

Sascha Dickel*

Im Imitationsspiel. Über die Kommunikation mit Maschinen und das Streben nach *Artificial General Intelligence*

Within the Imitation Game. On Communication with Machines and the Pursuit of Artificial General Intelligence

<https://doi.org/10.1515/zfsoz-2025-2011>

Zusammenfassung: Der Beitrag untersucht, wie Maschinen als Künstliche Intelligenzen kategorisiert werden und unter welchen Bedingungen diese Kategorisierung gesellschaftliche Akzeptanz finden kann. Er zeigt mithilfe differenzierungstheoretischer Ansätze, inwiefern Künstliche Intelligenz die Unterscheidung von Menschen und Maschinen infrage stellt und arbeitet heraus, welche Rolle dabei die Mediatisierung des Mensch-Technik-Verhältnisses spielt. Dazu rekonstruiert er den Turing-Test als Blaupause für eine kommunikative Lösung des Problems maschineller Intelligenz. Der Beitrag schließt mit der These, dass Künstliche Intelligenz sich aktuell in einer ontologischen Hybridposition zwischen Fiktion und Fakt befindet. Das von Technologieführern proklamierte Rennen hin zu einer Artificial General Intelligence (AGI) erfüllt diesbezüglich die Funktion einer Temporalisierung der Frage nach dem Realitätsstatus von KI.

Schlagworte: Künstliche Intelligenz; Humandifferenzierung; Systemtheorie; Grenzen des Sozialen; Interface; Turing-Test.

Abstract: This article examines how machines are categorized as artificial intelligence, and under which conditions this categorization gains societal acceptance. Drawing on theories of social and cultural differentiation, it shows how artificial intelligence blurs the line between humans and machines and underscores the role of mediatization. To this end, the paper reconstructs the Turing Test as a blueprint for a communicative solution to the problem of machine intelligence. The article suggests that artificial intelligence currently occupies an ontological hybrid position between fiction and fact. In this context, the race toward Artificial

General Intelligence (AGI), as proclaimed by technology leaders, serves to temporalize the question of AI's ontological status.

Keywords: Artificial Intelligence; Human Differentiation; Systems Theory; Borders of the Social World; Interface; Turing Test.

1 Einleitung

„By the time of the Singularity, there won't be a distinction between humans and technology. *This is not because humans will have become what we think of as machines today but, rather machines will have progressed to be like humans and beyond*“ (Kurzweil 2005: 69, Hervorhebung i.O.).

Dem wissenschaftlichen Diskurs ist Künstliche Intelligenz (KI) seit den 1950er Jahren ein Begriff. Auch der Soziologie ist KI nicht fremd (Rammert 2023; Suchman 2007). Frühe Arbeiten haben sich etwa der KI als sich herausbildendes Forschungsfeld (Ahrweiler 1995) oder neuer Hochtechnologie zugewandt (Rammert 1995). In Selbst- und Fremdbeschreibungen wird die Entwicklung des Feldes als eine Folge von Aufschwung- und Krisenphasen erzählt, so genannten ‚KI-Sommern‘ und ‚KI-Wintern‘. In KI-Sommern werden neue wissenschaftliche Ansätze entwickelt und technische Durchbrüche antizipiert. KI erfährt dann über die diesbezüglich aktiven *scientific communities* hinaus Aufmerksamkeit in Massenmedien, Wirtschaft und Politik. In KI-Wintern stoßen wissenschaftliche Paradigmen an ihre Grenzen, technische Durchbrüche scheinen in weite Fernen gerückt und die Aufmerksamkeit jenseits der mit KI befassten Expertengemeinschaften erlischt (Gonsalves 2019; Hirsch-Kreinsen & Krokowski 2023).

Seit der Veröffentlichung von ChatGPT kann von einem regelrechten Hochsommer der KI gesprochen werden.

*Correspondenzautor: Sascha Dickel, Johannes Gutenberg-Universität Mainz, Isaac-Fulda-Allee 2b, 55124 Mainz. E-Mail: dickel@uni-mainz.de

ChatGPT und verwandte Anwendungen werden auch als *generative KI* bezeichnet, da sie kulturelle Produkte wie Bilder, Videos, Audios und Programmcodes generieren können.¹ Die weitreichende Verbreitung und Nutzung von Anwendungen generativer KI in unterschiedlichen Bereichen – von der Kundenbetreuung über kreative Textproduktion bis zur wissenschaftlichen Forschung – verdeutlicht, wie selbstverständlich diese Technologien inzwischen in gesellschaftliche Prozesse eingebettet sind. KI erscheint damit nicht mehr als futuristisches Konzept, sondern zunehmend als Teil unserer alltäglichen Realität.

Das Leistungsspektrum aktueller Deep Learning-Systeme wird dabei von Entwicklerseite – und zunehmend auch in der medialen Öffentlichkeit – nur als eine Momentaufnahme betrachtet. Denn eine ‚echte‘ Künstliche Intelligenz ist auch nach der Auffassung von Expert:innen noch längst nicht erreicht: Das Ziel der aktuellen Technologieführer (vor allem: Open AI und Google) ist die Entwicklung von *Artificial General Intelligence* (AGI). AGI bezeichnet ein – aktuell nichtexistierendes – maschinelles System, das mit einer menschenähnlichen Intelligenz vergleichbar sein soll. Eben dies wurde ursprünglich zur Mitte des vergangenen Jahrhunderts eigentlich auch im wissenschaftlich-technischen Diskurs unter KI verstanden, bevor sich die Erwartungen über Jahrzehnte hinweg abkühlten. Eine solche menschenähnliche Maschine in den kommenden Jahren zu entwickeln, ist das öffentlich kommunizierte Ziel derjenigen Firmen, die sich an der vordersten Front der KI-Forschung positioniert haben (Altman 2023; Morris et al. 2024; Microsoft 2024; Patel & Hassabis 2023).

Bezüglich der Leistungsfähigkeit von KI war der soziologische Fachdiskurs lange eher von einer Skepsis bestimmt, die durch Verweise auf die offenkundigen Grenzen vergangener technischer KI-Paradigmen geprägt war (Collins 1990; Suchman 2007). Doch als Konsequenz der Anwendung von Technologien des Deep Learning wird auch von soziologischer Seite eine noch kaum absehbare Entwicklungsdynamik konstatiert. So diagnostizierte jüngst Fabian Anicker: „Es dürfte kaum einen kontemporären gesellschaftlichen Strukturwandel geben, der sich selbst im Rahmen der normalisierten Hypertrophie der Moderne mit höherem Tempo und größerer Breitenwirkung vollzieht, als die Transition zu Künstlicher Intelligenz (KI)“ (Anicker 2023:

79). Die neueren KI-Systeme seien nicht mehr länger nur Ergänzungen menschlicher Aktivitäten, „sondern potenziell substitutiv“ (Anicker 2023: 101, Hervorhebung i.O.).

Die Soziologie wird durch die aktuellen technologischen Durchbrüche – und mehr noch: die diesbezüglichen mittelfristigen Prognosen einer AGI – in entscheidender Weise herausgefordert. Sozial- und gesellschaftstheoretisch brisante Fragen nach den Graden der Handlungsträgerschaft von Technik (Rammert & Schulz-Schaeffer 2002), einer Symmetrie oder Asymmetrie von Menschen und Maschinen (Latour 2005) oder der Zugehörigkeit von nicht-menschlichen Entitäten zum gesellschaftlichen Personal (Lindemann 2018) erscheinen nicht mehr nur als fachspezifisch zu debattierende Problemstellungen. Sie werden vielmehr zu Fragen gesamtgesellschaftlicher Aushandlung und technologischer Realisierung (ähnlich Nassehi 2019: 257 f.): Was für Erwartungen muss Technik erfüllen, um gesellschaftlich als KI zu gelten? Inwiefern würde sich eine AGI von aktuellen KI-Anwendungen unterscheiden? Welcher Grad an Menschenähnlichkeit ist erforderlich und woran wird diese Ähnlichkeit festgemacht? Auch und gerade die öffentlichen Ausführungen der Technologieführer sind diesbezüglich vage. Ich argumentiere, dass diese technologische Unbestimmtheit kein zu behebendes Defizit der KI-Forschung ist. Denn es handelt sich beim Thema KI nicht um ein rein technisches Problem, sondern vielmehr auch um eine Frage gesellschaftlicher Aushandlung. Die Sozial- und Kulturwissenschaften können dabei einen Beitrag zur gesellschaftlichen Selbstaufklärung leisten. Wie Lucy Suchman betont, sollten sie dazu die öffentliche Rede von KI nicht unreflektiert reifizieren, sondern vielmehr die vermeintliche „thingness of AI“ (Suchman 2023: 2) selbst hinterfragen. Es gelte daher zu rekonstruieren, wie es zu der *Verhärtung der Kategorie* KI kommt (ebd.: 3).

Der folgende Beitrag kommt eben dieser Aufforderung nach. Sein Ziel besteht darin, einen nichtessentialistischen soziologischen Zugang zum Thema KI auszubuchstabieren, der es vermeidet, KI als stabiles Ding aufzufassen. Mein Ausgangspunkt ist, dass KI kein Begriff ist, der präzise bestimmte Artefakte beschreibt, sondern vielmehr ein grundsätzlich strittiger Begriff, der sowohl Informatiker:innen als auch Philosoph:innen immer wieder entgleitet (Morgan 2019). Er wird in Massenmedien, Wissenschaft, Wirtschaft und Politik mobilisiert, um bestimmte Erwartungen an Informationstechnologien zu wecken (Suchman 2023; Bareis & Katzenbach 2022). KI ist eine Kategorie zur Verbesonderung digitaler Technik. In den vergangenen Jahren ist diese Kategorie zur Bezeichnung zahlreicher technischer Anwendungen eingesetzt worden, z. B. für Technologien zur Mustererkennung bei medizinischen Bildern, zur Verarbeitung von Sensordaten für das autonome Fahren, für

¹ Generative KI basiert auf Fortschritten im Bereich des Deep Learning (Vaswani et al. 2017). Deep Learning ist ein spezifischer Ansatz des Machine Learning. Allgemein nutzen Machine-Learning-Methoden Algorithmen, um Muster in Daten zu erkennen und daraus neues Wissen abzuleiten. Deep Learning zeichnet sich dabei durch den Einsatz künstlicher neuronaler Netze aus, die insbesondere bei der Verarbeitung komplexer und umfangreicher Daten besonders leistungsfähig sind.

Empfehlungssysteme im E-Commerce und Streaming oder zur Gesichtserkennung im Kontext von Überwachung und Sicherheit. Es ist gleichwohl kulturell umstritten, ob diese Systeme *als KI* verbesondert werden sollten oder lediglich als gewöhnliche Fälle industriell-moderner Delegation von Funktionen an Technik zu betrachten sind (Latour 1992). Was gestern noch KI war, gilt heute mitunter einfach als Technik.²

Mit generativer KI beginnen sich die gesellschaftlichen Erwartungsmuster hinsichtlich dessen, *was KI ist und sein könnte* weiter zu verschieben. Denn die aktuell verfügbaren und zukünftig antizipierten Systeme scheinen sich von Technologien, die bislang mit dem Begriff KI belegt wurden, maßgeblich zu unterscheiden. Generative KI wird offensiv als Wegbereiter „echter“ Künstlicher Intelligenz, also einer AGI, gehandelt. Was also hat sich geändert? Was macht es gesellschaftlich plausibel, algorithmische Systeme als Artefakte zu betrachten, die als Prototypen einer „echten“ KI fungieren? Mit diesem Erkenntnisinteresse ergänzt der Beitrag die soziologische Erforschung zeitgenössischer digitaler Technologien, die auf Deep Learning basieren. Statt der Entwicklung und dem Gebrauch dieser Technologien techniksoziologisch nachzugehen (vgl. dazu aktuell Häufling et al. 2024), interessiert sich der Beitrag für deren Kategorisierung *als KI* und die Implikationen dieser Kategorisierung für die Grenzen der sozialen Welt.

Die These dieses Beitrags ist, dass erst die Verfügbarmachung von Algorithmen *als Zurechnungspunkten von Kommunikation* KI in der gesellschaftlichen Breite erfahrbar und erlebbar macht. Dies unterscheidet den aktuellen „KI-Sommer“ von seinen Vorläufern. Erst mit der infrastrukturell veralltäglichten „Kommunikativierung der Maschinen“ (Dogruel & Dickel 2022) hat KI eine Chance, sich als gesellschaftliche Realität jenseits spezifischer Expertengemeinschaften zu etablieren.

Der Beitrag wird zunächst den Hochsommer der KI skizzieren, der vom Technoimaginären der AGI geprägt ist (Kapitel 2). Im Anschluss daran werde ich mithilfe differenzierungstheoretischer Ansätze aufzeigen, wie Künstliche Intelligenz die Unterscheidung von Menschen und Maschinen infrage stellt (Kapitel 3). Danach widmet sich der Beitrag der spezifischen Mediatisierung von Maschinen, die der aktuellen Rede von KI ihre Plausibilität verleiht (Kapitel 4). Der Beitrag schließt mit einer theoriegeleiteten Technikfolgenabschätzung (Kapitel 5).

² Diesen Gedanken verdanke ich einem meiner anonymen Gutachter:innen.

2 Auf dem Weg zur AGI?

Der Enthusiasmus der frühen KI-Forscher:innen zur Mitte des 20. Jahrhunderts war von der Erwartung getragen, dass sich mit Computern das menschliche Denken nachbauen lassen könnte. Diese Position wurde von John Searle als *strong AI* bezeichnet: „[A]ccording to strong AI, [...] the appropriately programmed computer really *is* a mind, in the sense that computers given the right programs can be literally said to *understand* and have other cognitive states“ (Searle 1980: 417 Hervorhebung i.O.). Mit dem Abkühlen der frühen KI-Euphorie herrschte jedoch weitgehend Einigkeit innerhalb der KI-Community, dass es dem Feld eigentlich um *weak AI* gehen sollte, also den Einsatz von Computern zur Lösung konkreter Problemstellungen. In der Tat wurde der Begriff der KI in der innerwissenschaftlichen Kommunikation seit den 1990ern kaum noch verwendet: „[M]any researchers in fields currently associated with AI, such as natural language processing (NLP), machine learning (ML), and computer vision (CV), explicitly distanced themselves from the term ‘AI,’ in part because it became associated with unfulfilled grandiose promises“ (Gebru & Torres 2024: 12).

Demgemäß bevölkerten menschenähnliche Maschinen bis vor kurzem nur Bücher und Filme, philosophische Gedankenexperimente und futuristische Szenarien. Doch dies beginnt sich zu Beginn der 2020er Jahre zu ändern. Wissenschaftlich-technische Diskurse zur Entwicklung einer menschenähnlichen KI werden dabei weniger unter dem philosophisch vorbelasteten Begriff der *strong AI* geführt. Der neue diesbezügliche Leitbegriff ist nun vielmehr *Artificial General Intelligence* (AGI).

Der Begriff AGI wird seit den 2000er Jahren verwendet. Er markiert eine wachsende Unzufriedenheit transhumanistischer Enthusiast:innen (Bostrom 2003; Kurzweil 2005)³ mit der Fokussierung des Felds auf *weak AI* (oder: wie Ray Kurzweil (2005: 206) es nennt: *narrow AI*) bzw. der völligen Distanzierung vom KI-Begriff selbst und der damit ursprünglich verbundenen Zielsetzungen (Goertzel 2014: 1). Mit dem Begriff der AGI sollte das ambitionierte Ziel einer menschenähnlichen Maschine wieder fokussiert werden. In der Gruppe derjenigen, die das Konzept popularisierten, gilt AGI als Technologie, die den Weg zu einer „Superintelligenz“

³ Gebru und Torres (2024: 2) verweisen darauf, dass man von einer Melange aus sich überlappenden Ideologien sprechen sollte, welche den Diskurs zu AGI prägen und die sich als “transhumanism, Extropianism, singularitarianism, (modern) cosmism, Rationalism, Effective Altruism, and longtermism” beschreiben. Da der Transhumanismus diesbezüglich aber als maßgebliche ideologische Hintergrundfolie fungiert, soll auf eine genauere Differenzierung an dieser Stelle verzichtet werden.

(Bostrom 2014) bahnen soll, die der menschlichen Intelligenz nicht nur ebenbürtig, sondern grundsätzlich überlegen sein soll. Dies wurde (und wird) von Transhumanist:innen mit der technofuturistischen Hoffnung auf eine Beschleunigung des technologischen Fortschritts verbunden, die Menschen den Zugang zu übermenschlichen Fähigkeiten verschaffen könnte und/oder Menschen durch *mind uploading* selbst in eine übermenschliche KI verwandeln zu können (Kurzweil 2005; Bostrom & Sandberg 2008; kritisch Hauskeller 2012). Gegenüber diesen Zukunftshoffnungen wurde von Seiten des Mainstreams der KI-Forschung nachdrücklich „boundary work“ (Gieryn 1983) betrieben – sie wurden als außerhalb des seriösen wissenschaftlichen Diskurses stehend positioniert. Nichtsdestotrotz wurde das Thema in der transhumanistischen Diskursnische weiter intensiv verhandelt und sukzessiv durch diesbezügliche Konferenzen und Veröffentlichung popularisiert (Gebru und Torres 2024).

Während sich der Mainstream der KI-Forschung noch in den 2010er Jahren von dem Ziel einer AGI weitgehend distanzierte, wird der Begriff seit der Veröffentlichung von ChatGPT zunehmend populär.⁴ AGI stellt (bereits vor der Veröffentlichung von ChatGPT) das zentrale Forschungsziel von OpenAI dar. Die Firma beschreibt sich selbst als Pionier „on the path to AGI. We believe our research will eventually lead to artificial general intelligence“ (OpenAI 2024b). Die Firma Microsoft, die sich stark an OpenAI beteiligt hat, unterstützt diese Zielsetzung explizit. „Artificial general intelligence [...] has long captivated the imagination of scientists, technologists, and fiction writers alike. AGI would be able to perform a huge scope of tasks at or above the human level. It represents the next logical step from current AI technologies“ (Microsoft 2024). Forscher:innen von Microsoft sehen in der Transformer-Technologie von GPT-4 erste „Sparks of Artificial General Intelligence“ – so der Titel eines aktuellen Papers von Microsoft Research (Bubeck et al. 2023). Dabei kommuniziert OpenAI die Überzeugung, dass AGI eine Brücke zur Superintelligenz bildet: „The first AGI will be just a point along the continuum of intelligence. [...] Successfully transitioning to a world with superintelligence is perhaps the most important [...] project in human history“ (Altman 2023). Neben Open AI beteiligt sich auch Google (mit seinem Subunternehmen DeepMind) am Rennen zur AGI. Ebenso wie OpenAI (Cook 2024) hat auch DeepMind mittlerweile ein Stufenmodell entwickelt, um Fortschritte auf dem Weg zur AGI (inklusive einer „Superhuman Artificial Intelligence“ als höchster Stufe) zu veröffentlichen (Morris et al. 2024).

⁴ Dies lässt sich durch eine Google Trends-Abfrage gut beobachten: <https://trends.google.com/trends/explore?date=all&q=%2Fm%2F02s0k3>

Die Rede von AGI ist ein Symptom dafür, wie sich KI-Verständnisse im Laufe der Zeit wandeln und die Entwicklung ‚echter‘ Künstlicher Intelligenz ein *moving target* der informationstechnischen Forschung und Entwicklung bleibt. Die Unterscheidung von Künstlicher Intelligenz nach strong/weak wird zunehmend durch die Unterscheidung narrow/general ersetzt, obwohl die Begriffe *weak AI* und *narrow AI* sowie *strong AI* und *AGI* mitunter synonym verwendet werden. Während die Frage nach einem (menschenähnlichen) Bewusstsein von Maschinen in Diskursen um *strong AI* noch eine starke Rolle spielte, wird diese Frage im Diskurs um AGI an den Rand gedrängt. Was beim technischen Diskurs um AGI im Vordergrund steht, ist die *Leistungsfähigkeit* von Maschinen, die Menschen für eine Vielzahl von Aufgaben ersetzen können sollen. AGI wird demgemäß vor allem durch die *Breite an Tätigkeiten*, die sie (anstelle eines Menschen) übernehmen kann, definiert.

Der Begriff ist dabei als expliziter Gegenbegriff zu *narrow AI* angelegt, deren Kapazitäten sich auf eine *spezifische* Aufgabe beziehen, in der sie ggf. Menschen übertreffen kann (etwa im Schachspielen) (Kurzweil 2005, S. 206). Es geht also weniger darum, ob eine Maschine Menschen psychisch nahekommt, sondern eher, ob sie Menschen in sozialer Hinsicht grundsätzlich substituieren kann – unabhängig davon, wie es sich mit ihrer Innerlichkeit verhält. Die Rede ist von „highly autonomous systems that outperform humans at most economically valuable work“ (OpenAI 2024a), bzw. von Systemen, die zumindest mit „skilled adult humans on most cognitive tasks“ (Morris et al. 2024, S. 4) vergleichbar sein sollen.

Die Realisierung künstlichen Bewusstseins ist zwar bei den Autor:innen aus dem engeren transhumanistischen Umfeld weiterhin aktuell, da diese hoffen, ihr eigenes Bewusstsein künstlich replizieren zu können. Doch wie stark diese Zukunftshoffnung die technologische Entwicklung selbst antreibt und für führende Unternehmen des Feldes relevant ist, ist spekulativ. Hier liegt auch mein Kritikpunkt an ideologiekritischen Positionen zu Transhumanismus und AGI (Gebru & Torres 2024), die dazu tendieren, das aktuelle Rennen hin zur AGI eindeutig aus den Motiven transhumanistischer Technikenthusiast:innen abzuleiten. Es erscheint plausibler, dass dieses Rennen sich aus ökonomischen Gründen mittlerweile verselbständigt und von den Motiven dieser überschaubaren Gruppe gelöst hat. Der technologische Wettkampf hin zur AGI macht aus dieser auch einem Gegenstand, über dessen mögliche politische Regulierung debattiert wird (Naudé & Dimitri 2020) – bis hin zu Ideen einer Nationalisierung der AGI-Entwicklung nach dem Vorbild des Manhattan Projekts (Aschenbrenner 2024).

Wie kann nun die Soziologie dem neuen KI-Boom begegnen, ohne die vermeintlichen „thingness of AI“

(Suchman 2023) zu reproduzieren? Der Weg, der dafür in diesem Beitrag exploriert wird, ist der einer Beobachtung zweiter Ordnung (Luhmann 1990: 86 f.), welche die mit AGI implizierte Entdifferenzierung von Mensch und Maschine selbst zum Thema macht und nach den Bedingungen einer solchen Entdifferenzierung fragt.

3 Die Ent/Differenzierung von Mensch und Maschine

Die Frage nach den Möglichkeiten und Grenzen Künstlicher Intelligenz wird soziologisch typischerweise als Fortsetzung der Frage nach dem Status von Technik in der Gesellschaft verhandelt. Handlungstheoretische Ansätze unterscheiden dabei nach dem Grad der Handlungsträgerschaft, die sie der Technik zubilligen sowie nach der Differenziertheit der Handlungskonzepte (Rammert & Schulz-Schaeffer 2002; Schulz-Schaeffer et al. 2023). Die ANT (vgl. etwa Latour 1992, 2005) ist hier ein Beispiel für einen Ansatz, der von einer grundsätzlichen Symmetrie von Menschen und Dingen ausgeht und sich somit scharf von Ansätzen unterscheidet, die im Handeln von Technik nur eine Projektion menschlicher Subjekte erkennen (vgl. für einen Überblick Cerulo 2009; Muhle 2018).

Im Kontrast zu handlungstheoretischen Positionen, die sich an der Ent/Differenzierung von Mensch und Maschine selbst sozialtheoretisch beteiligen, möchte dieser Beitrag die Ent/Differenzierung von Mensch und Maschine reflexiv beobachten. Er versteht sich daher in einem breiten Sinne als *differenzierungstheoretisch*.⁵ Ihm geht es um die Rekonstruktion des Auf- und Abbaus von Unterscheidungen. Somit soll die Unterscheidung von Menschen und Maschinen im Folgenden nicht als gegebene Unterschiedlichkeit betrachtet werden. Vielmehr wird eine Perspektive eingenommen, die Differenzierung ebenso wie Entdifferenzierung als etwas betrachtet, das kulturell hervorgebracht wird. „[D]iscussions of nonhuman agency need to be reframed from categorical debates to empirical investigations of the concrete practices through which categories of human and nonhuman are mobilized and become salient“ (Suchman 2007: 1; vgl. auch Casper 1994). Differenzierungstheoretische Anknüpfungspunkte dafür liefern (1) Reflexionen zur Bedeutung der Unterscheidung von Mensch und

Maschine für die Konstitution der *Grenzen des Sozialen* in der Moderne (Lindemann 2018), (2) Überlegungen zu den kulturellen *Distinktionszonen des Humanen* aus der Perspektive der Humandifferenzierung (Hirschauer i.E.) sowie (3) systemtheoretische Thesen zur *Beteiligung von Maschinen an Kommunikation* (Esposito 2024).

3.1 Die Grenzen des Sozialen

Thomas Luckmann hat das Thema der „Grenzen des Sozialen“ (Luckmann 1980) in die Soziologie eingeführt, um die kulturell variierenden Definitionen dessen zu diskutieren, wer als soziales Gegenüber gilt. Diese Grenzen verschieben sich historisch und sind in unterschiedlichen Gesellschaften verschieden definiert. Zum Beispiel sehen animistische Kulturen viele Tiere und Pflanzen als beseelte Wesen, die als Teil der sozialen Gemeinschaft betrachtet werden (Descola 2022). Es lassen sich historisch variable „Grenzregime“ (Lindemann 2009; vgl. auch Luckmann 1980) diagnostizieren, welche die aktuelle Sonderstellung des Menschen als contingentes und für die europäische Moderne spezifisches Phänomen erscheinen lassen. Gesa Lindemann zufolge ist es für diese Gesellschaft konstitutiv, nur lebende Menschen als ihre Mitglieder zu begreifen und ihre normative Ordnung auf Basis grundlegender Menschenrechte zu entwerfen (Lindemann 2018). Sie sei „die erste Gesellschaft, die aus Menschen besteht“ (Lindemann 2018: 306). Im Anschluss an differenzierungstheoretische Überlegungen von Simmel, Durkheim und Luhmann betrachtet Lindemann die Hervorbringung des Menschen (in seiner heutigen Bedeutung) als institutionellen Stabilisierungsmechanismus intern hochdifferenzierter Gesellschaften.

Die europäische Moderne realisiert die besonders relevanten Abgrenzungen des Menschen, so Lindemann, in Form eines „anthropologischen Quadrats“, das den Menschen „durch eine vierfache Abgrenzung bestimmt: die beiden Grenzen am Lebensanfang und am Lebensende sowie die Grenze zum Tier und zur Maschine“ (Lindemann 2018: 28). Andere Unterscheidungen, etwa die zwischen Menschen und Engeln, verlieren hingegen an Relevanz. (Lindemann 2018: 133 f.). Doch wird diese Form der Grenzziehung des Humanen „nicht als eine allgemeine Voraussetzung jeder Vergesellschaftung verstanden, sondern als eine [...] gesellschaftliche Institution, die auch wieder vergehen kann“ (Lindemann 2018: 56).

Lindemann argumentiert, dass der Mensch mit dem wachsenden Einfluss der Naturwissenschaften auf das moderne Weltbild zunehmend als materielles Wesen verstanden wurde, das nicht nur als besonders hochentwickeltes Tier, sondern auch „im Prinzip als eine Maschine zu

⁵ Statt Differenzierungstheorie als Chiffre für die (Luhmannsche) Systemtheorie zu verwenden, schließt der Beitrag damit an Tendenzen einer Pluralisierung und Empirisierung differenzierungstheoretischer Ansätze an, welche die Systemtheorie freilich einschließen (vgl. Mölders et al. 2024).

begreifen ist. [...] Diese Entwicklung führte dazu, dass im Rahmen der modernen Begrenzung des Menschlichen die Mensch-Maschine-Differenz zentral wurde“ (Lindemann 2018: 96) – und zwar nicht deswegen, weil Maschinen dem Menschen so fern wären, sondern weil sie ihm nah genug sind, dass Verwechslungen, Vermischungen oder Verwandlungen nicht auszuschließen sind. Gerade diese (kulturell erzeugte) Nähe zur Maschine sorge dafür, dass die Mensch-Maschine-Differenz anfällig dafür ist, „diskursiv dramatisiert“ (Lindemann 2018: 122) zu werden, insbesondere dann, wenn in Aussicht gestellt wird, dass (1) der Mensch zur Maschine zu werden droht, (2) der Mensch durch die Maschine ersetzt wird, oder (3) die Maschine menschlich wird. Die mögliche Entdifferenzierung von Mensch und Maschine wird zum Faszinosum und Problem der europäischen Moderne.

3.2 Distinktionszonen des Humanen

Auch für Stefan Hirschauer gehört die Mensch/Maschine-Differenz zu den relevanten „Distinktionszonen des Humanen“ (Hirschauer i.E.). Er reiht ihre Erforschung in das Programm der Humandifferenzierung ein, also der Untersuchung derjenigen Unterscheidungen, die das gesellschaftliche Personal betreffen (Hirschauer 2021). Der Fokus der Humandifferenzierungsforschung lag bislang vor allem auf den *Binnendifferenzierungen* des Humanen – etwa in Frauen und Männer, schwarze und weiße, alte und junge Menschen (Hirschauer 2021). Die Unterscheidung von Menschen und Maschinen gehört hingegen zur „*Außendifferenzierung* der Menschen“ (Hirschauer 2021: 157, Hervorhebung im Original) also den „Unterscheidungen des Menschen von Nicht-Menschen“ (Hirschauer 2021: 171).

Hirschauer folgt grundsätzlich Lindemanns These einer Profilierung der Außengrenzen des Humanen, betont aber, dass diese Grenzen differenzierter zu betrachten seien. Lindemann fokussiere sich stark auf die Besonderheiten der europäischen Moderne und stütze sich auf wissenschaftliche und rechtliche Diskurse, die tendenziell klare, aber vereinfachte Grenzen ziehen. Sie tendiere somit dazu, intellektualistische Selbstbeschreibungen von Gesellschaften – mit Luhmann (1980) gesprochen: ihre „semantische Tradition“ – mit der Gesellschaftsstruktur zu verwechseln. Eine solche Fokussierung vernachlässige alltagsweltliche und massenmediale Diskurse, die oft fließendere Übergänge aufweisen, sowie konkrete Praktiken, mit denen Menschen materiell ent/differenziert werden (Hirschauer i.E.). Hirschauer und Nübling weisen darauf hin, dass nicht nur gelehrt Diskurse zur Differenzierung von Entitäten beitragen, sondern diese Differenzierung sich in vielfältigen *Sinn-*

schichten des Kulturellen vollzieht. Entitäten werden etwa auch durch materielle Strukturen (wie Zäune) auseinander gehalten, durch Zurichtungen des Körpers voneinander dissimiliert oder durch Kleidung voneinander unterscheidbar gemacht (Hirschauer & Nübling 2021; vgl. auch Hirschauer 2021). Differenzierung und Entdifferenzierung können auf verschiedenen Sinnsschichten in gegenläufige Richtungen und/oder unterschiedlichen Geschwindigkeiten verlaufen.

Ob Menschen oder andere Entitäten als soziale Gegenüber anerkannt werden, wodurch sich Menschen von Nicht-Menschen unterscheiden und wie schließlich ‚der Mensch‘ (im Singular) als Kategorie bestimmt und begrenzt wird, kann als kulturell contingent betrachtet werden. Der Mensch kann etwa als Homo Sapiens von seinen evolutionären Vorläufern differenziert werden, ebenso aber von Geistern und Monstrositäten, Über- und Untermenschen, Tieren und Maschinen. Zwischen vermeintlich eindeutigen Kategorien wie Mensch und Maschine lassen sich Hirschauer zufolge Hybridfiguren ausmachen, deren ontologischer Status strittig ist. Künstliche Intelligenzen betrachtet Hirschauer als ebensolche Hybridfiguren zwischen Menschen und Artefakten (Hirschauer i.E.).

KI ist eine Kategorie, die Maschinen mit einer Eigenschaft ausstattet, die wir üblicherweise Menschen (und einigen Tieren) zuschreiben, nämlich Intelligenz. Inwiefern KI aber dabei dem Menschen gegenüber verähnlicht oder verändert wird, variiert je nach der veranschlagten Begrifflichkeit. Weit weg von der Kategorie des Menschen erscheinen sowohl Systeme, denen lediglich Werkzeugcharakter (*weak/narrow AI*) zugeschrieben wird, als auch solche, die als gottähnliche Superintelligenzen imaginiert werden. Innerhalb des KI-Spektrums lassen sich ferner verschiedene Grade an Über- und Unterordnung registrieren. ‚Schwache‘ KI-Systeme rücken typischerweise in die Positionen von Helfer:innen, Assistent:innen und Diener:innen (Schmidt-Jüngst & Dickel 2021). Neuen generativen Systemen wie ChatGPT werden hingegen mitunter auch professionelle Rollen wie die der Autor:in oder der Copilot:in zugeschrieben. Zudem fällt auf, dass die Nähe oder Ferne zum Menschen typischerweise anhand von Fähigkeit und Leistung festgemacht wird, wodurch implizit die Kategorie Mensch ebenfalls als Fähigkeits- und Leistungskategorie erscheint. Das Bild des Menschen, an dem eine AGI gemessen wird, ist das einer autonomen erwachsenen Person, die ihre überdurchschnittliche geistige Leistungsfähigkeit in ökonomisch wertvolle Arbeit umzumünzen vermag (Morris et al. 2024).

Insgesamt scheinen aktuelle KI-Entwicklungen ein Indikator dafür zu sein, dass die Mensch/Maschine-Unterscheidung zumindest partiell fragil werden könnte. Die Benennung eines Artefakts als KI markiert zum einen den Beginn einer Differenzierung des Artefakts von anderen

Artefakten und zum anderen eine Entdifferenzierung von Mensch und Maschine. Doch wie weit reicht diese Entdifferenzierung? Dass KI – ebenso wie andere Grenzfiguren – *Thema* von Kommunikation werden kann, steht außer Frage. KI ist Leitmotiv der Technologiepolitik, Werbeslogan und Geschäftsmodell von Unternehmen, Gegenstand von Forschung und Forschungsförderung, Thema von populären Filmen und Serien, Problem rechtlicher Entscheidungs- und Verantwortungszuschreibung. Doch inwiefern wird sie auch als soziales Gegenüber behandelt? Um diese Frage trennscharf stellen zu können, kann die Betrachtung des kommunikativen Einbezugs von KI aufschlussreich sein (vgl. Muhle 2018).

3.3 Die Beteiligung von Maschinen an Kommunikation

Bei Lindemann (2018: 27) ist der Mensch im modernen Sinne eine kontingente „normative Institution“, aus der Perspektive der Humandifferenzierung ist er eine kontingente kulturelle Kategorie. Für die an Niklas Luhmann anschließenden Fassung der Systemtheorie ist der Mensch ein kontingenter Zurechnungspunkt der Kommunikation. Kommunikation wird bei Luhmann (1984: 191–200) als dreigliedriger Selektionsprozess verstanden: *Information* (was wurde gesagt) und *Mitteilung* (wie wurde es gesagt) werden durch *Verstehen* aufeinander bezogen und unterschieden. Keine dieser Selektionen ist als innerpsychischer Prozess zu betrachten, vielmehr bildet Kommunikation eine eigenständige Ordnung, die prozessual (und immer *ex post*) bestimmt, was von wem mitgeteilt wurde und ob ein Ereignis überhaupt als Kommunikation verstanden werden kann. Eine Entität wird zur Teilnehmer:in an Kommunikation, sofern ihr Mitteilungen zugerechnet werden. Autonomie und Handlungsfähigkeit von Entitäten werden systemtheoretisch nicht als deren Eigenschaften verstanden, sondern als Zurechnungen, die kommunikativ sichtbar, vorausgesetzt oder problematisiert werden können. Dieses Verständnis von Kommunikation ist posthumanistisch angelegt. Zwar geht Luhmann davon aus, dass Kommunikation in der modernen Gesellschaft typischerweise Menschen (und Organisationen) zugerechnet wird, dies ist aber grundsätzlich kontingent. Ob auch Maschinen als Teilnehmer an Kommunikation infrage kommen, ist damit letztlich eine empirische Frage (Muhle 2018: 155–161).

Bereits vor Aufkommen von ChatGPT argumentierten systemtheoretische Autor:innen wie Dirk Baecker (2011), Elena Esposito (1993, 2017) und Florian Muhle (2018) dafür, die mögliche Teilnahme von maschinellen Entitäten in Kommunikationsprozessen zu untersuchen. Und auch

wenn Baecker 2017 noch Zweifel an der Kommunikationsfähigkeit von Maschinen formulierte – „Maschinen kommunizieren nicht, noch nicht“ (Baecker 2017: 13), sah er „in der rasant zunehmenden *Beteiligung*, intelligenter‘ Maschinen an Kommunikation, [...] die eigentliche ‚Katastrophe‘, die den Wandel von der modernen Buchdruckgesellschaft zu einer nächsten Gesellschaft [...] auslöst“ (Baecker 2017: 17, Hervorhebung SD). Und auch wenn sich Armin Nassehi skeptisch gegenüber dieser Disruptionsthese zeigt, räumt er ein, dass neuere algorithmische Systeme das Potential haben „selbst zurechnungsfähig zu sein“ (Nassehi 2019: 245) da sie als intransparente Black Boxes eher dem Menschen ähneln würden als klassischer trivialer Technik. Künstliche Intelligenzen würden damit „Zurechnungspunkte in der Autopoiesis der Gesellschaft“ (Nassehi 2019: 257). Mit der Ankunft der KI hätte sich schließlich tatsächlich bewahrheitet, was Latour schon seit langem behauptet hätte, nämlich, „dass nicht-menschliche Akteure gleichsam auf Augenhöhe agieren“ (Nassehi 2019: 258).

In jüngerer Zeit plädiert Esposito daher dafür, „von einer neuartigen Form der ‚künstlichen Kommunikation‘“ zu sprechen. (Esposito 2024: 31; vgl. bereits Esposito 2017), in der algorithmische Systeme als Zurechnungspunkte von Kommunikation fungieren.⁶ Auch wenn es weiterhin „entscheidende qualitative Unterschied[e]“ zwischen natürlicher und künstlicher Intelligenz gäbe (Nassehi 2019: 260) und Menschen für die Fortsetzung gesellschaftlicher Autopoiesis weiterhin als unverzichtbar erschienen (Esposito 2024: 88 f.), ist die *Exklusivität* von Menschen als gesellschaftlichen Zurechnungspunkten für mehrere systemtheoretische Autor:innen damit infrage gestellt.

Während der Einbezug von Maschinen in Kommunikationsprozesse also bereits vor mehreren Jahren als soziologische Fremdbeschreibung verhandelt wurde, wird dieser Einbezug nun zunehmend Teil gesellschaftlicher Selbstbeschreibung. Generative KI ist heute auch und gerade als „kommunikative KI“ erfahrbar (Hepp et al. 2022). Während vergangene KI-Anwendungen vornehmlich im infrastrukturellen Hintergrund von Gesellschaften operierten, hat sich mit den aktuellen Anwendungen nicht nur ein informationstheoretischer, sondern auch ein kommunikations-theoretischer Paradigmenwechsel vollzogen.

⁶ In eine ähnliche Richtung weisen Konzepte einer „*Artificial Sociality*“ (Natale & Depounti 2024). Sie vermeiden ebenso wie der Begriff der „*Artificial Communication*“ (Esposito 2017) die Unterstellung von Intelligenz und betonen die soziale Kompetenzen maschineller Artefakte.

3.4 Zwischenbilanz: Entdifferenzierung und Inklusion?

Der Beitrag schließt an die angeführten drei Stränge differenzierungstheoretischer Überlegungen zur Mensch/Maschine-Differenz an. Er teilt mit Luckmann und Lindemann das gesellschaftstheoretische Interesse an den Grenzregimen der sozialen Welt. Er greift das von Hirschauer entworfene Programm der Humandifferenzierung auf, das dafür plädiert, den Auf- und Abbau von Differenzen granularer zu beschreiben, als dies binäre Kategorisierungen nahelegen. Und er folgt der sozialtheoretischen Intuition der Systemtheorie, dass eine Erweiterung kommunikativer Teilnehmerschaft sowohl grundsätzlich möglich als auch gesellschaftlich folgenreich ist.

Sofern die Diagnose stichhaltig ist, dass KI auf einen Differenzabbau zwischen Menschen und Maschinen hindeutet, die anhand ihrer kommunikativen Inklusion ablesbar wird, stellt sich die Frage nach den konkreten sozio-materiellen Bedingungen, die dafür verantwortlich sind, dass maschinelle Aktivitäten überhaupt als Kommunikation erscheinen können. Wie werden aus Artefakten ‚intelligente‘ Gegenüber gemacht, die in ein soziales Näheverhältnis zum Menschen gerückt werden?

Die soziologischen Antworten dazu oszillieren zwischen techno- und soziocentrischen Deutungen. So wird seit den frühen 2010er Jahren auf der einen Seite eine deutliche Beschleunigung im Forschungs- und Technologiefeld KI diagnostiziert, die durch die fortschreitende Entwicklung technischer Methoden, die massive Verfügbarkeit von Daten durch das Internet, die rapide Zunahme und Verbilligung der Rechnerleistung sowie die breite Digitalisierung der Gesellschaft angetrieben wird (Hirsch-Kreinsen & Krokowski 2023). Auf der anderen Seite wird die Bedeutung technologischer Narrative betont, die die gesellschaftliche Akzeptanz dieser neuen maschinellen Mitspieler:innen zu steigern versuchen (Bareis & Katzenbach 2022). Die Verhärtung der Kategorie KI (Suchman 2023) kann damit auch als rhetorische Strategie gelesen werden (Morozov 2023). Doch die partielle Entdifferenzierung von Mensch und Maschine im Fall von KI vollzieht sich nicht allein in der Sinnlichkeit der materiellen Infrastruktur oder der Sinnlichkeit der Sprache (Hirschauer & Nübling 2021). Sowohl die technischen Durchbrüche als auch die kulturellen Dynamiken von Hype and Hope spielen gewiss eine wichtige Rolle. Doch argumentiere ich, dass bei beiden Lesarten ein entscheidendes Puzzleteil fehlt. Dieses ist die Mediatisierung des Mensch-Technik-Verhältnisses, also der Art und Weise, wie Technik Menschen gegenübertritt. Diese ermöglicht es – so die im Folgenden entwickelte These – dass KI aktuell in der gesellschaftlichen Breite ein zunehmender Realitätsstatus

zugeschrieben wird und eine AGI als plausibles Ziel technowissenschaftlicher Entwicklung erscheinen kann.

4 Wie Maschinen kommunikationsfähig werden

4.1 Der Turing-Test

Um die Rolle medialer Vermittlung für KI zu rekonstruieren, geht der Beitrag zu einer Gründungszene des modernen KI-Diskurses zurück, nämlich dem Turing-Test. Dieser gilt als „the most well-known attempt to operationalize an AGI-like concept“ (Morris et al. 2024: 2). Der Ursprung des Tests liegt im 1950 veröffentlichten Aufsatz „Computing Machinery and Intelligence“ (1950) von Alan Turing, dem einflussreichen Pionier der Computerwissenschaften. Turings Ansatz steht in paradigmatischer Weise für eine doppelte Problemverschiebung: Die Frage nach der Denkfähigkeit der Maschinen wird durch ihn von einem technischen und philosophischen Problem in ein soziales Problem, genauer: ein kommunikatives Problem, verwandelt. Die Rekonstruktion dieser Problemverschiebung erlaubt es, zu verstehen, warum der Turing-Test bis heute als Prüfstein Künstlicher Intelligenz verhandelt wird (Collins 2021).

Turings Text beginnt mit der klassischen Frage, ob Maschinen denken können. Er verweist jedoch bereits im ersten Absatz darauf, dass es fruchtlos sei, die Frage in dieser Form empirisch klären zu wollen. Er argumentiert, dass es gewinnbringender sei, sich dem Thema indirekt zu nähern und die unentscheidbare Frage maschinellen Denkens in eine entscheidbare Problemstellung zu überführen (Turing 1950: 433). Diese nimmt bei ihm die Form eines Gedankenexperiments an, das Turing als das *imitation game* bezeichnet. In der von Turing zuvorderst präsentierten Variante wird das *imitation game* von drei Personen gespielt: einem Mann (A), einer Frau (B) und einer Person (C), die sowohl Frau als auch Mann sein kann. Die Aufgabe von C ist es, herauszufinden, wer der beiden anderen Mitspieler:innen der Mann und wer die Frau ist.

Entscheidend für den gesamten Versuchsaufbau ist, dass C die beiden anderen Mitspieler:innen weder sehen noch hören kann. Damit stehen C wichtige Marker, an denen Geschlechtszugehörigkeit in unserer Gesellschaft im Rahmen einer Interaktion unter Anwesenden festgemacht werden kann, nicht zur Verfügung. C sieht weder den Körperbau von A und B, noch hört sie deren Stimmen. C darf mit A und B lediglich schriftlich kommunizieren. Dabei sind laut Turing verschiedene Medien denkbar. So wäre etwa der Einsatz

eines Boten möglich, der zwischen den drei Akteuren hin- und herläuft und die schriftlichen Mitteilungen überbringt. Ideal erscheint Turing aber ein Modus der schriftlichen Kommunikation, die über einen Fernschreiber realisiert wird. C darf seinen Mitspieler:innen beliebige Fragen stellen. Die am Spiel beteiligte Frau B gewinnt das Spiel dann, wenn C seine Aufgabe der korrekten Geschlechtsidentifikation bewältigt. Das Spielziel von A besteht aber darin C zu täuschen. A gewinnt also, wenn C ihn für die Frau hält. Dadurch wird der Umgang mit der Geschlechtskategorie für alle drei Spielerrollen zu einer je unterschiedlich konturierten hermeneutischen Herausforderung, die darauf hinausläuft, dass alle Beteiligten verstehen müssen, was es bedeutet, ein Mann oder eine Frau zu *sein*, wie man Männer und Frauen *erkennt* und wie sich Männer und Frauen schriftlich *ausdrücken*.

Für Turing selbst ist diese erste Variante des Spiels der Auftakt, um sein eigentliches Thema vorzubereiten. „We now ask the question, ‘What will happen when a machine takes the part of A in this game?’ Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, ‘Can machines think?’“ (Turing 1950: 434). In dieser zweiten Variante des Spiels ist die Maschine herausgefordert, ebenso überzeugend zu lügen wie der Mann in der ursprünglichen Version. Wenn man nun beide Spielvarianten wiederholt durchführen und ihre Resultate vergleichen würde und dabei Mann und Maschine die Position von A gleich gut ausführen könnten, hätte man es Turing zufolge mit einer Art von Maschine zu tun, der man ihre Denkfähigkeit ebenso schwer absprechen könnte wie einem Menschen.

Schon bald nach der Publikation von Turings Artikel wurde eine dritte Variante des Spiels etabliert, die sich tiefer im gesellschaftlichen Bewusstsein verankern sollte als die von Turing ursprünglich erdachten Versionen. In dieser ist es die explizite Aufgabe von C, zu bestimmen, wer seiner Mitspieler:innen ein Mensch (A) und wer ein Maschine (B) ist. Hier besteht das Ziel von Maschine (und Mensch!) darin, C davon zu überzeugen, ein Mensch zu sein. Je besser es der Maschine A gelingt, Humanität vorzutäuschen, desto besser wird ihr Abschneiden im Test bewertet. Die Geschlechtskategorie ist hier verschwunden. An die Stelle der binären Unterscheidung Mann/Frau tritt die binäre Unterscheidung Mensch/Maschine. Es ist eben diese Variante des *imitation game*, die bis heute als Standardinterpretation des so genannten Turing-Test bekannt ist und die in den darauf folgenden Jahren zu einer Inspirationsquelle für die Erforschung, Entwicklung und Reflexion Künstlicher Intelligenz wurde (French 2000: 116).

Das Gedankenexperiment vom John Searle zum chinesischen Zimmer zielt darauf ab, den Turing Test zu

widerlegen. Dabei wird ein geschlossener Raum mit einem Menschen ohne Chinesisch-Kenntnisse vorgestellt, der mit Hilfe einer Anleitung in seiner Muttersprache sinnvolle Antworten auf chinesische Schriftzeichen gibt. Beobachter:innen außerhalb des Raums schließen fälschlicherweise, dass dieser Mensch Chinesisch verstehen würde. Das Experiment soll verdeutlichen, dass ein Computer mit Zeichen umgehen kann, ohne im mindesten ihre Bedeutung zu verstehen (Searle 1980: 417 ff.). Doch geht dieser Einwand am eigentlichen Witz von Turings Überlegung vorbei. Turing will das Problem denkender Maschinen nämlich genauso wenig philosophisch klären, wie er es für empirisch lösbar hält: „The original question, ‘Can machines think?’ I believe to be too meaningless to deserve discussion. Nevertheless I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted“ (Turing 1950: 442). Es geht ihm also um einen prognostizierten kulturellen Deutungswandel („use of words“, „general educated opinion“), der mit einer Transformation von Erwartungsstrukturen einhergehen würde.

Behandelt man die Frage „Can machines think?“ wie Searle als philosophisches Problem, dann geht es um die Klärung dessen, was man plausibel unter Denken, Intelligenz oder Bewusstsein verstehen kann. Was macht diese Phänomene aus? Finden wir sie nur in Menschen oder auch in Tieren oder Maschinen? Neben diesen ontologischen Problemen tritt die epistemologische Frage, ob und wie man diese Phänomene eigentlich beobachten kann. Gibt es eine Möglichkeit zu prüfen, ob das, was in Computern vorgeht, mit dem vergleichbar ist, was in Menschen vorgeht? Das *imitation game* weicht solchen Fragen aus. Turing bringt damit ein Paradigma zum Ausdruck, das bei den Pionieren der Künstlichen Intelligenz zur Mitte des letzten Jahrhunderts zunehmend einflussreich wurde. Statt nämlich die Realisierbarkeit ‚denkender Maschinen‘ allein als technisches oder philosophisches Problem zu betrachten, wurde sie zunehmend auch und gerade als Problem der *Anerkennung und Zuschreibung von Intelligenz durch die Nutzer:innen* behandelt. Die Entwicklung von KI war damit nicht mehr allein eine Frage des Designs von Schaltkreisen und Programmiertechniken, sondern auch eine Frage des Designs der Interaktionsbedingungen von Mensch und Maschine. Der Turing-Test ist damit eine Antwort auf die Frage, wie Interaktionen so gestaltet werden können, dass Maschinen eine Chance haben, als intelligente Gegenüber wahrgenommen und behandelt zu werden (Natale 2021: 71).

Das *imitation game* lässt sich diesbezüglich weniger als Blaupause für maschinelle Denkfähigkeit lesen, sondern als kommunikations- und medientheoretisches Lehrstück.

Es führt vor, wie die Konstruktion einer spezifischen Weise des Miteinander-in-Kontakt-Tretens, Mensch und Maschine entdifferenzieren kann. Die gesamte Anordnung des Tests funktioniert nur unter mediatisierten Sonderbedingungen (hier: von Intransparenz und Schriftlichkeit). Der Versuchsaufbau verschiebt das Spielfeld des Sozialen zugunsten der Maschine. Er erleichtert ihr ein *passing as human* indem er C sowohl den Körper von B als auch die Hard- und Software von A vorenthält. Die zentrale Intervention in das Mensch-Maschine-Verhältnis, die in den verschiedenen Varianten des *imitation game* vorgenommen wird, ist daher nicht die Entscheidung der Testperson über die Identität ihrer Interaktionspartner:innen, sondern die Einrichtung einer mediatisierten Situation, die den Körper von A unsichtbar macht und die Form der Interaktion auf das Medium der Schrift begrenzt (Hayles 1999: 14).

4.2 Chatbots

Ursprünglich war das *imitation game* lediglich ein Gedankenexperiment. Doch die Form der Mediatisierung, die mit diesem Gedankenexperiment imaginiert wurde, schuf ein Paradigma für das Design einer Kommunikation mit maschinellen Gegenübern, die menschenähnlich wirken sollten. Im Jahr 1966 entwickelte Joseph Weizenbaum mit ELIZA ein frühes Computerprogramm zur Verarbeitung natürlicher Sprache, das ein Gespräch mit einem Menschen simulieren sollte. Eine der berühmtesten Implementierungen von ELIZA war das Skript „DOCTOR“, das die Rolle eines/r Psychotherapeut:in übernahm. Das Skript vermittelte die Illusion des Verstehens, indem es die Aussagen der Benutzer:innen in Fragen umwandelte und sie zu weiteren Ausführungen ermutigte. ELIZA demonstriert, wie selbst einfache Rechenprozesse Antworten hervorbringen können, die den Nutzer:innen dialogisch und sinnvoll erscheinen. Dieses frühe Beispiel eines Chatbots verdeutlicht das Potenzial und die Grenzen früher Formen der maschinellen Kommunikation. ELIZAs Fähigkeit, an einer Konversation in natürlicher Sprache teilzunehmen, war wichtiger als die interne Komplexität des Programms (Suchman 2007: 47 f.).

Was in Turings Gedankenexperiment ein Fernschreiber war, ist hier ein Bildschirm und eine Tastatur. Das maschinelle Gegenüber wird dem Menschen gegenüber verähnlicht (und vice versa), da beide als Produzent:innen von Text fungieren, der auf einem Bildschirm erscheint. ELIZA spiegelt Weizenbaums Überzeugung wider, dass KI nur dann erfolgreich gemacht werden könne, wenn sie die natürliche menschliche Sprache beherrsche, verstehe und reproduziere. Die umgekehrte Verähnlichungsstrategie –

Menschen dazu zu bringen, Maschinensprache zu beherrschen – würde die Nutzung von KI hingegen auf engste Kreise von Expert:innen begrenzen (Weizenbaum 1976: 183). Obgleich hier (im Kontrast zu Turings Gedankenexperiment) die maschinelle Identität des Gegenübers gar nicht verschleiert wurde, überzeugte ELIZA als Dialogmaschine durch eine recht simple Funktionalität.

ELIZA stand Pate bei der Entwicklung von Chatbots, die sprachliche Äußerungen von Nutzer:innen verarbeiten können und selbst Zeichen produzieren, die von den Nutzer:innen als sinnhafte Texte verstanden werden können. Die Basiskonstellation solcher Systeme sieht vor, dass Nutzer:innen *qua* Textverarbeitung am Bildschirm mit einem augenscheinlich konversationsfähigen Gegenüber interagierten. Doch trotz innovativer Konzepte zum „Computer as a Communication Device“ (Licklider & Taylor 1968) waren die Fortschritte im Bereich konversationsfähiger Maschinen über Jahrzehnte hinweg begrenzt. Zum Ende des 20. Jahrhunderts wurde die Bedeutung des Computers als kommunikatives Werkzeug zwar zunehmend reflektiert. Dies führte jedoch immer noch nicht dazu „daß in dem individuellen Gebrauch des Computers die Maschine [...] als Kommunikationspartner betrachtet werden muß“ (Esposito 1993: 338). Auch im Jahr 2007 stellte Lucy Suchman noch fest: „[T]here is (still) no evidence for the achievement of conversation between humans and machines in the strong sense that we know it to go on between humans. [...] I would argue, we have yet to realize the creation of an interactive machine“ (Suchman 2007: 23). Unzulängliche sprachliche Kompetenzen von Chatbots führten immer wieder zu Erwartungsenttäuschungen auf Seiten der Nutzer:innen. Bis hin zu modernen Assistenzsystemen wie Amazons Alexa lässt sich diese Kluft aus Erwartung und Enttäuschung beobachten (Reeves et al. 2018).

Die zugeschriebenen Kommunikationsfähigkeiten von Maschinen steigerten sich erheblich mit der Veröffentlichung von ChatGPT. Mit der zeitgenössischen generativen KI wird die Erzeugung von Texten und Bildern in der gesellschaftlichen Breite als etwas beobachtbar, das Maschinen fast so gut, ebenso gut oder gar besser können als Menschen. Die Bedeutung von ChatGPT und ähnlichen Systemen (wie Claude oder Gemini) kann daher kaum überschätzt werden. Diese einfache Zugänglichkeit und Interaktion haben die Integration von KI in verschiedene soziale Kontexte befördert. Die weit verbreitete Verfügbarkeit der neuesten Generation von Chatbots ermöglicht es den Nutzer:innen, mit Maschinen auf eine dialogische Weise zu interagieren. Generative KI-Systeme werden auch zunehmend in Kunst und Unterhaltung, Journalismus und Öffentlichkeitsarbeit, Forschung und Lehre eingesetzt. Die kommunikativen Fähigkeiten von Menschen und Maschinen werden vergleichbar,

die Substituierbarkeit menschlicher Kommunikationsbeiträge wird sichtbar, kulturelle Praktiken werden herausgefordert (Pavlik 2023; Chu & Liu 2023; Borchardt et al. 2024; Elyoseph et al. 2023). Wenn generative KI zudem auch in Chatbots (wie *Replika*) integriert wird, die sich Menschen als Beziehungspartner:innen anbieten, kann der Austausch von Textnachrichten als etwas beschrieben werden, das der KI als emotionale Kompetenz zugeschrieben wird, die sie mit Menschen vergleichbar macht (Possati 2023; Depounti et al. 2023). Angesichts solcher Entwicklungen hat Elena Esposito darauf hingewiesen, dass Systeme wie ChatGPT in der Tat als Zurechnungspunkte von Kommunikation im Luhmannschen Sinne fungieren können (Esposito 2024).

4.3 Kommunikativierung am Interface

Esposito folgt damit Turings Intuition, dass die Eingemeindung der Maschinen in den Kreis der intelligenten Wesen im Alltag weder aufgrund technischen Detailwissens noch aufgrund philosophischer Argumente entschieden wird. Turings Vermutung war vielmehr, dass die Entdifferenzierung von Mensch und Maschine gesellschaftlich wirksam wird, wenn Maschinen (unter spezifischen medialen Bedingungen) Zeichen produzieren können, die von menschlichen Zeichen nicht mehr unterscheidbar sind. Dass diese Entdifferenzierung (erst) heute zunehmend möglich erscheint, ist zweifellos auf die Fortschritte der Deep Learning-Algorithmen zurückzuführen. Wie aber bereits im Turing-Test kommt es darüber hinaus auf die Einrichtung einer mediatisierten Kontaktzone an, durch die sich Mensch und Maschine auf spezifische Weise begegnen. Die Gestaltung eines Interface, das menschliche Kommunikation „nachahmt und den Output nicht mehr in abstrakter Form als Lochkarte oder als Band kaum lesbarer Zeichenskripte editiert“ (Rammert 2023: 44) wird neben der Programmierung von Systemen und deren diskursiver Plausibilisierung zu einem wichtigen Faktor, um KI in der gesellschaftlichen Realität zu verankern.

Dies lässt sich am Erfolg von ChatGPT zeigen. Seit dessen Einführung ist KI kein abstraktes Konzept oder ein futuristischer Traum mehr, der nur in Wissenschaft und Science-Fiction diskutiert wird. ChatGPT sieht auf verschiedenen Plattformen (z. B. Web-Interface, Desktop-Anwendung, mobile App) recht ähnlich aus: Das User Interface umfasst in der Regel ein Texteingabefeld, in das die Nutzer:innen Fragen oder Befehle eingeben können, sowie einen Antwortbereich, in dem der Gesprächsverlauf angezeigt wird, der die Eingaben der Nutzer:innen und die Antworten von ChatGPT nacheinander darstellt. Neuerdings bietet ChatGPT auch sprachbasierte Interaktionen an. Um mit ChatGPT

kompetent umgehen zu können, müssen User nicht verstehen, wie Deep Learning-Algorithmen funktionieren. Vielmehr kommt es darauf an, die Art des Schreibens und Sprechens mit ChatGPT zu erlernen, also das Formulieren von Prompts (Input) und die Nutzung der generierten Texte oder Bilder (Output). Die Operationen in der Black Box der Large Language Models können verborgen bleiben. „When a machine runs efficiently, [...] one need focus only on its inputs and outputs and not on its internal complexity. Thus, paradoxically, the more science and technology succeed, the more opaque and obscure they become“ (Latour 1999: 304). Diese Form des Kommunizierens kann sich auf kulturelle Skripte stützen, die jahrzehntelang gesellschaftlich eingeübt wurden. Das Schreiben mit einem unsichtbaren Gegenüber haben Menschen durch Chat, SMS und Instant Messaging erlernt, das Sprechen mit einem unsichtbaren Gegenüber durch das Telefon.

In der Gestaltung von kommunikativer KI radikalisiert sich eine bestimmte Gestaltungsform der Mensch-Maschine-Beziehung, die mit der Popularisierung der Computertechnik immer wirkmächtiger wurde, nämlich die Einrichtung von User Interfaces als Kontaktzonen zur Technik (im Deutschen auch als Benutzerschnittstellen oder Benutzeroberflächen bezeichnet). User Interfaces eröffnen selektive Wege des Miteinander-in-Kontakttretens, etwa durch Laute, Vibrationen, Wischgesten, Fingerabdrücke oder Texteingaben. Nutzer:innen bedienen Keyboards, schieben Mäuse über den Schreibtisch, betrachten den Mauszeiger auf einem Bildschirm und erleben den Rechner als Ansammlung grafischer Symbole, die ihnen etwa als Ordner oder App-Icons gegenüberstehen. Sie prägen die Vorstellung der Nutzer:innen, was ein Computer(programm) ist und wie man mit ihm interagieren kann (Wirth 2025: 79–85). User Interfaces transformieren nicht nur die mediale Sphäre im engeren Sinne. Vielmehr lässt sich eine Rekonfiguration von Mensch-Technik-Verhältnissen durch das Paradigma des Interface in nahezu allen gesellschaftlichen Bereichen feststellen. *Der Gebrauch von Technik verwandelt sich in eine Bedienung von User Interfaces* (Hookway 2014: 1).

In der Gestaltung solcher Interfaces wird nicht allein ein bestimmtes Artefakt entworfen. Vielmehr werden „Interaktionsarchitekturen“ (Mayer et al. 2020: 296) errichtet, die durch ihre Affordanzen Formen des erwarteten Gebrauchs von Computersystemen anbieten, nahelegen und vorstrukturieren (Distelmeyer 2021: 55). Ein tieferes Verständnis der durch das Interface verborgenen Teile der Hardware und Software von Computern ist für Nutzer:innen damit nicht mehr notwendig. Das Interface tritt aus Perspektive der Nutzer:innen an die Stelle der nicht mehr intelligiblen Technologie. Bereits in den 1980er Jahren wies Frieder Nake (2019 [1984]) darauf hin, dass Interfaces die Mensch-Tech-

nik-Beziehung tendenziell als *Kommunikationsverhältnis* inszenieren. Aus riesigen Geräten, die früher von Spezialist:innen zur Lösung mathematischer Aufgaben verwendet wurden, entwickelten sich durch die Verbindung von Tastatur und Bildschirm sowie der Integration von Textverarbeitungssoftware kommunikative Werkzeuge (Heilmann 2012). So wurden sukzessive Artefakte geschaffen, die sich als Quellen und Adressaten für Kommunikation behandeln lassen. Diese antworten auf (von Menschen produzierte) Zeichen in einer Weise mit Zeichen, die von Menschen nicht völlig durchschaubar ist und produzieren Zeichen, die sie nicht ohne weiteres von menschlich produzierten Zeichen unterscheiden können. Diese Zeichenhaftigkeit der Technik setzt eine langwelligere Konjunktur fort, die sich mit Hubert Knoblauch als *Kommunikativierung* bezeichnen lässt.⁷

„[Kommunikativierung] zeichnet sich zum einen durch die Umstellung der Technik auf Zeichen aus. Sie macht Technik kategorisch zu einer Kommunikationstechnik und jeden Umgang mit ihr zu einem zeichenhaften kommunikativen Handeln. Die Grundlage dafür bildet die Digitalisierung und die damit verbundene Kopplung von Zeichen und Dingen. Die Kommunikativierung zeichnet sich zum anderen durch die Interaktivierung der digitalisierten Technologie aus“ (Knoblauch 2017: 343).

Diese Kommunikativierung sorgt dafür, dass die Maschine dem Menschen materiell fremd bleiben kann, zu ihm aber in ein soziales Näheverhältnis gerückt wird. Ein Interface wie ChatGPT ist also nicht nur ein *Kanal* für KI. Es sorgt vielmehr dafür, dass ein algorithmisches System überhaupt für Nutzer:innen als KI erscheinen kann – nämlich als Gegenüber, das ‚intelligente‘ Antworten gibt. Menschen können weiterhin nicht unmittelbar mit einem Algorithmus sprechen (ebenso wenig, wie wir unmittelbar einen Kontakt zur Gedankenwelt eines Menschen aufnehmen können). Doch fungieren Interfaces wie ChatGPT nun als „social displays“ (Müller 2023). Sie sind funktionale Äquivalente für sprechende Münden und schreibende Hände.

Ein zentrales Moment der zeitgenössischen Entdifferenzierung von Mensch und Maschine im Fall kommunikativer KI besteht darin, dass die operativen Prozesse der Maschine vor Menschen verborgen werden, während die User Interfaces (wie Websites, Apps oder Sprachausgaben) die Resultate algorithmischer Operationen so darstellen, dass sie analog zu einer menschlichen Mitteilung (einem geschriebenen Text, einem gemalten Bild oder einer Stimme) erscheinen. Ob das, was in der Black Box geschieht, dem menschlichen Denken dann ähnlich oder unähnlich ist, kann somit im Alltag ausgeblendet und zurück an wissen-

schaftliche und literarische Diskurse delegiert werden. Es zeigt sich hier eine Gegenläufigkeit von Dynamiken der Differenzierung und Entdifferenzierung auf unterschiedlichen Sinnsschichten (Hirschauer & Nübling 2021). Zugespitzt formuliert: die innertechnischen Prozesse können radikal entähnlicht werden, wenn Ähnlichkeit auf der Ebene zeichenhafter Kommunikation in einem mediatisierten Setting hergestellt wird (Suchman 2007: 206 ff.). Die Deutung einer Technologie als KI entscheidet sich im Alltagsgebrauch nicht *im Algorithmus*, sondern *am Interface*, das Algorithmus und User vermittelt.

5 Schlussfolgerungen: KI zwischen Fiktion und Fakt

KI ist eine Kategorie zur Verbesonderung von Maschinen, denen menschenähnliche (Kommunikations-)Eigenschaften zugeschrieben werden (nicht etwa: menschenähnliches Denken oder menschenähnliche Körper). Dabei handelt es sich zunächst um eine Kategorisierung in der Sinnsschicht der Sprache, die aber zugleich als Ziel technologisch-materiellen Designs fungiert.

Systeme, die als *weak AI* oder *narrow AI* betrachtet werden, gelten gesellschaftlich als existierende Technik, zugleich aber ist ihr Status als ‚echte‘ Künstliche Intelligenz strittig. Erst *strong AI* oder AGI wären Entitäten, die das Versprechen Künstlicher Intelligenz im vollen Sinne einlösen würden, doch existieren diese Systeme bislang nur in imaginären Zukünften. Lediglich „Sparks of Artificial General Intelligence“ (Bubeck et al. 2023) werden in aktuellen Transformer-Modellen mitunter diagnostiziert.

Auch der aktuelle Hype um KI lässt daher nicht den Schluss zu, dass die von Lindemann (2018) für die Moderne als konstitutiv betrachtete Grenze von Mensch und Maschine gefallen ist. Wohl aber lässt sich mit Hirschauer (i.E.) argumentieren, dass wir es bei KI mit einer Hybridfigur zu tun haben, deren Realitätsstatus kulturell strittig ist und die erst durch mediale Vermittlungen real gemacht wird. Die Bedingung dafür, dass eine Entität tatsächlich im Alltag als KI anerkannt wird, ist die ihr zugeschriebene Kommunikationsfähigkeit. Damit sie diese erlangt, muss eine spezifische Kontaktzone für Mensch und Maschine geschaffen werden. Dies ist die zentrale experimentelle Intervention in Turings *imitation game*. Was bei Turing noch ein futuristisches Szenario war, beginnt mit kommunikativer KI heute sukzessive Realität zu werden. Turing prognostizierte, dass eine Gesellschaft dann eine Maschine als ‚intelligentes‘ Gegenüber sozial akzeptieren würde, wenn sie sich als kompetenter Kommunikationspartner bewährt

⁷ Vgl. zur Kommunikativierung bereits Dogruel & Dickel 2022.

und in dieser Hinsicht von Menschen nicht mehr unterschieden werden kann. Dafür imaginierte er eine Situation, in der Menschen und Maschinen durch ein Interface vermittelt werden, das eine selektive Kontaktzone schafft, die das Spielfeld des Sozialen zugunsten der Maschine gestaltet.

Ob Turings Prognose der vollständigen sozialen Eingemeindung der KI bereits eingetroffen ist, ist jedoch fraglich (Collins 2021). Eher ist die Gesellschaft zum Versuchslabor eines noch laufenden Turing-Tests geworden, der nicht unter kontrollierten Laborbedingungen, sondern als „Realexperiment“ (Krohn 2007) durchgeführt wird. KI befindet sich aktuell in einer ontologischen Hybridposition „located between fact and fiction, between the present and an imagined future“ (Schulz-Schaeffer & Meister 2017: 197). Auf der einen Seite ist KI nicht mehr nur ein Gegenstand fiktionaler Darstellungen oder populärwissenschaftlicher Spekulation, sondern auch Teil des gesellschaftlichen Alltags. Auf Basis massiv ausgebauter digitaler Infrastrukturen (Crawford 2021), die Nutzer:innen als User Interfaces wie Websites und Apps gegenüberstehen, wird Deep Learning nun „umfassend [...] mit Praktiken der alltäglichen Mediennutzung verwoben“ (Hepp et al. 2022: 464). Die Technologie erscheint dabei als sich textlich oder stimmlich äußerndes Gegenüber und ist damit gesellschaftlich wahrnehmbar und typisierbar geworden (Müller 2023). Um Künstliche Intelligenz über Nischendiskurse kulturell zu popularisieren, war es notwendig, Maschinen als kommunikationsfähige Wesen zu gestalten. Erst als kommunikative KI beginnt KI in der gesellschaftlichen Breite sozial plausibel zu werden. Bei den kommunikativen Outputs ist dabei längst nicht mehr nur um Antworten auf Fragen in synthetischen Interaktionssituationen zu denken (Knorr-Cetina 2009). Vielmehr können solche Outputs als Texte, Kunstwerke, Videos oder Programme erscheinen, über technische Verbreitungsmedien zirkulieren und rekombiniert werden (Hepp et al. 2022). Je mehr die „Kommunikationsgesellschaft“ (Knoblauch 2017: 329 ff.) dabei durch technische Infrastrukturen gestaltet ist, die Menschen als User Interfaces zugänglich gemacht werden, desto geräuschloser kann eine solche Integration maschineller Entitäten gestaltet werden – nicht zuletzt deshalb, da sich die Frage nach der Zurechnung in der Sozialdimension (‘Wer genau ist für dieses kommunikative Produkt verantwortlich?’) nur in spezifischen Ausnahmesituationen stellt, wie etwa Prüfungskontexten oder Urheberrechtsfragen.

Zugleich aber zeigen kritische Diskurse – z. B. ob ChatGPT nur ein „Stochastical Parrot“ (Bender et al. 2021; Arkoudas 2023) ist oder nicht –, dass der Status von generativer KI als KI weiterhin strittig ist. So plädiert Evgeny

Morozov für ein „retiring“ des „hackneyed label of ‘artificial intelligence’ from public debate. [...]. It survived the end of the cold war because of its allure for science fiction enthusiasts and investors. We can afford to hurt their feelings“ (Morozov 2023). Angesichts der weiterhin laufenden Debatte, ob mit der aktuell existierenden generativen KI nun eine ‚echte‘ KI erreicht ist und was dies kulturell bedeutet, kann das Rennen hin zu einer AGI als Temporalisierung (genauer: als Futurisierung) der Frage nach dem Realitätsstatus von KI betrachtet werden: Sie wird in eine Zukunft verschoben, die es anzustreben (oder abzuwehren) gilt.

Meiner Argumentation folgend ist es müßig, eine AGI anhand technischer Parameter zu definieren. Zwar werden Informatiker:innen diesbezüglich gewiss weiter an Definitionen, Konkretisierungen und heuristischen Stufenmodellen arbeiten. Ob aber eine Entität als ‚echte‘ Künstliche Intelligenz gesellschaftlich anerkannt wird, ist keine technische, sondern eine soziale Frage. Die von mir vorgeschlagene Antwort lautet, dass die Gesellschaft eine Entität erst dann als menschenähnliche KI (an)erkennen wird, wenn sie in mediatisierten Umwelten als Kommunikator:in funktioniert, die von Menschen *nicht mehr zu unterscheiden* ist. Erst dann würden Maschinen zu einer *generellen „social prosthesis“* werden (Collins 2024: 13).

Ich möchte abschließend einige Folgerungen des in diesem Beitrag dargelegten Argumentationsgangs für die soziologische Forschung ziehen. *Erstens* sollte KI von der Soziologie nicht unreflektiert als analytischer Begriff verwendet, sondern als kulturelle Kategorie untersucht werden. Es ist ohne weiteres möglich, die Entwicklung und den Gebrauch von Algorithmen zu untersuchen, ohne diese als KI zu verbesondern. Zugleich aber wäre es verkürzt, die Verhärtung der Kategorie KI (Suchman 2023) lediglich als Produkt von kontingennten Diskursstrategien zu begreifen. KI wird gesellschaftlich nicht *allein* dadurch wirkmächtig, dass man sie thematisiert. Vielmehr wird sie als Kommunikationsteilnehmerin durch das Design spezifischer Chat- und Audiointerfaces als erfahrbare Realität für Nutzer:innen inszeniert und gestaltet: Nicht allein das Reden *über* KI, sondern auch und gerade das Reden *mit* KI verleiht dieser eine gesellschaftliche Plausibilität. Eine soziologische Beforschung von KI sollte daher nicht zuletzt in der Analyse von User Interfaces bestehen – was eine intensivere Kooperation mit dem sich aktuell herausbildenden Feld der *Interface Studies* (Distelmeyer 2023, Lipp et al. 2024, Wirth 2025) sowie dem Feld der *Human Machine Communication* (vgl. dazu Hepp et al. 2024) erfordern könnte.

Zweitens verweist das Rennen hin zur AGI seitens der aktuellen Technologieführer auf die Temporalität von

KI. Was wann als ‚echte‘ KI gilt und wie nah oder fern diese erscheint, wandelt sich durch neue soziotechnische Kontexte. Der Diskurs um AGI demonstriert, dass sich KI aktuell in einer ontologischen Hybridposition zwischen Fakten und Fiktionen befindet. Solange man sich noch auf dem „path to AGI“ (OpenAI 2024b) verortet, kann offengelassen werden, wann eine Technik tatsächlich als ‚Künstliche Intelligenz‘ bezeichnet werden soll.

Drittens spiegelt sich die Hybridisierung von Fakten und Fiktionen auch in der Sozialdimension. AGI war in den 2000er Jahren ein Begriff, der nahezu ausschließlich von transhumanistischen Technikenthusiast:innen verwendet wurde. Er spielte im Mainstream der KI-Forschung keine Rolle. Dass sich dies nun ändert, sorgt auch für eine Entdifferenzierung zwischen fraglos anerkannten KI-Forschenden (wie dem Nobelpreisträger und DeepMind CEO Demis Hassabis) und Akteuren, deren Seriosität zumindest als fraglich gilt (wie dem Transhumanisten Ray Kurzweil). Der Gebrauch der Kategorie AGI allein taugt nicht mehr als Indiz für die Feststellung von Zugehörigkeit zu der einen oder der anderen Gruppe und das damit verbundene „boundary work“ (Gieryn 1983). Den Konsequenzen dieser sozialen Entdifferenzierung für das diesbezügliche technikwissenschaftliche Feld und seine öffentliche Glaubwürdigkeit, gilt es weiter nachzugehen.

Viertens hat der Beitrag gezeigt, dass *Leistung* zu einem Schlüsselkonzept zum Vergleich von Mensch und Maschine geworden ist. Die Frage nach der Realität von *strong AI* wurde als philosophische Frage verhandelt (Kann es ein menschenähnliches maschinelles Bewusstsein geben?). AGI soll hingegen nun anhand ihrer Leistungsfähigkeit in kommunikativen Kontexten identifiziert werden (In welchen Kontexten kann KI menschliche Aufgaben erfüllen? Wie gut erledigt sie diese Aufgabe im Vergleich zu Menschen?). Im Anschluss daran könnte nun genauer untersucht werden, was als menschenähnliche Leistung in verschiedenen sozialen Feldern und Kontexten gilt. Im professionellen Feld des Programmierens dürfte es darum gehen, ob KI ebenso gut Code erstellen kann, im Kapitalmarkt wird es um Investitionsentscheidungen gehen, in zwischenmenschlichen Beziehungen um die Darstellung von Empathie und Interesse am Gegenüber, in der Wissenschaft um Forschung und deren Darstellung in Publikationen. Hier könnten sich vergleichende Analysen zu Bewertungsmaßstäben und -Praktiken anschließen. Wird eine KI erst dann als Artificial *General Intelligence* gelten, wenn sie kontextübergreifend menschenähnliche Leistungen erbringt oder diese sogar übertrifft?

Sollte die gesellschaftliche Konstruktion von AGI gelingen, könnte die Verbesonderung des Menschen als *exklusives* Personal der Gesellschaft an ihr Ende gelangen.

Begreift man – wie Gesa Lindemann – die moderne Gesellschaft als eine Gesellschaft, die genau durch diese Zentralstellung des Menschen bestimmt ist, würde eine AGI in der Tat das Ende der Moderne einleiten. Auch die Rede vom Anthropozän als gesellschaftlicher Selbstbeschreibung einer durch den Menschen geprägten Epoche würde sich dann als Fall einer Übergangsssemantik erweisen. Aktuell aber sind die Grenzen von Menschen und Maschinen zwar fragil geworden, aber nicht gefallen. Somit ist bis auf Weiteres davon auszugehen, dass Gesellschaften Praktiken und Techniken des *doing difference* in Stellung bringen werden, um menschliche und maschinelle Kommunikation auch in tiefgreifend mediatisierten Umwelten weiterhin unterscheiden zu können – um herauszufinden, ob die Bilder, die Menschen online sehen, maschinengenerierte Fälschungen sind oder eine studentische Arbeit mithilfe von ChatGPT verfasst wurde. Ein Leitmotiv ist hier die Vermeidung von Täuschung und Betrug, die bereits im *imitation game* angelegt ist. Wenn Akteure solche Fragen stellen, begeben sie sich in die Position von C im *imitation game*. Sie suchen nach Markern, die auf Maschinenhaftigkeit hinweisen.

Eine *fünfte* soziologische Forschungsaufgabe besteht darin, spezifische Situationen und Kontexte zu differenzieren, in denen es Menschen auf das diesbezügliche Differenzieren (nicht) ankommt und wie sie dieses ggf. bewältigen. Der Gegenpol zur Reproduktion realexperimenteller Turing-Tests wäre eine Indifferenz gegenüber der menschlichen oder maschinellen Agency. Ein solches *undoing difference* könnte man daran erkennen, dass gesellschaftlich akzeptiert wird, dass kulturelle Artefakte wie dieser Text zunehmend hybride Mensch-Maschine-Artefakte sind und man darauf verzichtet, den maschinellen Teil zu bestimmen und ihn (etwa als Täuschung) zu verbesondern. Dies würde zugleich eine Abrüstung von Erwartungen hinsichtlich einer exklusiv zurechenbaren menschlichen Eigenleistung implizieren.

Förderhinweis: Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – Projektnummer 442261292 – SFB 1482

Literatur

- Ahrweiler, P., 1995: Künstliche Intelligenz-Forschung in Deutschland. Die Etablierung eines Hochtechnologie-Fachs. Zugl.: Berlin, Freie Univ., Diss., 1993. Münster: Waxmann.
- Altman, S., 2023: Planning for AGI and Beyond. <https://openai.com/index/planning-foragi-and-beyond/> (28.02.2025).
- Anicker, F., 2023: Sozialisierte Maschinen. Zur gesellschaftlichen Funktion von Künstlicher Intelligenz. Zeitschrift für Theoretische Soziologie 1: 79–105.

- Arkoudas, K., 2023: ChatGPT is no Stochastic Parrot. But it also Claims that 1 is Greater than 1. *Philosophy & Technology* 36: 54.
- Aschenbrenner, L., 2024: Situational Awareness: The Decade Ahead. <https://situational-awareness.ai/> (28.02.2025).
- Baecker, D., 2011: Who Qualifies for Communication? A Systems Perspective on Human and Other Possibly Intelligent Beings Taking Part in the Next Society. *Technikfolgenabschätzung in Theorie und Praxis* 20: 17–26.
- Baecker, D., 2017: Wie verändert die Digitalisierung unser Denken und unseren Umgang mit der Welt? Ausgangspunkte einer Theorie der Digitalisierung. S. 3–24 in: R. Gläß & B. Leukert (Hrsg.), *Handel 4.0. Die Digitalisierung des Handels – Strategien, Technologien, Transformation*. Berlin, Heidelberg: Springer Gabler.
- Bareis, J. & C. Katzenbach, 2022: Talking AI into Being: The Narratives and Imaginaries of National AI Strategies and Their Performative Politics. *Science, Technology, & Human Values* 47: 855–881.
- Bender, E. M., T. Gebru, A. McMillan-Major & S. Shmitchell, 2021: On the Dangers of Stochastic Parrots. S. 610–623 in: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. FAccT '21: Virtual Event Canada, 03.03.2021–10.03.2021. New York: Association for Computing Machinery.
- Borchardt, A., F. M. Simon, O. Zachrisson, K. Bremme, J. Kurczabinska, E. Mulhall & Y. Johann, 2024: Trusted Journalism in the Age of Generative AI. *EBU News Report*.
- Bostrom, N., 2005: Transhumanist Values. *Review of Contemporary Philosophy* 4: 87–101.
- Bostrom, N., 2014: Superintelligenz. Szenarien einer kommenden Revolution. Berlin: Suhrkamp.
- Bostrom, N. & A. Sandberg, 2008: Whole Brain Emulation. A Roadmap. Technical Report 3.
- Bubeck, S., V. Chandrasekaran, R. Eldan, J. Gehrke, E. Horvitz, E. Kamar, P. Lee, Y. T. Lee, Y. Li, S. Lundberg, H. Nori, H. Palangi, M. T. Ribeiro & Y. Zhang, 2023: Sparks of Artificial General Intelligence: Early experiments with GPT-4. <https://arxiv.org/abs/2303.12712> (26.03.2025).
- Casper, M., 1994: Reframing and Grounding Nonhuman Agency: What Makes a Fetus an Agent? *American Behavioral Scientist* 37: 839–856.
- Cerulo, K. A., 2009: Nonhumans in Social Interaction. *Annual Review of Sociology* 35: 531–552.
- Chu, Y. & P. Liu, 2023: Public Aversion Against ChatGPT in Creative Fields? *Innovation* 4: 100449. <https://www.cell.com/action/showPdf?pii=S2666-6758%2823%2900077-2> (26.03.2025)
- Collins, H., 1990: Artificial Experts. *Social Knowledge and Intelligent Machines*. Cambridge, MA: MIT Press.
- Collins, H., 2021: The Science of Artificial Intelligence and its Critics. *Interdisciplinary Science Reviews* 46: 53–70.
- Collins, H., 2024: Why Artificial Intelligence needs Sociology of Knowledge: Parts I and II. *AI & Society*. <https://doi.org/10.1007/s00146-024-01954-8>.
- Cook, J., 2024: OpenAI's 5 Levels Of 'Super AI' (AGI To Outperform Human Capability). *Forbes*, 16.07.2024. <https://www.forbes.com/sites/jodiecook/2024/07/16/openais-5-levels-of-super-ai-agi-to-outperform-human-capability/> (28.02.2025).
- Crawford, K., 2021: *Atlas of AI. Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven, London: Yale University Press.
- Depouinti, I., P. Saukko & S. Natale, 2023: Ideal Technologies, Ideal Women: AI and Gender Imaginaries in Redditors' Discussions on the Replika Bot Girlfriend. *Media, Culture & Society* 45: 720–736.
- Descola, P., 2022: *Jenseits von Natur und Kultur*. Berlin: Suhrkamp.
- Dickel, S. & M. Schmidt-Jüngst, 2021: Gleiche Menschen, ungleiche Maschinen. Die Humandifferenzierung digitaler Assistenzsysteme und ihrer Nutzer:innen in der Werbung. S. 342–367 in: D. Dizdar, S. Hirschauer, J. Paulmann & G. Schabacher (Hrsg.), *Humandifferenzierung. Disziplinäre Perspektiven und empirische Sondierungen*. Weilerswist: Velbrück.
- Distelmeyer, J., 2021: *Kritik der Digitalität*. Wiesbaden: Springer.
- Distelmeyer, J., 2023: A Case for Interface Studies. From Screenshots to Desktop/Screen Films. S. 317–331 in: W. Gerling, S. Möring, M. De Mutis (Hrsg.), *Screen images. In-Game Photography, Screenshot, Screencast*. Berlin: Kulturverlag Kadmos Berlin.
- Dogruel, L. & S. Dickel, 2022: Die Kommunikativierung der Maschinen. *Publizistik* 67: 475–486.
- Elyoseph, Z., D. Hadar-Shoval, K. Asraf & M. Lvovsky, 2023: ChatGPT Outperforms Humans in Emotional Awareness Evaluations. *Frontiers in Psychology* 14: 1199058.
- Esposito, E., 1993: Der Computer als Medium und Maschine. *Zeitschrift für Soziologie* 22: 338–354.
- Esposito, E., 2017: Artificial Communication? The Production of Contingency by Algorithms. *Zeitschrift für Soziologie* 46: 249–265.
- Esposito, E., 2024: Kommunikation mit unverständlichen Maschinen. Wien, Salzburg: Residenz Verlag.
- French, R. M., 2000: The Turing Test: The First 50 Years. *Trends in Cognitive Sciences* 4: 115–122.
- Gebru, T. & É. P. Torres, 2024: The TESCREAL bundle: Eugenics and the Promise of Utopia through Artificial General Intelligence. *First Monday* 29: 1–42.
- Gieryn, T. F., 1983: Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists. *American Sociological Review* 48: 781–795.
- Goertzel, B., 2014: Artificial General Intelligence: Concept, State of the Art, and Future Prospects. *Journal of Artificial General Intelligence* 5: 1–48.
- Gonsalves, T., 2019: The Summers and Winters of Artificial Intelligence. S. 168–179 in: M. Khosrow-Pour (Hrsg.), *Advanced Methodologies and Technologies in Artificial Intelligence, Computer Simulation, and Human-Computer Interaction*. Hershey, PA: IGI Global.
- Häußling, R., C. Härpfer & M. Schmitt (Hrsg.), 2024: *Soziologie der Künstlichen Intelligenz. Perspektiven der Relationalen Soziologie und Netzwerkforschung*. Bielefeld: Transcript.
- Hauskeller, M., 2012: My Brain, my Mind, and I. Some philosophical Assumptions of Mind-Uploading. *International Journal of Machine Consciousness* 4: 187–200.
- Hayles, N. K., 1999: *How We Became Posthuman. Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago, Illinois: University of Chicago Press.
- Heilmann, T. A., 2012: Textverarbeitung. Eine Mediengeschichte des Computers als Schreibmaschine. Bielefeld: Transcript.
- Hepp, A., W. Loosen, S. Dreyer, J. Jarke, S. Kannengießer, C. Katzenbach, R. Malaka, M. Pfadenhauer, C. Puschmann & W. Schulz, 2022: Von der Mensch-Maschine-Interaktion zur kommunikativen KI. *Publizistik* 67: 449–474.
- Hepp, A., G. Bolin, A. Guzman & W. Loosen, 2024: Mediatization and Human-Machine Communication: Trajectories, Discussions, Perspectives. *Human-Machine Communication* 7: 7–21.
- Hirschauer, S., (i.E.): Tiere, Götter, Dinge, Tote und andere Aliens. Eine vergleichende Kartierung der Distinktionszonen des Humanen in: S. Hirschauer, P. Hoffmann, A. Friedrichs & G. Schabacher (Hrsg.), *Komparatistik der Humandifferenzierung*. Weilerswist: Velbrück.

- Hirschauer, S., 2021: Menschen unterscheiden. Grundlinien einer Theorie der Humandifferenzierung. *Zeitschrift für Soziologie* 50: 155–174.
- Hirschauer, S. & D. Nübling, 2021: Sinnschichten des Kulturellen und die Aggregatzustände der Sprache. S. 58–83 in: D. Dizdar, S. Hirschauer, J. Paulmann & G. Schabacher (Hrsg.), *Humandifferenzierung. Disziplinäre Perspektiven und empirische Sondierungen*. Weilerswist: Velbrück.
- Hirsch-Kreinsen, H. & T. Krokowski, 2023: Technologieversprechen Künstliche Intelligenz. Vergangene und gegenwärtige Konjunkturen in der Bundesrepublik. *Berliner Journal für Soziologie* 33: 453–484.
- Hookway, B., 2014: *Interface*. Cambridge, London: MIT Press.
- Knoblauch, H., 2017: *Die kommunikative Konstruktion der Wirklichkeit*. Wiesbaden: Springer VS.
- Knorr-Cetina, K., 2009: The Synthetic Situation. *Interactionism for a Global World. Symbolic Interaction* 32: 61–87.
- Krohn, W., 2007: Realexperimente. Die Modernisierung der ‚offenen Gesellschaft‘ durch experimentelle Forschung. *Erwägen – Wissen – Ethik* 18: 343–356.
- Kurzweil, R., 2005: *The Singularity is Near. When Humans Transcend Biology*. New York: Viking.
- Latour, B., 1992: Where are the Missing Masses? *The Sociology of a Few Mundane Artifacts*. S. 225–258 in: W. E. Bijker & J. Law (Hrsg.), *Shaping Technology/Building Society. Studies in Sociotechnical Change*. Cambridge: MIT Press.
- Latour, B., 1999: *Pandora’s Hope. Essays on the Reality of Science Studies*. Cambridge: Harvard University Press.
- Latour, B., 2005: *Reassembling the Social. An Introduction to Actor-Network-Theory*. Oxford, New York: Oxford University Press.
- Licklider, J. C. & R. W. Taylor, 1968: The Computer as a Communication Device. *Science and Technology*: 21–41.
- Lindemann, G., 2009: Das Soziale von seinen Grenzen her denken. Weilerswist: Velbrück Wissenschaft.
- Lindemann, G., 2018: Strukturnotwendige Kritik: Weilerswist: Velbrück Wissenschaft.
- Lipp, B., P. Helm, A. Karafillidis & R. Pujadas, 2024: De-Centring the Interface. Towards the Integrated Study of Interfacial Relations. *Information Communication and Society*: 1–17.
- Luckmann, T., 1980: Die Grenzen der Sozialwelt. S. 56–92 in: *Lebenswelt und Gesellschaft. Grundstrukturen und geschichtliche Wandlungen*. Paderborn: Schöningh.
- Luhmann, N., 1980: *Gesellschaftliche Struktur und semantische Tradition*. S. 9–71 in: *Gesellschaftsstruktur und Semantik*. Frankfurt am Main: Suhrkamp.
- Luhmann, N., 1984: *Soziale Systeme. Grundriß einer allgemeinen Theorie*. Frankfurt am Main: Suhrkamp.
- Luhmann, N., 1990: *Die Wissenschaft der Gesellschaft*. Frankfurt am Main: Suhrkamp.
- Mayer, H., F. Muhle & I. Bock, 2020: Whiteboxing MAX. Zur äußeren und inneren Interaktionsarchitektur eines virtuellen Agenten. S. 295–322 in: E. Geitz, C. Vater & S. Zimmer-Merkle (Hrsg.), *Black Boxes – Versiegelungskontexte und Öffnungsversuche. Interdisziplinäre Perspektiven*. Berlin, Boston: De Gruyter.
- Microsoft, 2024: What Is Artificial General Intelligence? <https://www.microsoft.com/en-us/bing/do-more-with-ai/artificial-general-intelligence?form=MA13KP> (28.02.2025).
- Mölders, M., J. Renn & J. Siri (Hrsg.), 2024: Differenzierung und Integration. Zur Ausdifferenzierung der Differenzierungstheorie. 5. Sonderband der ZTS. Weinheim: Juventa Verlag.
- Morgan, J., 2019: Yesterday’s Tomorrow Today. Turing, Searle and the Contested Significance of Artificial Intelligence. S. 82–137 in: I. Al-Amoudi & J. Morgan (Hrsg.), *Realist Responses to Post-Human Society. Ex machina*. London, New York: Routledge.
- Morozov, E., 2023: The Problem with Artificial Intelligence? It’s neither Artificial nor Intelligent. *The Guardian*, 30.03.2023. <https://www.theguardian.com/commentisfree/2023/mar/30/artificial-intelligence-chatgpt-human-mind> (28.02.2025).
- Morris, M. R., J. Sohl-Dickstein, N. Fiedel, T. Warkentin, A. Dafoe, A. Faust, C. Farabet & S. Legg, 2024: Position: Levels of AGI for Operationalizing Progress on the Path to AGI. *Proceedings of ICML*. <https://arxiv.org/abs/2311.02462> (26.03.2025).
- Müller, M. R., 2023: Social Displays – Creating Accountability in Robotics. *Österreichische Zeitschrift für Soziologie* 48: 469–487.
- Muhle, F., 2018: Sozialität von und mit Robotern? Drei soziologische Antworten und eine kommunikationstheoretische Alternative. *Zeitschrift für Soziologie* 47: 147–163.
- Nake, F., 2019 [1984]: *Human-Machine Interface. Interface Critique* 2: 35–45.
- Nassehi, A., 2019: *Muster. Theorie der digitalen Gesellschaft*. München: C. H. Beck.
- Natale, S., 2021: *Deceitful media: Artificial Intelligence and Social Life after the Turing Test*. New York: Oxford University Press.
- Natale, S. & I. Depounti, 2024: Artificial Sociality. *Human-Machine Communication* 7: 83–98.
- Naudé, W. & N. Dimitri, 2020: The Race for an Artificial General Intelligence: Implications for Public Policy. *AI & Society* 35: 367–379.
- OpenAI, 2024a: OpenAI Charter. <https://openai.com/charter/> (28.02.2025).
- OpenAI, 2024b: Research. <https://openai.com/research/> (28.02.2025).
- Patel, N. & D. Hassabis, 2023: Inside Google’s Big AI Shuffle – and How it Plans to Stay Competitive, with Google DeepMind CEO Demis Hassabis. *The Verge*, 10.07.2023. <https://www.theverge.com/23778745/demis-hassabis-google-deepmind-ai-alpha-fold-risks> (28.02.2025).
- Pavlik, J. V., 2023: Collaborating with ChatGPT: Considering the Implications of Generative Artificial Intelligence for Journalism and Media Education. *Journalism and Mass Communication Educator* 78: 84–93.
- Possati, L. M., 2023: Psychoanalyzing Artificial Intelligence: The Case of Replika. *AI & Society* 38: 1725–1738.
- Rammert, W., 1995: *Soziologie und Künstliche Intelligenz. Produkte und Probleme einer Hochtechnologie*. Frankfurt/Main, New York: Campus.
- Rammert, W., 2023: Wie die Soziologie zur ‚Künstlichen Intelligenz‘ kam: Eine kurze Geschichte ihrer Beziehung. S. 31–66 in: F. Muhle (Hrsg.), *Soziale Robotik. Eine sozialwissenschaftliche Einführung*. Berlin, Boston: De Gruyter.
- Rammert, W. & I. Schulz-Schaeffer, 2002: Technik und Handeln. Wenn soziales Handeln sich auf menschliches Verhalten und technische Abläufe verteilt. S. 11–64 in: W. Rammert & I. Schulz-Schaeffer (Hrsg.), *Können Maschinen handeln? Soziologische Beiträge zum Verhältnis von Mensch und Technik*. Frankfurt/Main, New York: Campus.
- Schulz-Schaeffer, I. & M. Meister, 2017: Laboratory Settings as Built Anticipations – Prototype Scenarios as Negotiation Arenas Between the Present and Imagined Futures. *Journal of Responsible Innovation* 4: 197–216.
- Schulz-Schaeffer, I., M. Meister, T. Clausnitzer & K. Wiggert, 2023: Sozialität von Robotern aus handlungstheoretischer Perspektive.

- S. 165–192 in: F. Muhle (Hrsg.), Soziale Robotik. Eine sozialwissenschaftliche Einführung. Berlin, Boston: De Gruyter Oldenbourg.
- Searle, J., 1980: Minds, Brains and Programs. Behavioral and Brain Sciences 3: 417–424.
- Suchman, L., 2007: Human-Machine Reconfigurations. Plans and Situated Actions. Cambridge, New York: Cambridge University Press.
- Suchman, L., 2023: The Uncontroversial ‘Thingness’ of AI. Big Data & Society 10.
- Turing, A. M., 1950: Computing Machinery and Intelligence. Mind LIX 59: 433–460.
- Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser & I. Polosukhin, 2017: Attention Is All You Need. <http://arxiv.org/pdf/1706.03762> (28.02.2025).
- Weizenbaum, J., 1976: Computer Power and Human Reason. From Judgment to Calculation. San Francisco: Freeman.
- Wirth, S., 2025: Formierungen des Interface. Zur Mediengeschichte und-theorie des Personal Computing. Bielefeld: Transcript.

Autor

Sascha Dickel

Johannes Gutenberg-Universität Mainz
Isaac-Fulda-Allee 2b
55124 Mainz

Sascha Dickel, geb. 1978 in Gießen. Studium der Politikwissenschaft und Soziologie in Marburg und Frankfurt; Promotion 2010 (Universität Bielefeld); Habilitation 2019 (TU München); seit 2021 Universitätsprofessor für Mediensoziologie und Gesellschaftstheorie an der Johannes Gutenberg-Universität Mainz.
Forschungsschwerpunkte: Gesellschaftstheorie, Digitalisierung, Zukünfte. Wichtigste Publikationen: Prototyping Society. Zur vorauselenden Technologisierung gesellschaftlicher Zukunft. Bielefeld: Transcript 2019; Der kybernetische Blick und seine Grenzen. Zur systemtheoretischen Selbstbeschreibung der digitalen Gesellschaft. Berliner Journal für Soziologie 33 (2023): 197–226.