

Table 5: List of selected studies.

ID	Title	Year
ES01	Fairness in machine learning: definition, testing, debugging, and application	2024
ES02	Envisioning responsible quantum software engineering and quantum artificial intelligence	2025
ES04	Algorithmic bias, fairness, and inclusivity: a multilevel framework for justice-oriented AI	2025
ES10	Bias recognition and mitigation strategies in artificial intelligence healthcare applications	2025
ES15	Towards fair AI: a review of bias and fairness in machine intelligence	2025
ES20	Towards AI-Enabled Cyber-Physical Infrastructures—Challenges, Opportunities, and Implications for a Data-Driven eGovernment Theory, Policy, and Practice	2025
ES22	Logging requirement for continuous auditing of responsible machine learning-based applications	2025
ES27	Navigating Healthcare AI Governance: the Comprehensive Algorithmic Oversight and Stewardship Framework for Risk and Equity	
ES31	Ethical implications of using general-purpose LLMs in clinical settings: a comparative analysis of prompt engineering strategies and their impact on patient safety	
ES44	Fairness-aware machine learning engineering: how far are we?	2024
ES49	A comprehensive analysis of gender, racial, and prompt-induced biases in large language models	2024
ES51	Guidelines for designing visualization tools for group fairness analysis in binary classification	2025
ES56	Pediatrics 4.0: the Transformative Impacts of the Latest Industrial Revolution on Pediatrics	2025
ES58	Toward a Responsible Fairness Analysis: From Binary to Multiclass and Multigroup Assessment in Graph Neural Network-Based User Modeling Tasks	2024
ES67	A fair individualized polysocial risk score for identifying increased social risk in type 2 diabetes	2024
ES71	Human-Centric Multimodal Machine Learning: Recent Advances and Testbed on AI-Based Recruitment	2023
ES75	The Wells-Du Bois Protocol for Machine Learning Bias: Building Critical Quantitative Foundations for Third Sector Scholarship	2022
ES77	A bias detection tree approach for detecting disparities in a recommendation model's errors	2022
ES80	Social norm bias: residual harms of fairness-aware algorithms	2023

ES90	Evaluating causes of algorithmic bias in juvenile criminal recidivism	2021
ES95	Fairness in recommender systems: research landscape and future directions	2023
ES100	Managing bias and unfairness in data for decision support: a survey of machine learning and data engineering approaches to identify and mitigate bias and unfairness within data management and analytics systems	2021
ES102	Users' needs in interactive bias auditing tools introducing a requirement checklist and evaluating existing tools	2023
ES111	Unmasking bias in artificial intelligence: a systematic review of bias detection and mitigation strategies in electronic health record-based models	2024
ES120	Explainability and Fairness in Machine Learning: Improve Fair End-to-end Lending for Kiva	2020
ES122	Ethical AI in Charitable Systems: A Framework for Bias Mitigation in Volunteer Matching	2025
ES125	Bias Mitigation Methods: Applicability, Legality, and Recommendations for Development	2024
ES127	HIFI: Explaining and Mitigating Algorithmic Bias through the Lens of Game-Theoretic Interactions	2025
ES130	Impact of Imputation Strategies on Fairness in Machine Learning	2022
ES136	BiasNavi: LLM-Empowered Data Bias Management	2025
ES138	"It's Everybody's Role to Speak Up... But Not Everyone Will": Understanding AI Professionals' Perceptions of Accountability for AI Bias Mitigation	2024
ES143	A Classification of Feedback Loops and Their Relation to Biases in Automated Decision-Making Systems	2023
ES146	EDITS: Modeling and Mitigating Data Bias for Graph Neural Networks	2022
ES149	D-hacking	2024
ES157	Tiny, Always-on, and Fragile: Bias Propagation through Design Choices in On-device Machine Learning Workflows	2023
ES160	Can Fairness be Automated? Guidelines and Opportunities for Fairness-aware AutoML	2024
ES161	Fairness Evaluation in Text Classification: Machine Learning Practitioner Perspectives of Individual and Group Fairness	2023
ES163	Mitigating Bias in Algorithmic Systems—A Fish-eye View	2023
ES171	Fairness in Online Ad Delivery	2024
ES174	Fair Graph Mining	2021

ES178	Bias on Demand: A Modelling Framework That Generates Synthetic Data With Bias	2023
ES180	FRINGE: context-aware FaiRness engineerING in complex software systEms	2024
ES183	Towards an Ontology-Driven Approach to Document Bias	2025
ES202	InFoRM: Individual Fairness on Graph Mining	2020
ES204	Same Same, But Different: Conditional Multi-Task Learning for Demographic-Specific Toxicity Detection	2023
ES210	Auditing for Racial Discrimination in the Delivery of Education Ads	2024
ES212	Toward A Two-Sided Fairness Framework in Search and Recommendation	2023
ES215	Toward Fair Ultrasound Computing Tomography: Challenges, Solutions and Outlook	2024
ES220	Fairness in Ranking, Part II: Learning-to-Rank and Recommender Systems	2023