

**IMAGO · ATELIER**

Whitepaper 22 · Grenzthema

# Wer die Sprache kontrolliert, kontrolliert die Realität

## *Über semantische Hegemonie in der KI-Debatte*

**Silvia Mann-Kundt**Systems Thinker · Imago Atelier  
complexity-organizer.com · 2026**Abstract**

„Intelligenz“, „Autonomie“, „Sicherheit“, „Risiko“, „Ausrichtung“ — die zentralen Begriffe der KI-Debatte klingen neutral. Sie sind es nicht. Wer diese Begriffe definiert, setzt die Grenzen des Denkbaren: was als Problem gilt, wer als Experte anerkannt wird, welche Lösungen als legitim erscheinen. In der KI-Debatte sind es fast ausschließlich nordamerikanische Technologiekonzerne, anglophone Policy-Institute und eine kleine Gruppe von „AI-Safety“-Organisationen, die das Vokabular prägen — und damit die Bedingungen setzen, unter denen alle anderen über KI sprechen. Dieses Whitepaper analysiert die semantische Hegemonie der KI-Debatte mit den Werkzeugen der Framing-Theorie (Lakoff/Johnson), der Machtanalytik (Foucault, Lukes), der Sprachphilosophie (Wittgenstein) und der Systemtheorie (Luhmann). Es zeigt: Die wichtigste Intervention im KI-Diskurs ist keine technische und keine regulatorische. Es ist eine sprachliche. Wer die Begriffe nicht analysiert, die er benutzt, reproduziert die Machtstruktur, die er glaubt zu kommentieren. Das ist ein Grenzthema — weil es unbequem ist. Weil es den Boden des Diskurses in Frage stellt, nicht seine Oberfläche.

**Schlüsselwörter:** Framing · Semantische Hegemonie · KI-Diskurs · Lakoff · Foucault · Lukes · Wittgenstein · Macht · AI Safety · Sprachkritik · Technologiepolitik

## 01 · EINLEITUNG

## Der unsichtbare Rahmen

Es gibt Debatten, die verloren sind, bevor sie begonnen haben — nicht weil die Argumente schwach sind, sondern weil der Rahmen, in dem sie geführt werden, bereits die Antwort voraussetzt. Die KI-Debatte ist eine solche Debatte.

Wenn Technologiekonzerne von „AI Safety“ sprechen, rahmen sie KI als eine Kraft, deren primäre Gefahr von der KI selbst ausgeht — nicht von den Unternehmen, die sie entwickeln, nicht von den Marktstrukturen, in denen sie eingebettet ist, nicht von den geopolitischen Interessen, die ihre Verbreitung prägen. Wenn Regulatoren von „High-Risk AI“ sprechen, übernehmen sie eine Risikodefinition, die primär aus dem Blickwinkel der Systementwickler formuliert wurde. Wenn Journalisten von „kognitiver KI“ oder „lernenden Maschinen“ schreiben, anthropomorphisieren sie technische Prozesse auf eine Weise, die Entscheidungszurechnung strukturell verschleiert.

Keine dieser Formulierungen ist neutral. Jede ist ein Framing — eine Rahmung, die bestimmte Wahrnehmungen ermöglicht und andere verhindert, bestimmte Fragen stellt und andere ausschließt, bestimmten Akteuren Deutungshoheit verleiht und andere marginalisiert.

George Lakoff und Mark Johnson haben in „Metaphors We Live By“ (1980) gezeigt, dass Denken primär metaphorisch strukturiert ist — dass wir die Welt durch konzeptuelle Rahmen verstehen, die nicht neutral sind, sondern Wertungen, Annahmen und Machtverhältnisse einschreiben. Michel Foucault hat in seiner Diskursanalyse gezeigt, wie Wissen und Macht untrennbar verbunden sind: Wer bestimmt, was als Wahrheit gilt, bestimmt, wer im Diskurs sprechen darf. Steven Lukes hat in seiner dreidimensionalen Machttheorie gezeigt, dass die subtilste Machtausübung nicht Zwang ist, nicht Einfluss, sondern die Formung von Präferenzen: wenn Menschen wollen, was die Mächtigen wollen, weil sie es für selbstverständlich halten.

In der KI-Debatte ist die dritte Dimension von Lukes' Macht allgegenwärtig. Und sie ist nahezu unsichtbar.

**THESE I** *Die wichtigste Intervention im KI-Diskurs ist keine technische und keine regulatorische. Es ist eine sprachliche. Wer die Begriffe nicht analysiert, die er benutzt, reproduziert die Machtstruktur, die er glaubt zu kommentieren.*

## 02 · THEORETISCHE GRUNDLAGE

## Framing, Macht und Sprache: Vier Zugänge

### 2.1 Lakoff & Johnson: Das konzeptuelle Framing

George Lakoffs und Mark Johnsons Grundlagenwerk „Metaphors We Live By“ (1980) hat eine These entwickelt, die für die Analyse des KI-Diskurses unmittelbar relevant ist: Die Begriffe, durch die wir ein Phänomen verstehen, sind keine neutralen Bezeichnungen. Sie sind konzeptuelle Rahmen, die strukturieren, was wir wahrnehmen können.

Lakoffs späteres Werk „Don't Think of an Elephant!“ (2004) übersetzt diesen Befund in eine politische Theorie: Wer den Rahmen setzt, gewinnt die Debatte — unabhängig vom Inhalt der Argumente. Wer sich auf dem Boden des gegnerischen Framings bewegt und diesen Rahmen widerlegt, hat den Rahmen bereits aktiviert. Der Satz „KI ist kein Risiko“ setzt das Framing „KI als Risiko“ voraus. Er widerlegt es nicht. Er stärkt es.

Für die KI-Debatte bedeutet das: Wer über KI diskutiert, ohne die Rahmen zu analysieren, in denen er denkt, reproduziert diese Rahmen — auch wenn er ihnen widerspricht. Das ist keine rhetorische Spitzfindigkeit. Es ist ein strukturelles Phänomen, das erklärt, warum bestimmte Debatten über KI bestimmte Antworten nahezu unabdingbar machen.

### 2.2 Foucault: Diskurs als Machtregime

Michel Foucaults Diskursanalyse, insbesondere in „Die Ordnung des Diskurses“ (1970) und „Archäologie des Wissens“ (1969), entwickelt eine Theorie, die über Lakoff hinausgeht: Diskurse sind nicht nur Rahmungen, sie sind Machtregime. Sie regeln nicht nur, wie über ein Thema gesprochen wird, sondern wer sprechen darf, welche Aussagen als wahr gelten, welche Positionen als legitim anerkannt werden und welche marginalisiert, pathologisiert oder ignoriert werden.

In Foucaults Perspektive ist der KI-Diskurs ein Wissensregime, das bestimmte Experten produziert (AI-Forscher, Safety-Spezialisten, Data Scientists) und andere ausschließt (Philosophen, Sozialwissenschaftler, Gewerkschaften, Bürgerinnen), das bestimmte Probleme sichtbar macht (technische Sicherheit, algorithmische Effizienz) und andere unsichtbar (Machtkonzentration, demokratische Kontrolle, ökologische Kosten) — und das diese Sichtbarkeitsordnung als natürlich, als evident erscheinen lässt.

„Das Diskursive ist nicht das Supplement, es ist die Grundbedingung der Möglichkeit von Wirklichkeit.“ — Michel Foucault, Die Ordnung des Diskurses, 1970

### 2.3 Lukes: Die drei Dimensionen von Macht

Steven Lukes' „Power: A Radical View“ (1974) unterscheidet drei Dimensionen von Macht, von denen die dritte für die KI-Debatte die relevanteste ist. Die erste Dimension ist sichtbarer Konflikt: A setzt sich gegen B durch. Die zweite Dimension ist Agenda-Setting: A verhindert,

dass bestimmte Themen überhaupt auf die Tagesordnung kommen. Die dritte Dimension ist die subtilste: A formt die Präferenzen und Wahrnehmungen von B, so dass B nicht einmal bemerkt, dass seine Interessen verletzt werden – weil er die Situation als natürlich und unveränderlich erlebt.

Im KI-Diskurs operiert Lukes' dritte Dimension in Reinform. Wenn Gesellschaften KI-Regulierung primär als technische Compliance-Aufgabe verstehen (und nicht als politische Gestaltungsaufgabe), wenn sie „AI Safety“ als Schutz vor dem Systemrisiko lesen (und nicht als Schutz vor den Entwicklern), wenn sie „Pflicht zur Innovation“ als Selbstverständlichkeit akzeptieren (und nicht als interessen geleitetes Narrativ), dann operiert Lukes' dritte Dimension erfolgreich: Die Macht ist unsichtbar, weil ihre Ergebnisse als natürlich erscheinen.

#### 2.4 Wittgenstein: Bedeutung als Gebrauch

Ludwig Wittgensteins späte Philosophie („Philosophische Untersuchungen“, 1953) hat eine These formuliert, die für die Sprachkritik des KI-Diskurses fundamental ist: „Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache“ (§43). Nicht die Referenz, nicht die innere Repräsentation, sondern der soziale Gebrauch in konkreten Praxiskontexten.

Das bedeutet für die Analyse: „Intelligenz“ bedeutet in der KI-Debatte nicht das, was Psychologen unter Intelligenz verstehen, nicht das, was Neurowissenschaftler als Intelligenz beschreiben, und nicht das, was Philosophen als Vernunft konzeptualisiert haben. Es bedeutet das, was Technologieunternehmen in konkreten Deployment-Kontexten unter „Intelligenz“ verstehen und wie dieser Gebrauch sich sedimentiert hat. Die Bedeutung ist nicht intrinsisch. Sie ist durch Machtpraktiken geformt.

#### 2.5 Luhmann: Semantik und Gesellschaft

Niklas Luhmanns Unterscheidung zwischen Semantik und Kommunikation ist für die Diskursanalyse des KI-Feldes instruktiv. Luhmann beschreibt Semantik als den gesellschaftlichen Vorrat an vorgefertigten Sinnformen – als das Repertoire von Begriffen, Konzepten und Unterscheidungen, auf das Kommunikation zurückgreifen kann (Luhmann 1980: 19 ff.). Semantiken sind historisch geronnene Reduktionen von Komplexität: Sie ermöglichen Kommunikation, indem sie vorstrukturieren, was gesagt werden kann.

Die Semantik des KI-Diskurses – mit ihren Leitbegriffen Intelligenz, Autonomie, Sicherheit, Alignment, Verantwortung – ist eine historisch spezifische, machtgeladene Reduktion von Komplexität. Sie ermöglicht bestimmte Kommunikationen und verhindert andere. Wer die Semantik analysiert, analysiert die Bedingungen möglicher Kommunikation über KI.

**THESE II** *Die Semantik des KI-Diskurses ist kein neutrales Vokabular für die Beschreibung einer technischen Realität. Sie ist eine historisch geformte, interessen geleitete Sinnstruktur, die bestimmte Fragen stellt und andere verhindert. Das zu erkennen ist keine Verschwörungstheorie. Es ist analytische Grundkompetenz.*

## 03 · BEGRIFFSANALYSE

## Vier Schlüsselworte der KI-Debatte — und ihre verborgenen Rahmungen

Die folgende Analyse untersucht vier zentrale Begriffe des KI-Diskurses auf ihre Framing-Funktion. Die Methode ist die Anwendung des dialektischen Prüfrasters: Woher stammt die Bedeutung dieses Begriffs? Welches Interesse hat die definierende Quelle? Unter welchen Voraussetzungen gilt die Bedeutungszuschreibung? Und welche Alternativen werden durch den Begriff ausgeschlossen?

### „Künstliche Intelligenz“

**Framing:** Maschinen, die denken wie Menschen — oder es bald könnten.

**Wer definiert:** *Informatiker der 1950er Jahre (McCarthy 1956), übernommen von Tech-Industrie als Marketing-Label.*

**Alternative Rahmung:** *Statistische Mustererkennung auf großen Datensätzen. Keine Kognition, kein Verständnis, kein Bewusstsein.*

Der Begriff „Künstliche Intelligenz“ ist die folgenreichste PR-Umbenennung der Technologiesgeschichte. Was er bezeichnet, sind primär statistische Optimierungsverfahren auf großen Datenmengen — hocheffiziente Mustererkennung, keine Kognition im psychologischen Sinne. Der Begriff „Intelligenz“ aber aktiviert ein konzeptuelles Schema, das die Technologie mit menschlichen Fähigkeiten assoziiert und damit Fragen der Verantwortung, der Rechte und der Autonomie auf eine Weise stellt, die den Entwicklern nützt: als Neutralisierung von Verantwortungszuschreibung („der Algorithmus hat entschieden“) und als Anheizung von Zukunftsnarrativ („kognitiv überlegene Systeme“).

Gary Marcus und Ernest Davis haben in „Rebooting AI“ (2019) die Inflationierung des Intelligenzbegriffs systematisch dokumentiert. Melanie Mitchell schreibt in „Artificial Intelligence: A Guide for Thinking Humans“ (2019): „The term artificial intelligence tends to confuse rather than clarify — it creates unrealistic expectations and obscures what machines actually do.“ Der Begriff ist nicht deskriptiv. Er ist funktional.

### „AI Safety“

**Framing:** Schutz vor dem Risiko, das KI selbst darstellt — für die Menschheit.

**Wer definiert:** *Primär angloamerikanische Institute: Machine Intelligence Research Institute (MIRI), Future of Humanity Institute (Oxford), Anthropic, OpenAI Safety Teams.*

**Alternative Rahmung:** *Schutz vor den Risiken, die aus der Konzentration von KI-Macht in wenigen Unternehmenshaenden entstehen.*

Das Framing „AI Safety“ hat eine bemerkenswerte Eigenschaft: Es definiert das Risiko als aus der KI selbst kommend — als emergente Gefahr eines superintelligenten Systems, das

menschliche Kontrolle übertritt. Damit rückt die Bedrohung durch die Unternehmen, die KI entwickeln und kontrollieren, aus dem Blickfeld: Machtkonzentration, Marktmonopol, demokratische Unkontrollierbarkeit, geopolitische Instrumentalisierung.

Emily Bender, Timnit Gebru und Kolleginnen haben in ihrem einflussreichen Paper „On the Dangers of Stochastic Parrots“ (2021) gezeigt, wie das „AI Safety“-Framing konkrete, gegenwärtige Schäden — algorithmischer Bias, Verletzungen von Datenschutz und Persönlichkeitsrechten, Beschäftigungsverdrängung — durch die Konzentration auf hypothetische zukünftige Super-KI-Risiken systematisch unsichtbar macht. Sie nennen dieses Phänomen „future-washing the present harms“.

Das ist Lukes' dritte Dimension von Macht in Reinform: Die Rahmung lenkt Aufmerksamkeit und Ressourcen weg von den Akteuren, die profitieren, hin zu einer abstrakten, systemischen Gefahr.

### „Alignment“

**Framing:** KI-Systeme an menschliche Werte anpassen — damit sie tun, was wir wollen.

**Wer definiert:** *AI-Safety-Community, primär Stuart Russell, Paul Christiano, DeepMind Safety Team.*

**Alternative Rahmung:** *Wessen Werte? Welche Menschen? Wer entscheidet, was „wir“ wollen — und durch welchen Prozess?*

Das „Alignment“-Problem — wie KI-Systeme so gestaltet werden können, dass sie menschliche Werte widerspiegeln — stellt eine Frage, die suggeriert, es gebe einen kohärenten menschlichen Wertekonsens, an dem KI ausgerichtet werden könnte. Diese Suggestion ist empirisch falsch: Menschen haben fundamental verschiedene, oft konfligierende Werte. Die Frage „Alignment mit welchen Werten?“ ist eine politische Frage, keine technische.

Abeba Birhane und Iason Gabriel haben unabhängig gezeigt, dass „Alignment“ in der Praxis primär bedeutet: Alignment mit den Werten der Entwickler, der Investoren und der dominanten Nutzergruppen (Birhane 2021; Gabriel 2020). Die Rahmung als technisches Problem neutralisiert den normativen Gehalt und entzieht die Wertentscheidung dem demokratischen Diskurs.

### „Autonomie“

**Framing:** KI-Systeme handeln selbstständig, ohne menschliche Anweisung für jeden Schritt.

**Wer definiert:** *Robotik und KI-Forschung seit den 1980ern; popularisiert durch autonome Fahrzeuge und Agentic-AI-Diskurs.*

**Alternative Rahmung:** *Technische Automatisierung ohne menschliche Echtzeitsteuerung. Weder Willen noch Intentionalität.*

Der Begriff „Autonomie“ ist aus der politischen Philosophie und Ethik in den KI-Diskurs übertragen worden, wo er ein normatives Konzept bezeichnet: die Fähigkeit eines Subjekts,

nach eigenen Gründen zu handeln. In der KI-Anwendung beschreibt er eine technische Eigenschaft: die Fähigkeit eines Systems, ohne Echtzeitanweisung zu operieren. Diese Begriffsverschiebung ist folgenreich: Sie aktiviert das normative Schema des autonomen Subjekts für ein System, das kein Subjekt ist, und erzeugt damit Ambiguität bei der Verantwortungszuschreibung („das System hat autonom entschieden“).

Joanna Bryson hat wiederholt auf diese Begriffsverschiebung hingewiesen: „Robots Should Be Slaves“ (2010) ist ihr provokanter Gegenentwurf zum Autonomie-Framing — nicht normativ, sondern analytisch: KI-Systeme sind Werkzeuge, keine Agenten, und die Zuschreibung von Autonomie vernebelt Verantwortungslinien.

#### 04 · MACHTANALYSE

## Wessen Sprache dominiert — und warum

### 4.1 Die geografische und institutionelle Konzentration

Der globale KI-Diskurs ist geografisch und institutionell massiv konzentriert. Die Forschung, die Regulierung, die öffentliche Debatte — alle operieren überwiegend mit einem Vokabular, das in angloamerikanischen Kontexten entstand und dort dominant geblieben ist. Nicht weil dieses Vokabular präziser wäre, sondern weil die Produktionszentren der KI-Industrie — Silicon Valley, Seattle, New York, London — gleichzeitig die Produktionszentren der KI-Semantik sind.

**INSTITUTIONELLE KONZENTRATION DER KI-SEMANTIK** Die zehn einflussreichsten KI-Policy-Organisationen weltweit haben ihren Sitz überwiegend in den USA und Großbritannien. Ihre Publikationen, Frameworks und Begriffsangebote werden global zitiert, übersetzt und adaptiert — oft ohne die kulturellen, rechtlichen und politischen Voraussetzungen zu übernehmen, unter denen sie entstanden sind.

Der EU AI Act ist ein lehrreiches Beispiel: Er übernimmt die Hochrisikoklassifizierung und das Konzept der „transparency“ weitgehend aus dem angloamerikanischen Policy-Kontext. Das europäische Rechtsverständnis — mit seiner anderen Tradition der Grundrechte, der Daseinsvorsorge, des Verhältnismäßigkeitsprinzips — würde andere Begriffe, andere Rahmen, andere Schwerpunkte generieren. Es hat aber keinen gleichwertigen Produktionsapparat für KI-Semantik entwickelt.

### 4.2 Das Schnelligkeit-Asymmetrie-Problem

Tech-Unternehmen produzieren Semantik in einem Tempo, das öffentliche Institutionen strukturell nicht mithalten können. Wenn ChatGPT in wenigen Monaten zur meistgenutzten Anwendung der Geschichte wird, ist die semantische Infrastruktur — „Generative AI“, „Large Language Model“, „Halluzination“ — bereits gesetzt, bevor öffentliche Diskussion die Begriffe

überprüft hat. Diese Schnelligkeit ist kein Zufall. Sie ist Teil der Macht: Wer zuerst benennt, setzt den Rahmen.

Kate Crawford hat in „Atlas of AI“ (2021) dokumentiert, wie Tech-Unternehmen systematisch Sprache einsetzen, um Entscheidungen zu naturalisieren: „Feed“ für den algorithmischen Nachrichtenstrom, „Friends“ für Netzwerkverbindungen, „Cloud“ für physische Rechenzentren mit enormem ökologischem Fußabdruck. Jedes dieser Worte ist ein Framing-Entscheidung.

### 4.3 Die Ausschlusslogik des Expertentums

Foucaults Konzept der Diskursregimes erklärt ein Phänomen, das in der KI-Debatte gut beobachtbar ist: Wer als Experte gilt, und wer nicht. Die dominante KI-Expertise ist technisch definiert: Informatiker, Mathematiker, Machine-Learning-Spezialisten. Ihre Perspektiven gelten als sachlich, empirisch, relevant.

Perspektiven von Soziologen, Philosophen, Gewerkschafterinnen, Betroffenengruppen, nicht-westlichen Wissenschaftstraditionen gelten in vielen Foren als „nicht-technisch“ – und damit als weniger relevant für Entscheidungen über KI. Diese Ausschlusslogik ist keine bewusste Diskriminierung. Sie ist die strukturelle Wirkung einer Expertiserahmen, die technische Kompetenz mit sachlicher Urteilsfähigkeit gleichsetzt – und dabei die normativen, sozialen und politischen Dimensionen von KI-Entscheidungen systematisch untergewichtet.

Ruha Benjamin hat dieses Phänomen in „Race After Technology“ (2019) präzise beschrieben: „The most powerful designs are those that appear neutral, universal, and technical.“

**THESE III** *Der KI-Diskurs produziert eine spezifische Form von Expertentum, die soziale, politische und ethische Urteile als technische verkleidet. Wer nicht in der Sprache der Techniker spricht, gilt als nicht-sachverständig – auch wenn es um Fragen geht, die primär politische Antworten erfordern.*

## 05 · PERSPEKTIVE I

# Neuroleadership: Warum Framing so wirksam ist

## 5.1 Das Gehirn als Rahmungsmaschine

Lakoffs kognitionswissenschaftliche Grundlage ist die Theorie der „embodied cognition“: Konzeptuelle Strukturen sind körperlich verankert und metaphorisch organisiert. Das Gehirn verarbeitet abstrakte Konzepte durch Analogie zu gelebten Erfahrungen. „Kognitive KI“ aktiviert das neuronale Netz des menschlichen Denkens; „Statistisches Vorhersagemodell“ aktiviert etwas Kausales, Messbares, Begrenzteres.

Karl Fristons Predictive-Coding-Theorie ergänzt: Das Gehirn konstruiert Vorhersagen auf Basis bestehender Modelle und gleicht eingehende Signale damit ab. Begriffe, die mit bestehenden mentalen Modellen übereinstimmen, werden schnell integriert. Begriffe, die diesen Modellen widersprechen, erzeugen Prediction Error und kognitive Belastung. Das macht das Reframing — das bewusste Umformulieren von Begriffen — kognitiv aufwändig: Es erfordert das Aufgeben etablierter mentaler Muster, was das Gehirn als Bedrohung registriert.

Das erklärt, warum dominante Framings so stabil sind: Nicht weil sie überzeugender sind, sondern weil sie mit bestehenden kognitiven Mustern kompatibel sind und weniger Verarbeitungsaufwand erfordern. Alternative Rahmungen sind kognitiv anspruchsvoller — und werden daher unbewusst mit höherer epistemischer Hürde bewertet.

## 5.2 SCARF und die Verteidigung des Expertenstatus

Rocks SCARF-Modell erklärt ein weiteres Phänomen im KI-Diskurs: Warum Experten, die ihre Begriffe hinterfragt sehen, mit Abwehr reagieren. Begriffsinfrastruktur ist Statuspflege. Wer die Begriffe beherrscht, die im Diskurs gelten, hat Status. Wer diese Begriffe in Frage stellt, gefährdet diesen Status — und löst damit eine SCARF-Bedrohungsreaktion aus: Verteidigung statt Dialog.

Das ist keine persönliche Schwäche der Beteiligten. Es ist eine neurobiologisch erwartbare Reaktion, die erklärt, warum inter- und transdisziplinäre Sprachkritik im KI-Feld so selten ernsthaft aufgenommen wird: Nicht weil sie unzutreffend wäre, sondern weil sie die Statusgrundlage der Experten bedroht, die sie beurteilen.

**THESE IV** *Dominante Framings im KI-Diskurs sind nicht primär deshalb stabil, weil sie richtig sind. Sie sind stabil, weil sie kognitiv effizient und statuserhaltend sind. Das ist eine neurobiologische Tatsache, keine Kritik an den Beteiligten.*

## 06 · KONSTRUKTIVER AUSWEG

# Bedeutungsarchitektur als Sprachkritik-Werkzeug

## 6.1 Vom Analysieren zum Gestalten

Sprachkritik allein ist keine Intervention. Die Dekonstruktion dominanter Framings erzeugt Verstehen, aber keine Alternative. Was nach der Analyse kommt, ist die Frage, die dieses Whitepaper im Kontext des Imago-Atelier-Programms stellt: Wie gestaltet man Bedeutungsstrukturen bewusst — und nicht reaktiv?

Bedeutungsarchitektur (WP10 dieser Reihe) bezeichnet genau diese Gestaltungsaufgabe: die bewusste Schaffung von Strukturen, Prozessen und Kulturen, die es ermöglichen, über KI in Begriffen zu sprechen, die die eigene Wirklichkeitserfahrung repräsentieren — und nicht nur

die Wirklichkeitserfahrung der Unternehmen, die die Systeme entwickeln. Das ist keine Ablehnung technischer Expertise. Es ist die Forderung nach epistemischer Pluralität.

### 6.2 Drei konkrete Reframing-Strategien

Lakoffs Anleitung zum Reframing ist konkret: Nenne die Realität, die du beschreiben willst. Benutze Sprache, die mit existierenden Erfahrungsrahmen des Publikums verbunden ist. Vermeide es, den gegnerischen Rahmen zu wiederholen — auch um ihn zu widerlegen. Angewendet auf den KI-Diskurs ergeben sich drei konkrete Strategien:

Dominantes Framing	Alternatives Framing (Reframing-Vorschlag)
„AI Safety“ — Schutz vor der KI selbst	„KI-Governance“ — Kontrolle über die, die KI entwickeln und einsetzen
„Intelligente Systeme“ — Anthropomorphisierung	„Statistische Optimierungssysteme“ — deskriptive Präzision
„Alignment“ — technische Werteanpassung	„Wertekodierung“ — macht sichtbar: wessen Werte, durch welchen Prozess
„AI-Risiko“ — systemisches Zukunftsrisiko	„KI-Schaden“ — konkreter Gegenwartsbezug auf reale Betroffene
„Autonome KI“ — implizite Subjekthaftigkeit	„Automatisierte Systeme“ — klare Werkzeugcharakter und Verantwortungszuschreibung
„Pflicht zur Innovation“ — naturalisiertes Narrativ	„Gestaltungsoptionen“ — macht Handlungsspielräume sichtbar

### 6.3 Die vier Grundfragen als Sprachkritik-Instrument

Das dialektische Prüfraster aus den Arbeitsmitteln — die vier Grundfragen der Rationalität — ist das präziseste verfügbare Werkzeug zur systematischen Begriffsanalyse im KI-Diskurs:

- Woher wissen wir das? — Auf welcher empirischen Basis wurde dieser Begriff entwickelt? Welche Phänomene sind in die Definition eingegangen, welche wurden ausgeschlossen?
- Welches Interesse hat die Wissensquelle? — Wer hat den Begriff geprägt? Welche institutionellen, ökonomischen und politischen Interessen haben die Begriffsschöpfer?
- Unter welchen Voraussetzungen gilt das? — Welche kulturellen, rechtlichen, epistemischen Voraussetzungen muss man teilen, damit der Begriff sinnvoll erscheint?
- Wie lauten die Definitionen der zentralen Begriffe? — Erfüllt die Definition die Kriterien einer guten Definition: Geltungsbereich, nächsthöherer Oberbegriff, Abgrenzungsschärfe, keine Zirkel?

Angewendet auf „Künstliche Intelligenz“: Woher stammt die Definition? (Von Informatikern der 1950er, nicht von Kognitionswissenschaftlern oder Philosophen.) Welches Interesse hat die Quelle? (Forschungsförderung, Aufmerksamkeit, später kommerzielle Verwertbarkeit.) Unter welchen Voraussetzungen gilt es? (Wenn man Intelligenz als Aufgabenlösungskapazität

operationalisiert, nicht als Verstehen oder Bewusstsein.) Wie lautet die Definition? (Undefiniert – der EU AI Act definiert KI, nicht KI-Intelligenz.) Die Analyse zeigt: Der Begriff trägt keine wissenschaftliche Definition. Er trägt eine Erzählung.

## 07 · SYNTHETISCHE THESEN

### Fünf irritierende Thesen

#### **These A: „Künstliche Intelligenz“ ist der erfolgreichste PR-Begriff der Technologiesgeschichte**

Kein anderer Begriff hat es geschafft, ein statistisches Optimierungsverfahren mit menschlicher Kognition gleichzusetzen, ohne dass diese Gleichsetzung jemals wissenschaftlich belegt wurde. Der Begriff „Künstliche Intelligenz“ aktiviert ein Bedeutungsschema, das Anthropomorphisierung naturalisiert, Verantwortungszuschreibung diffusiert und Zukunftsangst mobilisiert – alles zugleich und zugunsten der Akteure, die KI-Systeme entwickeln und vermarkten.

John McCarthy, der den Begriff 1956 geprägt hat, soll später gesagt haben, wenn es „Künstliche Intelligenz“ wäre, würde es nicht als solche bezeichnet werden – weil die Assoziation mit menschlicher Intelligenz irreführend ist. Das hinderte die Industrie nicht daran, den Begriff zur dominanten Bezeichnung für Mustererkennung, Regression und neuronale Netze zu machen.

**THESE A** „Künstliche Intelligenz“ ist kein technischer Terminus. Es ist ein narratives Konstrukt, das die kommerziellen und politischen Interessen seiner Produzenten widerspiegelt. Wer ihn benutzt, ohne ihn zu problematisieren, wiederholt eine Erzählung – keine Beschreibung.

#### **These B: „AI Safety“ ist die beste Schutzversicherung, die sich eine Branche je selbst ausgestellt hat**

Die „AI-Safety“-Bewegung hat eine bemerkenswerte Doppeloperation vollzogen: Sie hat das Risiko von KI als systemisches Zukunftsrisiko gerahmt – und sich gleichzeitig als die einzige Instanz positioniert, die dieses Risiko versteht und managen kann. Die Folge: Regulatorische Aufmerksamkeit richtet sich auf hypothetische Super-KI-Szenarien und wird von den konkreten, gegenwärtigen Schäden durch bestehende KI-Systeme abgelenkt.

Timnit Gebru und Emily Bender haben dies „future-washing the present harms“ genannt. Meredith Whittaker, Präsidentin der Signal Foundation, formuliert es direkter: Die AI-Safety-Agenda dient primär dazu, die Debatte auf Terrain zu verlagern, auf dem die großen Labs die Expertisehoheit besitzen.

**THESE B** *Wer „AI Safety“ als neutrale wissenschaftliche Agenda betrachtet, übersieht, dass die Organisationen, die sie am lautesten vertreten, dieselben sind, die am meisten von einem regulatorischen Rahmen profitieren, der sich auf Zukunftsrisiken statt auf Gegenwartsschäden konzentriert.*

## **These C: Europäische Regulierung spricht angloamerikanisch**

Der EU AI Act ist die ambitionierteste KI-Regulierung der Welt. Er benutzt dennoch weitgehend das Vokabular, die Kategorien und die Risikorahmen, die in angloamerikanischen Policy-Kontexten entwickelt wurden. Das ist kein Vorwurf an die Autoren des Gesetzes — es reflektiert das strukturelle Ungleichgewicht in der globalen Produktion von KI-Semantik.

Was wäre ein genuines europäisches KI-Regulierungskonzept? Es würde wahrscheinlich die Verhältnismäßigkeit in den Mittelpunkt stellen (nicht nur Risiko). Es würde Daseinsvorsorge als KI-Regulierungsziel explizit machen. Es würde demokratische Partizipation nicht als Transparenzpflicht, sondern als Mitgestaltungsrecht konzeptualisieren. Es würde Grundrechte als nicht verhandelbaren Maßstab verwenden — nicht als Abwägungsposten gegen wirtschaftliche Interessen.

**THESE C** *Europa reguliert KI — aber in einer Sprache, die seine eigenen Rechtstraditionen, Werteverständnisse und politischen Erfahrungen nur unvollständig repräsentiert. Semantische Souveränität über KI ist eine Voraussetzung für politische Souveränität über KI.*

## **These D: Die Rede von „menschlicher Kontrolle“ ist eine Fiktion, die Verantwortung verschleiert**

Human Oversight — die Forderung, dass Menschen KI-Systeme überwachen — ist ein zentrales Element des EU AI Acts und vieler anderer Governance-Frameworks. Das Framing suggeriert, dass menschliche Anwesenheit im Entscheidungsprozess Verantwortung sichert. Das ist unter bestimmten Bedingungen falsch: nämlich dann, wenn der Mensch, der überwacht, nicht die Kompetenz hat, das Ergebnis zu beurteilen.

WP12 dieser Reihe hat das als Unterschied zwischen Sicherheitsmechanismus und Haftungsschild beschrieben. Das Framing „menschliche Kontrolle“ verschleiert diesen Unterschied: Es schreibt Verantwortung zu, ohne die Bedingungen zu spezifizieren, unter denen Verantwortungsübernahme möglich ist. In der Praxis kann „Human in the Loop“ bedeuten: Ein Mensch ist formal präsent und faktisch machtlos.

**THESE D** *Human Oversight als Governance-Prinzip ist dann sinnvoll, wenn der Mensch die Kompetenz hat, den Agenten zu beurteilen. Wenn nicht, ist es kein Sicherheitsinstrument. Es ist ein Verantwortungssymbol. Und das ist gefährlicher als explizite Automatisierung — weil es Sicherheit vortäuscht.*

## These E: Wer in der Sprache der Tech-Industrie über KI spricht, argumentiert immer schon auf ihrem Boden

Das ist die radikalste These dieses Papers, und sie verdient präzise Formulierung. Sie behauptet nicht, dass alle Begriffe des KI-Diskurses falsch sind oder dass man ohne die gängige Terminologie nicht sprechen könnte. Sie behauptet etwas Strukturelles: Wer die dominante KI-Semantik benutzt, ohne sie zu reflektieren, akzeptiert implizit die Rahmungen, die mit diesen Begriffen verbunden sind – und positioniert sich damit auf dem Boden der Diskursagenda derer, die die Begriffe gesetzt haben.

Das gilt auch für Kritiker. Wer „KI ist kein Risiko für die Menschheit“ sagt, hat das Framing „KI als Risiko für die Menschheit“ bereits aktiviert. Wer „KI ist nicht intelligent“ sagt, hat das Framing „KI könnte intelligent sein“ bereits akzeptiert. Lakoffs Diktum gilt ohne Ausnahme: Wenn man gegen einen Rahmen anargumentiert, stärkt man ihn.

Die Alternative ist nicht Schweigen. Sie ist eine bewusste Entscheidung für eine andere Sprache – und die Mühe, diese Sprache zu entwickeln, zu erklären und durchzuhalten.

**THESE E** *Die wichtigste Kompetenz im Umgang mit dem KI-Diskurs ist nicht technisches Wissen. Es ist die Fähigkeit, die eigene Sprache zu beobachten – und zu entscheiden, welche Rahmen man reproduziert und welche man verweigert. Das ist Bedeutungsarchitektur als Haltung.*

### 08 · AUSBLICK

## Sprachliche Intervention als politische Praxis

Dieses Paper ist ein Grenzthema – weil es keine einfachen Lösungen anbietet. Es kann keine Terminologie vorschlagen, die per Definition neutral wäre: Jede Sprache ist geformt. Die Frage ist nicht, ob Begriffe Rahmungen tragen, sondern welche Rahmungen man bewusst wählt und welche man unreflektiert übernimmt.

Was dieses Paper empfiehlt, ist keine terminologische Revolution. Es ist eine analytische Haltung: die Bereitschaft, jeden zentralen Begriff des KI-Diskurses durch die vier Grundfragen der Rationalität zu befragen, bevor man ihn verwendet. Das kostet Zeit. Es verletzt manchmal die Flüssigkeit der Argumentation. Es lädt zum Widerspruch ein.

Es ist jedoch die einzige Möglichkeit, die epistemische Souveränität zu wahren, ohne die gleichzeitig kein verantwortungsvoller Umgang mit KI möglich ist. Wer die Sprache nicht analysiert, in der er denkt, denkt in Kategorien, die andere für ihn gewählt haben.

„Präzise Sprache schafft Substanz.“ – Aus der Praxis dialektischer Führung

Dieser Satz aus den Arbeitsmitteln des Imago Ateliers hat eine Kehrseite, die dieses Paper entfaltet hat: Unpräzise Sprache zerstört Substanz – und schafft Raum für Interessen, die von der Unschärfe profitieren. Das gilt für die KI-Debatte mit einer Dringlichkeit, die kaum übertrieben werden kann: Die Entscheidungen, die in den nächsten Jahren über KI-Systeme getroffen werden, sind nicht technisch. Sie sind politisch. Und politische Entscheidungen werden in Sprache getroffen.

Wer diese Sprache nicht gestaltet, ist ihr ausgeliefert.

---

## AUSGEWÄHLTE LITERATUR

## Referenzen

- [1] Benjamin, R. (2019). *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity Press.
- [2] Bender, E. M., Gebru, T., McMillan-Major, A. & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *ACM FAccT* 2021.
- [3] Birhane, A. (2021). Algorithmic injustice: a relational ethics approach. *Patterns*, 2(2), 100205.
- [4] Bryson, J. J. (2010). Robots Should Be Slaves. *Close Engagements with Artificial Companions* (ed. Wilks). John Benjamins.
- [5] Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.
- [6] Foucault, M. (1969). *Archäologie des Wissens*. Suhrkamp [frz. *L'archéologie du savoir*].
- [7] Foucault, M. (1970). *Die Ordnung des Diskurses*. Hanser [frz. *L'ordre du discours*].
- [8] Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138.
- [9] Gabriel, I. (2020). Artificial Intelligence, Values, and Alignment. *Minds and Machines*, 30, 411–437.
- [10] Lakoff, G. (2004). *Don't Think of an Elephant! Know Your Values and Frame the Debate*. Chelsea Green.
- [11] Lakoff, G. & Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press.
- [12] Luhmann, N. (1980). *Gesellschaftsstruktur und Semantik. Studien zur Wissenssoziologie der modernen Gesellschaft*. Suhrkamp.
- [13] Lukes, S. (1974). *Power: A Radical View*. Macmillan.
- [14] Marcus, G. & Davis, E. (2019). *Rebooting AI: Building Artificial Intelligence We Can Trust*. Pantheon Books.
- [15] McCarthy, J. et al. (1956). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. Dartmouth College.
- [16] Mitchell, M. (2019). *Artificial Intelligence: A Guide for Thinking Humans*. Farrar, Straus and Giroux.
- [17] Rock, D. (2008). SCARF: A brain-based model for collaborating with and influencing others. *NeuroLeadership Journal*, 1, 44–52.
- [18] Whittaker, M. (2021). The Steep Cost of Capture. *Interactions*, 28(6), 50–55.
- [19] Wittgenstein, L. (1953). *Philosophische Untersuchungen*. Blackwell. [§43: Bedeutung als Gebrauch]