Who Cares About Fairness:
How Background Influences the Way Practitioners Consider ML Harms

Author: Pablo Biedma Nunez
(P.BiedmaNunez@student.tudelft.nl)

Supervisor: Balayn, A.M.A.
Responsible Professors: Gadiraju, U. & Yang, J.

1 INTRODUCTION & BACKGROUND

❏ Unfairness in machine learning (ML) is a frequent subject of discussion, both, in academia and popular media [1].

❏ Recent literature introduced algorithmic solutions to bias in ML [1]. However, there is a disconnect between these solutions and practitioners' needs [2].

❏ To fill this gap, we aim to study how practitioners perceive fairness and how this is shaped by their background. Therefore our research question becomes: To what extent does background influence how ML practitioners consider ML harms?

2 METHOD

❏ In-depth literature review and preparations:
 ❏ Elaboration of list of sources of harm in ML.
 ❏ Devising use cases.
 ❏ Preparing datasets for use cases with harms.
 ❏ Developing Jupyter Notebooks with use cases.

❏ Empirical study:
 ❏ 30 semi-structured interviews.
 ❏ Participants across 16 countries (see Table 1).
 ❏ Diff backgrounds: education, role, country, tech area...
 ❏ Observe how they deal with potential harms.

Practitioner	Technology / Role	Interviewee ID
Student	NA	P5, P9, P10 P11, P12, P13 P18, P28
Developer	Federated Learning Data Scientist	P4 P6, P13, P16 P23, P26, P27
	Data Engineer	P15
	ML Engineer	P7
	IT Assistant	P17
	Fintech	P21, P22, P30
	Mobility	P24
	Recommender Systems	P25
Researcher	Computer Vision	P8, P14
	Predictive Maintenance	P1
	Ethics	P2
	Information Systems	P3
	Trustworthy AI	P10
	Medical AI	P19, P20
	Speech Recognition	P24
	Cognitive Networks	P29

Table 1. Overview of participants.

3 FINDINGS & CONCLUSION

❏ Different educational backgrounds lead to different data exploration practices.

❏ Discrepancies in how data should represent the context in which it is used.

❏ Failed to find increased ethical sensitivity in students who received ethics training.

❏ Researchers focus on aspects that practitioners in the industry usually ignore.

❏ Practitioners stand against demographic parity.

❏ May be useful to use different tools for different backgrounds.

4 LIMITATIONS & FUTURE WORK

❏ Ethical sensitivity is not just a trait, risk of false correlation.

❏ Small sample size and only 2 toolkits used.

❏ Other techniques as focus groups or questionnaires can also be used.

❏ Future work can also:

 ❏ Study concerns brought up throughout the paper

 ❏ Devise different tools for different backgrounds

References:
[1] Springer, A. et al. Assessing and addressing algorithmic bias - but before we get there. ArXiv, 2018.
[2] Holstein, K. et al. Improving fairness in machine learning systems: What do industry practitioners need? ArXiv, 2018.

TUDelft