

Fair Al

Navigating Fairness Definitions in AL



Muiahid Pasha Mohsin Pasha Masters Thesis BEW 2022-2023 Promotor: Prof. dr. Koenraad VANHOOF **Mentor: Elisavet KOUTSOVITI-KOUMERI**



•

.

Selected: 30 relevant papers

- iluded: 25 papers not meeting AI fairness in industries criteria
- **Examined industry-specific AI fairness**
- Insights: Explored challenges and opportunities in fair AI implementation
- Keywords: Artificial Intelligence, Fairness, AI Fairness Metrics, Industry Applications, Fairness in Industries, Systematic Review, Algorithmic Bias, Case Studies, Ethical AI, Responsible AI, Equity in AI.

AI fairness definitions vary widely across different sources and contexts. Lack of differentiation among 80 types of fairness definitions in AI design poses challenges. Legal frameworks like EU AI ACT and Canada's AIDA align with proposed design agenda. Bias mitigation and transparent practices are pivotal for equitable AI decision-making. Fragmented legal approaches highlight the need for a comprehensive design agenda. The findings underscore the varying definitions and challenges of fairness across industries,

emphasizing the need for interdisciplinary collaboration, legal alignment, and the potential of fair AI in promoting inclusive decision-making in industries like Human Resources, Financial sector and Health Care.

Recommendations & Limitations

- Develop a unified theoretical framework for fair AI design for different industries.
- Enhance interdisciplinary collaboration between AI, Software Engineering, and Legal domains.
- Align AI design agenda with emerging legal initiatives for ethical and responsible development.
- Prioritize addressing sources of bias through data, algorithm, and user perspectives.
- Foster transparency, accountability, and trust-building practices in AI systems.
- Limited Literature: Scarcity of AI fairness literature in industries limits comprehensive insights.
- Exclusion of Technical Details: Focusing on ethics excludes technical AI nuances from discussion.
- Contextual Variation: Findings' applicability may vary in diverse industries and cultures.





- Diverse fairness definitions exist, requiring clarity and differentiation.
- Various contexts and legal frameworks influence fair AI implementation.
- A theoretical design agenda for each industry can guide industry practitioners effectively.
- Collaborative efforts across disciplines are crucial for fair AI advancement.
- Responsible AI practices are crucial for building trust and inclusivity.



References Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and machine learning. http://fairmlbook.org.

Zhang, Y., & Zhou, L. (2019). Fairness assessment for artificial intelligence in financial industry. arXiv preprint arXiv:1912.07211.

Holstein, K., Wortman Vaughan, J., Daumé III, H., Dudík, M., Wallach, H., & Vaughan, J. W. (2019). Improving fairness in machine learning systems: What do industry practitioners need? In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (pp. 1-14).