

KI-Folgenabschätzung für die Hochschule – verzichtbar oder notwendig?

Digitale Technologien nehmen seit langem Einfluss auf Forschung und Lehre. Wer bereits mehrere Technologieschübe auf dem digitalen Sektor erlebt hat, weiß um die dann regelmäßig einsetzenden Kontroversen zu Chancen und Risiken an Hochschulen. Diese Kontroversen sind aus mehreren Gründen richtig und wichtig: Sie machen explizit, was Hochschulangehörige und Gesellschaft aus je ihren Perspektiven für möglich, wahrscheinlich und wünschenswert halten; sie gehen mit Kritik als einem genuinen Modus der Wissenschaft einher, auf die zu verzichten einer Demontage wissenschaftlichen Denkens und Handelns gleichkäme; und sie dienen der Selbstaufklärung akademischer Praxis. Menschliche Handlungen mit digitalen Technologien haben nämlich Folgen: beabsichtigte und unbeabsichtigte, vorhersehbare und unvorhersehbare, primäre und sekundäre. Das gilt auch für Künstliche Intelligenz (KI) in der Hochschule, vor allem in ihrer Ausprägung als generative KI. Daraus wäre logisch zu folgern, dass wir eine möglichst unvoreingenommene Erforschung und Bewertung der Folgen von KI-Technologie auf Wissenschaft, Forschung und Hochschulbildung brauchen und praktizieren. Haben wir eine solche KI-Folgenabschätzung für die Hochschule? Nicht unbedingt.

In der Hochschulbildung hat der Launch von ChatGPT Ende 2022 zu großer Unsicherheit geführt und die eine oder andere überschießende Reaktion hervorgerufen – so etwa das Votum, auf schriftliche Hausarbeiten gänzlich zu verzichten. Der ersten Welle aufgeregter Debatten folgte erfreulich schnell die empirische Forschung, um bei Studierenden, Lehrenden und Forschenden Nutzungsweisen von und Einstellungen zu generativer KI zu erheben sowie Defizite im Umgang mit KI zu identifizieren. Neue Kompetenzrahmen und ein weiterer Aufschwung für „Future Skills“ ließen nicht lange auf sich warten. Politisch und wirtschaftlich geförderte Plattformen zur Hochschullehre und deren Weiterentwicklung engagieren sich mit Hochdruck für diese Themen und wirken darauf hin, die Chancen von KI ins Zentrum der Aufmerksamkeit zu rücken. Hochschulleitungen folgen dem vielfach und mahnen an, alle Beteiligten für die Arbeit mit KI „fit zu machen“. Wenn Risiken zur Sprache kommen, dann hauptsächlich unter rechtlicher, inzwischen auch finanzieller, Perspektive. Ist das ausreichend? Eher nicht.

Für die Bereiche Medizin, Schule sowie öffentliche Kommunikation und Verwaltung hat der Deutsche Ethikrat im März 2023 eine umfassende Stellungnahme zum Verhältnis zwischen Mensch und Maschine abgegeben und die Herausforderungen durch KI einer gründlichen philosophischen Analyse unterzogen. Erörtert werden die Chancen, die KI der Gesellschaft bieten kann, indem sie menschliches Handeln erweitert, ebenso wie die Risiken, die sich ergeben, wenn KI den Handlungsspielraum schmälert oder schließt. Die am Ende ausgesprochenen Empfehlungen richten sich im Kern daran aus, die Interessen des Menschen in den Mittelpunkt zu stellen und dehumanisierende Tendenzen infolge von KI zu vermeiden. Eine systematische Analyse für die Hochschule nach diesem Vorbild gibt es derzeit nicht; ernsthafte Bereitschaft, ein solches Vorhaben bis hin zu einer KI-Folgenabschätzung zu fördern, sehe ich bislang nicht.

Es ist für die Hochschulbildung naheliegend, dem humanistischen Grundgedanken zu folgen, mit dem der Deutsche Ethikrats arbeitet; jedenfalls sprechen die Zwecke von Hochschulbildung dafür: Ein akademisches Studium führt Studierende in eine Fachwissenschaft (oder mehrere) ein, bereitet diese (je nach Studiengang unterschiedlich direkt oder indirekt) auf berufliche Tätigkeiten vor und trägt im besten Fall zur Persönlichkeitsentwicklung bei, oft unter Berücksichtigung weiterer Werte wie Gerechtigkeit, Nachhaltigkeit, Diversität. Als Ideal (nicht zwingend das, was sich empirisch feststellen lässt) laufen Studium und Lehre also nach wie vor auf eine *humane Hochschulbildung* durch Wissenschaft hinaus. Diese kann und sollte den Referenzwert in der Einschätzung *aller*, positiver wie negativer, Folgen des Einsatzes von KI bilden.

Wie ließe sich das Humane in der Bildung konkreter fassen? Human ist eine Hochschulbildung, die *selbstbestimmtes Handeln* ermöglicht und befördert. Selbstbestimmtes Handeln setzt voraus, dass Menschen über dazu erforderliches Wissen und Können verfügen, prinzipiell frei entscheiden können und in der Lage sind, in einer Gemeinschaft ihren Platz zu finden und zu dieser etwas beizutragen. In seiner Stellungnahme führt der Deutsche Ethikrat aus, inwiefern philosophische Konzepte wie Handlung, Freiheit, Verantwortung und Vernunft das Verständnis selbstbestimmtes Leben ausmachen – man könnte ergänzen: auch das von selbstbestimmtem Forschen, Lehren und Lernen. Selbstbestimmung ist darüber hinaus Gegenstand psychologischer Forschung: So gelten etwa die Bedürfnisse nach Kompetenz- und Autonomieerleben sowie sozialer Eingebundenheit als grundlegend und deren Erfüllung als Bedingung für selbstbestimmtes Handeln. Schließlich gehören Selbstbestimmung und die Teilhabe an Gesellschaft bildungstheoretisch zu den Prämissen didaktischer Gestaltung – auch in der Hochschullehre.

Im Umkehrschluss wäre demnach festzuhalten: Alles, was die Entwicklung von als handlungsrelevant erachteten Kompetenzen behindert, autonomes Handeln unangemessen einschränkt und/oder soziale Kontakte und Beziehungen zerstört, ist mit einer humanen Hochschulbildung nicht vereinbar. Genau das aber sind mögliche Folgen eines unreflektierten und unkritischen KI-Einsatzes: KI an der Hochschule eröffnet nicht nur neue Handlungsspielräume, sondern kann auch Kompetenz-, Kontroll- und Sozialverluste nach sich ziehen und allmählich menschliche Handlungsmöglichkeiten reduzieren. Das gilt auch dann, wenn es sich um unbeabsichtigte Folgen und/oder Nebenfolgen an sich wünschenswerter Effekte infolge von KI handelt, die als solche möglicherweise gar nicht bewusst sind, implizit bleiben oder als irrelevant ausgeblendet werden. Was heißt das genau?

Kompetenzverluste sind möglich, wenn Menschen Wissen und Können nicht mehr einüben oder gar nicht erst aufbauen, weil Aufgaben, für die man dieses Wissen und Können benötigt, von KI erfüllt werden – unabhängig davon, ob das freiwillig oder erzwungenermaßen geschieht. Nun mag man einwenden, dass neue Technologien immer schon das für den Menschen und die Gesellschaft notwendige Kompetenzprofil verschoben haben. Dies ist ein berechtigter Einwand, der aber nur gelten kann, wenn sich guten Gewissens sagen lässt: Das brauchen wir (für den Beruf) nicht mehr zu wissen und zu können, oder: das wollen wir (als Menschen) auch nicht mehr wissen und können. Ob das etwa – um nur ein Beispiel zu nennen – bei der Fähigkeit, sich schriftlich angemessen artikulieren wie auch wissenschaftlich schreiben zu können, zutrifft, würde ich (als Didaktikerin) bezweifeln.

Kontrollverluste können eintreten, wenn Datafizierung in Lehr-, Lern- und Forschungsumgebungen dominant wird, sodass KI-induzierte Entscheidungen implizit oder explizit über den reflektierten Entscheidungen von Individuen stehen. Kontrolle verliert auch, wer nicht mehr über ausreichende kritische Urteilsfähigkeit verfügt, um die Qualität der Ausgaben von Sprachmaschinen einschätzen zu können. Urteilsfähigkeit setzt fachliches Wissen voraus; so gesehen kann Kontrollverlust eine Folge von Kompetenzverlusten sein. Vermutlich wird es in Zukunft erforderlich werden, verschiedene Grade von Kontrollverlust auf Seiten der Studierenden wie auch Lehrenden und Forschenden zu unterscheiden und abzuwägen, welche Verluste zugunsten welcher „Gewinne“ zu verkraften sind. Aus humanistischer Sicht nicht vertretbar aber dürfte eine Beziehung zwischen Mensch und Maschine sein, in der, bildlich gesprochen, nicht mehr der Mensch am Ruder sitzt, sondern ein Sprachmodell das Ruder übernimmt.

Sozialverluste können auftreten, wenn sich Menschen an interaktive und maßgeschneiderte KI-Systeme gewöhnen, deren Kommunikations- wie Feedbackangebote personalisiert sind und individuelle Optimierung ermöglichen. Studierende könnten im Zuge dessen die Interaktion und Zusammenarbeit mit Peers oder Lehrenden als weniger zielführend, zu anstrengend und unberechenbar empfinden und sukzessive reduzieren. Kommen KI-gesteuerte Avatare zum Einsatz, ist mit einer Illusion sozialer Beziehungen des Menschen zu Maschinen zu rechnen.

Sozialverluste können auch direkte Folge von Kontrollverlusten sein: Geben Studierende, Lehrende oder Forschende die Kontrolle nicht nur über Entstehungsprozesse, sondern ebenso über Endergebnisse maschinell erbrachter Leistungen aus der Hand, erodieren gleichzeitig gegenseitige Korrekturmöglichkeiten über soziale Beziehungen.

Das Nachdenken über potenzielle, in der Regel unbeabsichtigte, negative Folgen des KI-Einsatzes mag unattraktiv wirken im Vergleich zur Ausgestaltung positiver Szenarien für die Hochschule, die alte Probleme löst und neue Optionen aufzeigt. Potenzielle Verluste zu analysieren und ins „Gesamtkalkül“ mit aufzunehmen, wird schnell als konservative Reaktanz, dystopische Übertreibung oder Bremsklotz für innovative KI-Anwendungen gedeutet. Sollte aber eine humane Hochschulbildung noch unser Ziel sein, halte ich eine ernsthafte KI-Folgenabschätzung, die Chancen *und* Risiken für die Hochschule im Blick hat, für unverzichtbar. Eine solche würde nicht nur für dehumanisierende Effekte in Form von Kompetenz-, Kontroll- und Sozialverlusten sensibilisieren und diese eindämmen helfen; sie ist aus meiner Sicht auch eine Voraussetzung dafür, die immensen Potenziale von KI so einzusetzen, dass dies im Einklang mit dem Bildungsauftrag und Erkenntnisziel der Hochschule steht.

Eine ausführliche Befassung mit dem Thema einschließlich Literatur findet sich (im Open Access) in: Reinmann, G., Watanabe, A., Herzberg, D. & Simon, J. (2025). Selbstbestimmtes Handeln mit KI in der Hochschule: Forschungsdefizit und -perspektiven. Zeitschrift für Hochschulentwicklung, 20 (SH-KI-1), 33-50.
<https://doi.org/10.21240/zfhe/SH-KI-1/03>

Der Beitrag ist im Mai 2025 wie folgt erscheinen:

Reinmann, G. (2025) Verzichtbar oder notwendig? KI-Folgenabschätzung für die Hochschule. *Forschung & Lehre*, 5, 42-43.