

Teresa Forcades i Vila OSB

Sr. Dr. Dr. Teresa Forcades i Vila OSB, Jahrgang 1966, ist Ärztin und Theologin. Sie lehrte an der Humboldt-Universität Berlin und der Universidad Iberoamericana in Mexiko. Derzeit ist sie Direktorin der Sinclética Monastic School und der Zeitschrift für Christenheit und kritisches Denken Iglesia Viva.



Foto: privat

Teresa Forcades i Vila OSB

Daten sammeln – Daten nutzen

Ethische und moral-theologische Bewertung von KI

KI ist eine von vielen neuen und sich rasch entwickelnden Technologien, die in den letzten zehn Jahren zunehmend aus den abgeschirmten Forschungslabors in den Alltag der meisten Bürger in den reichen Ländern Einzug gehalten haben: Biotechnologie (synthetische Biologie, Gentherapie ...), Robotik (in der Pflege, Chirurgie, Armee, Schule, im Sport ...), Neue Stoffe (intelligentes Polymergel, Graphen ...), 3D-Druck (in Kunst, Architektur, Medizin ...), neue Methoden für Energieführung, -speicherung, -übertragung, virtuelle Realitäten mit AR-Funktionen (Metaversum in Schule und andere Ausbildungskontexte), Blockchain (Bitcoin, Kryptowährung ...), Computer der nächsten Generation (Supercomputer, Quantencomputer ...).

Die schiere Anzahl dieser technischen Innovationen und ihre Komplexität sind überwältigend, und wir laufen Gefahr,

unsere kritischen Fähigkeiten im Umgang mit ihnen zu verlieren. Wie können wir diese Techniken moralisch oder theologisch bewerten, wenn wir nicht einmal wissen, was sie sind oder wie sie funktionieren?

Innehalten und nachdenken

Und doch ist es unerlässlich, dass wir innehalten und so gründlich wie möglich über die Auswirkungen jeder dieser Innovationen und ihrer gemeinsamen Umsetzung in unserer Mitte nachdenken. In diesem Artikel werde ich einige kritische Fragen aufwerfen, die sich besonders auf generative KI beziehen, also jene Art von KI, die von ihren Nutzern lernen kann und in einigen Fällen darauf ausgelegt ist, mit ihnen zu interagieren und dabei zwischenmenschliche Beziehungen zu imitieren.

Das Saarbrücker Experiment oder: die manipulierte Partitur

Am 8. November 2022 fand in Saarbrücken ein innovatives Konzert statt¹. Das Saarländische Staatsorchester unter der Leitung von Benjamin Schäfer wollte Ives „Unanswered Question“ auf eine ungewöhnliche Weise aufführen: Jeder Musiker hatte die Partitur auf einem Tablet, und jedes Tablet war mit einem Computer verbunden, der die Partitur in Echtzeit entsprechend den physiologischen Informationen modifizierte, welche vom Publikum empfangen wurden. Pulsfrequenz, Gesichtserkennung, Blickgeschwindigkeit und Hautrötung – alles wurde über ein Armbandgerät und mehrere Kameras gemessen, anonym erfasst und ausgewertet. In der Werbung wurde dies als Experiment der „Co-Kreation“ bezeichnet.

Das Publikum trug jedoch nicht bewusst zu diesem Prozess bei. Auch wenn die physiologischen Reaktionen von den Zuschauern stammten, waren es nicht sie, sondern die Forscher, die entschieden, welche Informationen sie sammeln und wie sie diese verarbeiten würden, um die Partitur zu verändern. Die Forscher speicherten alle so aus der Öffentlichkeit gewonnenen Daten, und das Technologieunternehmen, das die Forschung finanzierte, bekam letztendlich das Eigentum an allen Daten aus diesem Konzert und aus vielen ähnlichen Experimenten, die in den letzten fünf Jahren weltweit durchgeführt wurden.

Die befragten Zuschauer fanden die Musik, die sie hörten, nicht besonders schön, aber sie waren froh, Teil einer so innovativen Erfahrung gemeinschaftlicher Kreation gewesen zu sein. Sie schienen sich nicht bewusst zu sein, wie stark ihre Erfahrung durch Forschungsalgorithmen

beeinflusst worden war, an deren Entwicklung sie nicht beteiligt waren, und wie lukrativ die biometrischen Informationen, für deren Bereitstellung sie bezahlt hatten, für einen Dritten sein konnten, von dessen Existenz die meisten nichts wussten.²

Das Publikum wurde um seine Zustimmung zur Datenerhebung gebeten. Die Unterzeichnung einer Einverständniserklärung, die die Forscher rechtlich absichert, ist jedoch nicht gleichbedeutend mit der Einhaltung der geltenden ethischen Standards. Der Nürnberger Kodex verlangt, „dass der Versuchsperson vor der Einholung ihrer Zustimmung das Wesen, die Länge und der Zweck des Versuches klargemacht werden; sowie die Methode und die Mittel, welche angewendet werden sollen“. Die Antworten, die die Teilnehmer am Ende des Konzerts gaben, scheinen zu zeigen, dass sie nicht vollständig verstanden hatten, woran sie teilgenommen hatten. Hätten sie gewusst, dass das eigentliche Ziel des Konzerts die Erhebung biometrischer Daten für weitere Forschungszwecke und die kommerzielle Verwertung war, hätten sie möglicherweise eine Bezahlung verlangt, anstatt für die Bereitstellung dieser wertvollen Informationen zu bezahlen.

Brain Transparency oder: Was weiß der Vorgesetzte?

Vom 16. bis 20. Januar 2023 fand in Davos die Jahrestagung des Weltwirtschaftsforums unter dem Titel „Zusammenarbeit in einer fragmentierten Welt“ statt, an der 2700 Führungskräfte aus 130 Ländern, darunter 52 Staats- und Regierungschefs, teilnahmen. Eine der Hauptrednerinnen war Nita A. Farahany, Professorin für Recht und Philosophie an

der Duke University, die 2010 von Präsident Obama in die Präsidentschaftskommission für die Untersuchung bioethischer Fragen berufen wurde.

Unter dem Titel „Bereit für Gehirntransparenz?“³ beschrieb Professorin Farahany neurotechnologische Innovationen, mit denen sich intime Geisteszustände durch die Messung der Gehirnwellenaktivität mit einem einfachen Gerät wie Kopfhörern identifizieren lassen. In den nächsten zwei Jahren werden Milliarden von Dollar von neurotechnologischen Unternehmen verdient werden. Diese präsentieren sich als interessiert daran, Tetraplegikern (Querschnittgelähmten, deren Arme und Beine betroffen sind) zu helfen, ihre Funktionen und Kommunikationsfähigkeiten wiederzuerlangen, aber in Wirklichkeit weniger altruistische Interessen verfolgen, wie z. B. Vorgesetzten in einem Büro die Möglichkeit zu geben, die Aufmerksamkeit ihrer Mitarbeiter zu kontrollieren, indem sie einfach die gesamte Sequenz der Gehirnströme der Mitarbeiter während der Arbeitszeit auf einem Computerbildschirm abrufen. Was wären die Folgen dieser sogenannten Gehirntransparenz im Arbeitsumfeld? Und im Bildungsbereich? Durch die Überwachung der Gehirnströme können diejenigen, die die Informationen über die Ohrstöpsel erhalten, zwischen vorgetäuschter und echter Aufmerksamkeit, zwischen Konzentration und Tagträumen unterscheiden und möglicherweise sogar erkennen, welche Mitarbeiter ein gutes Verhältnis zueinander haben, von wem ein Mitarbeiter sexuell erregt wird und ob diese Erregung erwidert wird. Die Mitarbeiter könnten über die Ohrstöpsel in Echtzeit Warnungen mit dem Befehl „Lass dich nicht ablenken!“ erhalten und je nach Anzahl der „Ablenkungsmomen-

te“, die die Gehirnwellenüberwachung während der Woche registriert hat, Gehaltsabzüge riskieren.

Science-Fiction? Eine ähnliche Technologie wird bereits bei Lkw-Fahrern eingesetzt: Sie haben eine Kamera mit Gesichtserkennung, die ihnen eine Nachricht sendet, wenn ihr Gesicht Anzeichen von Müdigkeit zeigt. Wenn sie die Nachricht ignorieren, übernimmt ihre Versicherung im Falle eines Unfalls nicht die Kosten. Die Kameras mit Gesichtserkennung und die Ohrstöpsel, die Gehirnströme lesen können, sind Geräte der generativen KI: Sie lernen von ihren Nutzern und werden durch die Speicherung und Verarbeitung von Daten mit jedem Einsatz immer leistungsfähiger.

„Abschaffung“ oder: Wohin Pokemon uns führt

2019 veröffentlichte die Sozialpsychologin Soshanna Zuboff, emeritierte Professorin an der Harvard Business School und erste Frau, die jemals eine Festanstellung an dieser Institution erhielt, ihr vielbeachtetes Buch „Surveillance Capitalism“, das seitdem in mehr als dreißig Sprachen übersetzt wurde⁴. Zuboff argumentiert, dass

„in einer Informationszivilisation Gesellschaften durch Fragen des Wissens definiert werden – wie es verteilt wird, welche Autorität seine Verteilung regelt und welche Macht diese Autorität schützt. Wer weiß? Wer entscheidet, wer weiß? Wer entscheidet, wer entscheidet, wer weiß? Die Überwachungskapitalisten haben jetzt die Antworten auf jede dieser Fragen, obwohl wir sie nie gewählt haben, um uns zu regieren. Das ist das Wesen des epistemischen Staatsstreichs. Sie be-

ansprechen die Autorität, zu entscheiden, wer weiß, indem sie Eigentumsrechte an unseren persönlichen Daten geltend machen, und verteidigen diese Autorität mit der Macht, kritische Informationssysteme und Infrastrukturen zu kontrollieren.“⁵

Laut Zuboff ermöglichte die Entwicklung des Internets nach dem Fall der Berliner Mauer (1989) eine massive Produktion von Informationen, die von einigen wenigen Unternehmen gesammelt und genutzt wurden. Dies war der Beginn des Big Data, und es wurden viele ethische Bedenken laut. 1997 wurden diese ethischen Bedenken von der Clinton/Gore-Regierung unter der Prämisse ignoriert, dass technologische Innovationen nicht aufgehalten werden könnten und der einzige Weg nach vorne darin bestehe, privaten Unternehmen die ungehinderte Weiterentwicklung zu ermöglichen. Ein dringend benötigter Rechtsrahmen für die digitale Welt wurde nicht geschaffen, und bereits bestehende Vorschriften, die die aufkommende Big-Data-Industrie behindern könnten, wurden abgeschafft. Es war der Beginn des epistemischen Staatsstreichs, eines Staatsstreichs, der nicht mit militärischen Mitteln, sondern mit einer anderen Art von Gewalt erreicht wurde, die viel subtiler, aber ebenso tödlich war. Dann kam der 11. September, und der Staat stellte fest, dass er für den Zugang zu allen für seine Sicherheit wichtigen Informationen von privaten Unternehmen abhängig war.

Zu spät? Nein, sagt Zuboff, aber radikale Maßnahmen seien erforderlich. Sie hält die derzeitige private Anhäufung von Big Data und Technologie für unvereinbar mit der Demokratie und fordert deren Abschaffung. Nicht die Abschaffung der Technologie, sondern die Abschaffung

von Big Data und Big-Data-Technologie in Privatbesitz, die in der Lage sind, menschliches Verhalten und Wahlergebnisse zu kontrollieren. Zuboff verwendet den Begriff „Abschaffung“ in Anlehnung an die auf Sklaverei basierende Wirtschaft des 19. Jahrhunderts: Die Lösung bestand nicht darin, die Sklaven besser zu behandeln, sondern die Sklaverei insgesamt abzuschaffen. Heute sollte es genauso sein: „Abschaffung“ des privaten Eigentums an Big Data.

Als Beispiel analysiert Zuboff das Augmented-Reality Pokémon Go, das 2016 erschien und schnell sehr populär wurde. Die Spieler nutzten ihr GPS-System auf dem Smartphone, um virtuelle Pokémons zu lokalisieren, zu fangen, zu trainieren und zu bekämpfen, die so aussahen, als befänden sie sich am realen Standort des Spielers. Das Spiel war kostenlos. Und dennoch erzielte Niantic (das Unternehmen, dem Pokémon Go gehörte) im Jahr seiner Einführung (2016) einen Gewinn von 500 Millionen Dollar, 2019 von einer Milliarde und 2020 von sechs Milliarden. Wie war das möglich? Die Antwort ist einfach: Werbung und Verhaltenskontrolle. Das Unternehmen platzierte seine virtuellen Pokémons in Geschäften und Unternehmen (Bars, Restaurants ...), die dafür bezahlten. Dann leitete das Spiel Tausende von Menschen auf der Suche nach Pokémons zu diesen Geschäften. Sobald die potenziellen Kunden dort waren, fügte das Pokémon-Programm zusätzliche praktische Werbung hinzu, die individuell auf jeden Kunden zugeschnitten werden konnte. Die Geschäfte und Unternehmen, die Pokémons unter Vertrag genommen hatten, verzeichneten einen Umsatzanstieg. Eine Win-Win-Situation? Sicher, außer für die manipulierten Spieler, die nichts davon wussten.

Die Gefahr, großen privaten Unternehmen zu erlauben, ihre KI-Algorithmen zu trainieren und zu perfektionieren, um individuelles und kollektives Verhalten durch manipulative Techniken zu kontrollieren, wird anhand weiterer Beispiele von Zuboff noch deutlicher, die ich hier nur kurz erwähnen möchte: die Rolle von Cambridge Analytica beim Brexit-Votum in Großbritannien und die Rolle von Facebook bei Trumps erster Präsidentschaftskampagne.

Zuboff ist der Ansicht, dass die Big-Data-Unternehmen und die Regierungen, die sie schützen, sich eines Diebstahls schuldig gemacht haben. Die Daten gehören uns (der Öffentlichkeit) und hätten niemals in Privatbesitz gelangen und zu Gewinnzwecken ausgebeutet werden dürfen. Noch weniger dürfen sie nach Algorithmen verarbeitet und genutzt werden, die der Öffentlichkeit, die die Daten produziert, nicht offengelegt werden und die niemals ihre Eigentumsrechte hätte verlieren dürfen.

Ethische Bewertung oder: ChatGPT ist eine Black Box

Die ethische Bewertung von KI ist für mich klar: Ich bin überzeugt, dass Zuboff und die vielen Intellektuellen und Aktivisten, die ihrer kritischen Analyse zustimmen, angehört werden müssen. Das private Eigentum an Big Data ist heute eine Realität, die eine klare aktuelle (nicht zukünftige) Gefahr für unsere politische Freiheit und Demokratie darstellt (Beeinflussung der Wahl politischer Vertreter und von Referenden), für unsere wirtschaftliche Freiheit und soziale Gerechtigkeit (massive Manipulation der Wünsche und Bedürfnisse der Menschen auf personalisierte Weise, indem die umfang-

Autoreninfo

siehe gedruckte Ausgabe

reichen und detaillierten Kenntnisse über jede Person und ihre Angehörigen ausgenutzt werden; dies wird als „Mikrotargeting-Werbung“ bezeichnet) und sogar für unsere kognitive Freiheit und unser Gefühl von Privatsphäre und Selbst (indem sie Vorgesetzten oder Lehrern hilft, Einblick in die intimsten Gedanken und Gefühle ihrer Mitarbeiter oder Schüler zu erhalten).

In ihrer derzeitigen Entwicklungsform stellt generative KI eine besondere Gefahr für Jugendliche und junge Erwachsene aus noch existierenden indigenen oder eingeborenen Stämmen dar, beispielsweise im Amazonasgebiet. KI bietet ihnen interaktive, personalisierte digitale Erfahrungen, für die sie keinen kulturellen Kontext haben. Die Nutzer werden existenziell entwurzelt und emotional und kognitiv überfordert, und die Selbstmordraten steigen.

Generative KI verändert das Bildungsumfeld. Wir werden sehen, wie DeepSeek, der chinesische Konkurrent von ChatGPT, die Spielregeln verändert. Am 8. Mai 2025 räumte Brad Smith, stellvertretender Vorsitzender und Präsident von Microsoft, in einer Anhörung vor dem US-Senat ein, dass seinen Mitarbeitern die Nutzung der DeepSeek-Anwendung aus Gründen der Datensicherheit und wegen Propaganda-Bedenken untersagt ist. Die

Datenschutzrichtlinie von DeepSeek besagt, dass die Daten der Nutzer auf chinesischen Servern gespeichert werden, die dem chinesischen Recht unterliegen, welches die Zusammenarbeit mit den Geheimdiensten des Landes vorschreibt. DeepSeek zensiert außerdem Themen, die von der chinesischen Regierung als sensibel eingestuft werden, stark⁶.

Aber wie steht es um die ethischen Standards von ChatGPT? Am 26. November 2024 wurde der junge KI-Forscher Suchir Balaji, einer der Leiter des Projekts, das heute als ChatGPT von OpenAI bekannt ist, tot in seiner Wohnung aufgefunden⁷. Offiziell wurde Selbstmord festgestellt, aber seine Familie und Freunde zweifeln daran. Zum Zeitpunkt seines Todes stand er als Zeuge unter dem Schutz der New York Times, da er OpenAI verlassen hatte, weil er Informationen weitergeben wollte. Er hatte bereits offen gesagt, dass KI-Tools mit urheberrechtlich geschützten Inhalten trainiert werden, also mit Inhalten, die effektiv gestohlen sind⁸. Die Öffentlichkeit wurde von OpenAI beraubt: Wir stellen Bilder, Musik und andere Inhalte zur Verfügung, die geschützt sein sollten. Suchir Balaji hat sie für das ChatGPT-Projekt verwendet, weil er glaubte, dass es nicht gewinnorientiert sei. Ein Irrtum. Daraufhin hat Balaji das Unternehmen aus moralischen Bedenken verlassen.

Künstliche Intelligenz interagiert mit uns, aber sie hat Programmiermerkmale, die nur Programmierer kennen und kontrollieren. Nehmen wir zum Beispiel an, dass die Zahl 1 bedeutet, dass alles, was in das Programm eingegeben wird, verfügbar ist und herauskommen kann, dass alles, was in die Maschine eingegeben wird, vom Programm ausgewählt werden kann, und dass 0 bedeutet, dass nichts,

was eingegeben wird, herauskommen kann. Balaji hat aufgedeckt, dass ChatGPT mit einem 0,5-Algorithmus arbeitet, was bedeutet, dass 50 Prozent der Inhalte nicht offen verfügbar sind. Es ist, als ob Subjektivität in die Maschine eingeführt worden wäre. Wozu? Um mehr Gewinn für das Unternehmen zu erzielen? Um eine bestimmte Ideologie zu verbreiten?

Es ist wichtig, dass die Menschen wissen, was sich in der algorithmischen Black Box verbirgt. Maschinen können uns nicht kontrollieren, aber andere Menschen können uns durch Maschinen kontrollieren. Wenn ich glaube, dass eine Maschine objektiv auf meine Anfragen reagiert, und mir nicht bewusst ist, dass Inhalte zensiert werden (bis zu 50 Prozent!), ist das gleichbedeutend damit, nicht von der Maschine manipuliert zu werden, sondern von denen, die sie programmiert haben.

Theologischer Begriff von Intelligenz oder: die innere Dreifaltigkeit des Augustinus

Meine theologische Kritik an KI deckt sich mit meiner ethischen Kritik daran (mangelnde Transparenz, Manipulation des menschlichen freien Willens, sozio-ökonomische Ungleichheiten), wobei ich einen spezifisch theologischen Punkt hinzufüge: KI verzerrt das in jedem Menschen eingeschriebene Bild Gottes (*imago Dei*). Gemäß diesem *Imago Dei* ist die menschliche Intelligenz untrennbar mit der menschlichen Liebe verbunden. Die Bezeichnung „Intelligenz“ für das logisch-probabilistische Unterfangen, immer komplexere vorprogrammierte Algorithmen in einer Maschine anzuwenden und weiterzuentwickeln, ist ein Symptom einer dualistischen Selbstwahrnehmung,

die im Widerspruch zur theologischen Anthropologie des *Imago Dei* steht. Ich werde diesen Punkt im Hinblick auf die Trinitätslehre und die Christologie kurz näher ausführen.

In seinem Werk über die Dreifaltigkeit verwendete der heilige Augustinus (5. Jahrhundert) die Analogie *Mens-Notitia-Amor*, um von den drei göttlichen Personen zu sprechen, wobei er sie voneinander unterschied und gleichzeitig ihre Untrennbarkeit bezeugte. *Mens* (auch Gedächtnis genannt) entsprach der Person des Vaters, *Notitia* (auch Idee oder *Nous-Intellect-Logos* genannt) entsprach der Person des Sohnes und *Amor* (auch *Voluntas* oder Wille genannt) entsprach der Person des Heiligen Geistes. Die sogenannte augustinische „innere Dreifaltigkeit“ ist ein Versuch, die Tätigkeit des menschlichen Geistes zu beschreiben, der durch die untrennbaren Funktionen von Verstand und Wille nach dem Bild Gottes geschaffen ist. Wie wahrhaft untrennbar sie sind, kann jeder Lehrer erleben, der auf einen unmotivierten Schüler trifft: Egal wie intelligent oder begabt der Schüler auch sein mag, wenn er nicht motiviert ist, wenn sein Wille oder sein Verlangen nicht mit der Sache übereinstimmt, wird er nichts verstehen und nichts lernen.

Aus christologischer Sicht kann KI mit der Inkarnation des Logos konfrontiert werden, d. h. mit der Tatsache, dass unser Intellekt untrennbar mit unserem Körper verbunden ist. Was das konkret bedeutet, veranschaulicht der Psychoanalytiker Jacques Lacan, wenn er über die Entstehung des Unbewussten als Ergebnis der zutiefst persönlichen und komplexen Wechselwirkungen reflektiert, die sich aus den konkreten kognitiven und emotionalen Inhalten ergeben, die mit

jedem einzelnen Wort verbunden sind, das die Grundlage unserer Denkfähigkeit bildet. Die Worte mögen für uns alle dieselben sein, aber es gibt keine zwei Menschen, für die auch nur ein einziges Wort dieselben kognitiven und emotionalen Resonanzen hat. Das ist das Ergebnis unserer Verkörperung: Jeder von uns nimmt einen einzigartigen Raum ein und hat somit eine einzigartige Sicht auf die Realität, eine einzigartige Intelligenz, die mit seiner Biografie und seinen inneren Sinneserfahrungen verbunden ist.

Zusammenfassend: Ja zu KI als nützliches Werkzeug für ganz konkrete Aufgaben, nein zu ihrem privaten Besitz und ihrer privaten Ausbeutung; nein zu KI als Ersatz für die menschliche Intelligenz als solche. Und ja zum Vorschlag, den Namen KI zu ändern, um Verwirrung zu vermeiden.

.....

- 1 Kessler, Tobias. Ungewöhnliches Konzert in Saarbrücken: Künstliche Intelligenz soll aus Emotionen der Zuschauer ein Musikstück schreiben. Saarbrücker Zeitung vom 4.10.2022.
- 2 In diesem Fall handelte es sich um eine akademische Einrichtung, die Helmholtz Information & Data Science Academy – HIDA – in Berlin.
- 3 Thompson, Nicholas & Farahany, Nita. Ready for Brain Transparency? World Economic Forum at Davos, January 19, 2023. www.weforum.org/meetings/world-economic-forum-annual-meeting-2023/sessions/ready-for-brain-transparency/ (abgerufen am 12.05.2025)
- 4 Zuboff, Shoshana: Das Zeitalter des Überwachungskapitalismus. Campus Verlag 2025
- 5 Zuboff, Shoshana (January 29, 2021). “The Coup We Are Not Talking About”. The New York Times. <https://www.nytimes.com/2021/01/29/> (abgerufen am 16.05.2025)

- 6 Rollet, Charles. Microsoft employees are banned from using DeepSeek app, president says. TechCrunch, May 8 2025. <https://techcrunch.com/2025/05/08/microsoft-employees-are-banned-from-using-deepseek-app-president-says/> (abgerufen am 12.5.2025)
- 7 Davies, Alys (December 14, 2024). "OpenAI whistleblower found dead in San Francisco apartment". BBC. Archived from the original on December 14, 2024. <https://www.bbc.com/news/articles/cd0e13r2nlko> (abgerufen am 29.01.2025)
- 8 Metz, Cade (October 23, 2024). "Former OpenAI Researcher Says the Company Broke Copyright Law". The New York Times. Archived from the original on December 15, 2024. <https://www.nytimes.com/2024/10/23/technology/openai-copyright-law.html> (abgerufen am 19.01.2025)

KI verzerrt das *imago Dei*,
das in jedem Menschen
eingeschriebene Bild Gottes.

Teresa Forcades i Vila OSB