

Von der Legalität zur Legitimität: Bias in KI zwischen Recht, Ethik und Technikgestaltung

Martina Gaisch, FH Oberösterreich, Hagenberg
Berthold Kerschbaumer, FH Oberösterreich, Hagenberg

Abstract:

Mit der zunehmenden Integration künstlicher Intelligenz in gesellschaftlich relevante Entscheidungsprozesse verschiebt sich Verantwortung von individuellen Akteur:innen hin zu datengetriebenen Systemen. Gleichzeitig zeigen zahlreiche Studien, dass algorithmische Systeme bestehende soziale Ungleichheiten nicht nur reproduzieren, sondern teilweise verstärken können. Diese Entwicklung stellt Recht, Ethik und Technikgestaltung vor neue Herausforderungen. Der vorliegende Beitrag untersucht, in welchem Verhältnis rechtliche Mindeststandards und ethische Reflexionsprozesse bei der Regulierung und Gestaltung von KI-Systemen stehen. Im Zentrum steht die analytische Unterscheidung zwischen Legalität, verstanden als formale Rechtskonformität, und Legitimität als normative Rechtfertigung algorithmischer Entscheidungen.

Die empirische Grundlage bilden acht leitfadengestützte Expert:inneninterviews mit Fachpersonen aus Rechtswissenschaft und Ethik. Die Interviews wurden mittels thematischer Inhaltsanalyse ausgewertet, um unterschiedliche Perspektiven auf Verantwortlichkeit, Fairness und gesellschaftliche Akzeptanz algorithmischer Systeme zu identifizieren. Die Ergebnisse zeigen, dass rechtliche Instrumente wie die Datenschutz-Grundverordnung und der europäische AI Act wichtige Mindeststandards für Transparenz, Risikomanagement und Verantwortlichkeit schaffen. Gleichzeitig weisen die Expert:innen darauf hin, dass rechtliche Konformität allein nicht ausreicht, um gesellschaftliche Legitimität sicherzustellen, insbesondere im Umgang mit strukturellen Verzerrungen wie Gender Bias.

Der Beitrag argumentiert daher für eine Governance-Perspektive, die rechtliche Regulierung durch institutionalisierte ethische Reflexionsprozesse ergänzt. Eine verantwortungsvolle KI-Praxis erfordert die systematische Integration technischer, organisatorischer und normativer Dimensionen. Auf dieser Grundlage entwickelt der Artikel einen konzeptionellen Rahmen, der Legalität und Legitimität als komplementäre Geltungsdimensionen algorithmischer Systeme versteht und damit einen Beitrag zur aktuellen Debatte um verantwortungsvolle und diskriminierungssensible KI-Governance leistet.

Keywords: KI-Ethik, Bias, AI Act, verantwortungsvolle KI, Legalität, Legitimität

1. PROBLEMSTELLUNG

Künstliche Intelligenz ist zu einem zentralen Bestandteil gesellschaftlich relevanter Entscheidungsprozesse in Verwaltung, Wirtschaft und Politik geworden. Gleichzeitig zeigen

zahlreiche empirische Studien, dass KI-Systeme bestehende soziale Ungleichheiten nicht nur reproduzieren, sondern unter bestimmten Bedingungen auch verstärken können (Adams, 2024; Kim, 2021; Özer et al., 2024). Besonders deutlich wird dies bei geschlechtsspezifischen Verzerrungen (Gender Bias), die nicht ausschließlich aus technischen Fehlfunktionen resultieren, sondern häufig in sozialen, kulturellen und institutionellen Strukturen verankert sind (Hall & Ellis, 2023).

Der Einsatz algorithmischer Systeme wirft daher nicht nur technische und regulatorische Fragen auf, sondern berührt grundlegende normative Dimensionen gesellschaftlicher Entscheidungsprozesse. Dazu zählen insbesondere Fragen der Fairness, Transparenz und Verantwortlichkeit. Im Zentrum steht die Frage, unter welchen Bedingungen algorithmische Entscheidungen als gerecht, nachvollziehbar und gesellschaftlich legitim gelten können, insbesondere dann, wenn sie tiefgreifende Auswirkungen auf individuelle Rechte, Zugangschancen oder gesellschaftliche Teilhabe haben.

Diese Problematik betrifft zentrale Diskussionsfelder der Datenethik, des Datenschutzrechts sowie der Demokratietheorie. Sie reicht über Fragen formaler Rechtskonformität hinaus und stellt die weitergehende Frage nach der normativen Rechtfertigung algorithmischer Entscheidungsmacht (Gray & Witt, 2021). Vor diesem Hintergrund untersucht der vorliegende Beitrag, wie rechtliche und ethische Geltungsansprüche im Kontext algorithmischer Systeme zusammenwirken und welche Rolle sie für eine verantwortungsvolle Regulierung und Gestaltung von KI spielen.

2. THEORETISCHER RAHMEN

Ein zentrales Anliegen dieses Beitrags besteht darin, zwischen Legalität und Legitimität als unterschiedlichen, jedoch komplementären Geltungsdimensionen algorithmischer Systeme zu unterscheiden. Diese Differenz bildet das theoretische Fundament für die Analyse der Expert:inneninterviews sowie für die normative Bewertung bestehender Regulierungsansätze.

Legalität bezeichnet im rechtswissenschaftlichen Verständnis die formale Übereinstimmung mit geltendem Recht. Sie bezieht sich auf die Einhaltung kodifizierter Normen, Verfahren und Zuständigkeiten, wie sie beispielsweise in der Datenschutz-Grundverordnung (DSGVO) oder im europäischen AI Act festgelegt sind (Van Bekkum & Borgesius, 2023). Legalität schafft damit einen institutionellen Rahmen, der bestimmte algorithmische Praktiken erlaubt, einschränkt oder verbietet und ihre Einhaltung durch rechtliche Durchsetzungsmechanismen absichert (Cromvelle, 2025).

Legitimität geht über diese formale Rechtskonformität hinaus. In Anlehnung an demokratietheoretische und rechtsphilosophische Traditionen (Habermas, 1992; Buchanan & Keohane, 2006) bezeichnet Legitimität die normative Rechtfertigung von Entscheidungs- und Herrschaftsmacht, insbesondere dann, wenn diese durch technologische Systeme ausgeübt wird. Legitimität ergibt sich nicht allein aus der Einhaltung rechtlicher Regeln, sondern aus dem Zusammenspiel von Rechtskonformität, gesellschaftlicher Akzeptanz und moralischer Plausibilität.

Ein KI-System kann somit formal legal sein, etwa wenn es den Anforderungen der DSGVO entspricht, ohne zugleich als legitim wahrgenommen zu werden. Dies kann beispielsweise dann der Fall sein, wenn algorithmische Entscheidungen systematische Diskriminierungseffekte erzeugen oder gesellschaftlich als unfair bewertet werden (UNESCO, 2021; Bietti, 2020).

Die Unterscheidung zwischen Legalität und Legitimität ist daher sowohl theoretisch als auch empirisch bedeutsam. Während Legalität die Grundlage rechtsstaatlicher Kontrolle bildet, bestimmt Legitimität, ob algorithmische Entscheidungen gesellschaftlich akzeptiert und normativ gerechtfertigt werden. Recht kann den Einsatz bestimmter Technologien erlauben. Ethik stellt die weitergehende Frage, ob und unter welchen Bedingungen dieser Einsatz gerechtfertigt ist.

Vor diesem Hintergrund argumentiert der Beitrag, dass verantwortungsvolle KI-Governance dort ansetzt, wo rechtliche Legalität durch ethisch reflektierte Legitimität ergänzt wird. Diese Perspektive ermöglicht eine kritisch-normative Bewertung algorithmischer Systeme, insbesondere in Situationen, in denen formale Rechtskonformität an die Grenzen ihrer gesellschaftlichen Akzeptanz stößt (Hintz, 2024; Gaisch & Mader, 2025).

Der Beitrag verbindet daher rechtliche und ethische Perspektiven auf KI-Governance und analysiert die Wechselwirkungen zwischen zentralen Regulierungsrahmen, insbesondere der DSGVO, dem AI Act sowie der UNESCO-Empfehlung zur Ethik der künstlichen Intelligenz (UNESCO, 2021). Auf dieser Grundlage wird argumentiert, dass ethische Prinzipien wie Gerechtigkeit, Verantwortung und Rechenschaft wichtige Ergänzungen zu rechtlichen Compliance-Strukturen darstellen und eine zentrale Rolle bei der Vermeidung systemischer Diskriminierung spielen. Zugleich wird diskutiert, inwiefern ethische Leitlinien gegenüber der Gefahr von Ethics Washing institutionell abgesichert werden müssen.

3. METHODISCHES VORGEHEN

Zur empirischen Fundierung der Analyse wurden acht leitfadengestützte Expert:inneninterviews durchgeführt. Die Auswahl der Interviewpartner:innen erfolgte nach dem Prinzip des purposive sampling, um unterschiedliche fachliche Perspektiven auf KI-Governance und algorithmische Verantwortung einzubeziehen. Vier Interviews wurden mit Expert:innen aus dem Bereich Datenschutz- und Technologierecht geführt, vier weitere mit Ethiker:innen, deren Forschungsschwerpunkte im Bereich KI- und Datenethik liegen.

Ziel der Interviews war es, unterschiedliche Deutungen von Verantwortlichkeit, Transparenz, Fairness und Gendergerechtigkeit im Kontext algorithmischer Systeme zu erfassen sowie mögliche Konvergenzen und Divergenzen zwischen rechtlicher und ethischer Argumentation sichtbar zu machen.

Die Interviews wurden auf Grundlage eines halbstrukturierten Leitfadens durchgeführt, der zentrale Themenfelder der KI-Regulierung, der algorithmischen Fairness sowie der institutionellen Verantwortung adressierte. Sämtliche Gespräche wurden audioaufgezeichnet und vollständig transkribiert.

Die Auswertung erfolgte mittels thematischer Inhaltsanalyse. Dabei wurden zunächst zentrale Themen und Argumentationsmuster identifiziert und anschließend entlang der Kategorien rechtlicher und ethischer Perspektiven systematisch vergleichend analysiert. Besonderes Augenmerk lag auf Passagen, in denen die Grenzen bestehender rechtlicher Instrumente thematisiert wurden oder in denen ethische Reflexion über formale Rechtskonformität hinausging.

Durch dieses Vorgehen ermöglicht die Studie eine empirisch fundierte Analyse der Wechselwirkungen zwischen rechtlichen und ethischen Perspektiven auf KI-Governance und liefert Einblicke in aktuelle Diskurse über die Legitimität algorithmischer Entscheidungsprozesse.

4. ERGEBNISSE UND ANALYTISCHE EINSICHTEN

Rechtliche Normen wie die Datenschutz-Grundverordnung (DSGVO) schaffen grundlegende Standards für Datenschutz, Nichtdiskriminierung und Verantwortlichkeit und etablieren zugleich klare Zuständigkeiten sowie rechtliche Durchsetzungsmechanismen. Ethische Leitlinien, etwa die UNESCO-Empfehlung zur Ethik der künstlichen Intelligenz, gehen über diese Mindeststandards hinaus und fordern eine umfassendere Bewertung der sozialen und geschlechterbezogenen Folgen algorithmischer Entscheidungen. Während das Recht vor allem institutionelle Ordnung und justiziable Grenzen gewährleistet, richtet sich der Fokus ethischer Ansätze stärker auf Fragen der Fairness, der sozialen Gerechtigkeit und der Sensibilität gegenüber strukturellen Ungleichheiten. Gerade im Kontext von Gender Bias zeigt sich, dass problematische Effekte algorithmischer Systeme formal rechtmäßig sein können, ohne im normativen Sinne als legitim zu gelten. In diesem Verständnis bildet das Recht das institutionelle Fundament, während ethische Reflexion einen Raum eröffnet, in dem gesellschaftliche Folgen, Wertkonflikte und Gerechtigkeitsfragen kritisch bewertet werden können.

Rechtliche Instrumente wie Datenschutz-Folgenabschätzungen gemäß Artikel 35 DSGVO sowie Governance-Anforderungen des AI Acts, etwa Dokumentationspflichten, Logging oder Risikomanagementprozesse, sind wichtige Elemente einer verantwortungsvollen KI-Regulierung. Sie schaffen Verfahren und institutionelle Anreize für Transparenz und Verantwortlichkeit. Gleichzeitig adressieren diese Instrumente strukturelle Formen geschlechtsspezifischer Verzerrungen nur begrenzt. Ergänzend gewinnen daher ethische Begleitprozesse an Bedeutung, etwa Ethical Impact Assessments, Gender Impact Assessments oder Diversity-by-Design-Ansätze. Die Interviews zeigen, dass solche Maßnahmen insbesondere dann wirksam sind, wenn sie institutionell verankert und mit ausreichenden Ressourcen ausgestattet sind, anstatt lediglich als freiwillige oder informelle Zusatzaufgaben zu fungieren. Daraus ergibt sich eine komplementäre Rollenverteilung zwischen rechtlichen und ethischen Instrumenten im Umgang mit algorithmischem Bias.

Neue empirische Erkenntnisse des Joint Research Centre der Europäischen Kommission aus dem Jahr 2025 weisen zudem darauf hin, dass menschliche Aufsicht allein Bias nicht zuverlässig verhindert. Fachkräfte neigen häufig dazu, algorithmischen Empfehlungen auch dann zu folgen, wenn diese als potenziell unfair oder verzerrt erkannt werden. Die Expert:inneninterviews

unterstreichen daher die Notwendigkeit eines systemischen Verständnisses von Aufsicht, das technische, organisatorische und kompetenzbezogene Dimensionen miteinander verbindet. Dazu gehören transparente Modelle, nachvollziehbare Entscheidungsroutrinen, klar definierte Eskalationswege sowie spezifische Kompetenzen im Bereich Gender- und Diversity-Sensibilität.

Aufsicht wird in diesem Zusammenhang nicht als individuelle Letztkontrolle verstanden, sondern als kollektive und organisationale Aufgabe. Insbesondere die Ethikexpert:innen betonen, dass wirksame menschliche Aufsicht nur dann entstehen kann, wenn sie in ein soziotechnisches Gesamtsystem eingebettet ist, in dem Verantwortung verteilt, Rollen klar definiert und Entscheidungsprozesse institutionell abgesichert sind. Menschliche Aufsicht bedeutet in diesem Verständnis nicht lediglich, algorithmische Entscheidungen punktuell zu überprüfen. Vielmehr geht es darum, organisatorische Rahmenbedingungen zu schaffen, die kritisches Hinterfragen ermöglichen und aktiv fördern. Dazu zählen ausreichende Zeitressourcen für Reflexion, institutionell legitimierte Möglichkeiten zum Widerspruch gegen algorithmische Empfehlungen sowie eine Organisationskultur, in der das Abweichen von automatisierten Entscheidungen nicht als Fehler, sondern als Ausdruck professioneller Verantwortung verstanden wird. Unter diesen Bedingungen kann Aufsicht ihre normative Funktion erfüllen, nämlich strukturelle Verzerrungen frühzeitig zu erkennen, ihre Ursachen zu analysieren und faire Entscheidungsprozesse langfristig zu sichern.

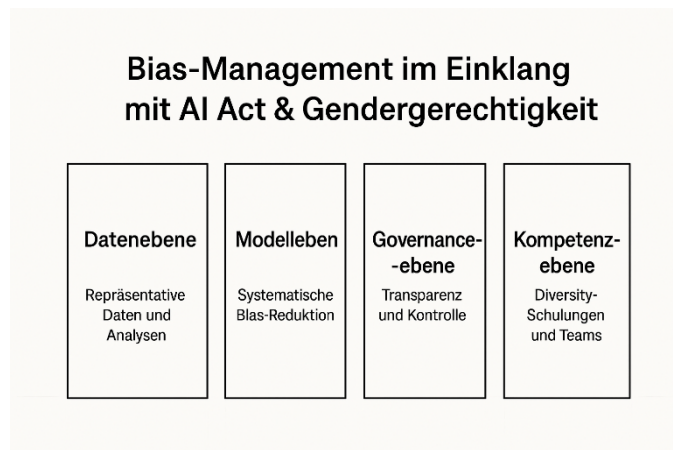


Tabelle 1. KI-Bias-Management entlang zentraler Gestaltungsebenen

Datenebene

Organisationen sollten Repräsentativität und potenzielle Verzerrungen in Datensätzen systematisch analysieren. Dazu gehört die transparente Dokumentation der Herkunft, Aufbereitung und Verwendung der Daten sowie die Sicherstellung ihrer Auditierbarkeit. Genderdifferenzierte und intersektionale Analysen ermöglichen es, Über- oder Unterrepräsentationen bestimmter Gruppen sichtbar zu machen. Verfahren wie Differential Privacy oder Pseudonymisierung können dabei im Einklang mit den Leitlinien des European Data Protection Board eingesetzt werden, ohne geschlechterbezogene Muster vollständig zu verdecken.

Modellebene

Auf Modellebene werden Fairness-Metriken eingesetzt, um diskriminierende Effekte algorithmischer Entscheidungen systematisch zu identifizieren. Methoden zur Bias-Reduktion, etwa Pre-Processing-, In-Processing- und Post-Processing-Ansätze, können kombiniert und regelmäßig evaluiert werden. Entsprechend den Anforderungen des AI Acts an Hochrisiko-Systeme sollten sowohl Performance als auch Fairness geschlechtsdifferenziert gemessen, dokumentiert und in regelmäßigen Reviews überprüft werden. Simulationen mit unterschiedlichen Gender-Profilen können zusätzlich dazu beitragen, systematische Benachteiligungen frühzeitig zu erkennen.

Governance-Ebene

Auf Governance-Ebene sind klare Rollen, Verantwortlichkeiten und Monitoring-Strukturen für die kontinuierliche Beobachtung von Bias erforderlich. Transparenzinstrumente wie Modellkarten oder Datasheets dokumentieren Annahmen, Grenzen und bekannte Verzerrungen von KI-Systemen. Für Anwendungen mit besonderer gesellschaftlicher Relevanz, insbesondere im Hinblick auf Geschlechtergerechtigkeit, können Gender Impact Assessments eingeführt werden, die mit bestehenden Instrumenten wie Datenschutz-Folgenabschätzungen verknüpft werden. Eine wirksame AI-Governance sollte zudem mit Gleichstellungs- und Diversity-Strategien der jeweiligen Organisation verbunden werden.

Kompetenzebene

Neben technischen und organisatorischen Maßnahmen spielt die Kompetenzentwicklung der beteiligten Akteur:innen eine zentrale Rolle. Schulungen zu Diversity, Bias und diskriminierungssensibler Technikgestaltung können dazu beitragen, die Sensibilität für strukturelle Verzerrungen zu erhöhen. Interdisziplinäre Teams, in denen Entwickler:innen, Data Scientists, Jurist:innen, Ethiker:innen sowie Vertretungen von Gleichstellungsstellen oder betroffenen Gruppen zusammenarbeiten, können Entwicklungs- und Implementierungsprozesse begleiten. Dadurch entsteht ein organisationales Lernsystem, das rechtliche Konformität mit kontinuierlicher ethischer Reflexion verbindet.

5. DISKUSSION UND THEORETISCHE EINORDNUNG

Die Ergebnisse der Interviewanalyse zeigen eine deutliche epistemische Differenz zwischen Rechtsexpert:innen und Ethikexpert:innen hinsichtlich der Bewertung der Legitimität algorithmischer Systeme. Jurist:innen orientieren sich primär an der rechtlichen Geltung bestehender Normen und institutioneller Regelwerke. In dieser Perspektive bilden Instrumente wie die Datenschutz-Grundverordnung (DSGVO) und der AI Act zentrale Referenzpunkte, um Datenschutz, Nichtdiskriminierung und Verantwortlichkeit zu gewährleisten (Van Bekkum & Borgesius, 2023; EU Council, 2025). Diese rechtliche Logik schafft einen Rahmen rechtsstaatlicher Kontrolle, definiert klare Zuständigkeiten und etabliert justiziable Verfahren, etwa in Form von Datenschutz-Folgenabschätzungen oder Risikomanagementprozessen. Bias wird innerhalb dieses Verständnisses primär als Risiko innerhalb eines regulierbaren Systems

betrachtet, das durch verfahrensorientierte Maßnahmen identifiziert, kontrolliert und reduziert werden kann.

Ethikexpert:innen vertreten demgegenüber ein erweitertes Verständnis von Verantwortung, das über formale Rechtskonformität hinausgeht und stärker auf gesellschaftliche Auswirkungen, strukturelle Gerechtigkeit und soziale Akzeptanz algorithmischer Entscheidungen fokussiert (UNESCO, 2021; Gaisch & Mader, 2025; Cromvelle, 2025). In dieser Perspektive wird Bias nicht nur als potenzieller Regelverstoß verstanden, sondern als Ausdruck und mögliche Verstärkung bestehender gesellschaftlicher Machtasymmetrien, insbesondere im Hinblick auf Geschlecht, ethnische Zugehörigkeit und soziale Herkunft (Hall & Ellis, 2023). Entsprechend betonen Ethikexpert:innen die Notwendigkeit intersektionaler Analysen, institutionalisierter Reflexionsräume sowie spezifischer Diversitätskompetenzen, um auch solche Effekte zu adressieren, die zwar formal rechtmäßig sein können, jedoch normativ problematische Folgen haben (Munarini, 2022; Gerards & Borgesius, 2022).

Diese unterschiedlichen Perspektiven werden insbesondere im Kontext geschlechtsspezifischer Verzerrungen deutlich. Entscheidungen, die aus juristischer Sicht als DSGVO-konform gelten, können aus ethischer Perspektive dennoch erhebliche Legitimitätsdefizite aufweisen, wenn sie systematische Benachteiligungen reproduzieren oder verstärken. Die Interviews verdeutlichen daher, dass Ethik als ein reflexiver Bewertungsraum fungiert, in dem nicht nur die Einhaltung rechtlicher Vorgaben, sondern auch die normativen Grundlagen der Systemgestaltung hinterfragt werden. Auf diese Weise trägt ethische Reflexion dazu bei, potenzielle blinde Flecken technokratischer Regulierung sichtbar zu machen und die gesellschaftlichen Folgen algorithmischer Entscheidungsprozesse stärker in den Blick zu rücken (Hintz, 2024; Bietti, 2020).

6. FAZIT UND AUSBLICK

Die Analyse zeigt, dass eine ausschließlich juristische Perspektive die Herausforderungen algorithmischer Verzerrungen nur teilweise erfassen kann. Regulatorische Instrumente wie die Datenschutz-Grundverordnung (DSGVO) und der europäische AI Act etablieren wichtige Mindeststandards für Datenschutz, Transparenz und Verantwortlichkeit und schaffen institutionelle Verfahren der Kontrolle (Van Bekkum & Borgesius, 2023; EU Council, 2025). Gleichzeitig orientieren sich diese Regelwerke primär an der Logik formaler Rechtskonformität und erfassen strukturelle Formen von Bias, insbesondere geschlechtsspezifische oder intersektionale Verzerrungen, nur begrenzt.

Die Expert:inneninterviews verdeutlichen daher die Bedeutung einer ergänzenden ethischen Perspektive. Während rechtliche Regulierung auf die Einhaltung kodifizierter Normen abzielt, richtet sich ethische Reflexion stärker auf gesellschaftliche Auswirkungen algorithmischer Entscheidungen sowie auf Fragen struktureller Gerechtigkeit, Machtasymmetrien und sozialer Akzeptanz (UNESCO, 2021; Gaisch, 2025; Cromvelle, 2025). Bias erscheint in dieser Perspektive nicht nur als technisches oder regulatorisches Risiko, sondern als mögliches Symptom gesellschaftlicher Ungleichheiten.

Die zentrale Erkenntnis des Beitrags besteht darin, dass Legalität und Legitimität unterschiedliche, aber komplementäre Geltungsdimensionen algorithmischer Systeme

darstellen. Rechtliche Regulierung schafft verbindliche Normen und Kontrollmechanismen. Ethische Reflexion ermöglicht hingegen eine kritische Bewertung der gesellschaftlichen Folgen algorithmischer Entscheidungen und macht normative Grenzen rechtlicher Regelwerke sichtbar (UNESCO, 2021; Balakrishnan, 2025; Paternolli et al., 2024).

Vor diesem Hintergrund erscheint eine Governance-Perspektive notwendig, die rechtliche Regulierung systematisch mit institutionalisierter ethischer Reflexion verbindet. Interdisziplinäre Ethik-Review-Boards sowie strukturierte Ethical Impact Assessments könnten dazu beitragen, rechtliche Compliance mit kontinuierlicher normativer Bewertung zu verknüpfen (UNESCO, 2021; Gaisch & Mader, 2025).

Eine nachhaltige Regulierung algorithmischer Systeme entsteht daher erst dort, wo rechtliche Legalität und gesellschaftliche Legitimität zusammengeführt werden und KI nicht nur regelkonform, sondern auch normativ verantwortbar gestaltet wird.

Literaturverzeichnis

Adams, R. (2024). *The new empire of AI: the future of global inequality*. John Wiley & Sons. Balakrishnan, A. (2024). *Ethical and legal implications of AI judges: Balancing efficiency and the right to fair trial.* Utrecht University. <https://studenttheses.uu.nl/handle/20.500.12932/48242>

Balakrishnan, S., Thongprayoon, C., Wathanavasin, W., Miao, J., Mao, M. A., Craici, I. M., & Cheungpasitporn, W. (2025). Evaluating artificial intelligence bias in nephrology: the role of diversity, equity, and inclusion in AI-driven decision-making and ethical regulation. *Frontiers in Artificial Intelligence*, 8, 1525937.

Bietti, E. (2020, January). From ethics washing to ethics bashing: a view on tech ethics from within moral philosophy. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 210-219).

Cromvelle, D. S. (2025). *Algorithmic discrimination and equal protection law: Legal remedies for AI bias in automated decision-making.* *International Journal of Law, Policy & Social Research*. * <https://ijlpsr.com/index.php/ijlpsr/article/download/1/1>

Dinker, N. (2024). Artificial intelligence and inequality: examining the social divides created by technological advancements. *International Journal of Innovations in Science, Engineering And Management*, 228-236.

European Data Protection Board. (2025). Guidelines on fairness under the GDPR. <https://edpb.europa.eu>

European Council. (2025). Council conclusions on gender equality and AI <https://www.consilium.europa.eu/en/press/press-releases/>

Farahani, M., & Ghasemi, G. (2024). Artificial intelligence and inequality: Challenges and opportunities. *Int. J. Innov. Educ*, 9, 78-99.

Gaisch, M. & Mader, I. (2025). *AI Ethics & Human Factors in AI. Excellence Edition*.

Gaisch, M. (2025). Systemic Ethical Dilemmas in Machine Learning: From Predictive Accuracy to Collective Fairness. *International Journal of Environmental Sciences*, 127-133.

Gerards, J., & Borgesius, F. Z. (2022). Protected grounds and the system of non-discrimination law in the context of algorithmic decision-making and artificial intelligence. *Journal of Technology and Human Rights*. <https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=3723873>

Gray, J., & Witt, A. (2021). A feminist data ethics of care for machine learning: The what, why, who and how. *First Monday*.

Hall, P., & Ellis, D. (2023). A systematic review of socio-technical gender bias in AI algorithms. *Online Information Review*, 47(7), 1264-1279.

Hintz, A. (2024). AI, big data and bias: governing datafication through a data justice lens. In *Handbook of Media and Communication Governance* (pp. 526-537). Edward Elgar Publishing. <https://www.elgaronline.com/book/9781800887206>

Kim, P. (2021). AI and Inequality. *Forthcoming in The Cambridge Handbook on Artificial Intelligence & the Law*, Kristin Johnson & Carla Reyes, eds.(2022), *Washington University in St. Louis Legal Studies Research Paper*, (21-09), 03.

Munarini, M. (2022). New perspectives on the mitigation of gender bias in AI by EU regulations. **Peace Human Rights Governance Journal.** <https://phrg.padovauniversitypress.it/system/files/papers/PHRG-2022-2-2.pdf>

Özer, M., Perc, M., & Suna, H. E. (2024). Artificial intelligence bias and the amplification of inequalities in the labor market. *Journal of Economy Culture and Society*, (69), 159-168.

Paternolli, V., Dalla Preda, M., & Giacobazzi, R. (2024, September). Fairness of AI Systems in the Legal Context. In *European Conference on Computer Vision* (pp. 53-67). Cham: Springer Nature Switzerland. https://link.springer.com/chapter/10.1007/978-3-031-92648-8_4

UNESCO. (2021). Recommendation on the ethics of artificial intelligence. <https://unesdoc.unesco.org/ark:/48223/pf0000380455>

Van Bekkum, M., & Borgesius, F. Z. (2023). Using sensitive data to prevent discrimination by artificial intelligence: Does the GDPR need a new exception? **Computer Law & Security Review.** <https://www.sciencedirect.com/science/article/pii/S0267364922001133>

Van Maanen, G. (2022). AI ethics, ethics washing, and the need to politicize data ethics. *Digital Society*, 1(2), 9.

Qing, Y. J., & Lim, E. (2022). A legal framework for artificial intelligence fairness reporting. **Cambridge Law Journal.** <https://www.cambridge.org/core/journals/cambridge-law-journal/article/legal-framework-for-artificial-intelligence-fairness-reporting/C2D73FBE9BB74E5D41DDA6>