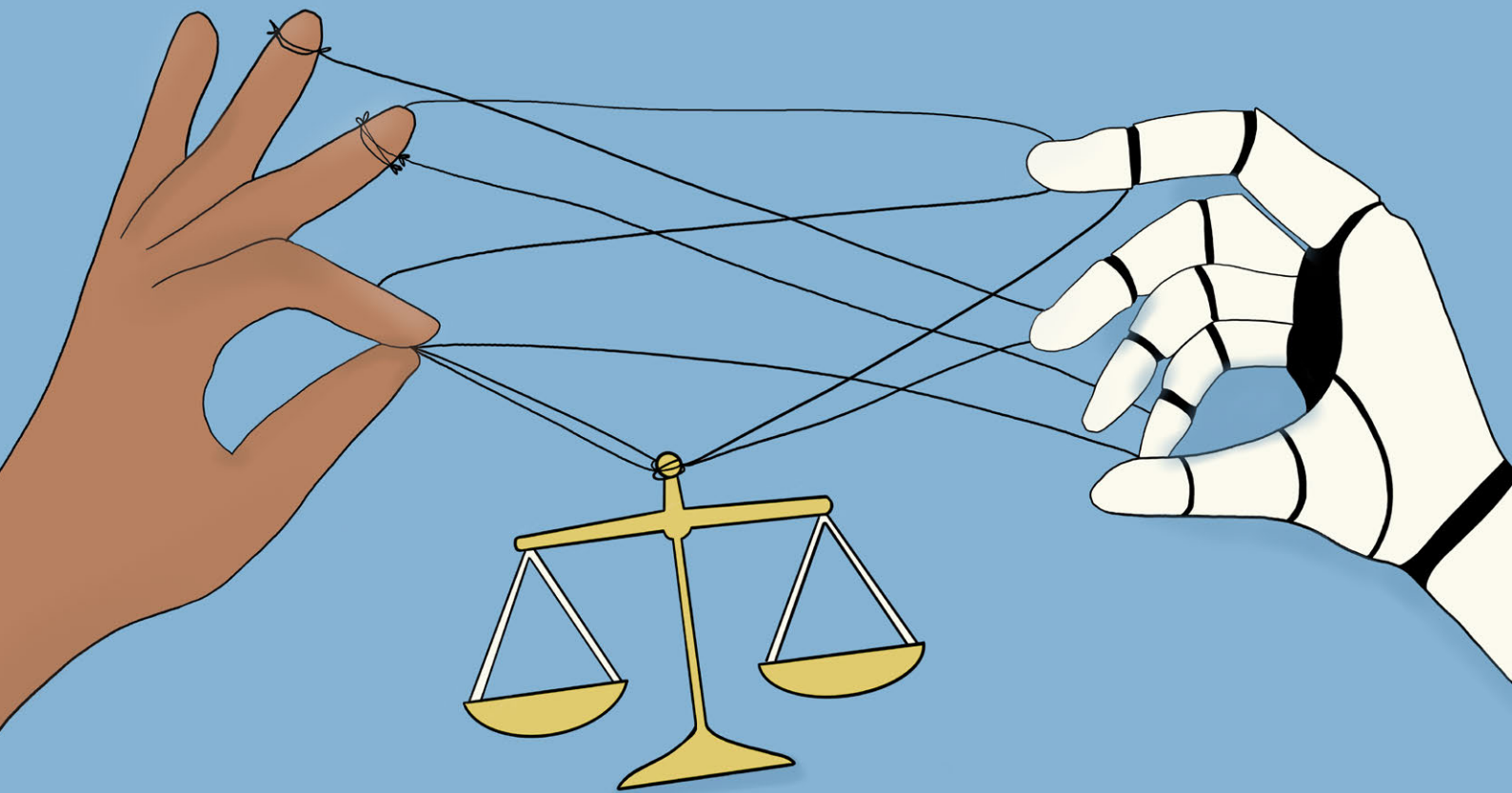


Denken bleibt Pflicht

Ein Interview mit Judith Simon



Generative KI verändert, wie wir lernen, lehren und denken. Die Ethikprofessorin Judith Simon sieht darin nicht nur technische, sondern auch fundamentale bildungspolitische Fragen: Wie erhalten wir zentrale Kompetenzen wie kritisches Denken, wenn Tools wie ChatGPT alltäglich werden? Was bleibt vom Sinn universitärer Bildung? Theresa Sommer (HFD) sprach mit ihr über Prüfungsformate, ethische Herausforderungen und die Frage, warum Hochschulen sich nicht in Effizienzlogik verlieren sollten.



Im Gespräch mit:

Prof. Dr. Judith Simon

ist Professorin für Ethik in der Informationstechnologie an der Universität Hamburg. Sie beschäftigt sich mit der Verschränkung ethischer, erkenntnistheoretischer und politischer Fragen im Kontext von Künstlicher Intelligenz und Digitalisierung im Allgemeinen sowie der Philosophie des Vertrauens. Judith Simon ist stellvertretende Vorsitzende des Deutschen Ethikrates und war dort Sprecherin der Arbeitsgruppe „Mensch und Maschine - Herausforderungen durch Künstliche Intelligenz“.

strategie digital: Als Professorin für Ethik in der Informationstechnologie und Mitglied des Deutschen Ethikrats beschäftigen Sie sich intensiv mit den gesellschaftlichen Auswirkungen generativer KI. Rund um generative KI gibt es viele Annahmen und Debatten – von überzogenen Ängsten bis hin zu übertriebenen Hoffnungen. Welches zentrale Missverständnis über KI-Sprachmodelle begegnet Ihnen immer wieder?

Judith Simon: Oft bestehen Missverständnisse darüber, wie KI funktioniert, was sie kann, was sie aber auch nicht kann. Bei vielen KI-Systemen tendieren Menschen dazu, diesen zu sehr zu vertrauen – einfach weil Software, weil Zahlen so neutral und objektiv daherkommen. Bei Sprachmodellen besteht darüber hinaus das Problem, dass diesen Systemen manchmal Fähigkeiten zugeschrieben werden, die diese gar nicht haben – Verständnis oder gar Bewusstsein. Und das schürt dann sowohl übertriebene Ängste als auch Hoffnungen.

strategie digital: Wie lässt sich dieses Missverständnis ausräumen?

Judith Simon: Insgesamt ist es wichtig, über die Funktionsweisen, aber auch die Limitationen verschiedener KI-Systeme aufzuklären. KI Systeme lernen Muster aus Daten zur Klassifikation oder Vorhersage, oder – im Fall von Sprachmodellen – zur Produktion neuer Texte basierend auf diesen Mustern. Das kann sehr nützlich sein, ist aber eben nur Statistik – und die hängt wie alle Statistik von der Qualität und Angemessenheit der Daten und Methoden ab.

strategie digital: Sie haben die Einführung von ChatGPT einmal als „riesiges Sozialexperiment“ bezeichnet. Was meinen Sie damit und was sind Ihrer Meinung nach die wichtigsten Erkenntnisse dieses noch andauernden Experiments?

Judith Simon: Im November 2022 wurde mit ChatGPT ein noch unfertiges System auf den Markt geworfen und dann geschaut, welche Auswirkungen dies weltweit in den unterschiedlichsten Bereichen bewirkt, ohne dass die Unternehmen sich dafür verantwortlich gefühlt hätten. Im Gegenteil – es wurden dann noch Ängste vor KI von genau denen geschürt, welche diese Systeme auf den Markt geworfen haben. Das halte ich für doppelt fahrlässig.

strategie digital: Das Feld der generativen KI entwickelt sich rasant weiter – es entstehen neue Modelle, die immer leistungsfähiger und vielseitiger werden. Ihr Einsatz in Hochschulen eröffnet zahlreiche Möglichkeiten, wirft aber gleichzeitig komplexe ethische Fragen auf. Welche ethischen Fragestellungen halten Sie aktuell für die drängendsten, wenn es um den Einsatz generativer KI an Hochschulen geht?

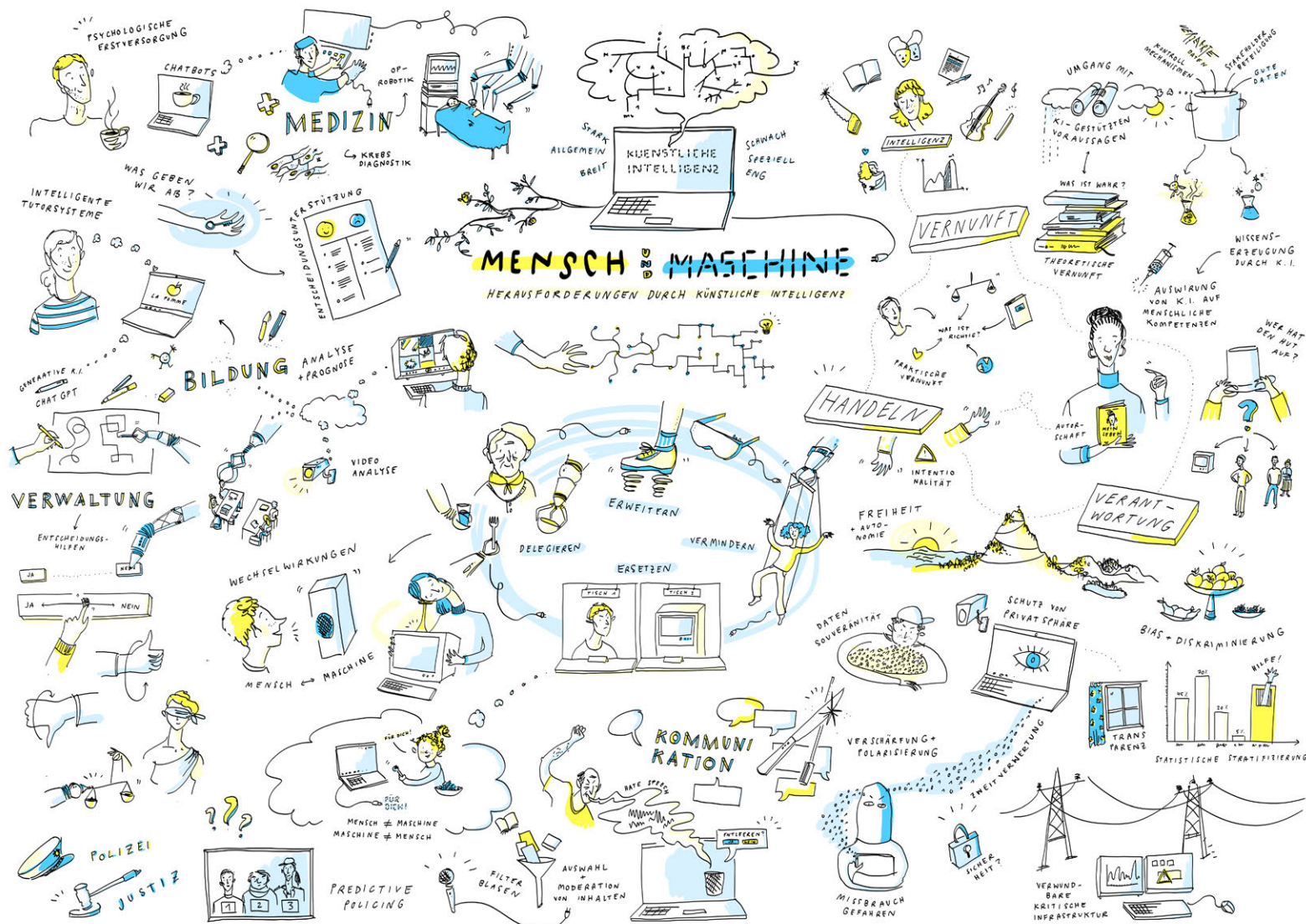
Judith Simon: Im Kontext von KI im Allgemeinen gibt es ja eine ganze Reihe von ethischen Fragen, welche sich in vielen Kontexten stellen – vom Schutz der Privatsphäre und vor Diskriminierung über Probleme mit Transparenz und Nachvollziehbarkeit bis zu Nachhaltigkeit und den Arbeitsbedingungen in der Herstellung von KI-Systemen. Auch stellt sich die Frage, was es mit uns als Individuen und als Gesellschaft macht, wenn wir die Logik der Entscheidung

über Menschen basierend auf statistischen Vorhersagen auf immer mehr Lebensbereiche ausweiten.

Darüber hinaus stellt sich für mich im Kontext der Hochschulen insbesondere die Frage des Zieles, Wertes und Sinns von Lernen und Lehren. Als ChatGPT auf den Markt kam, ging es ja erst einmal darum, wie wir einerseits Leistungen überhaupt noch prüfen können und wie dies fair geschehen kann, wenn wir nicht wissen, wer zum Schreiben der Hausarbeit ChatGPT verwendet hat und wer nicht. Mittlerweile durchdringt generative KI aber viele Prozesse des Lernens und Lehrens. Viele Studierende delegieren zumindest in Teilen Lesen und Schreiben, aber auch andere Tätigkeiten wie Recherche-

ren oder Programmieren an generative KI. Das hat meines Erachtens sehr grundlegende Auswirkungen auf Bildung, auf den Erwerb von Fähigkeiten und Fertigkeiten. Wir vergessen oft, dass wir Expertise zumeist erst durch mühsames Wiederholen und Einüben erlangen. Anders formuliert: Wie viel muss ich programmiert haben, um Fehler schnell und sicher finden zu können, sodass ich ChatGPT zum Programmieren nutzen kann und nur noch die Kontrolle selbst machen muss? Wie viele Texte muss ich gelesen und geschrieben haben, um ChatGPT zur Verbesserung meines Schreibprozesses zu nutzen? Wie können wir also diesen Expertise-Erwerb unter der Bedingung generativer KI sicherstellen, d.h. in einem Kontext, in dem die Anreize für Studierende nicht anders als

Welche Herausforderungen der zunehmende Einsatz Künstlicher Intelligenz mit sich bringt. Illustration zur Stellungnahme des Deutschen Ethikrats.
Quelle: Hannah Robold/Berliner Ideenlabor.



für den Rest der Bevölkerung darin bestehen, KI zur Effizienzsteigerung zu nutzen und den mühsamen Weg des Selbermachens zu überspringen oder zumindest abzukürzen.

Und dahinter steht dann die noch grundlegendere Frage, wofür wir Lesen, Schreiben, Fremdsprachenkenntnisse eigentlich überhaupt wollen? Warum ist das wichtig, was sollte weiter erlernt werden und was kann vielleicht weg?

Und damit sind wir bei Fragen des Ziels, Sinns und Werts von Bildung, die über den Erwerb von Skills hinausgehen, und sich um kritisches Denken, Persönlichkeitsentwicklung und die Relevanz von Bildung für Individuen, Gesellschaften und Demokratie drehen.

strategie digital: Sie sprechen davon, dass generative KI uns dazu bringt, Hochschulbildung neu zu überdenken. Wie sollten Hochschulen und Lehrende Ihrer Ansicht nach damit umgehen – gerade im Hinblick auf Prüfungsformate und die Vermittlung zentraler Kompetenzen?

Judith Simon: Ich glaube, der positivste Effekt generativer KI ist, dass sie uns dazu zwingt, uns erneut zu fragen, was Wert, Ziel und Sinn von Bildung sein sollen und mittels welcher Methoden wir diesen Idealen unter den gegebenen Bedingungen möglichst nahekommen können. Wir haben lange Lehr- und Prüfungsmodi wie z. B. Hausarbeiten verwendet, unter der Annahme, dass Studierende dadurch zeitgleich Lesen, Schreiben und die Inhalte erlernen. Heute müssen wir uns ggf. neu überlegen, welche Prioritäten wir in welchem Seminar setzen. D. h., wenn der Fokus auf Lesekompetenz und Inhalten liegt, baue ich das Seminar und die Prüfung anders auf, als wenn er auf dem Erwerb von Schreibkompetenz liegt. Im ersten Fall greife ich im Moment auf mündliche Prüfungen und handgeschriebene Spickzettel zurück, im zweiten wären In-class-Writing und Peer-Feedback zum Schreibprozess

eine Möglichkeit. Das ist alles vielleicht nicht neu, aber die Verführung generativer KI zwingt uns neu über Lehrmethoden nachzudenken.

strategie digital: Warum sind Bias und Verzerrungen in den Trainingsdaten von KI problematisch, und welche Folgen hat das für die Nutzung von KI-Sprachmodellen an Hochschulen?

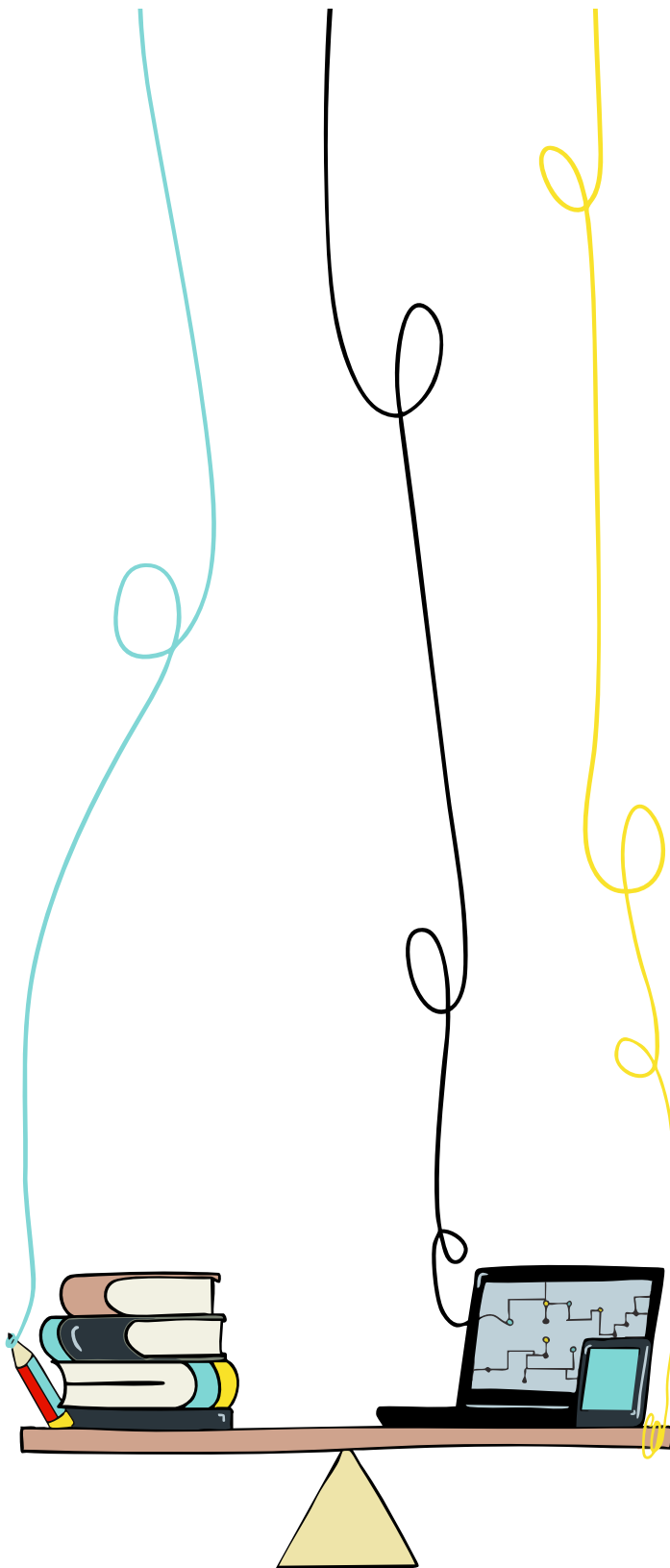
Judith Simon: Die meisten KI-Anwendungen, von denen wir heute reden, sind datenbasierte Systeme. Als solche lernen sie Muster aus alten Daten, um daraus Klassifikationen, Prognosen oder neue sprachliche oder visuelle Inhalte zu erstellen.

Denken Sie an ChatGPT:

gefüttert mit all den Daten, Texten und Bildern aus dem Internet, hat es alle Stereotypen, Ungleichheiten und Ungerechtigkeiten aus dem Netz gelernt und reproduziert diese dann in stereotypen oder diskriminierenden neuen Bildern oder Texten. Dieses Problem von sogenannten Biases und daraus möglicherweise resultierender Diskriminierung ist für alle möglichen KI-Systeme vielfach nachgewiesen. In der Informatik gibt es vielfältige Möglichkeiten, diesen Verzerrungen methodisch zu begegnen, aber diese sind alles andere als trivial und lassen sich für gewöhnlich nur kontextspezifisch beheben.

Sprachmodelle haben aber ja nun genau den Vorteil, dass man sie in jedem Kontext einsetzen kann – das macht es umso schwieriger, Verzerrungen sinnvoll zu adressieren. Dennoch müssen wir bei KI-Systemen – also nicht nur bei Sprachmodellen, sondern auch bspw. bei KI-Systemen zur Analyse und Vorhersage von Lernerfolgen, ganz zu schweigen von Gesichtserkennungssoftware oder Ähnlichem – Sorge tragen, Biases zu vermeiden, um sicherzustellen, dass bestimmte Personen und Gruppen nicht systematisch benachteiligt werden.

„Meine Rolle als Lehrende besteht darin, Studierende auf ihrem Bildungsweg, dem Weg zur Bildung ihrer Persönlichkeit und zur Ausbildung eines kritisch reflektierenden Geistes und zu verantwortlichen Mitmenschen zu unterstützen.“



Wie wandelt sich Bildung zwischen
Tradition und digitalem Fortschritt?

strategie digital: Glauben Sie, dass Hochschulen bereit sind, mit diesen ethischen Herausforderungen umzugehen?

Judith Simon: Bereit sein beinhaltet eine doppelte Voraussetzung, mit den Herausforderungen umgehen zu können und es zu wollen. Ich sehe im Moment in vielen Bereichen einen totalen Push hin zur Nutzung generativer KI in der Hochschule. Dies wird oft verbunden mit den Argumenten, man müsse Absolvent:innen auf den Arbeitsmarkt vorbereiten oder dass die Studierenden es ja ohnehin nutzen würden und man dann generative KI auch in der Lehre einsetzen müsse. Ich teile die Annahme, dass Studierende generative KI ohnehin nutzen und auch, dass im Arbeitsmarkt oft Druck herrscht, effektiver zu werden. Dadurch ist es in vielen Bereichen schwer bis unmöglich, sich der Nutzung generativer KI zu entziehen. Aber ich teile nicht die Schlussfolgerung, dass wir als Lehrende deswegen generative KI in der Lehre einsetzen müssen.

Zum einen sind da all die bereits zuvor angerissenen ethischen und erkenntnistheoretischen Probleme mit KI-Systemen – vom unzureichendem Schutz der Privatsphäre über Fragen von Autonomie, Gerechtigkeit und Diskriminierung, mangelnder Transparenz und Nachvollziehbarkeit bis hin zum nicht-existenten Wahrheitsbezug generativer KI. Auch in Bezug auf Nachhaltigkeit kann ich nicht einerseits den enormen Energie- und Ressourcenverbrauch von KI kritisieren und dann zeitgleich KI für alles Mögliche in der Lehre einsetzen. Ich glaube, es bedürfte da einer größeren Sensibilisierung, weil viele dieser Probleme für Nutzer:innen unsichtbar sind.

Zum Anderen sehe ich meine Rolle auch nicht darin, Studierende schon vorab auf etwaige ökonomische Effizienzerwägungen hin zu trimmen. Meine Rolle als Lehrende besteht darin, Studierende auf ihrem Bildungsweg, dem Weg zur Bildung ihrer Persönlichkeit und zur Ausbildung eines kritisch reflektierenden Geistes und zu verantwortlichen Mitmenschen zu unterstützen. Als Professorin für Ethik in der Informatik bedeutet dies für mich, sie über die Funktionsweisen und Limitationen von KI aufzuklären. Es kann aber auch bedeuten, mit Ihnen Leibniz oder Marx zu lesen und über Wahrscheinlichkeit, Notwendigkeit und Kontingenz sowie Macht und Markt zu sprechen.

strategie digital: Hochschulen stehen vor der Herausforderung, generative KI sinnvoll in Lehre, Forschung und Verwaltung zu integrieren. Hochschulleitungen müssen dabei ein Potpourri von verschiedenen Positionen, Einstellungen, Erwartungen und Forderungen berücksichtigen und abwägen. Können ethische Grundsätze hier als Navigationsinstrument dienen? Und welche konkreten Maßnahmen würden Sie Hochschulleitungen ans Herz legen?

Judith Simon: In der Stellungnahme des Deutschen Ethikrates „Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz“ haben wir angeregt, beim Einsatz von KI dafür Sorge zu tragen, dass der Einsatz von KI immer die Handlungsmöglichkeiten – wir haben es auch Autorschaft genannt – aller Beteiligten erweitert und nicht vermindert. Darüber hinaus haben wir zwei Dinge betont. Erstens, dass die Auswirkungen für unterschiedliche Personengruppen unterschiedlich sind – und meist am negativsten für jene, die ohnehin schon marginalisiert sind. Als Daumenregel kann daher dienen, dass wir immer prüfen sollten, wie unser Einsatz von Technik jene betrifft, welche vielleicht ohnehin marginalisiert sind und am wenigsten Macht haben. Zweitens gilt, dass der Teufel oft im Detail steckt. Dies bedeutet, dass man sich

nicht nur die Technik selbst, d.h. die Daten, Algorithmen und Methoden, genau anschauen muss, sondern auch den Einbau von KI in Institutionen und organisationale Abläufe, weil dies einen großen Einfluss darauf hat, ob wir KI angemessen vertrauen – oder vielleicht zu viel. Eine weitere Frage, welche sich natürlich nicht nur, aber auch für Hochschulen stellt, betrifft die Abhängigkeit von wenigen Tech-Unternehmen, nicht zuletzt vor dem Hintergrund geopolitischer Veränderungen. Wie wir digitale Souveränität herstellen können, ist sowohl eine institutionelle Frage als auch eine individuelle Frage. Institutionelle digitale Souveränität beinhaltet Resilienz und eine möglichst geringe Abhängigkeit von Techfirmen, individuelle Souveränität betrifft hingegen Fragen der Autonomie, welche z.B. durch invasives Tracking gefährdet ist. Zu guter Letzt würde ich mir wünschen, dass grundlegende Fragen über Wert und Ziele von Bildung in einem umfassenden Sinn bei allen Überlegungen und Entscheidungen zum Einsatz von KI an den Hochschulen im Vordergrund stehen – mit einem breiten Bildungsbegriff, welcher sich nicht nur um Effizienz, den Erwerb von Future Skills oder Ähnliches dreht, sondern Bildung als Persönlichkeitsentwicklung, als Weg zu kritischem und verantwortlichem Denken und Handeln im demokratischen Miteinander versteht. // Layout: Theresa Sommer

Mehr zur Stellungnahme des Deutschen Ethikrats

Der Deutsche Ethikrat untersucht in der Stellungnahme „Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz“ [2023], wie digitale Technologien und insbesondere KI das Selbstverständnis des Menschen, gesellschaftliches Zusammenleben und zentrale Entscheidungsprozesse verändern. Auf philosophischer Grundlage werden in der Stellungnahme Empfehlungen für einen verantwortungsvollen Einsatz in verschiedenen Bereichen formuliert – von Medizin und Bildung bis hin zu öffentlicher Kommunikation und Verwaltung.

