

# THESENPAPIER ZU DEN SOZIALEN IMPLIKATIONEN VON KÜNSTLICHER INTELLIGENZ

GEFÖRDERT VOM

# INHALT

Inhalt.....	2
Einleitende Bemerkungen.....	4
Diskurse, Macht und Ungleichheit.....	6
Künstliche Intelligenz ist bestimmt durch Interessen und wirkmächtige Bilder.....	6
Künstliche Intelligenz wird von verheißungsvollen Zukunftsnarrativen gerahmt.....	6
Der öffentliche Diskurs um Künstliche Intelligenz folgt teils überhöhten, interessengeleiteten Bildern und braucht mehr Perspektivenvielfalt sowie wissenschaftliche Fundierung.....	7
Künstliche Intelligenz muss als Gegenstand von Ideologiekritik gesehen werden.....	7
Künstliche Intelligenz birgt die Gefahr, bestehende Verzerrungen und Ungleichheiten zu reproduzieren und zu verschärfen.....	7
Künstliche Intelligenz verschärft den Kampf um gesellschaftliche Macht- und Herrschaftsverhältnisse.....	8
Welt- und Menschenbilder.....	8
„Machine learning“ ist ein technischer Fachbegriff, nur Menschen lernen.....	8
Der Mensch ist keine flexible Maschine, die es technisch nachzubilden gilt.....	9
Der Mensch ist nicht bloß ein Stör- oder Risikofaktor, den es herauszuhalten gilt.....	9
Die Funktionen Künstlicher Intelligenz sind keine angeeigneten Kompetenzen.....	9
Menschliche Intelligenz und soziales Handeln beinhalten mehr, als Künstliche Intelligenz zu erfassen und umzusetzen in der Lage ist.....	9
Gesellschaftliche und soziale Grenzen.....	10
Künstliche Intelligenz erfasst nur das, was sich in Zahlen umwandeln und somit formalisieren lässt.....	10
Wenn die Grenzen von KI-Lösungen nicht gesehen und berücksichtigt werden, entstehen zahlreiche soziale Risiken.....	10
Künstliche Intelligenz ersetzt menschliche Tätigkeiten nicht, sondern transformiert Handlungszusammenhänge.....	11
Künstliche Intelligenz ist mit einer unterschätzten Fülle an sichtbarer und unsichtbarer (vor-, parallel und nachgelagerter) Arbeit verbunden.....	11
Soziale Entwicklungs- und Nutzungskontexte.....	11
Künstliche Intelligenz forciert bei ihrer Umsetzung und in ihrer Nutzung einen Anpassungsprozess an die Bedarfe des Systems.....	12
Künstliche Intelligenz standardisiert Wissen und Verhalten.....	12
Künstliche Intelligenz strukturiert gesellschaftliche Erfahrungs- und Wissensräume auf latente Weise.....	12

Gesellschaftliche Risiken und Gefahren.....	13
Künstliche Intelligenz birgt das Risiko, die Prinzipien von Demokratie und politischer Teilhabe zu unterlaufen.....	13
Künstliche Intelligenz bedroht Autonomie durch Scheinentscheidungen in algorithmisch vorbestimmten Entscheidungsräumen.....	13
Künstliche Intelligenz birgt das Risiko, durch die Verengung von Sinnbildungsprozessen zu einem Verlust gesellschaftlicher Diversität zu führen. ....	13
Praktische Gestaltung.....	14
Künstliche Intelligenz lässt sich nicht auf technologische Funktionen reduzieren.....	14
Künstliche Intelligenz am Arbeitsplatz braucht den Blick auf die organisationalen Praktiken. ....	14
Künstliche Intelligenz bietet mehr Möglichkeiten, menschliches Handeln zu ergänzen als zu ersetzen.....	14
Ziel der KI-Entwicklung und des KI-Einsatzes sollte ein komplementäres und koevolutionäres Verhältnis von Mensch und Künstlicher Intelligenz sein. ....	15
Künstliche Intelligenz verliert ihre Mystik, wenn entsprechende Anwendungen als Alltagstechnologie angeeignet werden.....	15
Gesellschaftspolitische Gestaltung.....	15
Künstliche Intelligenz erfordert einen neuen Modus gesellschaftlicher Innovation. ....	15
Künstliche Intelligenz bedarf einer integrativen und innovationsfreundlichen Normierung. ....	16
Künstliche Intelligenz benötigt neue Formen ethischer Verantwortung und rechtlicher Haftung. ....	16
Technikreflexion.....	16
Künstliche Intelligenz bedarf einer gesellschaftstheoretischen Rahmung. ....	16
Künstliche Intelligenz fordert zu einer Reflexion gesellschaftlicher, sozialer und anthropologischer Kategorien auf. ....	17
KI-Forschung kann von geistes-, sozial- und kulturwissenschaftlichen Begriffen und Konzepten stark profitieren. ....	17
Künstliche Intelligenz benötigt geistes-, sozial- und kulturwissenschaftliches Methodenwissen. ....	17
Künstliche Intelligenz macht eine Neuausrichtung des Bildungssystems notwendig. ....	18
Ausblick.....	19
Impressum.....	20

# EINLEITENDE BEMERKUNGEN

Künstliche Intelligenz (KI) stellt als komplexes Phänomen einen wesentlichen Faktor des gesellschaftlichen Wandels im 21. Jahrhundert dar. Mittlerweile werden zahlreiche Anwendungen entwickelt und genutzt, die auf maschinellem Lernen, mitunter hochentwickelten „deep-learning“-Algorithmen und künstlichen neuronalen Netzen aufbauen oder neue Verfahren der symbolischen KI und statistischer Auswertung von großen Datenmengen anwenden. Auf diese Weise ergeben sich neue Möglichkeiten des Umgangs mit und der Nutzung von Daten: Der Einsatz und die Entwicklung von KI zielen darauf ab, große und unstrukturierte oder strukturierte Datenmengen mittels statistischer und mathematischer Methoden zu bearbeiten, um dort verborgene Zusammenhänge zu erkennen und nutzbar zu machen. Dabei können KI-Systeme überwacht bzw. angeleitet (supervised learning) oder nicht überwacht bzw. angeleitet (unsupervised learning) oder indirekt durch gewichtete Zielsetzungen („Belohnungen“) bestärkt lernen (reinforcement learning). KI-Algorithmen können unbekannte Daten schrittweise (teils eigenständig) strukturieren und fortlaufend ihre Strukturierungsverfahren weiterentwickeln. Aber der Begriff KI weist auch weit über maschinelles Lernen hinaus. KI eint die Eigenschaft, neue Grundlagen des Handelns und Entscheidens sowie der Steuerung von Maschinen und Prozessen zu schaffen – sei es im Privatleben, in der Arbeit und Wirtschaft oder in der Politik.

Offen bleibt jedoch die Frage, welche tiefgreifenden, teilweise latenten und schleichenden Folgen für den Menschen und das gesellschaftliche Zusammenleben mit der Entwicklung und Nutzung von KI verbunden sind. KI steht unbestreitbar für einen großen technischen Fortschritt. Doch wie wandelt sich das Verhältnis von Mensch und Technik mit KI und wie ist dieser Wandel zu bewerten? Welche Chancen, aber auch Risiken eröffnen sich durch den Einsatz und die Entwicklung von KI für Mensch und Gesellschaft? Mit welchen mittel- und langfristigen sozialen Implikationen gehen die neuen Nutzungsmöglichkeiten einher? Welchen Grenzen und Mustern unterliegt der KI-getriebene Wandel und welche Gestaltungsmöglichkeiten bieten sich? Und nicht zuletzt: Wer und was bestimmt die Entwicklungspfade, die KI nimmt – mit welchen Folgen für wen?

Dieses Thesenpapier möchte den Blick für soziale und gesellschaftliche Folgen von KI schärfen. Es nimmt geistes-, sozial- und kulturwissenschaftliche sowie juristische Standpunkte ein, berücksichtigt dabei aber auch mediale und künstlerische Darstellungen und die damit verbundenen Zuschreibungen. Das bedeutet: Es begreift KI nicht bloß als Technologie, die lediglich angewandt und benutzt wird. KI wird vielmehr als eine menschengemachte, das heißt, als eine spezifischen Interessen unterworfenen und sozial geformte Technologie begriffen, deren Folgen und Wirkungen teils entwicklungsseitig vorbereitet sind, sich teils jedoch erst in der konkreten Nutzung manifestieren. KI ist demnach – wie jede andere Technologie auch – eine soziale Technologie, die für den Menschen und die Gesellschaft, in der er lebt, spezifische Chancen eröffnet, aber auch eigene Risiken birgt. Dazu gehört, dass die (Neben-)Folgen und (Wechsel-)Wirkungen von KI, Mensch und Gesellschaft nicht immer abzusehen sind und durch KI soziale Prozesse und Strukturen nachhaltig verändert werden können – auch ohne dass dies auf explizite Motive zurückzuführen ist. Für ein Verständnis von KI, das als Ergänzung ethischer und rechtlicher Einordnungen und als Korrektiv naiver technizistisch-positivistischer Diskurse gesehen werden kann, möchte dieses Thesenpapier werben.

KI ist nicht allgemein gültig definiert. Es gibt keine eigentliche Bedeutung von KI und keine konsensfähige Menge an Elementen, die zu KI notwendig oder hinreichend gehören. KI ist ein wissenschaftsstrategischer Sammelbegriff, der je interessengeleitet besetzt wird. Die Bedeutung des Begriffs wird in weiten Teilen in bildhaften, übertragenen, uneigentlichen Hinsichten konstruiert.

Zum Phänomen KI gehören verschiedenste Elemente wie Technologien und Strukturen, aber auch Narrative, Imaginationen, Verheißungen etc. KI ist bezüglich der technischen Elemente des Gesamtphänomens einerseits eine Technologie wie andere auch, andererseits weist KI Besonderheiten auf, die nur für sie gelten. Die Grenze zwischen Überlegungen, die Technologien im Allgemeinen betreffen, und solchen, die auf die Spezifika von KI fokussieren, ist jedoch fließend. KI muss daher sowohl in ihrer technologischen Allgemeinheit als auch in ihrer technologischen Spezifität betrachtet werden.

Die nachfolgenden Thesen, die dem BMBF-geförderten Forschungsverbund „KI – Mensch – Gesellschaft“ ([www.kimege.de](http://www.kimege.de)) entstammen, beziehen sich folglich immer auf beide Aspekte von KI: KI als Beispiel oder Zuspitzung typischer Aspekte der Technisierung allgemein und KI als neues (auch, aber nicht nur technisches) Phänomen. In diesem Verbund befassen sich 26 Expert\*innen aus der sozialwissenschaftlichen und philosophischen Technikforschung gemeinsam mit gesellschaftlichen und sozialen Fragestellungen, die – teils prospektiv – die Folgen von KI für Mensch und Gesellschaft ausloten. Soziale, kulturelle, philosophische, politische und rechtliche Aspekte werden dabei gleichermaßen beleuchtet, sodass ein umfassendes Bild der gesellschaftlichen und sozialen Implikationen, Chancen und Risiken von KI gezeichnet wird.

Die hier versammelten Thesen geben einen Einblick in die Arbeit des Forschungsverbundes und dokumentieren einen Zwischenstand der Debatte. Sie sind *work in progress* und dienen dazu, geistes- und sozialwissenschaftliche Zugänge zu KI zu erschließen und erste Erträge dieser Zugänge zu markieren. Dazu gehört, unterbelichtete Aspekte und Lücken in der öffentlichen und wissenschaftlichen Auseinandersetzung mit KI zu identifizieren und inhaltlich zu füllen. Das Thesenpapier beansprucht infolgedessen nicht, sämtliche Aspekte von KI aufzugreifen und vollumfänglich zu beschreiben. Dokumentiert werden die Gesichtspunkte und Facetten, die sich im Forschungsverbund als politisch und gesellschaftlich drängend und für bestehende Forschungskontexte als anschlussfähig erwiesen haben. Die Darstellung der Thesen erfolgt in thematischen Blöcken, die diese Relevanz widerspiegeln: Es geht um Fragen von Macht und sozialer Ungleichheit, um mit KI verbundene Vorstellungen vom Menschen und der Welt, um gesellschaftliche und soziale Grenzen von und für KI, um soziale Entwicklungs- und Nutzungskontexte und vieles mehr.

Veröffentlicht werden die Thesen mit dem Ziel, einen Anstoß zu geben für die Reflexion und Diskussion dieser Technologie mitsamt ihren gesehenen und ungesehenen, erwünschten und unerwünschten, manifesten und latenten Folgen. Im Diskurs über diese Thesen, so die Absicht dieses Papiers, schärft sich der Blick weiter. Technikorientierte Rechts-, Geistes-, Sozial- und Kulturwissenschaften können Impulse für eine sozial und gesellschaftlich nachhaltige Entwicklung und Nutzung von KI geben. Denn letztlich geht es um eine Frage, die nur gemeinsam beantwortet werden kann: Mit welchen Technologien wollen wir auf welche Weise leben?

# DISKURSE, MACHT UND UNGLEICHHEIT

Technik und Maschinen sind im Laufe der Jahrhunderte zu einem integralen Bestandteil unserer Gesellschaft geworden. Sie prägen heutzutage mehr denn je unsere Kultur und unsere Lebensweise und spiegeln diese wider. Doch obwohl technologische Innovationen weitreichenden sozialen Wandel auszulösen vermögen, steht der Zusammenhang von Technik, Herrschaft und sozialer Ungleichheit jenseits offensichtlicher militär-technischer Kontexte selten im Fokus der öffentlichen und wissenschaftlichen Reflexion. Die Entwicklung und Nutzung von KI birgt das Potenzial, gesellschaftliche, praktische und diskursive Macht- und Ungleichheitsverhältnisse auf vielfältige Weise zu verändern, aber auch zu reproduzieren oder zu bestärken. Darauf nehmen die folgenden Thesen Bezug.

## KÜNSTLICHE INTELLIGENZ IST BESTIMMT DURCH INTERESSEN UND WIRKMÄCHTIGE BILDER.

Als Imaginationsfeld geht KI weit über die Sphäre des Technischen hinaus. Strategische Interessen prägen die Definitionen dieses Geflechts aus Technologien und Technologieimaginationen auf einem umkämpften Diskursfeld. Es werden vielfache und sehr verschiedene Ansprüche auf Definitionshoheit erhoben. Deren Äußerungen können nicht einfach für „bare Münze“ genommen und als unverrückbarer Sachstand akzeptiert werden. In die Bezeichnung und Imagination von KI gehen implizit sehr alte und wirkmächtige Bilder von Mensch und Technik ein, die zu erkennen und zu hinterfragen sind. Wenn man die mit KI verwobenen kulturell-historischen Bedeutungstraditionen missachtet, kann es zu folgenreichen Missverständnissen und Fehleinschätzungen im Umgang mit ihr kommen. So verwischt die Unterscheidung zwischen „künstlich intelligent“ und „nicht künstlich intelligent“ den Intelligenzbegriff. KI sollte daher als ein Inbegriff verstanden werden, der es erlaubt, heterogene Elemente zusammenzudenken und KI jenseits technizistischer Verkürzungen in reflexionsbegrifflicher Hinsicht als etwas zugleich Technisches, Natürliches und Kultürliches zu analysieren. Hierbei ist auch zu beachten, dass Intelligenz als Leistungsbegriff immer nur in sektoralen Anwendungsfeldern für Vergleiche (zwischen Mensch und Technik) dienlich ist.

## KÜNSTLICHE INTELLIGENZ WIRD VON VERHEISSUNGSVOLLEN ZUKUNFTSNARRATIVEN GERAHMT.

Ein tiefergehendes Verständnis zeitgenössischer Verheißungen zu KI braucht Perspektivenvielfalt, um den mit KI verbundenen schleichenden Wandel und die gesellschaftliche Funktion von Zukunftserzählungen einzufangen. Das zentrale Bezugsproblem von Verheißungserzählungen ist hierbei nicht die Plausibilität von Aussagen über KI, sondern die gesellschaftlichen Funktionen von Verheißungen. Verheißungserzählungen sind Fortsetzungsgeschichte zwischen Gegenwart und Zukunft – einerseits schon latent prognostische und spekulative Zukunftsszenarien, andererseits noch empirisch und theoretisch fundierte Gegenwartsdiagnose. Verheißungserzählungen werden in einem heterogenen narrativen Wissensraum zwischen Wissenschaft, Kunst, Medien und Öffentlichkeiten (re-)produziert und rezipiert. Sie symbolisieren die Flucht ins Technische, die Flucht ins Mythische sowie die Flucht ins Utopische. Gleichwohl können sie als Katalysator bzw. Generator von entlastender Zukunftseuphorie beschrieben werden, die in säkularisierten Gesellschaften als Trostersatz wirken. Zukunftsnarrative bzw. Technologieversprechen haben eine eminent soziale Funktion: Sie fungieren als Koordinationsmechanismen interessierter, aber heterogener Akteure, begründen kollektive Erwartungen, präzisieren Innovationsschritte und Entwicklungsrichtungen und tragen zur Genese eines sozio-technischen Feldes bei. Zukünftig sollte es daher darum gehen, verheißungsvolle Zukunftsnarrative nicht (mehr) als unplausibel abzuwerten, sondern als notwendigen Diskursbeitrag zu erkennen, zuzulassen und das damit verbundene spekulative Denken (im Sinne der Erzeugung experimentellen Wissens) im Bildungssystem zu fördern.

## **DER ÖFFENTLICHE DISKURS UM KÜNSTLICHE INTELLIGENZ FOLGT TEILS ÜBERHÖHTEN, INTERESSEGELEITETEN BILDERN UND BRAUCHT MEHR PERSPEKTIVENVIELFALT SOWIE WISSENSCHAFTLICHE FUNDIERUNG.**

Die öffentliche Wahrnehmung von KI orientiert sich nur zum Teil an faktischen technischen Neuerungen und dem Stand der Forschung. Wirkmächtig sind auch massenmedial verbreitete Visionen, die sich nicht nur in Science-Fiction-Filmen und -Literatur finden, sondern auch von populärwissenschaftlichen Akteuren strategisch platziert werden. Die damit verbundenen Narrative und Erwartungen prägen das Verständnis von KI in vielen Kontexten stark und beeinflussen auch die wissenschaftliche Reflexion. Auch wenn populäre KI-Propheten (z.B. der Singularität) und Autoren fiktionaler Formate sich mit überzogenen Zukunftsversprechen und emotionalisierter Rhetorik an der Zukunftsschau über KI beteiligen und dies auf den ersten Blick als unseriös kritisiert werden kann, muss anerkannt werden, dass Bilder, Metaphern und Narrative Systemgrenzen überspringen und teils stark handlungsleitend (z.B. im politischen System) wirken. Nur das Festhalten an Perspektivenvielfalt und die Nutzung von „scientific fiction“ (statt „science fiction“) kann für eine angemessene Diskurserneuerung jenseits von Stereotypen sorgen. Hierzu braucht es neue Dialogformate und -plattformen.

## **KÜNSTLICHE INTELLIGENZ MUSS ALS GEGENSTAND VON IDEOLOGIEKRITIK GESEHEN WERDEN.**

KI hat das Potenzial zum technologischen Fetischismus. Sie bezieht sich in ihrer Wirkweise auf mentale Konzepte und Handlungsmöglichkeiten, die wiederum als ideologisch überformt angenommen werden müssen; zum Beispiel die Idee der Überwindung des „Mängelwesens Mensch“ (Gehlen) mit Hilfe der intelligenten Maschine. KI muss daher umfassend vor dem Hintergrund der jeweiligen bestehenden gesellschaftlichen Verhältnisse (nicht nur der kapitalistischen Produktionsweise, sondern auch autoritärer Regime) verstanden werden – als ihr Verstärker oder als Möglichkeit der Überwindung. Für eine entsprechend ausführliche kritische Analyse sind sowohl die konkrete Gestaltung, Entwicklung und Nutzung von KI als auch die damit verbundenen Aushandlungs- und Aneignungsprozesse, Bilder und Imaginationen relevant. Gleichzeitig braucht es die Aufarbeitung der ikonografischen, narrativen und argumentativen Traditionslinien, die die jeweilige kulturelle Einbettung von KI (z.B. den zeitgenössischen technologischen Solutionismus oder die Akzeptanz technologisch vermittelter Herrschaft) als Kulturelement erst versteh- und einordenbar machen.

## **KÜNSTLICHE INTELLIGENZ BIRGT DIE GEFAHR, BESTEHENDE VERZERRUNGEN UND UNGLEICHHEITEN ZU REPRODUZIEREN UND ZU VERSCHÄRFEN.**

Es ist offen, ob KI ein Lösungsangebot für Inklusion und einen Ort der Neuverhandlung von Diversität darstellen kann oder ob KI bestehende Ungleichheits- und Diversitätskonstruktionen, -stereotype und -verhältnisse und die damit verbundenen Ungleichheiten bestätigt bzw. sogar noch verstärkt. Dabei besteht nicht nur die Gefahr, dass Verzerrungen in dem Datenmaterial, auf das KI-Lösungen zugreifen, reproduziert oder verschärft werden. Sondern bereits die auf unhintergehbaren, ihr qua Programmierung eingeschriebenen Selektivitäten beruhende Eigenlogik von KI (wie KI grundsätzlich funktioniert) birgt Gefahren der Ungleichheitsverstärkung und -erzeugung. Auch bleiben in der Entwicklung und im gesellschaftlichen Diskurs um KI weiterhin systematisch Gruppen unterrepräsentiert, die aktive Gestalter\*innen und Nutzer\*innen von KI darstellen. Diese Schieflage manifestiert sich bei der Programmierung von KI (coded bias), in der gesamten Wertschöpfung rund um KI-Lösungen, bei KI-Einsatzformen sowie in öffentlichen Diskursen über KI. Ein potenziell für alle Menschen offenes Tätigkeits- und Wissensfeld droht somit faktisch vor allem Chancen für

politische und ökonomische Eliten zu bieten. Insofern diese Verzerrungen im Diskurs über und in der Gestaltung von KI nicht hinreichend reflektiert werden, wird KI weiterhin exkludierend wirken und bestehende Ungleichheiten reproduzieren oder gar verstärken.

## **KÜNSTLICHE INTELLIGENZ VERSCHÄRFT DEN KAMPF UM GESELLSCHAFTLICHE MACHT- UND HERRSCHAFTSVERHÄLTNISSE.**

Aus dem Besitz großer Datenvolumina, der Verfügung über sie und ihrer Verwertung („Big Data“) erwachsen neuartige Machtpotenziale und Herrschaftstechniken, die die Machtpositionen bereits mächtiger (Wirtschafts- und/oder politischer) Akteure erheblich stärken. Zugleich werden neue Machtzentren hervorgebracht, die Wissensvorräte ansammeln und sich in ihrer Art und Weise der gesellschaftlichen und demokratischen Legitimation entziehen. Dies gilt sowohl für demokratische als auch nicht-demokratische Gesellschaften. Die Kämpfe um Macht und Herrschaft sowie um ihre Legitimationsgrundlagen werden mit hoher Wahrscheinlichkeit zunehmen und sich in gesellschaftlichen Spannungen und Konflikten niederschlagen. Zudem entstehen im Kontext von Big Data und KI vollkommen neue Machttypen (korrelative Macht, metrische Macht), die sich mit einer hohen Innovationsdynamik immer wieder neu konstellieren und die sich (zumindest derzeit) nicht mehr ausreichend mit bestehenden soziologischen Begriffen beschreiben lassen. Zu prüfen ist zukünftig auch, welche Emanzipations- und Entlastungsmöglichkeiten (z.B. im Kontext kultureller Gegenbewegungen) im positiven Sinne damit verbunden sein können.

## **WELT- UND MENSCHENBILDER**

Die Entwicklung und Nutzung von KI geht einher mit und ist eingebettet in spezifische Bilder vom Menschen und der Welt. Diese Bilder sind mehr als reine Vorstellungen davon, was KI in der Welt bewirken soll und welche Rolle der Mensch im Verhältnis zu KI spielt. Sie geben vielmehr Orientierungen vor, die reflektiert oder unreflektiert, manifest oder latent sein können und an denen sich konkretes Handeln ausrichtet. Den Imaginationen der KI-Entwicklung und -Nutzung wohnt damit eine normative Kraft inne, die soziale Prozesse der Entwicklung, Nutzung und Einbettung von KI zu prägen vermag und gesellschaftliche Strukturen verändern kann. Dies zu reflektieren, haben sich die folgenden Thesen zur Aufgabe gemacht.

### **„MACHINE LEARNING“ IST EIN TECHNISCHER FACHBEGRIFF, NUR MENSCHEN LERNEN.**

Die Vermenschlichung von Technik durch Begriffe wie „maschinelles Lernen“ oder „künstliche Intelligenz“ kann zu problematischen Fehleinschätzungen führen, die weder der Technik noch dem Menschen gerecht werden. Wenn Technik im Sprechen über sie ihres genuin technischen Charakters entkleidet wird, dann hat dies Folgen dafür, wie die Möglichkeiten, Grenzen und Risiken von KI bei ihrer Gestaltung, Entwicklung und Nutzung wahrgenommen werden. Ist im Zusammenhang mit KI von „Intelligenz“ oder „Lernen“ die Rede, dann muss immer gesehen werden, dass dahinter – im Unterschied zum Menschen – IT-Infrastrukturen und -Systeme, mathematische und statistische Verfahren sowie Algorithmen, Modelle und Programme stehen, die ihr gewisse Flexibilität und Optimierungsmöglichkeiten verschaffen. Gleichwohl spielt KI als soziale Technologie in Lehr-, Lern- und Bildungsprozessen zunehmend eine große Rolle. Hierbei gilt es zu klären, welches Rollenverständnis diesem Einsatz zugrunde liegt und welche sozialen, ethischen und epistemischen Herausforderungen damit verbunden sind.



## **DER MENSCH IST KEINE FLEXIBLE MASCHINE, DIE ES TECHNISCH NACHZUBILDEN GILT.**

So sehr der Mensch der technischen Entwicklung immer wieder als Inspirationsquelle dient und es Sinn ergibt, Technologien so zu gestalten, dass sie sich der Umgebung und den Interaktionsformen der Menschen anpassen (insbesondere mit Blick auf die Schnittstellengestaltung) – so wenig ergibt es entwicklungstechnisch Sinn, die objektiv vorhandenen Unterschiede zwischen Mensch und Technik auszublenden. Denn werden die Unterschiede nicht reflektiert und berücksichtigt, drohen KI-Lösungen nicht ihre vollen Potenziale auszuschöpfen und zugleich die Entfaltung menschlicher Potenziale einzuengen. Eine Humanisierung von KI durch eine Anpassung der Oberfläche an den Menschen führt nicht immer zu einer Humanisierung durch KI, sondern kann auch das Gegenteil bewirken. Die Risiken eines Humanisierungsbias sollten bei der KI-Entwicklung reflektiert werden.

## **DER MENSCH IST NICHT BLOSS EIN STÖR- ODER RISIKOFAKTOR, DEN ES HERAUSZUHALTEN GILT.**

Komplexe technische und soziale Zusammenhänge benötigen zunehmend soziotechnische Lösungen, um Sicherheit und Stabilität, aber auch Anpassungs- und Innovationsfähigkeit erreichen zu können. Hierzu braucht es ein Verständnis für die Stärken und Schwächen von KI-Lösungen einerseits und menschlichen Denkens und Handelns andererseits sowie eine hierauf beruhende Funktions- und Aufgabenzuteilung. Die Steuerung sozialer und ökonomischer Prozesse allein über KI erzeugt eigene Systemrisiken mit möglicherweise selteneren, dann jedoch umso gravierenderen Folgen. Die Nichtberücksichtigung dieser andersartigen Risiken kann als Kontrollbias bezeichnet werden.

## **DIE FUNKTIONEN KÜNSTLICHER INTELLIGENZ SIND KEINE ANGEEIGNETEN KOMPETENZEN.**

Wenn KI neue Funktionalitäten bietet, die vormals menschlichem Handeln vorbehalten waren, oder wenn mit KI-Tätigkeiten automatisiert wird, müssen hierfür zuerst die Bedingungen geschaffen werden. Denn auch wenn die Ziele die gleichen sind, unterscheiden sich die Lösungswege. Dies betrifft alle mit dem Prozess verbundenen Ressourcen: technische, organisationale, soziale und materielle. Insofern kann aus der Automatisierung einer Tätigkeit mittels KI nicht gefolgert werden, dass sich KI die mit der Tätigkeit vormals verbundenen menschlichen Handlungskompetenzen „angeeignet“ hat. KI wird also nicht dadurch „intelligent“, „kreativ“, „intuitiv“ oder „emotional“, dass mit ihr entsprechende Aufgaben technisch gelöst werden. Die vereinfachte Gleichsetzung einer mit KI gelösten Aufgabe mit den hierfür notwendigen menschlichen Kompetenzen kann als Äquivalentenbias bezeichnet werden.

## **MENSCHLICHE INTELLIGENZ UND SOZIALES HANDELN BEINHALTEN MEHR, ALS KÜNSTLICHE INTELLIGENZ ZU ERFASSEN UND UMZUSETZEN IN DER LAGE IST.**

KI kann so aussehen, als ob sie soziale Prozesse „verstehen“ könnte. Doch auch wenn sie menschliche Eigenschaften simuliert, geschieht dies nicht, indem sie menschliches Handeln als sinnhaftes Handeln inkorporiert, reproduziert oder nachvollzieht. Neben der Sinnhaftigkeit menschlichen Handelns ist auch Irrationalität als

Fundament und Restkategorie des Menschlichen zu beachten. KI erfasst menschliches Handeln und menschliche Intelligenz nur im Horizont bloßer Daten und Rechenoperationen. Wenn verkürzte Vorstellungen von Intelligenz und Handeln die Entwicklung und Nutzung von KI befördern, dann entsteht zwar eine hohe Innovationsdynamik, es kommt aber auch zu vielfältigen, rechtlichen und ethischen Problemen.

## GESELLSCHAFTLICHE UND SOZIALE GRENZEN

Wie jede andere Technik auch verschiebt und schafft KI Grenzen, indem sie menschliche Fähigkeiten erweitert und in bestimmten Bezügen auch übertrifft. Nicht übersehen werden darf dabei jedoch, dass in diesem Zuge wieder neue Grenzen entstehen – und zwar nicht nur der technischen Machbarkeit, sondern auch der Einbettung und Umsetzbarkeit in gesellschaftlichen und sozialen Zusammenhängen. Diese Zusammenhänge sind komplex und eigendynamisch, von Unsicherheiten und Unwägbarkeiten geprägt und dementsprechend kontingent. Es geht somit nicht so sehr darum, was KI aus technischer Perspektive noch nicht kann, sondern welche Grenzen sich mit Blick auf die Beschaffenheit der sozialen Strukturen, in denen sie eingesetzt wird, und auf die Prozesse ihrer Entwicklung und Nutzung zeigen. Diesen Fokus nehmen die folgenden Thesen ein.

### KÜNSTLICHE INTELLIGENZ ERFASST NUR DAS, WAS SICH IN ZAHLEN UMWANDELN UND SOMIT FORMALISIEREN LÄSST.

Die soziale Welt ist voller Dinge, die sich nicht vollständig oder ohne Verluste messen und objektivieren lassen – man denke nur an komplexe sinnlich-leibliche Wahrnehmungen, an situative Atmosphären oder an implizites Wissen. Auch besteht ein großer Teil physischer und sozialer Realität aus komplexen Zusammenhängen und Unbekanntem und lässt sich nicht in Gänze erfassen bzw. objektivieren. Zudem gibt es wichtige Unterschiede zwischen der Art und Weise, wie technische Sensorik als Schnittstelle der Technik zur Welt funktioniert und wie menschliche Wahrnehmung explizit und implizit abläuft. Und letzten Endes muss KI ihren Gegenstand (z.B. Töne, Bilder, Texte) rechenbar machen, um ihn zu bearbeiten. Für einen gesellschaftlich nachhaltigen Umgang mit KI ist es wichtig, einen Formalisierungsbias zu vermeiden und auch das in den Blick zu nehmen, was sich noch nicht oder nicht formalisieren, also in Daten übersetzen und rechenbar machen lässt. Dies betrifft sowohl die Erwartungen an KI und die KI-Entwicklung wie auch den KI-Einsatz.

### WENN DIE GRENZEN VON KI-LÖSUNGEN NICHT GESEHEN UND BERÜCKSICHTIGT WERDEN, ENTSTEHEN ZAHLREICHE SOZIALE RISIKEN.

KI ist eine Schlüsseltechnologie, die mit hohen Erwartungen verbunden ist, da sie stark erweiterte Möglichkeiten bietet, Komplexität zu bearbeiten. Dennoch unterliegt KI systematischen Grenzen der technischen Komplexitätsbearbeitung in sozio-technischen Umwelten auf drei Ebenen: Erstens erhöht sich die Komplexität enorm vom Schritt der KI-Entwicklung im Labor hin zur Umsetzung in der realen physischen Umwelt und der sozialen Anwendungspraxis. Zweitens erzeugt der Einsatz von KI selbst immer wieder neue Anforderungen und (Folge-)Problematiken, die nicht allein durch KI gelöst werden können (Automatisierungsdilemma). Und drittens kann KI nur das erfassen, was sich messen und in eine rechenbare Form bringen bzw. „objektivieren“ lässt. Werden diese drei Grenzen beim Einsatz von KI nicht berücksichtigt, wird die Beherrschbarkeit sozio-technischer Kontexte überschätzt und es kommt zu Folgeproblematiken (Objektivierungsfalle; vgl. auch Solutionism oder Techfix).

Aus der Perspektive eines kritischen Verbraucherschutzes ist darauf zu achten, dass im Zusammenspiel dieser Problematiken die „rationale Diskriminierung“ von Menschen, d.h. die Kopplung numerischer Werte an soziale Chancen, vermieden wird.

## **KÜNSTLICHE INTELLIGENZ ERSETZT MENSCHLICHE TÄTIGKEITEN NICHT, SONDERN TRANSFORMIERT HANDLUNGSZUSAMMENHÄNGE.**

Mit Blick auf die Folgen und den Gestaltungsbedarf von KI reicht es nicht aus, nur die von KI-Automatisierung erfassten Tätigkeitsschritte und die damit verbundenen Zielsetzungen zu betrachten. Denn wenn eine Tätigkeit mittels KI automatisiert wird, verändert sich auch die Tätigkeit an sich – mit der Folge, dass auf andere Lösungswege, Voraussetzungen und Ressourcen zurückgegriffen wird, das Resultat oftmals nicht genau das gleiche ist und andere Wege der weiteren Anschlussfähigkeit nötig sind. Insofern transformiert KI den gesamten Handlungszusammenhang. Es macht also einen (gesellschaftlichen) Unterschied, ob dieselbe Tätigkeit durch einen Menschen oder durch eine Künstliche Intelligenz durchgeführt wird.

## **KÜNSTLICHE INTELLIGENZ IST MIT EINER UNTERSCHÄTZTEN FÜLLE AN SICHTBARER UND UNSICHTBARER (VOR-, PARALLEL UND NACHGELAGERTER) ARBEIT VERBUNDEN.**

Nicht nur die IT-seitige Implementation, Integration und Wartung von KI-Systemen, sondern auch die Schulung der Nutzenden im Umgang mit den Systemen sind integraler Teil der Transformationen im Kontext von KI. Zudem sind im Zusammenhang mit der Anpassung an neue Prozesse und Arbeitsabläufe Unterstützungsbedarfe und systematische Fehler zu erwarten, die in der Folge behoben werden müssen. All dies verweist auf eine Fülle an Tätigkeiten, die erforderlich sind, um KI einzuführen, am Laufen zu halten und sozial anschlussfähig zu machen. Im Diskurs über KI wird diese unsichtbare Arbeit, die parallel, vor- oder nachgelagert zu ihrer Implementation nötig ist, jedoch kaum beachtet (vgl. Diskussionen zu Ghost Work).

## **SOZIALE ENTWICKLUNGS- UND NUTZUNGSKONTEXTE**

Als Technologie ist KI menschengemacht und von ihrer Entwicklung bis hin zu ihrer Nutzung in soziale Kontexte und Praktiken eingebettet. Zugleich werden diese Kontexte und Praktiken in der Auseinandersetzung mit KI verändert – als Technologie wirkt KI zurück auf die sozialen Prozesse und Strukturen, innerhalb derer sie entwickelt und genutzt wird. Sie tut dies auf spezifische Weise, indem sie Daten erzeugt, diese Daten menschlichem Wissen und Handeln zur Verfügung stellt und es auf diese Weise verändert. In diesem Sinne unterscheidet sich KI nicht grundlegend von anderen Technologien, ihre Folgen sind jedoch (auch mit Blick auf den Grad ihrer Auswirkung) andere. Mit dieser Wechselwirkung zwischen sozial entwickelter und genutzter KI und ihren sozialen Erzeugungs- und Nutzungszusammenhängen setzen sich die folgenden Thesen auseinander.

## **KÜNSTLICHE INTELLIGENZ FORCIERT BEI IHRER UMSETZUNG UND IN IHRER NUTZUNG EINEN ANPASSUNGSPROZESS AN DIE BEDARFE DES SYSTEMS.**

Aus technischer Sicht lassen sich soziotechnische KI-Lösungen umso leichter umsetzen, je stärker sich die soziale Umwelt an die Bedarfe des KI-Systems (d.h. an seine Logik, Kategorien, Sensorik etc.) anpasst. Auch der Nutzungsprozess lässt sich auf diese Weise einfacher und reibungsloser funktionierend modellieren. Hierin begründet sich eine permanente Tendenz zur Standardisierung des Sozialen (Formierungs- bzw. Standardisierungsbias) – sowohl in der Nutzung von bzw. Interaktion mit KI wie auch durch die Schaffung entsprechender struktureller Voraussetzungen durch gesellschaftliche Anpassungen (z.B. rechtlich-formale Regelungen, alltagspraktische Verhaltensnormen). Dies hat Auswirkungen auf Kompetenzen, Wissen, Handlungsräume, Offenheit und Vielfalt. Auch erschwert dies die individuelle wie gesellschaftliche Fähigkeit, mit permanenten Unwägbarkeiten und Unsicherheit anders umzugehen als nur zu versuchen, sie zu reduzieren. In letzter Konsequenz erhöht KI den Konformitätsdruck, erzeugt vorauseilenden Gehorsam und fördert Internalisierung vorstrukturierter Identitäten.

## **KÜNSTLICHE INTELLIGENZ STANDARDISIERT WISSEN UND VERHALTEN.**

Der Fokus der im KI-Diskurs stattfindenden Normierungsdebatte liegt vor allem auf der Normierung von KI. Das Phänomen der Normierung von Wissen und Verhalten durch KI findet hingegen deutlich weniger Erwähnung. Diese Normierung findet latent statt, indem KI in sozialen Praktiken Standardisierungseffekte von Wissen (z.B. Filterblasen) und Handeln (z.B. erwünschtes Verhalten gegenüber KI-Anwendungen) erzeugt. Im Schatten des Normierungsdiskurses entsteht somit die Gefahr weitgehend ungesehener, gleichwohl tiefgreifender und weitreichender sozialer und gesellschaftlicher Folgen von KI-Anwendungen. Umgekehrt stellen KI-Anwendungen auch selbstverständliche Wissensformen, soziale und ästhetische Kriterien in Frage, was zu einer Dynamisierung von Wissens- und Erkenntnisformen und zu neuen Gestaltungs- und Nutzungsspielräumen führen kann.

## **KÜNSTLICHE INTELLIGENZ STRUKTURIERT GESELLSCHAFTLICHE ERFAHRUNGS- UND WISSENSRÄUME AUF LATENTE WEISE.**

In welchen Anwendungen KI steckt, was KI tut und wie KI in gesellschaftlichen Zusammenhängen wirkt, ist für die Akteur\*innen, die mit ihr zu tun haben, nicht immer ersichtlich. Indem KI-Daten aus dem jeweiligen Anwendungskontext verarbeitet und von ihr generierte Daten in diesen zurückspielt, findet eine unterschwellige Strukturierung der Erfahrungs- und Wissensräume statt, in denen gesellschaftliche Akteur\*innen mit KI (inter-)agieren. Auf diese Weise schleifen sich ungesehene und unreflektierte Formen der Wissensgenese ein: Was erfahren und gewusst werden kann, hängt maßgeblich davon ab, wie KI eine soziale Praxis beeinflusst und sich zu dieser verhält.

# GESELLSCHAFTLICHE RISIKEN UND GEFAHREN

Gesellschaften sind mit ihren komplexen Dynamiken, Strukturen, Institutionen und Akteur\*innen resiliente und vulnerable Gebilde zugleich. KI entsteht und wirkt innerhalb gesellschaftlicher und sozialer Zusammenhänge und bietet auf vielfältige Weise sinnvolle und produktive Anwendungs- und Einsatzmöglichkeiten. Zugleich entstehen mit ihrer Entwicklung und Nutzung aber auch Risiken und Gefahren, die sich aus ihren technologischen Eigenheiten ergeben. Einige davon werden bereits unter dem Stichwort „Diskriminierung durch KI“ diskutiert. Die folgenden Thesen hingegen adressieren Risiken und Gefahren, die in der Nutzung und Entwicklung von KI latent angelegt sind und für Menschen und Gesellschaften negative Folgen haben können.

## KÜNSTLICHE INTELLIGENZ BIRGT DAS RISIKO, DIE PRINZIPIEN VON DEMOKRATIE UND POLITISCHER TEILHABE ZU UNTERLAUFEN.

Die Pluralität von Meinungen und Perspektiven im demokratischen Diskurs läuft Gefahr, durch KI-basierte Schlussfolgerungen ersetzt zu werden. KI verleiht den Folgen der Digitalisierung dabei eine neue Qualität: Indem die Muster vergangener Entscheidungen zur Grundlage für Entscheidungen über die Zukunft werden, verstärken sich die Pfadabhängigkeiten politischer Entscheidungen. Dies kann dazu führen, dass politische Handlungsmöglichkeiten verengt und soziale Ungleichheiten reproduziert – oder neu eingeführte perpetuiert – werden. Widerständigkeiten, moralische oder innovative Brüche und Wendungen werden damit potenziell unwahrscheinlicher und die Spielräume politischer Autonomie verringert.

## KÜNSTLICHE INTELLIGENZ BEDROHT AUTONOMIE DURCH SCHEINENTSCHEIDUNGEN IN ALGORITHMISCH VORBESTIMMTEN ENTSCHEIDUNGSRÄUMEN.

Die Komplexität des Zusammenwirkens von menschlichem Handeln und algorithmischen Aktivitäten kann zu fehlgeleiteten öffentlichen Debatten führen und in Kontrollbemühungen seitens der mit KI befassten Industrie münden. In zahlreichen Handlungsfeldern kann es zu einem Souveränitätstransfer vom Menschen zur Maschine (Algorithmus) kommen. Die Transformationsketten und Datenproduktionen der großen KI-Unternehmen sind extrem kleinteilig im Alltagshandeln mit KI-Systemen eingelassen (u.a. Tesla-Autopilot, Motorsoftware, Google-Suche, Sprachübersetzer). Es besteht die Gefahr, die „Scheinpartizipation“ im Umgang mit KI zu übersehen: Die Autonomie über die eigene Datenproduktion entpuppt sich schnell als eine Illusion, da die Entscheidungs- und Handlungsalternativen durch das KI-System oftmals bereits vorstrukturiert sind – teilweise ohne dass man dies in der Nutzung bewusst bemerken würde.

## KÜNSTLICHE INTELLIGENZ BIRGT DAS RISIKO, DURCH DIE VERENGUNG VON SINNBILDUNGSPROZESSEN ZU EINEM VERLUST GESELLSCHAFTLICHER DIVERSITÄT ZU FÜHREN.

Die Abhängigkeit von und Orientierung an KI-Daten kann – in Fortsetzung historisch bereits stattgefundener Prozesse der Medialisierung und Digitalisierung – verstärkt zu Diversitätsreduktion, Polarisierung und Monopolisierung individueller und institutioneller Sinnkonstruktionen führen. Statt sich im sozialen Raum an anderen Menschen zu orientieren, wird sozialer Sinn durch KI stochastisch aggregiert und als technisch-autoritativ ausgewiesen. Die Qualität großer Datenmengen kann kaum gesichert werden. KI-Systeme entfalten vor

diesem Hintergrund eine nur schwer zu kontrollierende handlungskonstitutive Eigenlogik. Dies kann zu einer Gefahr für soziales Handeln und Entscheiden werden.

## PRAKTISCHE GESTALTUNG

Wie alle technischen Errungenschaften haben KI-Anwendungen bestimmte Gebrauchseigenschaften und in ihnen angelegte Verwirklichungstendenzen der intendierten Folgen und Wirkungen. Blickt man näher auf die Interaktion mit und Aneignung von KI, dann zeigen sich Aspekte, die wichtig sind für ihre Entwicklung, Implementierung, Nutzung und Analyse. Die folgenden Thesen beleuchten aus diesem Blickwinkel Fragen der praktischen Gestaltung von KI.

### KÜNSTLICHE INTELLIGENZ LÄSST SICH NICHT AUF TECHNOLOGISCHE FUNKTIONEN REDUZIEREN.

Wie bei jeder anderen Technik auch vermittelt die Vorstellung von KI als Werkzeug ein Bild ihrer Einbettung und Nutzung in der sozialen Praxis, das sich empirisch nicht halten lässt. Die Aneignung und Nutzung von KI findet immer in komplexen und vielschichtigen Praxiszusammenhängen statt, die einer eigenen Logik folgen. Diese Logik wird durch die Wechselwirkungen von KI-Systemen mit den sozialen Zusammenhängen, in denen sie genutzt werden, jedoch irritiert und verändert. Auf diese Weise entstehen neue und vielfältige Prozesse und Strukturen im Umgang mit KI, die nicht planbar sind – auch dort nicht, wo es sich um hochstandardisierte und -kontrollierte Umgebungen handelt. Dies ist kein Makel einer „unfertigen“ oder „unreinen“ Praxis, sondern ein notwendiger Prozess der sozialen Aneignung und Nutzbarmachung von KI.

### KÜNSTLICHE INTELLIGENZ AM ARBEITSPLATZ BRAUCHT DEN BLICK AUF DIE ORGANISATIONALEN PRAKTIKEN.

Die Debatte um KI in der Arbeitswelt fokussiert zu eng auf die Mensch-Maschine-Schnittstelle und negiert die notwendige Einbindung in organisationale Praktiken. KI-Systeme müssen nicht nur in ihrer Implementierungsphase, sondern über ihre Nutzungsdauer hinweg entwickelt und gestaltet werden. Konkrete Gestaltungsanforderungen an KI am Arbeitsplatz aus Sicht der Beschäftigten sind in groß angelegten, repräsentativen Studien ein bislang unterrepräsentiertes Thema.

### KÜNSTLICHE INTELLIGENZ BIETET MEHR MÖGLICHKEITEN, MENSCHLICHES HANDELN ZU ERGÄNZEN ALS ZU ERSETZEN.

Nimmt man die Komplexität sozialer Praxis und menschlichen Handelns umfassend in den Blick (weit über Kreativität und Empathie hinaus), verschiebt sich der Fokus von der Substitutionsproblematik (dem Ersatz menschlicher Arbeitskraft durch Technik) zur Frage der Transformation des Mensch-Technik-Verhältnisses und dessen permanenter Neugestaltung. So zeigt sich bislang auch kein wesentlicher Effekt auf Beschäftigung oder auf die Steigerung der Produktivität; jedoch auf die Verteilung von Ressourcen. Arbeit bzw. Wertschöpfung ist ein sehr komplexer, sich permanent wandelnder, aber auch immer wieder stabilisierender sozialer Zusammenhang, in den KI (wie andere Technologien auch) eingeeht wird.

## **ZIEL DER KI-ENTWICKLUNG UND DES KI-EINSATZES SOLLTE EIN KOMPLEMENTÄRES UND KOEVOLUTIONÄRES VERHÄLTNIS VON MENSCH UND KÜNSTLICHER INTELLIGENZ SEIN.**

Für ein sozial und gesellschaftlich nachhaltiges Mensch-KI-Verhältnis ist es essenziell, dass die technische und die soziale bzw. gesellschaftliche Entwicklung eine sich (ggf. wechselseitig) bestärkende Dynamik gewinnen. Heruntergebrochen bedeutet das, dass KI – neben arbeitsorganisatorischen und Bildungsmaßnahmen – so gestaltet werden muss, dass sie Menschen nicht dequalifiziert und entmächtigt, sondern durch den Nutzungsprozess in ihrer Entwicklung fördert. Ein ideales Gestaltungsziel wäre eine sich wechselseitig bestärkende, koevolutionäre Mensch-KI-Interaktion, die auf einem komplementären Zusammenspiel von Mensch und KI beruht.

## **KÜNSTLICHE INTELLIGENZ VERLIERT IHRE MYSTIK, WENN ENTSPRECHENDE ANWENDUNGEN ALS ALLTAGSTECHNOLOGIE ANGEEIGNET WERDEN.**

Die tatsächlichen Potenziale und Auswirkungen von KI sind in der praktischen Anwendung zu interpretieren und zu bewerten. Was KI kann oder nicht kann, aber auch wie sich KI als technisches Prinzip der Datenerfassung und -verarbeitung, als technisches Artefakt, das bedient und gesteuert werden muss, und als Erwartung an eine neue Technologie auf eine soziale Praxis auswirkt, erweist sich im Moment ihrer konkreten Nutzung. In den ex-ante-Zuschreibungen an aktuelle und künftige KI lassen sich oftmals überhöhte Erwartungen (bis hin zu Utopien und Dystopien) identifizieren, die sich in der Praxis aufgrund der ihr eigenen Komplexität nicht bewahrheiten. Zugleich zeigen sich auch erst in der Praxis die langfristigen sozialen Folgen und Wirkungen von KI. Beides führt dazu, dass die individuellen, organisationalen und gesellschaftlichen Kompetenzen, KI-basierte Technologien im Rahmen eines technisch induzierten Wandels anzueignen, umzuformen und praktisch einzuhegen (z.B. durch Anpassungsleistungen), oftmals unterschätzt werden.

## **GESELLSCHAFTSPOLITISCHE GESTALTUNG**

Innovatorische Neuerungen bilden Fortschritt und Entwicklung ab, schaffen etwas Neues, bisher nicht Dage-wesenes und können kreative Ideen und alternative Lösungswege fruchtbar machen. KI kann als eine solche Innovation gesehen werden, die stetig verändert und weiter entfaltet wird und mit deren Hilfe sich neue Wege, Ziele, Funktionen und Inhalte erschließen lassen. Mit den Freiheiten in diesem Geschehen entsteht gleichzeitig Verantwortung, die auch auf gesellschaftlicher Ebene wahrgenommen werden muss. Dieses Thema haben die folgenden Thesen zum Inhalt.

## **KÜNSTLICHE INTELLIGENZ ERFORDERT EINEN NEUEN MODUS GESELLSCHAFTLICHER INNOVATION.**

KI erscheint im weiten Bereich gesellschaftlicher Innovation als „promising technology“ und unterliegt daher mit Blick auf ihre (Weiter-)Entwicklung konkreten Risiken der Über- bzw. Unterschätzung ihrer Potenziale. Problematisch werden Erwartungen und/oder Befürchtungen hinsichtlich KI, wenn sie von konkreten Anwendungsszenarien abstrahieren oder reale Widersprüche und Grenzen ausblenden. Daher bedarf es eines

neuen Innovationsmodus (u.a. im Sinne von Transdisziplinarität, Dynamik und fluider Interaktion), an den sich bestehende Forschungs- und Innovationsstrukturen anpassen sollten.

## **KÜNSTLICHE INTELLIGENZ BEDARF EINER INTEGRATIVEN UND INNOVATIONSFREUNDLICHEN NORMIERUNG.**

Die Normierung von KI-Systemen wird von allen Akteur\*innengruppen begrüßt – allerdings in ganz unterschiedlicher Form, Tiefe und Ausrichtung. Zugleich erweist sie sich im Spannungsfeld zwischen einer zu umfassenden Einbeziehung KI-verwandter Technologien einerseits und einer ungenügenden Berücksichtigung denkbarer KI-Anwendungen andererseits als schwierig. Die Normierung von KI muss daher unterschiedliche Perspektiven einbinden und möglichst viele Anwendungen integrieren, ohne dabei innovationshemmend zu wirken.

## **KÜNSTLICHE INTELLIGENZ BENÖTIGT NEUE FORMEN ETHISCHER VERANTWORTUNG UND RECHTLICHER HAFTUNG.**

Die neuen Formen der Haftung und Verantwortung sollten sich nach den Geltungsbedingungen unterschiedlicher sozio-digitaler Institutionen richten. Deswegen wird die folgende Differenzierung von Haftungsregimes vorgeschlagen. Für die Institution der digitalen Assistenz dürfte es erforderlich sein, Algorithmen als handlungsfähige Akteure zu definieren, um die Prinzipal-Agenten-Haftung zu ermöglichen. Für die Institution digitaler Hybridität wären die Handlungen der emergenten Mensch-Algorithmen-Assoziation selbst zuzurechnen. Für die Institution der digitalen Interkonnektivität (im Sinne eines eigenständigen Zusammenwirkens mehrerer KI-Systeme) sollte die Verantwortung für Schäden, die auf kollektives Maschinenverhalten zurückgehen, vergesellschaftet und durch kollektive Fondslösungen abgesichert werden. Während die ersten beiden Haftungsregimes durch die Gerichte eingeführt werden können, ist für das dritte Haftungsregime die Gesetzgebung gefordert.

# **TECHNIKREFLEXION**

Um ein umfassendes Bild von KI als sozialer Technologie zu zeichnen, braucht es eine tiefgehende Reflexion, die ihre gesellschaftlichen und sozialen Zusammenhänge, Verhältnisse, Folgen, Risiken und Notwendigkeiten aus unterschiedlichen Blickwinkeln bedenkt und auslotet. Insbesondere geistes-, sozial- und kulturwissenschaftliche Perspektiven kommen im KI-Diskurs bisher wenig zum Zuge, bieten aber wertvolle Erkenntnisse und eigenes Wissen. Darum geht es in den folgenden Thesen.

## **KÜNSTLICHE INTELLIGENZ BEDARF EINER GESELLSCHAFTSTHEORETISCHEN RAHMUNG.**

KI kann nicht als wertneutral angenommen, analysiert und verstanden werden. Vielmehr ist KI als „gesellschaftliches Potenzial“ zu begreifen, dessen Hebung innerhalb realer gesellschaftlicher Perspektiven und Bedeutungszusammenhänge permanent verhandelt wird. Dies geschieht derzeit zwischen verschiedenen Polen (technologisch-funktional vs. gesellschaftlich-sozial, Ent- vs. Wiederverzauberung, utopisch vs. dystopisch),



die jedoch nicht als Gegensätze, sondern als Horizonte dialektisch verschränkter Interpretationsräume verstanden werden können.

## **KÜNSTLICHE INTELLIGENZ FORDERT ZU EINER REFLEXION GESELLSCHAFTLICHER, SOZIALER UND ANTHROPOLOGISCHER KATEGORIEN AUF.**

Geistes-, sozial- und kulturwissenschaftliche Konzepte beziehen sich auf Vorstellungen menschlichen Handelns, Imaginierens und Wissens. In der Regel werden die anthropologischen Grundlagen der jeweiligen Wissenschaften von diesen nicht weiter problematisiert. Wenn Begriffe wie „Akteur\*in“, „Entscheidung“, „Kommunikation“ und „Interaktion“ oder auch wie „Intelligenz“, „Lernen“, „Autonomie“ oder „Wissen“ auf KI bezogen werden, geschieht dies daher meist ohne Klärung, auf welchen Grundlagen dies beruht oder beruhen kann. KI kann in diesem Sinne als Anlass und Inspiration gesehen werden, die grundlegenden Kategorien, Begriffe und Konzepte der Geistes-, Sozial- und Kulturwissenschaften in einer soziotechnischen Gesellschaft zu reflektieren und zu schärfen.

## **KI-FORSCHUNG KANN VON GEISTES-, SOZIAL- UND KULTURWISSENSCHAFTLICHEN BEGRIFFEN UND KONZEPTEN STARK PROFITIEREN.**

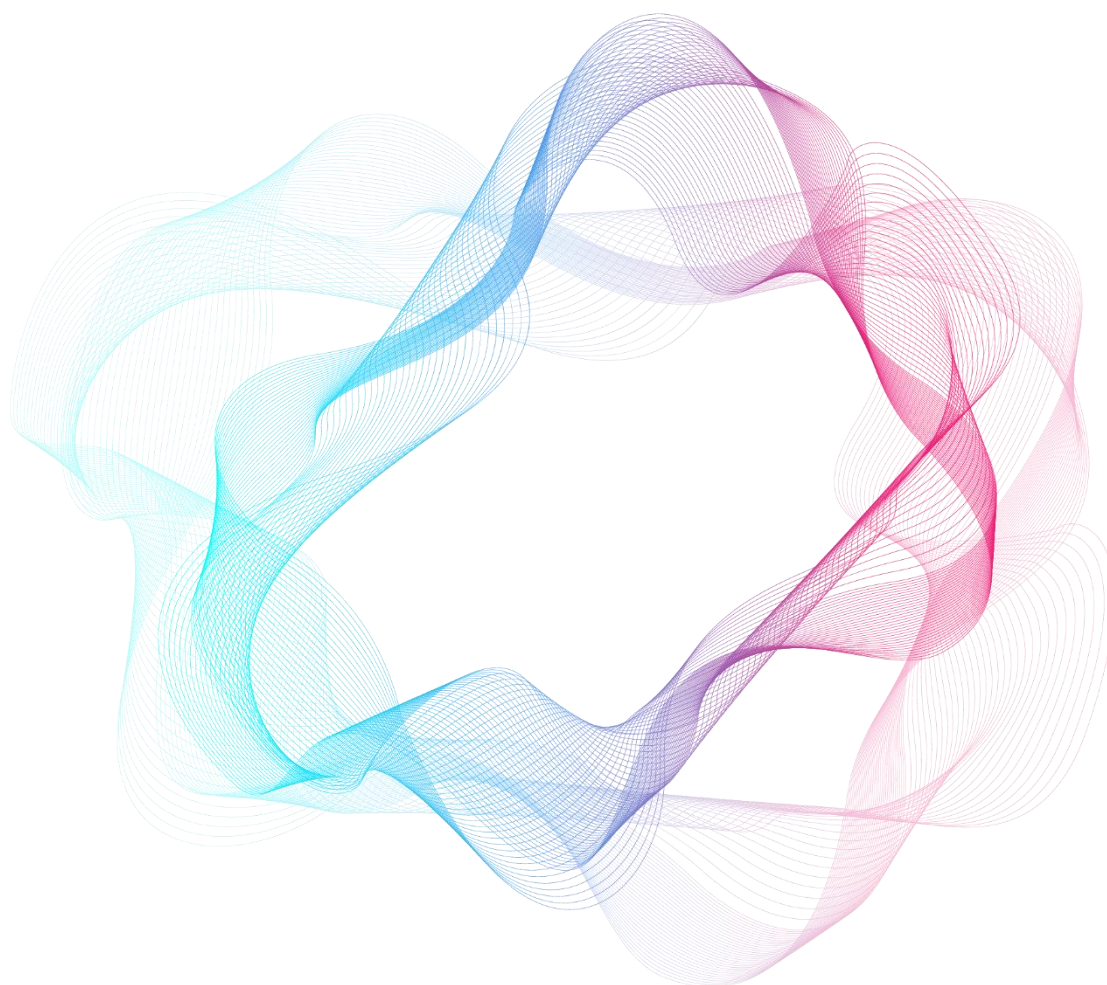
Wie bereits die Sozionik-Debatten der 1990er Jahre gezeigt haben, werden fundiert ausgearbeitete und verhandelte Begriffe aus der Philosophie und den Geistes- und Sozialwissenschaften auch für die Informatik und die Technikwissenschaften relevant (z.B. Interaktion, Kommunikation, Handlung, Bewusstsein). Dies gilt umso mehr für die Entwicklung und Nutzung von KI-Systemen. Eine unreflektierte Nutzung mensch-, sozial- und gesellschaftsbezogener Begriffe in den Technikwissenschaften und der Informatik läuft Gefahr, den Mehrwert begrifflicher Differenzierungen und die damit verbundenen Erkenntnisse außen vor zu lassen und so bereits erreichte Erkenntnisse wieder folgenreich zu unterschreiten.

## **KÜNSTLICHE INTELLIGENZ BENÖTIGT GEISTES-, SOZIAL- UND KULTURWISSENSCHAFTLICHES METHODENWISSEN.**

Geistes-, sozial- und kulturwissenschaftliche Perspektiven brechen mit der Vorstellung von KI als einem monolithischen Objekt, indem sie die komplexen gesellschaftlichen Anwendungs- und Nutzungsweisen historisch gewachsener KI-Systeme nachzeichnen und erhellen. Ansprüche an Gütekriterien für Datensätze, Reflexivität, Bias-Sensitivität sowie Erfahrung im Umgang mit unreinen und unvollständigen Daten oder Uneindeutigkeit (z.B. Korrelation vs. Kausalität) sind für das Verständnis von KI als Datenverarbeitungsmaschinen in einer durch und durch sozialen (Um-)Welt unerlässlich. Sozialwissenschaftliche Methodologie ist multiparadigmatisch: Sie enthält mannigfaltige Perspektiven und Ansätze, die für verschiedene Fälle und Forschungsinteressen ausdifferenziert sind. Dadurch sind sie besonders für den Umgang mit der Heterogenität und Partikularität von Imaginationen, Systemen und Anwendungen von KI in unseren Lebens- und Alltagswelten geeignet.

## KÜNSTLICHE INTELLIGENZ MACHT EINE NEUAUSRICHTUNG DES BILDUNGSSYSTEMS NOTWENDIG.

Gerade weil im Kontext von KI Unsicherheiten über zentrale anthropologische (Leben, Intelligenz), soziale (Rollen, Erwartungen) und ethische Begriffe zunehmen, sollte das Bildungssystem auf allen Ebenen KI als Kulturphänomen und Kulturtechnik thematisieren, anstatt KI durch Vermittlung von Informatikkenntnissen „beherrschbar“ zu machen. Statt nur instrumentelles Wissen (*know how*) über KI zu vermitteln (Technikkompetenz, digital literacy), braucht es mehr reflexives Wissen (*know why*), darüber hinaus aber vor allem die Vermittlung transformativen Wissens (*know what*), das sich auf Gestaltungsoptionen und Wunschbilder zukünftiger Gesellschaften bezieht und den latent politischen Gehalt von KI betont. Als Querschnittsthema sollte KI daher in der Breite in das Bildungssystem integriert werden.



# AUSBLICK

Künstliche Intelligenz wird aktuell gesellschaftlich geformt. Dabei ist ein breites Spektrum unterschiedlicher Entwicklungspfade denkbar und möglich. Mit dem Blick auf mögliche, eher latente soziale Folgen von Künstlicher Intelligenz wurden hier Risikofelder identifiziert, die es im Blick zu halten und noch weiter zu eruieren gilt. Dahinter steht die Idee, dass KI in ihren sozialen Chancen und Risiken gesellschaftlich gestaltet werden kann. Ein aufgeklärter Blick auf die Gefahren ermöglicht erst die sozial nachhaltige Nutzung und den Ausbau der Kapazitäten von KI. Insofern geht es nicht um ein Ja oder Nein zur Entwicklung und zum Einsatz von KI, sondern vielmehr um Baustellen für eine sozial nachhaltige Gestaltung.

Das Thesenpapier ist anschlussfähig an aktuelle Diskussionen rund um eine humane Gestaltung von KI und kann unter anderem zu folgenden Punkten beitragen:

Zunächst zeigen die Thesen komplexe und vor allem in ihrer gesellschaftlichen Wirkung latente soziale Risiken auf, die bislang zu sehr nur an der Oberfläche diskutiert werden. So sollte die zu eng geführte Diskussion um die Problematik des Daten- bzw. Diskriminierungsbias auf Verzerrungen und Engführungen bzw. KI-Selektivitäten insgesamt bezogen werden und dabei nicht nur das Datenmaterial in den Blick nehmen, sondern auch KI-Verfahren (z.B. Mustererkennung) selbst und Einhegungspraxen. Die Aufrechterhaltung von Vielfalt und Offenheit bzw. Unbestimmtheit scheint eine wesentliche Herausforderung im gesellschaftlichen Umgang mit KI-Technologien zu sein. Insgesamt braucht es ein differenzierteres Verständnis bzw. eine systematische Erforschung der Stärken und Schwächen von KI als Technik bzw. Methode, um KI als Instrument der Erfassung und Bearbeitung von Realität zu verstehen und nachhaltig gesellschaftlich einbetten zu können. Dies ist auch eine Voraussetzung für ein differenzierteres Verständnis des Zusammenwirkens von KI, Mensch und Gesellschaft und Entwicklungsdynamiken. Eine noch ausstehende Grundlage wäre die eingehendere Erforschung und Fassung der bestehenden Unbestimmtheit sowie der meist bildhaften und uneigentlichen Begrifflichkeit KI sowie deren Kopplungen mit kulturellen Schemata, Narrativen, Mythen und Imaginationen, ohne die KI-Phänomene inklusive ihrer Diskurse in Öffentlichkeit und Wissenschaft nicht verstanden werden können.

KI macht Komplexität auf neue Weise bearbeitbar. Damit verbunden sind spezifische KI-Selektivitäten, neue Handlungspraxen sowie gesellschaftliche Aushandlungs- und Einbettungsprozesse, die neue Optionen eröffnen, aber auch andere verschließen. Wie KI also insgesamt oder im Einzelfall gesellschaftlich wirkt, wird konkret gestaltet. Für eine humane Gestaltung ist die Erklärungsfähigkeit von KI ein wichtiger Aspekt der Aufrechterhaltung von Handlungsmacht und Entscheidungskompetenzen in der KI-Nutzung. Jedoch muss wesentlich breiter auf verschiedene Aneignungschancen und Befähigungsmöglichkeiten (wie z.B. Erfahrungs- und Lernförderlichkeit) bzw. Dequalifizierungsgefahren und Entfremdungspotenziale im praktischen Nutzungsprozess geblickt werden. So sind auch Vertrauen und Vertrauenswürdigkeit komplexe soziale Phänomene, die weit über Zuverlässigkeit, Robustheit, Vorhersehbarkeit etc. hinausreichen. Die Überlegungen zu einer vertrauenswürdigen KI können von den hier aufgestellten Thesen profitieren, wenn sie sich mit den aufgezeigten Risikofeldern explizit auseinandersetzen. Auch aktuelle Einordnungsversuche verschiedener KI-Systeme nach Kritikalitätsstufen sollten noch einmal entlang der hier dargelegten latenten sozialen Implikationen überprüft werden.

Insgesamt braucht es eine fallspezifische und multidisziplinäre Forschung, um die Entwicklungen rund um KI in ihren jeweiligen Besonderheiten erfassen und öffentlichen Debatten und Willensbildungsprozessen zugänglich zu machen. Insbesondere die Normierung individuellen wie gesellschaftlichen Handelns als Folge der Diskussion über KI sowie deren Existenz und Einsatz sollten verstärkt in den Blick sowohl geistes-, sozial- und kulturwissenschaftlicher Studien, des Rechts wie auch der politischen Entscheidungsfindung genommen werden. Notwendig ist nicht zuletzt eine Innovationspolitik, die vor den hier aufgezeigten Zusammenhängen ihre eigene steuernde Wirkung reflektiert und ihre bisherigen KI-orientierten Schwerpunkte neu ausrichtet in Richtung: einer Ausweitung des Wissenstransfers und die Förderung von Transdisziplinarität, ein erweitertes und auf KI ausgerichtetes Spektrum von grundlagen- und anwendungsorientierter Forschung, Technologieoffenheit und Anschluss an ein breites Methodenverständnis, Transparenz und Erklärbarkeit sowie die Förderung ethisch orientierter Machine-Learning-Entwicklungen, die Unterstützung flexibler innovativer Ecosysteme und der Entwurf neuer agiler Modelle der Forschungsförderung.

# IMPRESSUM

**HERAUSGEBER:**

Institut für Sozialwissenschaftliche Forschung e.V. – ISF München  
Dr. Michael Heinlein & Dr. Norbert Huchler  
Jakob-Klar-Straße 9, 80796 München  
[www.isf-muenchen.de](http://www.isf-muenchen.de)

**VERANTWORTLICH FÜR DEN INHALT:**

Dr. Michael Heinlein & Dr. Norbert Huchler, ISF München

**LAYOUT:**

Torsten Royère, ISF München

Juli 2022

„KI – Mensch – Gesellschaft. Den Wandel des Mensch-Technik-Verhältnisses durch Künstliche Intelligenz ganzheitlich verstehen und bewerten“ (KI.Me.Ge) ist ein vom BMBF gefördertes Projekt. Beteiligt waren die folgenden Personen:

Dr. Anna Beckers, Universität Maastricht

Dr. Bruno Gransche, Karlsruher Institut für Technologie (KIT)

Dr. Michael Heinlein, Institut für Sozialwissenschaftliche Forschung e.V. (ISF München)

Prof. Dr. i.R. Hartmut Hirsch-Kreinsen, TU Dortmund/Sozialforschungsstelle

Prof. Dr. Jeanette Hofmann, Alexander von Humboldt Institut für Internet und Gesellschaft (HIIG)

Dr. Norbert Huchler, Institut für Sozialwissenschaftliche Forschung e.V. (ISF München)

Dr. Karin Hutflötz, Katholische Universität Eichstätt-Ingolstadt

Dr. Clara Iglesias Keller, Wissenschaftszentrum Berlin für Sozialforschung (WZB)

Prof. Dr. Peter Imbusch, Bergische Universität Wuppertal

Peter Kahlert (MA), Europa-Universität Viadrina

Nadine Kleine (MA), Ostbayerische Technische Hochschule Regensburg

PD Dr. Reinhard Kreissl, Wiener Zentrum für Sozialwissenschaftliche Sicherheitsforschung (VICESSE)

Prof. Dr. Arne Manzeschke, Evangelische Hochschule Nürnberg

Lea-Sophie Natter (MA), Bergische Universität Wuppertal

Prof. Dr. Jan-Hendrik Passoth, Europa-Universität Viadrina

Prof. Dr. Sabine Pfeiffer, Friedrich-Alexander-Universität Erlangen-Nürnberg

Dr. Heike Raab, Independent Scholar

Dr. Valentin Rauer, Türkisch-Deutsche Universität Istanbul

Samuel Rieger (MA), Institut für Sozialwissenschaftliche Forschung e.V. (ISF München)

Dr. Irmhild Rogalla, Institut für praktische Interdisziplinarität

Prof. Dr. Stefan Selke, Hochschule Furtwangen

Dr. Joris Steg, Bergische Universität Wuppertal

Prof. Dr. Jur. Dr. h.c. mult. Gunther Teubner, Goethe Universität Frankfurt am Main

Roger von Laufenberg, PhD, Wiener Zentrum für Sozialwissenschaftliche Sicherheitsforschung (VICESSE)

Prof. Dr. phil. habil. Karsten Weber, Ostbayerische Technische Hochschule Regensburg

Regina Wittal (MA), Institut für Sozialwissenschaftliche Forschung e.V. (ISF München)

GEFÖRDERT VOM