

Die Interdependenz von Learning Analytics und Studierenden

Implikationen studentischer Wahrnehmungen für eine sozialverantwortliche KI-Implementierung¹

Birte Keller,² Marco Lünich³ und Frank Marcinkowski⁴

Ein sozialverantwortlicher Einsatz von auf Künstlicher Intelligenz (KI) basierender Learning Analytics an Hochschulen erfordert die Einbindung aller relevanten Stakeholder, insbesondere der Studierenden als direkt Betroffene des Systems. Konzeptionelle Überlegungen und empirische Erkenntnisse aus fünf Jahren Forschung zu studentischen Wahrnehmungen von KI an deutschen Hochschulen liefern wichtige Impulse für die Gestaltung eines iterativen Prozesses zur fairen und sozialverantwortlichen Entwicklung von LA-Systemen. Dabei werden studentische Schadens- und Fairnesswahrnehmungen gegenüber KI-basierten Leistungsprognosesystemen sowie individuelle Anforderungen an das Design dieser soziotechnischen Systeme und Präferenzen der Verteilung von Fördermaßnahmen untersucht. Die gewonnenen Erkenntnisse werden vor dem Hintergrund des internationalen Forschungsstands kritisch diskutiert, um Implikationen für eine sozialverantwortliche KI-Implementierung abzuleiten.

The interdependence of learning analytics and students: Implications of student perceptions for socially-responsible AI implementation

The socially responsible use of artificial intelligence (AI)-based learning analytics (LA) at universities requires the involvement of all relevant stakeholders, especially students as those directly affected by the system. Conceptual considerations and empirical findings from five years of research on student perceptions of AI at German universities

1 Basiert auf einem Impulsbeitrag im Rahmen der Tagung.

2 ORCID-ID: 0000-0002-3145-5206

3 ORCID-ID: 0000-0002-0553-7291

4 ORCID-ID: 0000-0001-6497-9324

provide important impulses for the design of an iterative process for the fair and socially responsible development of LA systems. Student perceptions of harm and fairness towards AI-based performance prediction systems as well as individual requirements for the design of these socio-technical systems and preferences for the distribution of support measures are examined. The findings are critically discussed against the background of the international state of research in order to derive implications for a socially responsible AI implementation.

Einleitung

Ein sozialverantwortlicher Einsatz von KI-basierten Learning Analytics (LA) im tertiären Bildungsbereich, der fair, transparent und inklusiv gestaltet ist, erfordert die Einbindung aller relevanten Stakeholdergruppen (Keller/Lünich/Marcinkowski 2022; Lang/Davis 2023; Tsai/Whitelock-Wainwright/Gašević 2021). Neben Entwickler:innen und Hochschulverantwortlichen sind insbesondere die Studierenden von Bedeutung, da ihre Bildungswege durch Entscheidungen der vermeintlich objektiven Computerverfahren unmittelbar beeinflusst werden. Ethische Bedenken, wie Risiken der Diskriminierung (Simis/Weydner-Volkman 2024; Slade/Prinsloo 2013), machen iterative Prozesse zur sozialverantwortlichen Gestaltung solcher Systeme unverzichtbar. Andernfalls drohen die Reproduktion gesellschaftlicher Ungleichheiten und die Ablehnung des KI-Systems, wie ein Benotungsalgorithmus in Großbritannien zeigte, der Schüler:innen sozial schwächerer Familien benachteiligte (Smith 2020). Dieses Negativbeispiel verdeutlicht, dass die Einbettung von LA in den sozialen Kontext der Hochschulen und die Lebensrealität der Studierenden essenziell ist. LA-Systeme verarbeiten studentische Daten und beeinflussen gleichzeitig ihre Zukunft, was soziale Konflikte hervorrufen kann. Die Einbeziehung studentischer Wahrnehmungen in die Entwicklung fairer und sozialverantwortlicher Systeme ist daher zentral, insbesondere angesichts zentraler Herausforderungen der strategischen Planung, des Implementierungsmonitorings und der Evaluation pädagogischer Ziele (Tsai/Gasevic 2017; Tsai et al. 2020).

Im Folgenden werden Ergebnisse aus fünf Jahren empirischer Forschung zu studentischen Wahrnehmungen von KI-Anwendungen mit Fokus auf Leistungsprognosesysteme (Academic Performance Prediction Systems, kurz: APP) an deutschen Hochschulen vorgestellt. APP zählt zu den am meistbeachteten KI-Systemen im Bildungswesen (Zawacki-Richter et al. 2019; Alyahyan/

Düstegör 2020) und soll den Studienerfolg durch individuelle Unterstützungs- und Beratungsangebote auf Basis ihrer Prognosen fördern (Attaran/Stark/Stotler 2018; Daniel 2015). Die vorgestellten Studien sind Teil der Forschungsprojekte *Fair Artificial Intelligence Reasoning* (FAIR, Volkswagenstiftung) und *Responsible Academic Performance Prediction* (RAPP, BMBF [16DHB4020]). In interdisziplinärer Zusammenarbeit wurden studentische Bewertungen des Studienerfolgsbegriffs untersucht (Ülpenich et al. 2023), diskriminierungsbehaftete Einflussfaktoren zur Studienerfolgsvorhersage ermittelt (Junghänel et al., im Review) sowie Fairness- und Schadenswahrnehmungen analysiert (Keller/Lünich/Marcinkowski 2022; Lünich/Keller 2024). Auf dieser Basis wurde ein Prototyp einer sozialverantwortlichen APP-Anwendung entwickelt (Duong et al. 2023).

Das komplexe Wechselspiel von Studierenden und Learning Analytics

Stellt die Datenverfügbarkeit im digitalen Zeitalter, in dem jeder Log-in auf Lernplattformen Datenspuren hinterlässt, keine besondere Herausforderung dar (Liebowitz 2017), sind an die tatsächliche Nutzbarkeit der Daten neben technischen auch soziale und ethische Fragen geknüpft (Nguyen et al. 2023; Simis/Weydner-Volkman 2024). So bietet der Einsatz von LA Chancen wie personalisierte Lernwege und Warnsysteme zur Steigerung des Studienerfolgs (Alyahyan/Düstegör 2020; Arnold/Pistilli 2012; Chiu et al. 2023), birgt aber auch Risiken wie Benachteiligungen durch unreflektierte Entwicklungen oder fehlerhafte Dateninterpretationen (Fazelpour/Danks 2021). Durch die Einhaltung ethischer Anforderungen an *Fairness*, *Accountability* und *Transparency* sollen diese minimiert werden (Memarian/Doleck 2023), unbeabsichtigte Folgen bleiben jedoch möglich (Smuha 2022). Die Partizipation betroffener Personen im Entwicklungsprozess ist daher unerlässlich (Radtko et al. 2024), um Machtungleichgewichte gegenüber Entwickler:innen und einführenden Institutionen zu vermeiden (Kalluri 2020; Kasy/Abebe 2021; Weinberg 2022).

Eine Studie der Heinrich-Heine-Universität Düsseldorf zeigt, dass wahrgenommene Unfairness gegenüber LA die Protestbereitschaft steigert und der Reputation der Hochschule schadet (Marcinkowski et al. 2020). Vertrauen in Hochschule und System ist entscheidend für eine erfolgreiche und sozialverantwortliche Nutzung von LA (Jones/Rubel/LeClere 2020; Tsai/Whitelock-Wainwright/Gašević 2021), während Unfairnesswahrnehmungen und Datenschutzbedenken die Freigabe persönlicher Daten behindern (Keller/Lünich/

Marcinkowski 2022; Sun et al. 2019) und so die Leistungsfähigkeit von LA einschränken können (Li/Brooks/Schaub 2019).

Konzeption eines sozialverantwortlichen Academic-Performance-Prediction-Systems

Eine idealtypische Betrachtung des ethischen LA-Einsatzes (Brossi/Castillo/Cortesi 2022; Cerratto Pargman/McGrath 2021; Fazelpour/Danks 2021; Holmes/Porayska-Pomsta 2022; The Institute for Ethical AI in Education 2021; Nguyen et al. 2023; Pardo/Siemens 2014) ermöglicht die Identifikation von Stellschrauben des soziotechnischen Systems auf Input-, Throughput- und Output-Ebene, die studentische Fairnesswahrnehmungen eines sozialverantwortlichen APP-Systems adressieren (Keller/Lünich/Marcinkowski 2022; Marcinkowski/Starke 2019).

Auf der Input-Ebene sind Problemdefinition, Wahl eines Algorithmus, Auswahl der Inputdaten sowie Freiwilligkeit und Rechtmäßigkeit ihrer Nutzung essenziell. Zielsetzungen und Funktionen von APP sollten klar definiert sein (Fazelpour/Danks 2021), während White-Box-Algorithmen Transparenz und informationelle Fairness fördern (Lünich/Keller 2024; Perez Vallejos et al. 2017; Rudin 2019). Definition von soziodemografischen und leistungsbezogenen Daten sowie die Freiwilligkeit der Nutzung sind zentral für Diskriminierungsfreiheit (Keller/Lünich/Marcinkowski 2022). Auf der Throughput-Ebene beeinflussen Systemsicherheit und Fairnessnormen die Wahrnehmung. Das APP-System muss vor unbefugtem Zugriff geschützt sein (Wirtz/Weyerer/Geyer 2019), und Fairnessnormen sollten so gewählt werden, dass eine diskriminierungsfreie algorithmische Kategorisierung entsprechend der Gerechtigkeitsvorstellung im jeweiligen Kontext gewährleistet wird (Dunkelau/Duong 2022; Lepri et al. 2018; Wong 2020). Eine respektvolle Interaktion zwischen soziotechnischem System und Mensch erfordert den »human-in-the-loop«, sodass der Mensch in die Entscheidungsfindung eingebunden wird (Amershi et al. 2015; Holzinger 2016; Starke/Lünich 2020). Die Output-Ebene wird durch die Output-Kategorisierung, Transparenz und Erklärbarkeit geprägt. Während APP stets transparente Entscheidungsprozesse und klar definierte Verantwortlichkeit bei Fehlentscheidungen erfordert (Memarian/Doleck 2023), bedarf es einer Entscheidung für eine Output-Kategorisierung z.B. entlang einer Rangordnung nach Studienabbruchswahrscheinlichkeit oder anhand der Bedürftigkeit der Studierenden. Auch die auf APP beruhende

Verteilung von Unterstützungsmaßnahmen kann die Fairnesswahrnehmungen der Studierenden beeinflussen (Keller/Lünich/Marcinkowski 2022; Lünich/Keller/Marcinkowski 2024a).

Kritische Haltung Studierender gegenüber Academic Performance Prediction

Die empirischen Erkenntnisse der Autor:innen zur KI-Entwicklung an deutschen Hochschulen zeigen, dass diese im internationalen Vergleich vor wenigen Jahren noch am Anfang stand (Keller et al. 2019). Dennoch äußerten Studierende der Universität Düsseldorf bereits Bedenken gegenüber hochschulrelevanten KI-Systemen. Während »Learning Analytics« von 83,3 % befürwortet und lediglich von 16,7 % abgelehnt wurden, stießen KI-Anwendungen mit potenziellen Auswirkungen auf die persönliche Zukunft auf höhere Ablehnung, darunter automatisierte Benotungssysteme (55,3 %), KI-basierte Bewerberauswahl (60,8 %) und besonders APP-Systeme (78,8 %) (Kieslich et al. 2019).

Ein Grund für die kritische Haltung gegenüber APP könnte im wahrgenommenen Risiko liegen. Eine repräsentative Befragung von Studierenden (n = 751) und der Bevölkerung (n = 1.008) in Deutschland zeigt, dass 40,2 % der Studierenden und 37,4 % der Bevölkerung APP entsprechend der Klassifizierung des AI-Acts als Hochrisikooanwendung einordnen (Europäische Union 2024). 11 % der Studierenden halten das Risiko für inakzeptabel. Gleichzeitig zeigen 48,8 % der Studierenden und 55,2 % der Bevölkerung eine ambivalente Risikoeinschätzung, indem sie das Risiko als minimal oder begrenzt bewerten (Lünich/Keller/Marcinkowski 2024b). Studierende erwarten im Vergleich zur restlichen Bevölkerung häufiger gesellschaftliche und individuelle Schäden sowie Diskriminierung durch APP. Besonders jene, die Diskriminierung im Alltag erfahren haben, neigen dazu, größere Bedenken zu äußern (Lünich/Keller/Marcinkowski 2024b).

Eine qualitative Analyse individueller Schadenswahrnehmungen gegenüber APP zeigt vielfältige studentische Befürchtungen, die sowohl die Fairnesswahrnehmung als auch die faktische Fairness eines Systems beeinträchtigen könnten. Die Ergebnisse der Onlinebefragung von n = 1.378 deutschen Studierenden verweisen auf eine technische und soziale Dimension der wahrgenommenen Schäden. In der technischen Dimension stehen Bedenken über Fehler im KI-System im Vordergrund, wie unvollständige oder

selektive Inputdaten, die zu falschen Vorhersagen führen könnten. Technische Sicherheitsbedenken wie Hackerangriffe oder Datenschutzprobleme wurden seltener genannt. In der sozialen Dimension äußern Studierende Bedenken zu individuellen und gesellschaftlichen Folgen von APP, wie Stigmatisierung und unfaire Chancenverteilung bei Studienplätzen oder Stipendien. Zudem werden ethische Fragen zur Quantifizierung von Leistung und Vereinheitlichung von Studienverläufen aufgeworfen. Individuell wird ein Potenzial für psychische Schäden betont, da APP-Systeme Motivation, Emotionen und Leistungsdruck negativ beeinflussen könnten (Lünich/Maxhuni/Keller zur Einreichung vorbereitet; Maxhuni et al. 2023).

Differenzierte Anforderungen der Studierenden an faire Academic Performance Prediction

Angesichts der Schadenswahrnehmungen, ambivalenten Risikoeinschätzungen und identifizierten Stellschrauben eines idealtypischen APP-Systems stellt sich die Frage, wie ein faires APP-System aus Studierendensicht gestaltet werden kann und welche Anforderungen es erfüllen muss. Eine repräsentative Befragung von Studierenden ($n = 751$) zeigt differenzierte Anforderungen an ein sozialverantwortliches APP-System. Studierende legen großen Wert auf Selbstbestimmung und möchten entscheiden, ob und welche ihrer Daten für APP genutzt werden (Lünich/Keller/Marcinkowski 2024b). Dies entspricht Forderungen amerikanischer Studierender nach Kontrolle und informierter Zustimmung bei der Datennutzung (Sun et al. 2019). Die geringe Bereitschaft deutscher Studierender, persönliche Daten preiszugeben, unterstreicht die Relevanz ethisch und fair bewerteter Inputdaten (Keller/Lünich/Marcinkowski 2022). Internationale Forschung zeigt, dass Vertrauen in den ethischen Umgang mit Daten und die Abwägung zum individuellen Nutzen die Datenfreigabe fördert (Li et al. 2022). Studierende verlangen Transparenz und Erklärbarkeit. So soll ihnen ihre individuelle Leistungsprognose verständlich erklärt und jederzeit (nur) für sie persönlich einsehbar sein (Lünich/Keller/Marcinkowski 2024b).

Ein Vergleich der Anforderungen zeigt signifikante Unterschiede zwischen Studierenden mit hoher und niedriger Schadenswahrnehmung. Letztere sind häufiger bereit, ihre Daten freiwillig bereitzustellen, und befürchten seltener, dass Dozierende sie bei Einsicht in ihrer APP schlechter bewerten könnten (Lünich/Keller/Marcinkowski 2024b). Gerade die mangelnde Bereitschaft zur Da-

tenfreigabe einzelner Studierendengruppen kann sich dabei als problematisch für die Diskriminierungsfreiheit des Systems erweisen.

Neben Anforderungen an das Systemdesign fordern Studierende ebenso eine Selbstbestimmung in Bezug auf die Verteilung von Förderangeboten auf Basis von APP. Verpflichtende Maßnahmen bei schlechten Prognosen lehnen sie ab und bevorzugen unabhängige Studienberatungen gegenüber Beratungen durch eigene Dozierende (Lünich/Keller/Marcinkowski 2024b). Die Zuteilung von Fördermaßnahmen kann nach verschiedenen Verteilungsnormen erfolgen. Bei *Equality* erhalten alle Studierenden gleiche Unterstützung, bei *Equity* wird nach Leistung, und bei *Need* nach Bedürftigkeit zugeteilt (vgl. Deutsch 1975). So könnten etwa nur besonders leistungsstarke (*Equity*) oder leistungsschwache (*Need*) Studierende gezielt gefördert werden. Im Rahmen der studentischen Onlinebefragung ($n = 1.378$) zeigt sich jedoch eine klare Bevorzugung der *Equality*-Norm, die signifikant fairer bewertet wird als Verteilungen nach dem Leistungs- oder Bedürftigkeitsprinzip (Lünich/Keller/Marcinkowski 2024a).

Implikationen für einen sozialverantwortlichen Einsatz von Learning Analytics

Konzeptionelle und empirische Erkenntnisse aus fünf Jahren Forschung zu studentischen Wahrnehmungen von LA und APP liefern wichtige Implikationen für einen sozialverantwortlichen KI-Einsatz. Die Studien zeigen eine kritische, mitunter aber auch ambivalente Haltung der Studierenden. Sorgen vor Schäden und die Ablehnung der Datenfreigabe betonen die Notwendigkeit umfassender Leitlinien seitens der Hochschulen (Scheffel et al. 2024). Angesichts unterschiedlicher Wissensstände zwischen Studierenden, Entwickler:innen und Institutionen entstehen Fragen zu politisch-partizipatorischer und informationeller Fairness. Eine sozialverantwortliche LA-Einführung erfordert Einbezug und Transparenz, ohne Studierende zu überfordern (Lünich/Keller 2024).

Die Studienergebnisse zeigen Unterschiede in den Schadenswahrnehmungen Studierender und Forschender und unterstreichen die Relevanz der Betroffenenperspektive. Herausforderungen wie Transparenz, Fairness und Verantwortlichkeit stehen bereits im Fokus ethischer KI-Richtlinien (Hagendorff 2020; Jobin/Ienca/Vayena 2019). Allerdings fokussieren Studierende stärker auf individuelle Folgen wie Demotivation und Druck und kritisie-

ren die soziale und ethische Logik der KI. Sie hinterfragen die Kosten der Zielsetzungen, etwa die Quantifizierung von Leistung, und zweifeln die generelle Sinnhaftigkeit von LA-Systemen an (Ülpenich et al. 2023). Effizienz und Optimierung durch LA fördern eine Ethik der ständigen Verbesserung, die abweichende Perspektiven verdrängt und Handlungs- und Definitionsmacht an Maschinen überträgt. Hegemoniale Kräfteverhältnisse in der Problem- und Output-Definition können demokratischen Bildungsidealen zuwiderlaufen.

Neben den Designfaktoren der LA-Systeme müssen auch darauf basierende Interventionen und pädagogische Ansätze reflektiert werden (Tsai/Gasevic 2017; Tsai et al. 2020). Die Bewertung der Verteilung von Fördermaßnahmen wird zentral, da Studierende absolute Gleichbehandlung fordern, was die Sinnhaftigkeit der APP-Einführung infrage stellt. Individuelle Fördermaßnahmen auf Basis von APP sind zwar wichtig zur Leistungsverbesserung, doch unbegrenzte Zugänglichkeit könnte den Zweck der Implementierung untergraben. Unabhängige Beratungsangebote können hingegen die Bereitschaft zur Nutzung von Unterstützungsleistungen fördern. APP-Systeme, die aus Sicht der Studierenden primär Informationszwecken dienen sollten, werden kritisch gesehen, da sie potenziell Demotivation statt Studierenerfolg fördern. Angesichts der Abhängigkeit von freiwilliger Datenfreigabe müssen studentische Bedenken sowie Fairness- und Risikowahrnehmungen ernst genommen und integraler Bestandteil von Entwicklungs- und Implementierungsprozessen sein. Eine sozialverantwortliche LA-Entwicklung erfordert:

- **Gemeinsame Problemspezifizierung:** Akkurate LA-Anwendungen erfordern erhebliche Ressourcen. Daher muss die Definition des Problems, der Ziele, Funktionen und Grenzen gemeinsam mit allen Stakeholdern erfolgen, um Ablehnung zu vermeiden und Diskriminierung zu verhindern. Alternative Problemdefinitionen und nicht-algorithmische Lösungen sollten in Betracht gezogen werden.
- **Bedenken im Mittelpunkt:** Studentische Bedenken sind oft vielfältiger und unterscheiden sich von etablierten ethischen Richtlinien und müssen für eine verantwortungsvolle KI-Implementierung ernst genommen werden.
- **Sicherung der Diskriminierungsfreiheit:** Die Vermeidung der Reproduktion von Stereotypen und Diskriminierung in Inputdaten, algorithmischer Verarbeitung und Fördermaßnahmen muss oberste Priorität haben und darf nicht durch Trial-and-Error auf Kosten der Studierenden erreicht werden.

- **Transparenz und Verständlichkeit:** Transparenz muss gewährleistet werden, damit der KI-Einsatz nachvollziehbar bleibt, ohne die Betroffenen zu überfordern. Eine forschungsgestützte, klare Vermittlung der Anwendungsfelder und der Einsatz von White-Box-KI-Verfahren fördern Nachvollziehbarkeit und Verständnis.
- **Unabhängige Beratungsinstanzen:** Freiwillige, unabhängige Beratungsangebote und ein vertrauensvolles Umfeld können die Bereitschaft zur Nutzung von Unterstützungsleistungen steigern.
- **Fortlaufende Evaluation:** Der Einsatz von LA sollte sorgfältig geplant, verständlich erklärt und kontinuierlich unter Einbezug aller universitären Stakeholder evaluiert sowie an neue Erkenntnisse angepasst werden.

Die Entwicklung sozialverantwortlicher LA-Systeme und insbesondere akademischer Leistungsprognosen stellt eine anspruchsvolle Aufgabe dar, die ethische, technische und soziale Aspekte berücksichtigen muss. Die vorgeschlagenen Maßnahmen bieten eine Grundlage, um Akzeptanz und Wirksamkeit von LA-Systemen zu fördern. Ihre Umsetzung erfordert kooperatives Engagement aller Beteiligten – von Hochschulverantwortlichen und politischen Entscheidungsträger:innen bis zu den Studierenden. Nur ein integrativer Ansatz stellt sicher, dass KI-basierte LA-Systeme nicht nur technologische Fortschritte, sondern auch faire und inklusive Bildungschancen für alle Studierenden ermöglichen.

Literatur

- Alyahyan, Eyman/Düştüğü, Dilek (2020): »Predicting academic success in higher education: literature review and best practices. literature review and best practices«, in: International Journal of Educational Technology in Higher Education 17, pp. 1–21.
- Amershi, Saleema/Cakmak, Maya/Knox, William B. et al. (2015): »Power to the People: The Role of Humans in Interactive Machine Learning. The Role of Humans in Interactive Machine Learning«, in: AI Magazine 35, pp. 105–120.
- Arnold, Kimberly E./Pistilli, Matthew D. (2012): »Course signals at Purdue. Using Learning Analytics to Increase Student Success«, in: Shane Dawson/Caroline Haythornthwaite/Simon Buckingham Shum et al. (Eds.), Pro-

- ceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK 2012), New York, NY, USA: ACM, pp. 267–270.
- Attaran, Mohsen/Stark, John/Stotler, Derek (2018): »Opportunities and challenges for big data analytics in US higher education«, in: *Industry and Higher Education* 32, pp. 169–182.
- Brossi, Lionel/Castillo, Ana M./Cortesi, Sandra (2022): »Student-centred requirements for the ethics of AI in education«, in: Holmes/Porayska-Pomsta (Eds.), *The Ethics of Artificial Intelligence in Education*, pp. 91–112.
- Casanova, Joana R./Gomes, Cristiano M. A./Bernardo, Ana B. et al. (2021): »Dimensionality and reliability of a screening instrument for students at-risk of dropping out from Higher Education«, in: *Studies in Educational Evaluation* 68, p. 100957.
- Cerratto Pargman, Teresa/McGrath, Cormac (2021): »Mapping the Ethics of Learning Analytics in Higher Education: A Systematic Literature Review of Empirical Research«, in: *Journal of Learning Analytics* 8, pp. 123–139.
- Chiu, Thomas K./Xia, Qi/Zhou, Xinyan et al. (2023): »Systematic Literature Review on Opportunities, Challenges, and Future Research Recommendations of Artificial Intelligence in Education«, in: *Computers and Education: Artificial Intelligence* 4, pp. 1–15.
- Costa-Mendes, Ricardo/Cruz-Jesus, Frederico/Oliveira, Tiago et al. (2021): »Machine Learning Bias in Predicting High School Grades: A Knowledge Perspective. A Knowledge Perspective«, in: *Emerging Science Journal* 5, pp. 576–597.
- Daniel, Ben (2015): »Big Data and analytics in higher education: Opportunities and challenges«, in: *British Journal of Educational Technology* 46, pp. 904–920.
- Deutsch, Morton (1975): »Equity, Equality, and Need: What Determines Which Value Will Be Used as the Basis of Distributive Justice?«, in: *Journal of Social Issues* 31, pp. 137–149.
- Dunkelau, Jannik/Duong, Manh K. (2022): »Towards Equalised Odds as Fairness Metric in Academic Performance Prediction«, in: *FATED'22: 2nd Workshop on Fairness, Accountability, and Transparency in Educational Data.*, pp. 1–6.
- Duong, Manh K./Dunkelau, Jannik/Cordova, José A. et al. (2023): »RAPP: A Responsible Academic Performance Prediction Tool for Decision-Making in Educational Institutes«, in: Birgitta König-Ries/Stefanie Scherzinger/Wolfgang Lehner et al. (Eds.), *BTW 2023. Software Engineering 2023: Fachtagung vom 06. – 10. März 2023, Dresden, Bonn: Köllen.*

- Europäische Union (2024): Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828.
- Fazelpour, Sina/Danks, David (2021): »Algorithmic bias: Senses, sources, solutions«, in: *Philosophy Compass* 16, pp. 1–16.
- Hagendorff, Thilo (2020): »The Ethics of AI Ethics: An Evaluation of Guidelines. An Evaluation of Guidelines«, in: *Minds and Machines* 30, pp. 99–120.
- Holmes, Wayne/Porayska-Pomsta, Kaška (Eds.) (2022): *The Ethics of Artificial Intelligence in Education. Practices, Challenges, and Debates*, New York: Routledge.
- Holzinger, Andreas (2016): »Interactive machine learning for health informatics: when do we need the human-in-the-loop?«, in: *Brain informatics* 3, pp. 119–131.
- Jobin, Anna/Ienca, Marcello/Vayena, Effy (2019): »The global landscape of AI ethics guidelines«, in: *Nature Machine Intelligence* 1, pp. 389–399.
- Jones, Kyle M. L./Rubel, Alan/LeClere, Ellen (2020): »A Matter of Trust: Higher Education Institutions as Information Fiduciaries in an Age of Educational Data Mining and Learning Analytics«, in: *Journal of the Association for Information Science and Technology* 71, pp. 1227–1241.
- Junghänel, Luisa/Ülpenich, Bettina/Krause, Johannes et al. (im Review): *Diskriminierungsfreie Vorhersage des Studienerfolgs – Eine empirische Untersuchung möglicher Leistungsindikatoren*.
- Kalluri, Pratyusha (2020): »Don't ask if artificial intelligence is good or fair, ask how it shifts power«, in: *Nature* 583, p. 169.
- Kasy, Maximilian/Abebe, Rediet (2021): »Fairness, Equality, and Power in Algorithmic Decision-Making«, in: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, New York, NY, United States: Association for Computing Machinery, pp. 576–586.
- Keller, Birte/Baleis, Janine/Starke, Christopher et al. (2019): *Machine Learning and Artificial Intelligence in Higher Education. A State-of-the-Art Report on the German University Landscape. Working Paper No. 1*, Düsseldorf, https://www.phil-fak.uni-duesseldorf.de/fileadmin/Redaktion/Institute/Sozialwissenschaften/Kommunikations-_und_Medienwissenschaft/KMW_I/Working_Paper/Keller_et_al._2019_-_AI_in_Higher_Education.pdf from November 9, 2021.

- Keller, Birte/Lünich, Marco/Marcinkowski, Frank (2022): »How Is Socially Responsible Academic Performance Prediction Possible? Insights From a Concept of Perceived AI Fairness«, in: Fernando Almaraz-Menéndez/Alexander Maz-Machado/Carmen López-Esteban et al. (Eds.), *Strategy, Policy, Practice, and Governance for AI in Higher Education Institutions*, IGI Global, pp. 126–155.
- Kieslich, Kimon/Lünich, Marco/Marcinkowski, Frank et al. (2019): *Hochschule der Zukunft. Einstellungen von Studierenden gegenüber Künstlicher Intelligenz an der Hochschule. Précis*, https://diid.hhu.de/wp-content/uploads/2019/10/DIID-Precis_Kieslich-et-al_Fin.pdf.
- Lang, Charles/Davis, Laura (2023): »Learning Analytics and Stakeholder Inclusion: What do We Mean When We Say »Human-Centered«?«, in: Isabel Hilliger/Hassan Khosravi/Bart Rienties et al. (Eds.), *LAK23: 13th International Learning Analytics and Knowledge Conference*, New York, NY, United States: Association for Computing Machinery, pp. 411–417.
- Lepri, Bruno/Oliver, Nuria/Letouzé, Emmanuel et al. (2018): »Fair, Transparent, and Accountable Algorithmic Decision-making Processes. The Premise, the Proposed Solutions, and the Open Challenges«, in: *Philosophy & Technology* 31, pp. 611–627.
- Li, Warren/Brooks, Christopher/Schaub, Florian (2019): »The Impact of Student Opt-Out on Educational Predictive Models«, in: *Proceedings of the 9th International Conference on Learning Analytics & Knowledge (LAK '19)*, New York, NY: ACM, pp. 411–420.
- Li, Warren/Sun, Kaiwen/Schaub, Florian et al. (2022): »Disparities in Students' Propensity to Consent to Learning Analytics«, in: *International Journal of Artificial Intelligence in Education* 32, pp. 564–608.
- Liebowitz, Jay (2017): »Thoughts on Recent Trends and Future Research Perspectives in Big Data and Analytics in Higher Education«, in: Ben Kei Daniel (Ed.), *Big Data and Learning Analytics in Higher Education. Current Theory and Practice*, Cham: Springer International Publishing, pp. 7–17.
- Lünich, Marco/Keller, Birte (2024): »Explainable Artificial Intelligence for Academic Performance Prediction. An Experimental Study on the Impact of Accuracy and Simplicity of Decision Trees on Causability and Fairness Perceptions«, in: *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, pp. 1031–1042.
- Lünich, Marco/Keller, Birte/Marcinkowski, Frank (2024a): »Fairness of Academic Performance Prediction for the Distribution of Support Measures

- for Students: Differences in Perceived Fairness of Distributive Justice Norms», in: *Technology, Knowledge and Learning* 29, pp. 1079–1107.
- Lünich, Marco/Keller, Birte/Marcinkowski, Frank (2024b): »Diverging Perceptions of Artificial Intelligence in Higher Education. A Comparison of Student and Public Assessments on Risks and Damages of Academic Performance Prediction in Germany«, in: *Computers and Education: Artificial Intelligence*.
- Lünich, Marco/Maxhuni, Albina/Keller, Birte (zur Einreichung vorbereitet): Focusing on Student Harm Perceptions in Academic Performance Prediction: A Participatory Technology Assessment Framework to Prevent Reductionist and Hegemonic AI Implementation in Higher Education [Arbeitsstiel].
- Marcinkowski, Frank/Kieslich, Kimon/Starke, Christopher et al. (2020): »Implications of AI (Un-)Fairness in Higher Education Admissions. The Effects of Perceived AI (Un-)Fairness on Exit, Voice and Organizational Reputation«, in: Mireille Hildebrandt/Carlos Castillo/Elisa Celis et al. (Eds.), *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* 20)*, New York, NY, USA: ACM, pp. 122–130.
- Marcinkowski, Frank/Starke, Christopher (2019): »Wann ist Künstliche Intelligenz (un-)fair? Ein sozialwissenschaftliches Konzept von KI-Fairness«, in: Jeanette Hofmann/Norbert Kersting/Claudia Ritzi et al. (Eds.), *Politik in der digitalen Gesellschaft. Zentrale Problemfelder Und Forschungsperspektiven*, Bielefeld, Germany: transcript Verlag, pp. 269–288.
- Maxhuni, Albina/Lünich, Marco/Keller, Birte et al. (2023): *Hegemoniale Technologieimplementierung an der Hochschule – Eine qualitative Analyse der Schadenswahrnehmung betroffener Studierender bei der Einführung von Dropout Detection*, Bremen, Deutschland.
- Memarian, Bahar/Doleck, Tenzin (2023): »Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI) and Higher Education: A Systematic Review«, in: *Computers and Education: Artificial Intelligence* 5, pp. 1–12.
- Nguyen, Andy/Ngo, Ha N./Hong, Yvonne et al. (2023): »Ethical principles for artificial intelligence in education«, in: *Education and Information Technologies* 28, pp. 4221–4241.
- Niyogisubizo, Jovial/Liao, Lyuchao/Nziyumva, Eric et al. (2022): »Predicting Student's Dropout in University Classes Using Two-Layer Ensemble Machine Learning S Approach: A Novel Stacked Generalization«, in: *Computers and Education: Artificial Intelligence* 3, pp. 1–12.

- Pardo, Abelardo/Siemens, George (2014): »Ethical and privacy principles for learning analytics«, in: *British Journal of Educational Technology* 45, pp. 438–450.
- Perez Vallejos, Elvira/Koene, Ansgar/Portillo, Virginia et al. (2017): »Young People's Policy Recommendations on Algorithm Fairness«, in: *Proceedings of the 2017 ACM Web Science Conference*, pp. 247–251.
- Radtke, Anna/Osinski, Meike/Serova, Katja et al. (2024): »Learning Analytics in der Erziehungswissenschaft: Lerndatenbasierte Förderung von Selbstregulation in einem Statistikkurs«, in: Salden/Leschke (Eds.), *Learning Analytics und Künstliche Intelligenz in Studium und Lehre*, pp. 45–61.
- Rudin, Cynthia (2019): »Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead«, in: *Nature Machine Intelligence* 1, pp. 206–215.
- Salden, Peter/Leschke, Jonas (Eds.) (2024): *Learning Analytics und Künstliche Intelligenz in Studium und Lehre. Erfahrungen und Schlussfolgerungen aus einer hochschulweiten Erprobung*, Wiesbaden: Springer Fachmedien Wiesbaden.
- Scheffel, Maren/Simis, Christos/Leschke, Jonas et al. (2024): »Learning Analytics-Policies im Hochschulkontext«, in: Salden/Leschke (Eds.), *Learning Analytics und Künstliche Intelligenz in Studium und Lehre*, pp. 169–185.
- Simis, Christos/Weydner-Volkman, Sebastian (2024): »KI-gestützte Learning Analytics: Geschenk oder Falle für die Hochschulbildung? Ein ethischer Exkurs«, in: Salden/Leschke (Eds.), *Learning Analytics und Künstliche Intelligenz in Studium und Lehre*, pp. 127–148.
- Slade, Sharon/Prinsloo, Paul (2013): »Learning Analytics. Ethical Issues and Dilemmas«, in: *American Behavioral Scientist* 57, pp. 1510–1529.
- Smith, Helen (2020): »Algorithmic bias: should students pay the price?«, in: *AI & SOCIETY* 35, pp. 1077–1078.
- Smuha, Nathalie A. (2022): »Pitfalls and Pathways for Trustworthy Artificial Intelligence in Education«, in: Holmes/Porayska-Pomsta (Eds.), *The Ethics of Artificial Intelligence in Education*, pp. 113–145.
- Starke, Christopher/Lünich, Marco (2020): »Artificial intelligence for political decision-making in the European Union: Effects on citizens' perceptions of input, throughput, and output legitimacy«, in: *Data & Policy* 2, pp. 1–17.
- Sun, Kaiwen/Mhaidli, Abraham H./Watel, Sonakshi et al. (2019): »It's My Data! Tensions Among Stakeholders of a Learning Analytics Dashboard«, in: Stephen Brewster/Geraldine Fitzpatrick/Anna Cox et al. (Eds.), *Proceed-*

- ings of the 2019 CHI Conference on Human Factors in Computing Systems, New York, NY, USA: ACM, pp. 1–14.
- The Institute for Ethical AI in Education (2021): The Ethical Framework for AI in Education, <https://www.buckingham.ac.uk/research-the-institute-for-ethical-ai-in-education/>.
- Tsai, Yi-Shan/Gasevic, Dragan (2017): »Learning analytics in higher education—challenges and policies«, in: Alyssa F. Wise (Ed.), Proceedings of the Seventh International Learning Analytics & Knowledge Conference, New York, NY: ACM, pp. 233–242.
- Tsai, Yi-Shan/Rates, Diego/Moreno-Marcos, Pedro M. et al. (2020): »Learning analytics in European higher education—Trends and barriers«, in: Computers & Education 155, pp. 1–16.
- Tsai, Yi-Shan/Whitlock-Wainwright, Alexander/Gašević, Dragan (2021): »More Than Figures on Your Laptop: (Dis)trustful Implementation of Learning Analytics«, in: Journal of Learning Analytics 8, pp. 81–100.
- Ülpenich, Bettina/Junghänel, Luisa/Krause, Johannes et al. (2023): »Studien-erfolg und -misserfolg aus institutionenund studierendenzentrierter Perspektive: Eine qualitative Untersuchung«, in: Beiträge zur Hochschulfor-schung 45, pp. 54–74.
- Weinberg, Lindsay (2022): »Rethinking Fairness: An Interdisciplinary Survey of Critiques of Hegemonic ML Fairness Approaches«, in: Journal of Artificial Intelligence Research 74, pp. 75–109.
- Wirtz, Bernd W./Weyerer, Jan C./Geyer, Carolin (2019): »Artificial Intelligence and the Public Sector—Applications and Challenges«, in: International Journal of Public Administration 42, pp. 596–615.
- Wong, Pak-Hang (2020): »Democratizing Algorithmic Fairness«, in: Philosophy & Technology 33, pp. 225–244.
- Zawacki-Richter, Olaf/Marín, Victoria I./Bond, Melissa et al. (2019): »System-atic review of research on artificial intelligence applications in higher edu-cation – where are the educators?«, in: International Journal of Educational Technology in Higher Education 16, pp. 1–27.