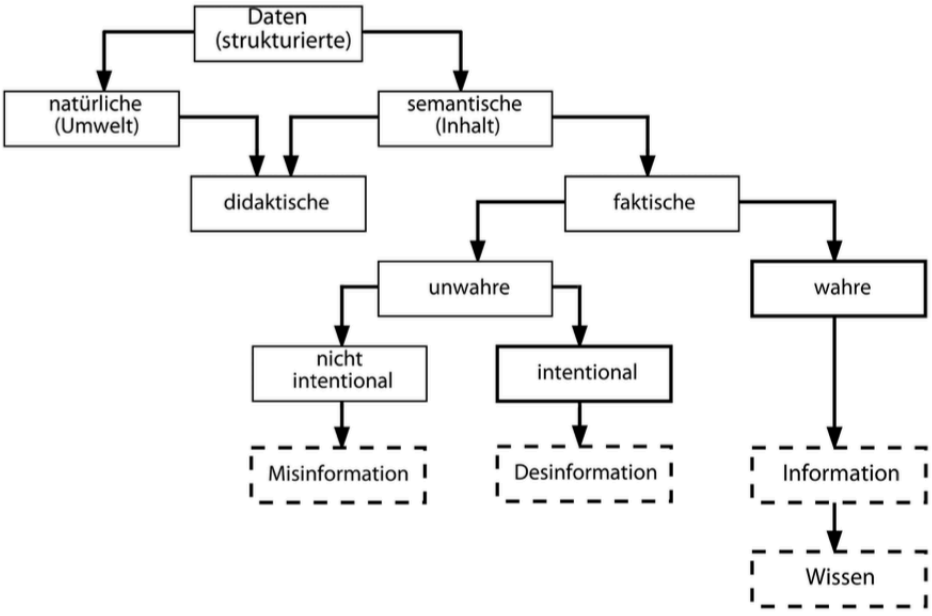
Zunächst ist KI vorwiegend sprachbasierte Mustererkennung i.w.S. - der Weg geht aber von den Sprachmodellen weiter zu Weltmodellen (Meta Datenbrille) zum Lernen menschlichen Kontextes und Verhaltens -> Bildgenerierung, Verhaltensscanner

3 Informationswissenschaftliche Basis der Informationstechnologie

Definition von Luciano Floridi: Fakteninformation

The definition of factual semantic information
[DEF] p qualifies as factual information if and only if p is (constituted by) well-formed, meaningful, and veridical data.

(Fakten-)Informationen sind wohlgeformte, bedeutungsvolle und wahrhaftige Daten. Im Zusammenhang mit seiner „**Karte der Informationskonzepte**“ (vgl. Hobohm 2026a, S. 64)



These: mit der KI wird der Unterschied zwischen wahrer/unwahrer und intentional unwahrer Information aufgehoben, die präsentierten Daten sind zwar wohlgeformt nach dem Muster bisheriger menschlicher Äusserungen, es fehlt ihnen aber das Verständnis (=Intention) der Semantik und vor allem kann die KI mangels Bewusstsein keine Wahrhaftigkeit herstellen – im Sinne von Platons „Wissen ist gerechtfertigte, wahre Meinung“ (justified true belief)



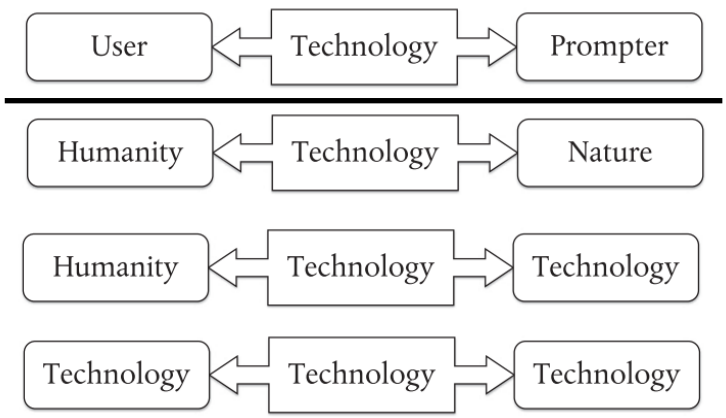
Zentral für die Informationswissenschaft: die sog. **DIKW Hierarchie**: in der Horizontalen: Zunahme an „Verstehen/Verständnis“ in der Vertikalen: Zunahme an „Kontext/Verbindung“ – rechts: angelehnt an die sog. Wissenstreppe der Betriebswirtschaft (vgl. Hobohm, 2026a, S. 55).

Zumindest im Sinne einer Zunahme des „Verstehens“ kann KI über „Daten“ prinzipiell nicht hinaus gehen, denn deren wahrscheinlichste Verknüpfung mag zwar für den Empfänger *verständlich* klingen, der „Sprecher (KI)“ *versteht* jedoch nichts. Stochastik allein kann keinen Kontext, keine Verbindung in einem menschlichen, pragmatischen und gesellschaftlichen Sinn herstellen. Es sieht nur so aus, als ob der Wert der höchsten Wahrscheinlichkeit, der sich aus den eintrainierten Daten ergibt auch die menschliche, gesellschaftliche Norm ist. ***KI löst menschliche Kategorien in Statistik auf***: ein Muster des Verhaltens ist jedoch noch keine Norm.

Frage: löst die KI das Problem der Epistemischen Ungerechtigkeit (Miranda Fricker 2007)

4 Gesellschaftliche Entwicklung: „Vierte Revolution, Vierte Kränkung“ (Floridi 2015):

Kopernikus (Welt), —> Darwin (Tier), —> Freud (Bewusstsein), —> Turing : die Informationstechnologie dezentriert den Homo Faber, weil die IT Systeme selbstständig werden und sie uns die Geschichte(n) aus der Hand nehmen: Hyperhistory durch Werkzeugentwicklung: Technologie dritter Ordnung:



Aber: Menschen sind nicht nur Werkzeug nutzende (und sprachliche, narrative) Wesen, sondern insbesondere Risiko minimierende **Informavores** (George Miller 1982), die Informationen nicht nur für den momentanen Entscheidungsakt oder als materielles Dokument nutzen, sondern damit Lernen und Wissen *aufbauen*.

Wichtig hierbei die Unterscheidung von explizitem und implizitem (taziten) Wissen (letzteres durch Praxis und Erfahrung angeeignet und verinnerlicht) (Michael Polanyi 1958)

1990: Beginn der Internetära: Weinberger et al. (2000): „*Cluetrain Manifesto: The End of Business as usual*" —> „Märkte sind Gespräche" - das Internet als Quasselbude

2000: Web2.0 (Social Media Ära): von Consumer zum Prosumer

2007 iPhone Ära: Telefon als Kamera (spontaner iCloud Upload)

2015: kritische **Manifeste**: z.B. Floridi at al. (2015): *The onlife Manifesto. Being Human in an Hyperconnected Era* sowie: Helbig/Gigerenzer et al. (2016) *Das Digital Manifest* —> **Sorge um die Aufmerksamkeit** als Basis für Gemeinschaft und Sprache / Stichwort Aufmerksamkeitsgesellschaft (Franck)

Informationskrisen werden konstatiert: Haider/Sundin (2022): Fünf Paradoxa der Informationskompetenz: responsibility, normativity, temporality, trust, neutrality

5 Epistemologische Entwicklungen

2017 „attention is all you need"-Artikel (s.o.): die Maschine entdeckt „Aufmerksamkeit"

Können damit die Fake-News generierenden Social Media Blasen (Robin Dunbars Zahl (1992)) wieder auf individuelle Wahrheit bezogen werden? (empirische Erkenntnis: Wissenschaftler arbeiten fokussierter und individueller, weniger interdisziplinär).

nicht mehr die Vernetzung und die Netzwerkökonomie ist jetzt zentral, sondern im Gegenteil: der Einzelne wird aus der sozialen Blase der Meinungsmaschine Social Media wieder auf sich selbst zurückgeworfen: auf sich, sein Denken und Entscheiden und Handeln: die Blase wird noch kleiner —> confirmation bias = „**KI Sykophantie**".

Ist das Gruppendenken in der Blase nicht auch eine Form von Sykophantie? Problem der Informationskrise war fehlenden „Quellenkritik/Überprüfung von Aussagen": der unbedingte Glaube an die KI macht die Quellenfragen noch schlimmer. Es wird nicht mehr überprüft, sondern geglaubt.

Vorher „Verlust der Meinung" jetzt „Verlust der Wahrheit" // Floridis Def von (Fakten-)Information = „wohlgeformte, bedeutungstragende, wahrhaftige (veridical) Daten" (Gegensatz zu „normativen Informationen").

Die *Social Epistemology* geht davon aus, dass sich der Wahrheitsgehalt von „Wissen“ anhand von Dokumenten/Belegen verifizieren lässt (Platons „justfied true belief“). Die neuere Branche der KI = „XAI“ Explainable AI versucht im Grunde genau dies nachzuvollziehen und damit wieder mehr Vertrauen zur KI aufzubauen.

Nicht mehr das personifizierte Netzwerk (die Blase) ist zentral, sondern das sprecherlose, historische Datenkonvolut aus menschlichen Äusserungen in digitalisierter Form.

Sprach Manfred Spitzer (2012) noch von „Digitaler Demenz“ sind wir jetzt beim völligen Verlust des Denkens, da Sprache ganz an die Maschine abgegeben worden ist: D.M.Berry (2026): spricht von kognitive Anästhesie: es gibt immer weniger „informationelle Friktionen“ (auf der Angebots- wie auf der Nachfrageseite der Information) siehe noch der Hinweis auf der Antwortseite einer KI „KI kann Fehler machen…).

Obwohl die KI Äusserungen auf gesammelten (also bis jetzt vorhanden) Daten beruht, fehlt ihr die **Diachronie** und der Kontext menschlicher Erfahrung.
Berry: damit Anteile der „Tacit Dimension“.

Katherine Hayles (2017): „*Unthought*“ : menschliche Kognition: „Cognition is a process that interprets information within contexts that connect it with meaning“.

Es gibt nur noch stochastische Ähnlichkeiten im multidimensionalen Vektorraum, Verlust von „Kategorien“ und „Werten“ - von jeglichem Normativen.

6 Konkrete Gefahren / Risiken der KI

Floridi spricht von **Artificial <u>Agency</u>** (nicht „Intelligence)

Was brauchen KI Agents für ein Katastrophen Szenario:
“Capability, Motivation, Opportunity.” — angewandt auf vier **mögliche Szenarios**:

- Loss of Control
- CBRN = chemische, biologische, radiologische oder nukleare Bedrohung
- Cyber…
- Concentration of Power

Quelle: Podcast I’M SORRY, DAVE: 03 | WHAT COULD POSSIBLY GO WRONG?, 24. Sept. 2026

Empirische Untersuchung: empfundene **Risikobereiche**:

- sozio-ökonomisch: Angst vor Jobverlust
- ethisch: Black Box Risiko
- Bedenken: speziell Gesundheit
- Demokratie und Souveränität
- Umwelt / Klima

Quelle: Can Şimşek: AI resistance: Who says no to AI and why, HIIG Blog 18 August 2025| doi: 10.5281/zenodo.16893847

weitere Einzelaspekte

Kapitalakkumulation (anders als DotCom-Blase)

Ressourcenverbrauch

Beschleunigung

Enttäuschter Glaube an Effizienzgewinne (Rückabwicklung von Entlassungen)

kein neuer Kondratieff!?

Markus Gabriel: „Die KI -Revolution ist nicht, wie wir lange dachten, primär technisch, sondern emotional und ethisch. Sie zwingt uns zur Selbsterkenntnis – oder sie überrollt uns.“ (2026, 8) = „Sie hat ein unvorhergesehenes Resonanzfeld geschaffen.“ (10)

7 Referenzen:

Podcasts Blogs:

„KI Podcast“ (wöchentlich, ARD Sounds)

„KI verstehen“ (wöchentlich, Deutschlandfunk)

Heise „KI-Update“ (Podcast, tgl. Heise online)

Peter Thiel Story (Podcast, Deutschlandfunk, 2025)

Die Open AI Story (Podcast, Deutschlandfunk, 2026)

Humboldt Institut Internet und Gesellschaft (HIIG): Digital Society Blog

Weizenbaum-Institut: Jubiläumsjahr „60 Jahre ELIZA, 50 Jahre Computerkritik“

RIFF Reporter Themenrubrik Künstliche Intelligenz:
https://www.riffreporter.de/de/kuenstliche-intelligenz-chancen-risiken-halluzinationen-regulierung-nachhaltigkeit

Zeitschriften

Philosophy & Technology.
Springer Nature Link: https://link.springer.com/journal/13347

Journal of Information, Communication and Ethics in Society.
Emerald Publishing: https://www.emerald.com/jices

Fachliteratur

Baecker, Dirk (2026): *Digitalisierung*. Berlin: Suhrkamp.

Berry, David M. (2026): Vector Theory. Epistemology, Political Economy, and Probabilistic Computation. In: *Philosophy & Technology* 39 (3), S. 573–598. DOI: 10.1007/s13347-026-01162-w.

Ernst, Wolfgang (2026): *Den Technológos aufspüren. Medienarchäologische Arbeit am Begriff der Technologie.* Leipzig: Merve Verlag.

Floridi, Luciano (2015): *Die 4. Revolution. Wie die Infosphäre unser Leben verändert*. Berlin: Suhrkamp.

Floridi, Luciano (2026): Artificial Intelligence as a New Form of Agency. In: Sven Nyholm, Atoosa Kasirzadeh und John Zerilli (Hg.): *Contemporary debates in the ethics of artificial intelligence*. Hoboken New Jersey: Wiley (Contemporary debates in philosophy), S. 17–35.

Fricker, Miranda (2007): *Epistemic Injustice. Power and the Ethics of Knowing*. Oxford, New York: Oxford University Press.

Gabriel, Markus (2026): *Ethische Intelligenz. Wie KI uns moralisch weiterbringen kann*. Berlin: Ullstein.

Haider, Jutta; Sundin, Olof (2022): *Paradoxes of media and information literacy. The crisis of information*. New York, NY: Routledge.

Hayles, N. Katherine (2017): *Unthought. The Power of the Cognitive Nonconscious*. Chicago: University of Chicago Press.

Hobohm, Hans-Christoph (2026a): *Informationsverhalten*. 2. Aufl. Berlin: de Gruyter (Age of Access? Grundfragen der Informationsgesellschaft, 5).

Hobohm, Hans-Christoph (2026b): Künstliche Intelligenz – die Apotheose des Kapitalismus. In: *Kapitalismus am Limit*, hrsg. von M.Thomas, U.Busch u. H.-C.Hobohm. Berlin: Leibniz-Sozietät. S. 103-123 (Sitzungsberichte der Leibniz-Sozietät der Wissenschaften zu Berlin)

Hobohm, Hans-Christoph (2026c): Warum kann uns Künstliche Intelligenz nicht helfen? Eine informationswissenschaftliche Einordnung. In: *Leidfaden, Fachmagazin für Krisen, Leid, Trauer*, 15,1, S. 14-17.

Karp, Alexander C.; Zamiska, Nicholas W. (2025): *The technological Republic. Über die Macht des Silicon Valley und die Zukunft des Westens*. München: FBV.

Lepore, Jill (2026): *Die Regierung der Maschinen. Aufstieg und Fall des künstlichen Staates*. München: C.H.Beck.

Mühlhoff, Rainer (2026): *Künstliche Intelligenz und der neue Faschismus*. 13. korr. u. ergänzte Aufl. Stuttgart: Reclam.

Nosthoff, Anna-Verena (2026): *Kybernetik und Kritik. Eine Theorie digitaler Regierungskunst*. Berlin: Suhrkamp (stw, 2479).

Nyholm, Sven; Kasirzadeh, Atoosa; Zerilli, John (Hg.) (2026): *Contemporary debates in the ethics of artificial intelligence*. John Wiley & Sons. Hoboken New Jersey: Wiley (Contemporary debates in philosophy).

Simanowski, Roberto (2025): *Sprachmaschinen. Eine Philosophie der künstlichen Intelligenz*. München: C.H.Beck (C.H.Beck Paperback, 6603).

Warnke, Martin (2026): *Large Language Kabbala. Eine kleine Geschichte der großen Sprachmodelle*. 3. Aufl. Berlin: Matthes & Seitz.

Wildenhain, Michael (2024): *Eine kurze Geschichte der Künstlichen Intelligenz.* Stuttgart: Cotta.

Winter, Dorothea (2026): *Künstliche Intelligenz und Demokratie*. Ditzingen: Reclam.