

Ethische Perspektiven auf Künstliche Intelligenz im Kontext der Hochschule

Dirk Ifenthaler

Abstract: Im Kontext der Hochschule werden vermehrt Daten und Algorithmen zur Unterstützung von Lernen und Lehren, für Assessments, zur Weiterentwicklung von Curricula sowie zur Optimierung von Hochschulserviceangeboten eingesetzt. Frühzeitig wurde das Spannungsfeld von Künstlicher Intelligenz (KI) in der Hochschule zwischen Potentialen und ethischen Grundsätzen erkannt. Vorliegende konzeptionelle und empirische Beiträge zu Ethik und KI im Kontext der Hochschule zeigen, dass Datenschutz und Persönlichkeitsrechte einen zentralen Problembereich bei der Implementation von KI darstellen. Aus holistischer Sicht eröffnet der vorliegende Beitrag ethische Perspektiven auf KI im Hochschulbereich. Hochschulen müssen sich der Datenschutzthemen und ethischen Leitprinzipien, die in Verbindung mit KI stehen, annehmen. Ein Kernproblem beim Einsatz von KI in Hochschulen ist die Kontextabhängigkeit, Fragmentierung und Verzerrung verfügbarer Daten. Ziel der aktuellen Forschung sind Systeme mit KI, die theoretisch fundierte und transparente Datenanalysen mit pädagogisch relevanten Indikatoren und verlässlichen Interventionen ermöglichen. Es wird ein Diskurs um ethische Leitprinzipien im Zusammenhang mit KI im Kontext der Hochschule angeregt. Daraus sollen auf KI basierte Fehlentscheidungen vermieden und Schäden für Beteiligte der Hochschulen abgewendet werden.

In the context of higher education, data and algorithms are increasingly being used to support learning and teaching, for assessments, for the further development of curricula and to optimize university service offers. The tension between the potential of artificial intelligence (AI) and ethical principles was recognized at an early stage. Existing conceptual and empirical contributions on ethics and AI in the context of higher education show that data protection and personal rights represent a key concern in the implementation of AI. From a holistic point of view, this contribution opens up ethical perspectives on AI in the context of higher education. Universities need to address the privacy issues and guiding ethical principles associated with AI. Key issues interrelated with the use of AI in higher education are the contextual idiosyncrasies and dependency as well as fragmentation and distortion of available data. The aim of current research are AI systems that enable theoretically sound and transparent data analyzes with pedagogically relevant indicators and reliable interventions. A discourse on ethical guiding principles in connection with AI in the context of higher education is encouraged. This

should avoid wrong decisions based on AI and prevent damage to those involved in the higher education arena and beyond.

Keywords: Hochschule, Ethik, Künstliche Intelligenz, Learning Analytics / Higher Education, Ethics, Artificial Intelligence, Learning Analytics.

Ethische Perspektiven auf Künstliche Intelligenz im Kontext der Hochschule

Video: <https://link.transcript-open.de/5769/video/003>

© Dirk Ifenthaler



1 Einleitung

Künstliche Intelligenz (KI) hält mit einer atemberaubenden Geschwindigkeit und nahezu unbegrenzten Möglichkeiten Einzug in das Alltagsleben der Menschen. Tägliche Berührungspunkte mit KI sind u.a. Muster-, Bild- und Spracherkennung, Autovervollständigung oder Korrekturvorschläge bei digitalen Suchanfragen. Seit den 1950er Jahren wird KI nicht nur in der Ursprungswissenschaft Informatik, sondern auch interdisziplinär in Fachrichtungen wie Philosophie, Kognitionswissenschaft, Neurowissenschaft sowie Ökonomie rezipiert (Tegmark, 2018). KI bezeichnet den Versuch Maschinen zu entwickeln, welche Dinge tun können, die zuvor nur durch menschliche Kognition möglich waren (Zeide, 2019). Im Gegensatz zu Menschen können KI-Systeme jedoch eine um vielfaches größere Datenmengen in Echtzeit verarbeiten (De Laat, Joksimovic & Ifenthaler, 2020). Seit Beginn der 2010er werden auch im Kontext der Hochschule vermehrt Daten und Algorithmen zur Unterstützung von Lernen und Lehren, für Assessments, zur Weiterentwicklung von Curricula sowie zur Optimierung von Hochschulservicangeboten eingesetzt (Pinkwart & Liu, 2020). Eine systematische Übersichtsarbeit von Zawacki-Richter, Marín, Bond und Gouverneur (2019) zeigt verschiedene Anwendungsfelder von KI im Kontext der Hochschulbildung auf: (a) die Modellierung von Studieren-

dendaten, um Vorhersagen über den Studienerfolg bzw. den Studienabbruch zu treffen; (b) intelligente Tutorensysteme, welche Lernartefakte präsentieren oder Hilfestellungen und Rückmeldungen geben; (c) adaptive Systeme, welche Lernprozesse unterstützen und im Bedarfsfall Vorschläge für Lernunterstützungen anbieten; (d) automatisierte Prüfungssysteme zur Klassifikation von Lernleistungen. Darüber hinaus werden (e) Unterstützungsfunktionen im Bereich pädagogischer Entscheidungen von Lehrenden (Arthars et al., 2019) sowie der (f) Weiterentwicklung von Kursinhalten und Curricula implementiert (Ifenthaler, Gibson & Dobozy, 2018a). Zu den Potentialen von KI im Kontext der Hochschule liegen jedoch nur wenige belastbare empirische Studien hinsichtlich deren Wirkung vor (Ifenthaler & Yau, 2020). Auch stehen systemweite Implementierungen der unterschiedlichen KI-Anwendungsfelder im Hochschulkontext derzeit nach aus (Gibson & Ifenthaler, 2020; Ifenthaler, 2020a). Darüber hinaus sind KI Anwendungen im Alltag von Stakeholdern der Hochschule denkbar, wie z.B. die Unterstützung bei Bewerbungs- und Verwaltungsprozessen.

Frühzeitig wurde das Spannungsfeld von KI in der Hochschule zwischen Potentialen und ethischen Grundsätzen erkannt (Slade & Prinsloo, 2013). Ifenthaler und Tracey (2016) führten den Diskurs um ethische Fragen, Datenschutz und Privatheit von Daten im Rahmen von KI-Anwendungen weiter. Die vorliegenden konzeptionellen und empirischen Beiträge zu Ethik und KI im Kontext der Hochschule zeigen, dass Datenschutz und Persönlichkeitsrechte einen zentralen Problembereich in der Implementierung von KI darstellen.

Der vorliegende Beitrag eröffnet ethische Perspektiven auf KI im Kontext der Hochschule. Zunächst werden KI und deren Anwendungsfelder in der Hochschule beleuchtet. Darauf aufbauend werden Ethikansätze im Zusammenhang mit KI in der Hochschule diskutiert und mit praktischen Ansätzen verknüpft. Zuletzt folgt ein Ausblick und damit verbundene Forschungsperspektiven.

2 Künstliche Intelligenz in der Hochschule

Dem Einsatz von KI wird ein disruptiver Innovationsschub innerhalb der digitalen Transformation von Hochschulen beigemessen. Dabei gilt es, das Konstrukt KI näher zu bestimmen und Anwendungsfelder der KI im Kontext der Hochschule herauszuarbeiten.

2.1 Das Konstrukt Künstliche Intelligenz

Aus konzeptueller Sicht bezeichnet KI die Abfolge und Anwendung von Algorithmen, welche bestimmte Befehle für die Transformation eines Daten-Inputs zu einem Daten-Output ermöglichen. Folgt man der Definition von Graf Ballestrem, Bär,

Gausling, Hack und von Oelffen (2020) ist mit KI ein System gemeint, welches intelligentes Verhalten zeigt, indem es die Umgebung analysiert und mittels gewisser Freiheitsgrade gezielte Maßnahmen ergreift, um spezifische Ziele zu erreichen. In diesem Zusammenhang wird intelligentes Verhalten mit der menschlichen Kognition in Verbindung gebracht. Dabei stehen die menschlichen kognitiven Funktionen wie Entscheidungsfindung, Problemlösen und Lernen im Vordergrund (Bellman, 1978). KI sind folglich von Menschen entwickelte Maschinen, die komplexe Ziele (teil-)autonom erreichen können. Durch Anwendung von Techniken des maschinellen Lernens werden diese Maschinen immer mehr in die Lage versetzt, die Anwendungsumgebung und deren Kontext zu analysieren und sich an veränderte Rahmenbedingungen anzupassen (De Laat et al., 2020).

Daugherty und Wilson (2018) analysieren das Zusammenwirken von Menschen und KI. Dabei identifizierten sie drei Tätigkeitsfelder: (a) Tätigkeiten, welche von Menschen ausgeführt werden, z.B. Teams führen, Standpunkte verdeutlichen, Dinge erschaffen oder Situationen beurteilen. (b) Tätigkeiten, welche von Maschinen ausgeführt werden, z.B. Prozesse durchführen und beliebig wiederholen, Zielzustände prognostizieren oder Abläufe anpassen. Dazwischen befinden sich die (c) Mensch-Maschine-Allianzen. Menschen werden in dieser Allianz benötigt, um KI-Systeme zu entwickeln, zu trainieren und zu managen – zu befähigen. Maschinen erweitern in dieser Allianz die Fähigkeiten der Menschen, um große Datenmengen aus unzähligen Quellen in (nahezu) Echtzeit zu analysieren. In diesen Allianzen sind Menschen und Maschinen keine Konkurrenten, vielmehr werden sie zu symbiotischen Partnern, welche sich gegenseitig zu höheren Leistungen antreiben. Der Paradigmenwechsel von Computern als Werkzeuge hin zu Computern als Partnern differenziert sich in unterschiedlichen Anwendungsfeldern immer weiter aus (Wesche & Sonderegger, 2019), auch im Kontext der Bildung.

2.2 Künstliche Intelligenz im Kontext der Hochschule

Folgt man den Analysen von Bates, Cobo, Mariño und Wheeler (2020) verharrt KI im Kontext der Hochschule als schlafender Riese. Trotz der großen Aufmerksamkeit für das Thema KI in Hochschulen bleibt die praktische Anwendung von KI hinter den antizipierten Potentialen weit zurück (Buckingham Shum & McKay, 2018). Als Gründe werden Defizite in organisatorischen Strukturen sowie mangelnde personelle und technologische Ausstattungen der Hochschulen genannt (Ifenthaler, 2017). Abbildung 1 stellt für KI vier Reifegrad-Stufen und deren zusammenhängende Wirkkraft für Hochschulen vor:

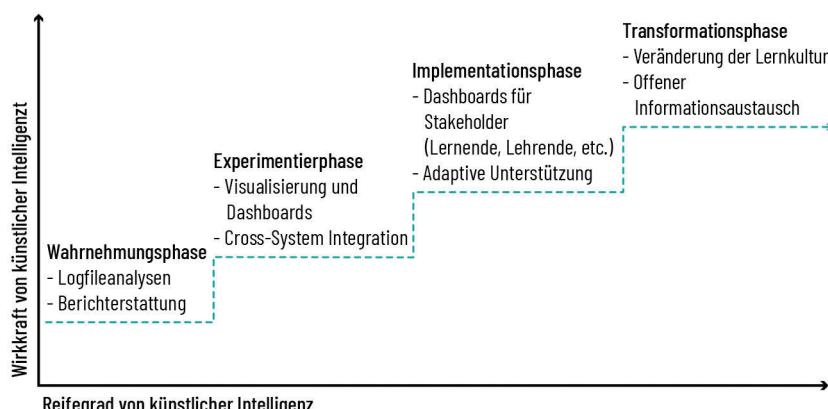
- **Wahrnehmungsphase:** Aktuell befindet sich die Mehrzahl von Hochschulen in der Wahrnehmungsphase von KI. Analysen aus Datenbanken zu Lernergebnissen

und Lernverhalten werden durchgeführt und einfache Berichte aus diesen Analysen für ausgewählte Stakeholder erstellt.

- *Experimentierphase*: In der Experimentierphase zu KI befinden sich eine ständig wachsende Zahl an Hochschulen. Dabei werden die manuell erstellen Berichte aus der Wahrnehmungsphase in Dashboards überführt und einem erweiterten Stakeholderkreis zur Verfügung gestellt. Für umfassendere Analysen werden Daten aus mehreren Systemen zusammengeführt und berichtet.
- *Implementationsphase*: Hochschulen in der Implementationsphase von KI sind nur wenige aufzufinden. In dieser Phase wird die Wirkkraft für mehrere Stakeholder (Lernende, Lehrende etc.) mittels spezieller Dashboards erreicht. Darüber hinaus werden adaptive Unterstützungen in nahezu Echtzeit (d.h. wenn Unterstützungsbedarf besteht) angeboten.
- *Transformationsphase*: Die Transformationsphase setzt eine organisationsweite Implementation von KI voraus, welche zu einer Veränderung der Lernkultur in der Hochschule führt und einen offenen Informationsaustausch unter allen Stakeholdern fördert.

Trotz der zögerlichen Implementierung birgt KI weit mehr Potentiale um die Hochschulen zu verändern, als jede Technologie zuvor. Durch KI ermöglichte Potentiale für Hochschulen sind u.a. die Erweiterung von Hochschulzugängen, die Steigerung von Studienerfolg, die Verbesserung des Studienverbleibs, die Senkung von Kosten und die Verringerung der Studiendauer. Die Anwendung von KI-Systemen im Kontext der Hochschule kann mindestens zwei Ebenen zugeordnet werden (Bates et al., 2020).

Abbildung 1: Reifegrad und Wirkkraft von künstlicher Intelligenz im Kontext der Hochschule (eigene Darstellung).



Die erste Ebene zielt auf institutionelle Prozesse. Hierzu zählen skalierbare Anwendungen zur Bewältigung von Bewerbungs- und Zulassungsverfahren (Adekitan & Noma-Osaghae, 2019) sowie KI-basierte Unterstützung für Studienberatungen und -services (Jones, 2019a). Ein weiteres Anwendungsfeld zielt auf die Identifikation von Risikostudierenden und die Vermeidung von Studienabbrechenden ab (Azcona, Hsiao & Smeaton, 2019; Hinkelmann & Jordine, 2019; Russell, Smith & Larsen, 2020).

Die zweite Ebene zielt auf die Unterstützung von Lern- und Lehrprozessen ab. Hierzu zählt die Empfehlung relevanter nächster Lernschritte und Lernmaterialien (Schumacher & Ifenthaler, 2021), die Automatisierung von Assessments und Feedback (Ifenthaler, Greiff & Gibson, 2018b), Förderung der Reflektion und des Bewusstseins über den Lernprozess (Schumacher & Ifenthaler, 2018), soziales Lernen zu unterstützen (Gašević, Joksimović, Eagan & Shaffer, 2019), unerwünschtes Lernverhalten und -schwierigkeiten aufzuspüren (Nespereira, Vilas & Redondo, 2015), den aktuellen Gefühlszustand der Lernenden ausfindig zu machen (Taub et al., 2020) sowie den Lernerfolg vorherzusagen (Glick et al., 2019). Auch werden KI-Systeme für die Qualitätssicherung von Curricula und den dazugehörigen didaktischen Arrangements (Ifenthaler et al., 2018a) sowie zur Unterstützung von Lehrenden verwendet (Arthars et al., 2019).

Um diese und weitere Implementierungen entwickeln zu können, müssen KI-Systeme auf umfangreiche Daten (Big Educational Data) aus verschiedenen Kontexten bzw. Datenquellen der Hochschulen zugreifen. Yau und Ifenthaler (2020) konnten als Ergebnis einer systematischen Übersichtsarbeit ein umfangreiches Spektrum an Indikatoren für spezifische KI-Anwendungen wie folgt identifizieren:

- Eine Gruppe von KI-Indikatoren basiert auf Daten, die durch Online-Verhalten gesammelt wurden, hauptsächlich Logfiles und Trace-Data. Dies sind Forumsinteraktionen (z.B. Beiträge, Antworten, Länge der Beiträge), der Zugriff auf Lernartefakte (z. B. ePortfolio, Vortragsfolien, Videos, Aufgaben, Selbsteinschätzungen) und die allgemeine Interaktion mit einer digitalen Lernumgebung auf der Grundlage von Logfiles.
- In ähnlicher Weise werden Daten von Websites oder Lernmanagementsystemen (z.B. ereignisbasierte Zeitstempel) in Kombination mit Noten zur Vorhersage des Abbrecherrisikos von Lernenden verwendet, wobei die detaillierte Analyse von Clickstream- oder Trace-Daten auch zur Vorhersage von Studienabbruch, Lernprozessen oder Lernleistungen verwendet wird.
- Andere Faktoren, welche für die Identifikation und Vorhersage von Lernerfolg verwendet werden basieren auf Informationen über die Lernenden wie z.B. demografische Daten (z.B. Alter, Geschlecht), sozioökonomischer Status (z.B.

Familieneinkommen, Hintergrund, Ausgaben) oder vorherige akademische Erfahrung und Leistung.

- Zusätzlich zu demografischen Variablen werden das akademische Selbstkonzept von Lernenden, die akademische Historie und arbeitsbezogene Daten zur Vorhersage der Lernleistung verwendet.
- Andere KI-Systeme analysieren den Notendurchschnitt (GPA), die akademische Belastung und den Zugang zur Beratung, den finanziellen Hintergrund der Lernenden oder vorhergehende akademische Leistungen.
- Daten, welche durch Umfragen erhoben wurden, wie z.B. die Selbstauskunft der Lernenden über erwartete Noten, die (Lern-)Motivation sowie die akademische und technologische Kompetenz bilden eine weitere Gruppe der Datenanalysen.
- Schließlich konnten Faktoren identifiziert werden, welche auf multimodaler Basis gebildet werden, d.h. sie verwenden Daten aus verschiedenen Quellen, wie z.B. Logfiles oder Trace-Data (nicht reaktive Datenerhebung), Bewertungen und Umfragedaten (reaktive Datenerhebung), sowie aus aggregierten Informationen oder historischen Daten.

Im Hinblick auf die Vielzahl und Vielfältigkeit der notwendigen Daten für KI-Systeme, müssen sich Stakeholder der Hochschulen der Themen annehmen, die in Verbindung mit KI stehen (Ifenthaler & Tracey, 2016; West, Huijser & Heath, 2016). Dazu gehören Zugriffsrechte, Speicherdauer, Analysen und Schlussfolgerungen. KI-Systeme greifen dabei auf Daten aus verschiedenen Kontexten zu, etwa aus der Studierenden-, Studien- und Prüfungsverwaltung, der Lernumgebung oder sozialen Interaktionen. Die Bereitschaft Daten preiszugeben, kann in den jeweiligen Kontexten unterschiedlich ausgeprägt sein. Entsprechend den Annahmen der ‚kontextuellen Integritätstheorie‘ sind Informationen, die in einem spezifischen Kontext preisgegeben wurden, nicht übertragbar auf einen anderen Kontext, ohne die Bedeutung zu beeinträchtigen oder die Privatsphäre zu verletzen (Ifenthaler & Schumacher, 2016). Datenschutzprinzipien für KI-Systeme unterstreichen dabei die aktive Rolle der Studierenden in ihren Lernprozessen, den temporären Charakter und die Unvollständigkeit von vorhandenen Daten sowie insbesondere Transparenz hinsichtlich Nutzung, Analysen, Zweck, Zugriff, Kontrolle und Eigentumsverhältnisse der anfallenden Daten (Hoel & Chen, 2018; Jones, 2019b). Neben Standards zur Sicherung der Privatsphäre und zum Datenschutz unter Einhaltung der EU-DSGVO sowohl für Einzelpersonen als auch für die Hochschule werden Richtlinien zur Einhaltung von ethischen Gesichtspunkten benötigt.

3 Ethische Perspektiven auf künstliche Intelligenz in der Hochschule

Neben datenschutzrechtlichen Anforderungen an künstliche Intelligenz ethische Aspekte eine notwendige Basis für aktuelle und zukünftige Implementierungen im Kontext der Hochschule. Zum ethischen Grundverständnis im Umgang KI gehören Rechenschaft, Verantwortung und Transparenz. Auch Bedarf es einfacher Instrumente für Hochschulen, um das Bewusstsein der involvierten Stakeholder rund um Ethik in Verbindung mit KI zu steigern.

3.1 Ethikansätze

Ethikansätze fanden ihren Ursprung bei den antiken griechischen Philosophen, welche sich mit der Frage beschäftigt haben, wie durch moralisch verantwortungsvolles Handeln ein gutes Leben geführt werden kann. Dabei bezieht sich diese grundsätzliche ethische Fragestellung sowohl auf die Einzelperson als auch auf Personengruppen, Organisationen bzw. die Gesellschaft. Folglich ist Ethik die Theorie der Moral und Moral die praktische Anwendung von Ethik (Hasenbein, 2020). Ethische Grundsätze und moralische Entscheidungen finden sich im gesamten Spektrum menschlichen Handelns. Eine sehr vereinfachte Konzeption von Ethik berücksichtigt moralische Entscheidungen, welche von einzelnen Personen bzw. der Gesellschaft als gut bzw. richtig betrachtet werden. Demgegenüber stehen schlechte bzw. falsche moralische Entscheidungen. Die konsequentialistische Ethik bevorzugt diejenigen Handlungen, welche für einen möglichst großen Personenkreis positive Auswirkungen erzeugen. Die deontologische Ethik bemüht sich hingegen um Regeln und Pflichten für spezifische Handlungssituationen. Dennoch entstehen für Beteiligte in moralischen Entscheidungssituationen immer wieder Dilemmata, deren Auflösung nicht immer vollumfänglich möglich ist (Kvalnes, 2015).

Seit geraumer Zeit beschäftigen sich Ethikansätze auch mit dem digitalen Wandel (Berberich, 2019). Von zentralem Interesse ist das moralisch angemessene Verhalten mit und von digitalen Systemen. Die sogenannte digitale Ethik bzw. KI-Ethik fokussiert Fragen zum guten und richtigen Leben und Zusammenleben in einer von digitalen Technologien und KI-Systemen geprägten Welt (BVDW, 2019). Im Zentrum ethischer Fragen im Zusammenhang mit KI steht die Fähigkeit der Maschine, moralisch zu handeln. Durch moralisch geprägte Algorithmen soll den Maschinen die Möglichkeit gegeben werden, Prozesse und Analysen entlang ethischer Prinzipien zu regulieren. Die umfassende Fragestellungen der KI-Ethik werden von interdisziplinär Forschenden und zukünftig nutzenden Personen aus den Anwendungsbereichen diskutiert. Ziel des Diskurses muss es sein, ethische Leitlinien für moralisch akzeptiertes Handeln mit und für KI-Systeme zu entwickeln (Hasenbein, 2020).

Mit der Anwendung von KI-Systemen bleiben deren Ergebnisse und Vorhersagen für Einzelpersonen oder die Gesellschaft nicht immer erklärbar und im Einzelnen vollständig nachvollziehbar. Dellermann, Ebel, Söllner und Leimeister (2019) plädieren daher für eine erklärbare, faire, sichere und transparente KI (engl. Explainable Artificial Intelligence). Aus diesem Anspruch werden KI-Systeme gefordert, welche für die Beteiligten jederzeit nachvollziehbar sind und somit eine Diskriminierung ausschließen. Der Diskurs um ethische Fragen zu KI lässt sich auf spezifische Anwendungsfelder und -situationen übertragen. Dazu zählen KI-basierte Alltagsanwendungen wie die Wahl eines TV-Programms für Kinder, Alltagssituationen wie die eines autonomen Fahrzeugs im Straßenverkehr, komplexe medizinische KI-Anwendungen im Intensivbereich aber auch vielfältige KI-Anwendungen und Situationen in der Hochschulbildung.

3.2 Ethisch reflektierte Anwendung künstlicher Intelligenz im Kontext der Hochschule

Charakteristisch für KI-Systeme im Kontext der Hochschule sind deren Autonomie, Interaktivität und Adaptabilität. Diese Eigenschaften ermöglichen eine effektive Bewältigung der dynamischen und vielfach unvollständig erfassbaren Lern-Lehr-Prozesse. Jedoch sind KI-Systeme mit diesen Eigenschaften schwer einzuschätzen und deren Vorhersagen bzw. Empfehlungen können zu unerwartetem Verhalten bzw. ungewollten Aktivitäten führen (Black-Box). Slade und Prinsloo (2013) haben in diesem Zusammenhang holistische Kriterien für KI-Anwendungen im Kontext der Hochschule formuliert. Dazu zählen Transparenz, Besitz, Zugang und Kontrolle über Daten, Validität und Reliabilität von Daten, institutionelle Kommunikation, Verantwortung und Verbindlichkeit, Inklusion und Einverständnis sowie kulturelle Werte und Handlungskompetenz der Stakeholder. Dignum (2017) integriert die Vielzahl ethischer Kriterien in den sogenannten ART-Prinzipien (Accountability, Responsibility, Transparency).

- Rechenschaft (Accountability) bezieht sich auf die Anforderung, KI-Entscheidungen und KI-Handlungen allen Beteiligten zu erklären und zu rechtfertigen. Dafür müssen KI-Entscheidungen aus den verwendeten Algorithmen ableitbar und nachvollziehbar sein. Die Rechenschaft ist notwendig, um Vertrauen in die Verwendung von KI zu fördern und wird auch im Rahmen der EU-DSGVO gefordert. Die Rechenschaft sollte auf moralischen und sozialen Prinzipien beruhen, einschließlich geltender Werte, sozialer Normen, organisatorischer Gebräuche und individueller Motive und Ziele (Miller, 2019).
- Verantwortung (Responsibility) bezieht sich auf die Rolle der Beteiligten selbst im Umgang mit KI-Systemen. Verantwortung meint dabei jedoch nicht nur das Regulieren von intelligenten Maschinen, vielmehr muss geklärt werden, wie mit

aus KI-Systemen erzeugten Daten umgegangen wird und in welchem Umfang diese geteilt werden können. Verantwortung bedeutet nachhaltige Strukturen für Mensch-Maschine-Allianzen in einer KI-geprägten Gesellschaft zu etablieren.

- Transparenz (Transparency) bezieht sich auf die Notwendigkeit, die Mechanismen zu beschreiben, zu prüfen und zu reproduzieren, mittels derer KI-Systeme Lernen und Entscheidungen treffen. Um Klarheit über Daten und Algorithmen zu schaffen, bedarf es speziellen Methoden, welche KI-Prozesse und deren Ergebnisse für die Beteiligten analysierbar machen. Transparenz erfordert die Minimierung bzw. Ablösung der algorithmischen Black-Box.

Hochschulen müssen sich der ART-Prinzipien annehmen und einen verantwortungsvollen, transparenten sowie erklärbaren Umgang mit KI-Systemen gewährleisten. Erste Studienergebnisse weisen darauf hin (Howell, Roberts, Seaman & Gibson, 2018; Ifenthaler & Schumacher, 2016; West et al., 2016), dass Studierende trotz antizipierter Vorteile nicht bereit sind, alle Daten für KI-Anwendungen preiszugeben. Es wird zwar Bereitschaft signalisiert, lernbezogene Daten zu teilen, nicht aber persönliche Informationen oder soziale Nutzerpfade. Insbesondere bei der Implementierung der vielseitig geforderten adaptiven KI-Systeme, welche auf eine Vielzahl an Daten angewiesen sind, bleibt dies ein kritischer Aspekt.

Richards und Dignum (2019) schlagen einen wertefokussierten Design-Ansatz vor, welcher ethische Prinzipien in jeder Phase der Entwicklung und Verwendung von KI-Systemen für Hochschulen berücksichtigt. Diesem Ansatz folgend müssen für KI-Systeme an Hochschulen (a) relevante Stakeholder bestimmt werden; (b) Werte und Anforderungen der Beteiligten ermittelt werden; (c) Möglichkeiten zur Aggregation der Werte und Werteinterpretation aller Beteiligten bereitgestellt werden; (d) Verknüpfung von Werten und Systemfunktionalitäten zur Unterstützung von Implementierungsentscheidungen sowie einer nachhaltigen Anwendung gewährleistet werden; (e) Unterstützung bei der Auswahl von Systemkomponenten (von innerhalb oder außerhalb der Organisation) vor dem Hintergrund ethischer Prinzipien angeboten werden.

Werden zukünftige KI-Systeme in Entscheidungsprozesse der Hochschulen involviert, z.B. in Zulassungsverfahren oder für automatisierte Prüfungsverfahren, erfolgt eine Übertragung der Entscheidungsprozesse auf KI-Systeme (Prinsloo & Slade, 2014; West et al., 2016; Willis, Slade & Prinsloo, 2016). Die Verantwortung wird an Algorithmen abgegeben. Hochschulen und deren Beteiligte werden folglich vor dem Hintergrund ethischer Prinzipien entscheiden, ob diese Verantwortung an KI delegiert werden kann. Gleichzeitig müssen die Beteiligten an der jeweiligen Hochschule beurteilen, in wieweit KI-Systeme in der Lage sind, Verantwortung für die getroffenen Entscheidungen zu übernehmen.

3.3 Checkliste für Ethik und Datenschutz an Hochschulen

Schließlich liegt es an den einzelnen Hochschulen, ethische und datenschutzrechtliche Richtlinien und Maßnahmen für die Nutzung von KI einzuführen (Hoel & Chen, 2018; Jones, 2019b). Die sogenannte DELICATE (Determination, Explain, Legitimate, Involve, Consent, Anonymise, Technical, External) Checkliste (Drachsler & Grelle, 2016; Ifenthaler & Drachsler, 2020) bietet ein einfaches Instrument für Hochschulen, um das Bewusstsein der Stakeholder einer Hochschule rund um Ethik und Datenschutz in Verbindung mit KI zu steigern. DELCIATE möchte helfen, das Thema zu demystifizieren und aus der komplexen Welt von juridischen Texten zu extrahieren, um es in der Hochschule verständlich und praktikabel zu machen (Ifenthaler & Drachsler, 2020). In Anlehnung an die DELICATE Checkliste sollen Hochschulen mit KI-Systemen folgende Prinzipien umsetzen (Ifenthaler, 2020b):

- *Begründung* (Determination): Das Potential von KI und Datenanalysen im Allgemeinen sowie Learning Analytics im Besonderen wird aus Sicht der Organisation und der einzelnen Stakeholdergruppen verdeutlicht. Aktuell geltende Datenschutzrechte werden umgesetzt.
- *Erklärung* (Explain): Eine transparente Speicherung und Analyse von Daten und Klarstellung derer Verwendungszwecke wird sichergestellt. Auskünfte bezüglich Speicherdauer und Zugriffsrechten werden in allen Fällen der Datenspeicherung bzw. -analyse gewährleistet.
- *Legitimation* (Legitimate): Die Notwendigkeit der Datennutzung sowie die Erlaubnis auf Daten zugreifen zu können sind legitimiert.
- *Einbeziehung* (Involve): Jegliche Datenschutzbedenken werden transparent kommuniziert und vollständig geklärt. Individueller Zugang zu gespeicherten Daten wird ermöglicht. Informations- und Weiterbildungsangebote für alle Stakeholder sind verfügbar.
- *Einverständnis* (Consent): Individuelle Stakeholder müssen aktiv der Datennutzung, -speicherung und -analyse zustimmen (opt-in). Das Einverständnis der Datennutzung kann jederzeit widerrufen werden.
- *Anonymisierung* (Anonymise): Daten werden im höchstmöglichen Maß anonymisiert, pseudonymisiert oder aggregiert.
- *Aktualisierung* (Technical): Der Datenzugriff wird permanent geprüft. Bei technischen oder organisatorischen Veränderungen wird ein erneutes Einverständnis bei den Stakeholdern eingeholt. Technische Systeme sind auf aktuellem technologischen Stand.
- *Distribution* (External): Externe Zugriffe auf Daten und Analyseergebnisse sind zweckgebunden und eindeutig geregelt. Die geltenden Regularien und Gesetze für die Hochschule müssen auch von externen Stakeholdern eingehalten werden.

4 Ausblick

KI in der Hochschule verbindet sozio-technologische Data-Mining-, Analyse- und Interventionspraktiken mit dem Ziel, individuelle und systemische Bildungsprozesse zu unterstützen. Wie in systematischen Übersichtsarbeiten dokumentiert (Ifenthaler & Yau, 2020; Larrabee Sønderlund, Hughes & Smith, 2018; Zawacki-Richter et al., 2019), sollten KI-Systeme aktives und selbstorganisiertes Lernen ermöglichen, z.B. durch adaptive Hilfestellungen oder durch die Unterstützung von Lehrpersonen beim Kuratieren und Bearbeiten von Daten über die Lernenden und Lerninhalte.

Ein Kernproblem des Einsatzes von KI in Hochschulen ist die Kontextabhängigkeit, Fragmentierung und Verzerrung verfügbarer Daten. KI-basierte Entscheidungen, welche aus umfangreichen Trainingsdaten hervorgehen, können bei unzureichender Datengrundlage und mangelnder Transparenz zu Einseitigkeiten oder Befangenheiten führen. Eine Diskriminierung von Personen durch KI-Systeme ist somit nicht mehr auszuschließen.

Folglich sind holistische KI-Systeme, die theoretisch fundierte und transparente Datenanalysen mit pädagogisch relevanten Indikatoren und verlässlichen Interventionen ermöglichen, Ziel der aktuellen Forschung. Dabei ist zu erwarten, dass neben bereits bestehenden datenschutzrechtlichen Standards auch ethische Leitprinzipien zum Austausch und der Analyse von Daten aus dem Bildungskontext weiterentwickelt werden. Aus Sicht des Konsequentialismus sind KI-Systeme im Kontext der Hochschule ethisch vertretbar, sollten diese positive Ergebnisse für die Beteiligten erzielen. Folgt man der Deontologie sind KI-Systeme und deren autonome Entscheidungen über Bildungsprozesse grundsätzlich unethisch.

Der Diskurs um ethische Leitprinzipien im Zusammenhang mit KI im Kontext der Hochschule muss breit und tief geführt werden. Nur dann können KI-basierte Fehlentscheidungen vermieden und Schäden für Beteiligte der Hochschulen abgewendet werden. Es gilt, die Potentiale des Zusammenwirkens von Menschen und KI kritisch zu reflektieren. Während die KI dem Menschen in Kapazität und Geschwindigkeit bei Datenanalysen und -prognosen überlegen ist, überzeugt der Mensch aufgrund seiner Empathiefähigkeit bei ethischen Urteilen. Daraus kann ein Konsens für die ethisch verantwortungsvolle Verwendung von KI in der Hochschule etabliert werden. Dieser Konsens befindet sich irgendwo zwischen übertriebener Vorsicht und unkalkulierbaren Risiken.

Dirk Ifenthaler: Universität Mannheim, Chair of Economic and Business Education – Learning, Design and Technology; Curtin University, UNESCO Deputy Chair of Data Science in Higher Education Learning and Teaching; dirk@ifenthaler.info, ifenthaler@uni-mannheim.de

Literatur

- Adekitan, A. I. & Noma-Osaghae, E. (2019). Data mining approach to predicting the performance of first year student in a university using the admission requirements. *Education and Information Technologies*, 24, 1527–1543.
- Arthars, N., Dollinger, M., Vigentini, L., Liu, D. Y., Kondo, E. & King, D. M. (2019). Empowering teachers to personalize learning support. In D. Ifenthaler, D.-K. Mah & J. Y.-K. Yau (Hg.), *Utilizing learning analytics to support study success* (S. 223–248). Cham: Springer.
- Azcona, D., Hsiao, I. & Smeaton, A. F. (2019). Detecting students-at-risk in computer programming classes with learning analytics from students' digital footprints. *User Modeling and User-Adapted Interaction*, 29, 759–788.
- Bates, T., Cobo, C., Mariño, O. & Wheeler, S. (2020). Can artificial intelligence transform higher education? *International Journal of Educational Technology in Higher Education*, 17(42), 1–12.
- Bellman, R. (1978). *An introduction to artificial intelligence: can computers think?* San Francisco, CA: Boyd & Fraser.
- Berberich, N. (2019). Künstliche Intelligenz und Ethik. KI oder nicht KI? Das ist hier nicht die Frage. In K. Kersting, C. Lampert & C. Rothkopf (Hg.), *Wie Maschinen lernen* (S. 229–239). Wiesbaden: Springer.
- Buckingham Shum, S. & McKay, T. A. (2018). Architecting for learning analytics. Innovating for sustainable impact. *EDUCAUSE Review*, 53(2), 25–37.
- BVDW. (2019). *Mensch, Moral, Maschine. Digitale Ethik, Algorithmen und künstliche Intelligenz*. Berlin: Bundesverband Digitale Wirtschaft (BVDW) e.V.
- Daugherty, P. R. & Wilson, H. J. (2018). *Human + machine: Reimagining work in the age of AI*. Boston, MA: Harvard Business Review Press.
- De Laat, M., Joksimovic, S. & Ifenthaler, D. (2020). Artificial intelligence, real-time feedback and workplace learning analytics to support in situ complex problem-solving: a commentary. *International Journal of Information and Learning Technology*, 37(5), 267–277.
- Dellermann, D., Ebel, P., Söllner, M. & Leimeister, J. M. (2019). Hybrid intelligence. *Business & Information Systems Engineering*, 61(5), 637–643.
- Dignum, V. (2017). *Responsible autonomy*. Paper presented at the Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, Melbourne, VIC, AUS.
- Drachsler, H. & Grelle, W. (2016). Privacy and analytics – it's a DELICATE issue. A checklist for trusted learning analytics, *Sixth International Conference on Learning Analytics & Knowledge*. Edinburgh, UK.
- Gašević, D., Joksimović, S., Eagan, B. R. & Shaffer, D. W. (2019). SENS: Network analytics to combine social and cognitive perspectives of collaborative learning. *Computers in Human Behavior*, 92, 562–577.

- Gibson, D. C. & Ifenthaler, D. (2020). Adoption of learning analytics. In D. Ifenthaler & D. C. Gibson (Hg.), *Adoption of data analytics in higher education learning and teaching* (S. 3–20). Cham: Springer.
- Glick, D., Cohen, A., Festinger, E., Xu, D., Li, Q. & Warschauer, M. (2019). Predicting success, preventing failure. In D. Ifenthaler, D.-K. Mah & J. Y.-K. Yau (Hg.), *Utilizing learning analytics to support study success* (S. 249–273). Cham: Springer.
- Graf Ballestrem, J., Bär, U., Gausling, T., Hack, S. & von Oelffen, S. (2020). *Künstliche Intelligenz. Rechtsgrundlagen und Strategien in der Praxis*. Wiesbaden: Springer Gabler.
- Hasenbein, M. (2020). Ethik in Zeiten von Digitalisierung und künstlicher Intelligenz. In M. Hasenbein (Hg.), *Der Mensch im Fokus der digitalen Arbeitswelt* (S. 183–200). Berlin: Springer.
- Hinkelmann, M. & Jordine, T. (2019). The LAPS project: using machine learning techniques for early student support. In D. Ifenthaler, J. Y.-K. Yau & D.-K. Mah (Hg.), *Utilizing learning analytics to support study success* (S. 105–117). Cham: Springer.
- Hoel, T. & Chen, W. (2018). Privacy and data protection in learning analytics should be motivated by an educational maxim—towards a proposal. *Research and Practice in Technology Enhanced Learning*, 13–20.
- Howell, J. A., Roberts, L. D., Seaman, K. & Gibson, D. C. (2018). Are we on our way to becoming a »helicopter university«? Academics' views on learning analytics. *Technology, Knowledge and Learning*, 23(1), 1–20.
- Ifenthaler, D. (2017). Are higher education institutions prepared for learning analytics? *TechTrends*, 61(4), 366–371.
- Ifenthaler, D. (2020a). Change management for learning analytics. In N. Pinkwart & S. Liu (Hg.), *Artificial intelligence supported educational technologies* (S. 261–272). Cham: Springer.
- Ifenthaler, D. (2020b). Gelingensbedingungen zum Einsatz von Learning Analytics. In K. Wilbers (Hg.), *Handbuch E-Learning* (85, S. 1–16). Köln: Wolters Kluwer.
- Ifenthaler, D. & Drachsler, H. (2020). Learning Analytics. In H. M. Niegemann & A. Weinberger (Hg.), *Lernen mit Bildungstechnologien* (S. 515–534). Heidelberg: Springer.
- Ifenthaler, D., Gibson, D. C. & Dobozy, E. (2018a). Informing learning design through analytics: Applying network graph analysis. *Australasian Journal of Educational Technology*, 34(2), 117–132.
- Ifenthaler, D., Greiff, S. & Gibson, D. C. (2018b). Making use of data for assessments: harnessing analytics and data science. In J. Voogt, G. Knezek, R. Christensen & K.-W. Lai (Hg.), *International handbook of IT in primary and secondary education* (2, S. 649–663). New York, NY: Springer.
- Ifenthaler, D. & Schumacher, C. (2016). Student perceptions of privacy principles for learning analytics. *Educational Technology Research and Development*, 64(5), 923–938.

- Ifenthaler, D. & Tracey, M. W. (2016). Exploring the relationship of ethics and privacy in learning analytics and design: implications for the field of educational technology. *Educational Technology Research and Development*, 64(5), 877–880.
- Ifenthaler, D. & Yau, J. Y.-K. (2020). Utilising learning analytics to support study success in higher education: a systematic review. *Educational Technology Research and Development*, 68(4), 1961–1990.
- Jones, K. M. L. (2019a). Advising the whole student: eAdvising analytics and the contextual suppression of advisor values. *Education and Information Technologies*, 24, 437–458.
- Jones, K. M. L. (2019b). Learning analytics and higher education: a proposed model for establishing informed consent mechanisms to promote student privacy and autonomy. *International Journal of Educational Technology in Higher Education*, 16–24.
- Kvalnes, Ø. (2015). Moral dilemmas. In Ø. Kvalnes (Hg.), *Moral reasoning at work: re-thinking ethics in organizations* (S. 9–17). London: Palgrave Macmillan.
- Larrabee Sønderlund, A., Hughes, E. & Smith, J. (2018). The efficacy of learning analytics interventions in higher education: A systematic review. *British Journal of Educational Technology*, 50(5), 2594–2618.
- Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1–38.
- Nespereira, C., Vilas, A. & Redondo, R. (2015). Am I failing this course?: risk prediction using e-learning data, *Conference on Technological Ecosystems for enhancing Multiculturality* (S. 271–276).
- Pinkwart, N. & Liu, S. (Hg.). (2020). *Artificial intelligence supported educational technologies*. Cham: Springer.
- Prinsloo, P. & Slade, S. (2014). Student data privacy and institutional accountability in an age of surveillance. In M. E. Menon, D. G. Terkla & P. Gibbs (Hg.), *Using data to improve higher education. Research, policy and practice* (S. 197–214). Rotterdam: Sense Publishers.
- Richards, D. & Dignum, V. (2019). Supporting and challenging learners through pedagogical agents: Addressing ethical issues through designing for values. *British Journal of Educational Technology*, 50(6), 2885–2901.
- Russell, J.-E., Smith, A. & Larsen, R. (2020). Elements of Success: Supporting at-risk student resilience through learning analytics. *Computers & Education*, 152.
- Schumacher, C. & Ifenthaler, D. (2018). The importance of students' motivational dispositions for designing learning analytics. *Journal of Computing in Higher Education*, 30(3), 599–619.
- Schumacher, C. & Ifenthaler, D. (2021). Investigating prompts for supporting students' self-regulation – A remaining challenge for learning analytics approaches? *The Internet and Higher Education*, 49, 100791.

- Slade, S. & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, 57(10), 1510–1529.
- Taub, M., Azevedo, R., Rajendran, R., Cloude, E. B., Biswas, G. & Price, M. J. (2020). How are students' emotions related to the accuracy of cognitive and metacognitive processes during learning with an intelligent tutoring system? *Learning and Instruction*.
- Tegmark, M. (2018). *Life 3.0: Being human in the age of artificial intelligence*. London: Penguin Books.
- Wesche, J. S. & Sonderegger, A. (2019). When computers take the lead: The automation of leadership. *Computers in Human Behavior*, 101, 197–209.
- West, D., Huijser, H. & Heath, D. (2016). Putting an ethical lens on learning analytics. *Educational Technology Research and Development*, 64(5), 903–922.
- Willis, I. J. E., Slade, S. & Prinsloo, P. (2016). Ethical oversight of student data in learning analytics: a typology derived from a cross-continental, cross-institutional perspective. *Educational Technology Research and Development*, 64(5), 881–901.
- Yau, J. & Ifenthaler, D. (2020). Reflections on different learning analytics indicators for supporting study success. *International Journal of Learning Analytics and Artificial Intelligence for Education*, 2(2), 4–23.
- Zawacki-Richter, O., Marín, V. I., Bond, M. & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16(39), 1–27.
- Zeide, E. (2019). Artificial intelligence in higher education: applications, promise and perils, and ethical questions. *EDUCAUSE Review*, 54(3), 21–39.