





Accountability In AI WS 2023/2024

Kordian Gontarska, Weronika Wrazen, Felix Grzelka

<https://osm.hpi.de/aiai>

 Operating Systems and Middleware

Teaching Research Seminar Research IoT Lab Theses Publications People Posts 

Verantwortung der Informatik – Accountability In AI (AIAI 2023)

Prof. Dr. Andreas Polze

Kordian Gontarska, Weronika Wrazen, Felix Grzelka

Dates: Tuesday, 13:30-15:00 Uhr

In this seminar, we will talk about the accountability of computer science in the area of artificial intelligence. Each week a student will give a presentation, in which different perspectives of accountability, ethics, fairness, transparency, auditability, explainability, interpretability, and regulation are introduced. After each presentation, a group discussion about the presented topic will take place. The presentation should be based on literature and statements from recognized domain experts, however, it should also include an assessment of the arguments and the opinion of the presenter. Each 45-minute presentation should be single-handedly prepared by a participant using primary- and secondary literature. In preparation for the presentation, each participant will schedule a consultation with the supervisors and email a draft of the slides one week before the date of the presentation.

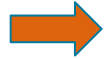
Verantwortung der Informatik – Accountability In AI

WS 2023/24

Introduction

Chart 2

Schedule



24.10.2023		Topic Deadline & Ethics 101
31.10.2023		Holiday (Reformationstag)
07.11.2023	Isabel Kurth	Fairness/Bias/Data diversity
14.11.2023	Jan Reppien	Recommender Systems – Data Collection / Privacy
21.11.2023	Laura Spies	How could a misuse of AI services to hijack democracy be prevented?
28.11.2023	Daphna Beljavskij	Generative AI vs Artists (SAG/AFTRA strike)
05.12.2023	Cara Scheuringer	Synthetic Media - Do DeepFakes Endanger Democracy?
12.12.2023	Jan Vincent Hoffbauer	State of Regulation
19.12.2023	Dilem Kaya	GDPR vs Big Data (in Medicine)
26.12.2023		Holiday (X-Mas)
02.01.2023		Holiday (New Year)
09.01.2023	Paula Marten	Dual-use
16.01.2023	Paula Klinke	Training on the Internet
23.01.2023	Judith Herrmann	Energy consumption
30.01.2023	Jessica Dierking	Recommender Systems – Social Implications / Filter Bubble
06.02.2023		TBD

Verantwortung der Informatik – Accountability In AI
WS 2023/22

Introduction

Chart 3

Intro



Source: <https://youtu.be/bOpf6KcWYyw>



The field of Ethics explained on one Slide

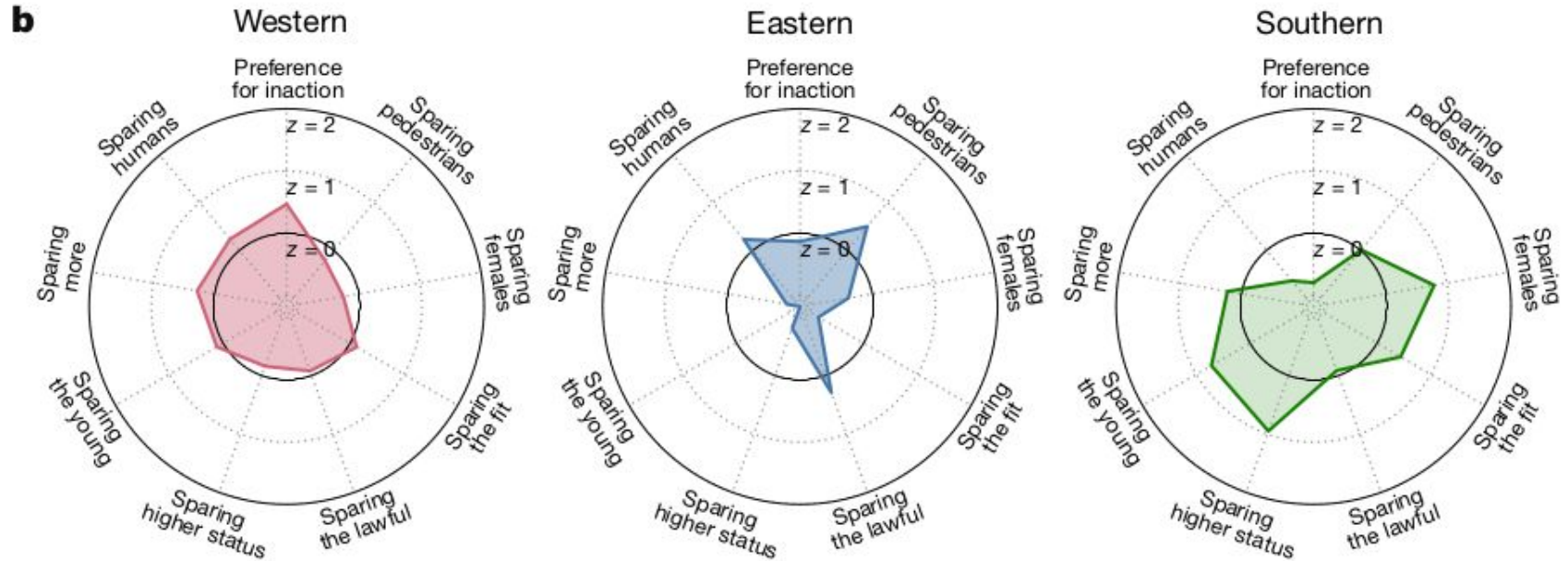
Philosophy Professors hate this Trick!

Ethics	
Deontological Ethics	Teleological Ethics (Consequentialism)
δέον, deon (Greek) = obligation, duty	τέλος, telos (Greek) = aim, goal
judge based on actions NOT their result	judge based on results
Kantianism: "Act only according to that maxim by which you can also will that it would become a universal law."	Utilitarianism: maximize wellbeing & minimize suffering

Moral Machine



Source: <https://www.moralmachine.net/>





Source: <https://youtu.be/1sl5KJ69qiA>

Your thoughts

- Was this experiment ethical?
- Was it useful?
- Is the trolley problem itself useful?

Bonus: GI Guidelines

<https://gi.de/ethicalguidelines>



Thanks for your attention!
Website: <https://osm.hpi.de/aiai>

kordian.gontarska@hpi.de

felix.grzelka@hpi.de

weronika.wrazen@hpi.de

Image References

- IBM course on 'Mitigating fairness'
- <https://www.businessdisabilityinternational.org/when-is-equality-not-equality/>
- Awad, E., Dsouza, S., Kim, R. et al. The Moral Machine experiment. Nature 563, 59–64 (2018). <https://doi.org/10.1038/s41586-018-0637-6>

**Verantwortung der
Informatik –
Accountability In
AI**

WS 2023/22

Introduction

Chart **13**