

Fairness und Diskriminierungsfreiheit aus Sicht von Ethik und Informatik

Seminar im Sommersemester 2019

Abschlussdiskussion | 15. Oktober 2019

INSTITUT FÜR THEORETISCHE INFORMATIK & INSTITUT FÜR PHILOSOPHIE



This image is a modified version of "Scales of Justice - Frankfurt Version" by Michael Gouffon (CC BY-NC 3.0 from 22 Sep 2012 via Flickr).

Diskriminierung durch (maschinell-gelernte) Entscheidungsverfahren

Diskriminierung durch (maschinell-gelernte) Entscheidungsverfahren

Hierzu Fragestellungen an der Verbindung zwischen . . .

- **Praktischer Philosophie:** Ethik
- und **Theoretischer Informatik:** Formale Logik & Kryptographie

Diskriminierung durch (maschinell-gelernte) Entscheidungsverfahren

Hierzu Fragestellungen an der Verbindung zwischen . . .

- **Praktischer Philosophie:** Ethik
- und **Theoretischer Informatik:** Formale Logik & Kryptographie

Beispielhafte Fragestellungen:

- Unter welchen Bedingungen spricht man von Diskriminierung?
- Sind faire Entscheidungen überhaupt möglich bzw. nützlich?
- Wie kann man sich sicher sein, dass ein Algorithmus fair handelt?
- Was bedeutet Fairness genau? Ist das eindeutig?
- Wie kann man Unfairness abmildern?

#	Thema	Studieng.
1.	Discrimination in the Philosophical Debate	EUKLID/SQ
2.	Discrimination as a Legal Term	EUKLID/SQ
3.	Racial Profiling	EUKLID/SQ
4.	Techniques to Discover and Evaluate Algorithmic Discrimination	INF
5.	Formalizing Fairness and Non-Discrimination	INF
6.	Counterfactual Fairness	INF
7.	Trade-Offs and the Cost of Fairness	INF
8.	Fairness and Social Equality (Economic Models)	INF
9.	Approximate Fair Treatment and Sanitizing Classifiers	INF