



**Wertorientierung in der angewandten
Forschung**
Beispiel: Faire Algorithmen

ZHAWARE 29.03.2023
CHRISTOPH HEITZ

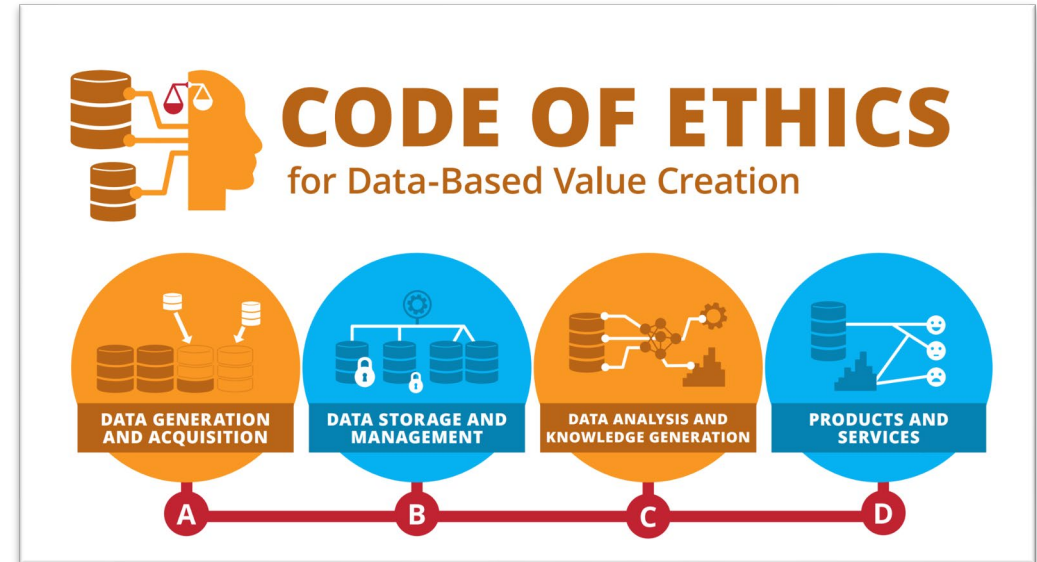
Mein Weg von der Physik zur Ethik in der Forschung

- Studium&Promotion in Physik, 4 Jahre in Industrie
- seit 2000 Dozent an der ZHAW, School of Engineering
- Gründungsmitglied Institut für Datenanalyse und Prozessdesign IDP (2001)
- Forschung: Datenbasierte Entscheidungssysteme
 - Seit >10 Jahren: Fokus auf Nutzung von Daten von Menschen (v.a. Kunden) im Dienstleistungssektor
 - «Was können wir aus Daten über unsere Kunden lernen? Und wie können wir damit neuen Mehrwert schaffen?»
- 2017: Gründung/Leitung Expert Group «Data Ethics»
 - Firmen & Hochschulen
 - Innerhalb des Schweizer Innovationsnetzwerkes data innovation alliance
- Seit 2019: Fokus auf ethische Fragen von KI, insbesondere Algorithmic Fairness

2019-20: Code of Ethics

Was sind die ethischen Fragen bei der kommerziellen Nutzung von Daten?

- Projektteam von Praktikern und Akademikern, verschiedene Disziplinen
- Kodex heute adaptiert von verschiedenen Firmen und Verbänden



Swiss Alliance for Data-Intensive Services, 2020.

<https://data-innovation.org/data-ethics/>

Algorithmen und ihre sozialen Kollateralschäden: Diskriminierung?

Zurich University
of Applied Sciences



PRO PUBLICA

Facebook Twitter Comment Donate

Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica
May 23, 2016

RETAIL · OCTOBER 11, 2018 / 1:04 AM / UPDATED 2 YEARS AGO

Amazon scraps secret AI recruiting tool that showed bias against women

NEWS · 24 OCTOBER 2019 · UPDATE 26 OCTOBER 2019

Millions of black people affected by racial bias in health-care algorithms

Study reveals rampant racism in decision-making software used by US hospitals – and highlights ways to correct it.

Apple Card Investigated After Gender Discrimination Complaints

A prominent software developer said on Twitter that the credit card was “racist” against women applying for credit.

UK ditches exam results generated by biased algorithm after student protests

Protesters chanted ‘Fuck the algorithm’ outside the country’s Department for Education

By Jon Porter | @JonPorty | Aug 17, 2020, 12:16pm EDT

Diskriminierung? - Ja, wir haben ein Problem



- Erschienen 2019 im Auftrag des Antidiskriminierungsstelle des Bundes (D).
- 47 dokumentierte Beispiele von implementierter Diskriminierung durch Algorithmen
- Lebensbereiche
 - › Arbeitsleben
 - › Medizin
 - › Verkehr
 - › Staatliche Sozialleistungen
 - › Bildung
 - › Polizei und Strafvollzug
 - › ...



Eigene Forschung: Algorithmic Fairness

Verschiedene Projekte in den letzten Jahren
(> 1 MCHF Fördermittel)

Interdisziplinäres Team: Philosophie,
Sozialwissenschaften, Computer Science

Unser Ziel: Faire Algorithmen

Weg: Konsistente Verknüpfung von

- Ethik und Philosophie (Was ist Fairness?)
- Soziale Entscheidungsprozesse (Was für eine Fairness wollen wir?)
- Technologische Umsetzung (Wie baut man konkret faire KI?)



Socially acceptable and fair
artificial intelligence

Projektteam

Zurich University
of Applied Sciences



Corinna Hertweck
PhD Stud, Computer Science
ZHAW



Michele Loi
Postdoc, Philosophie
Politechnico di Milano



Ulrich Leicht-Deobald
Prof., Sozialwissenschaft
University of St. Gallen



Markus Christen
PD, Ethik
University of Zurich



Serhiy Kandul
Postdoc,
University of Zurich
ZHAWARE 29.3.2023



Eleonora Viganò
Postdoc, Philosophie
University of Zurich

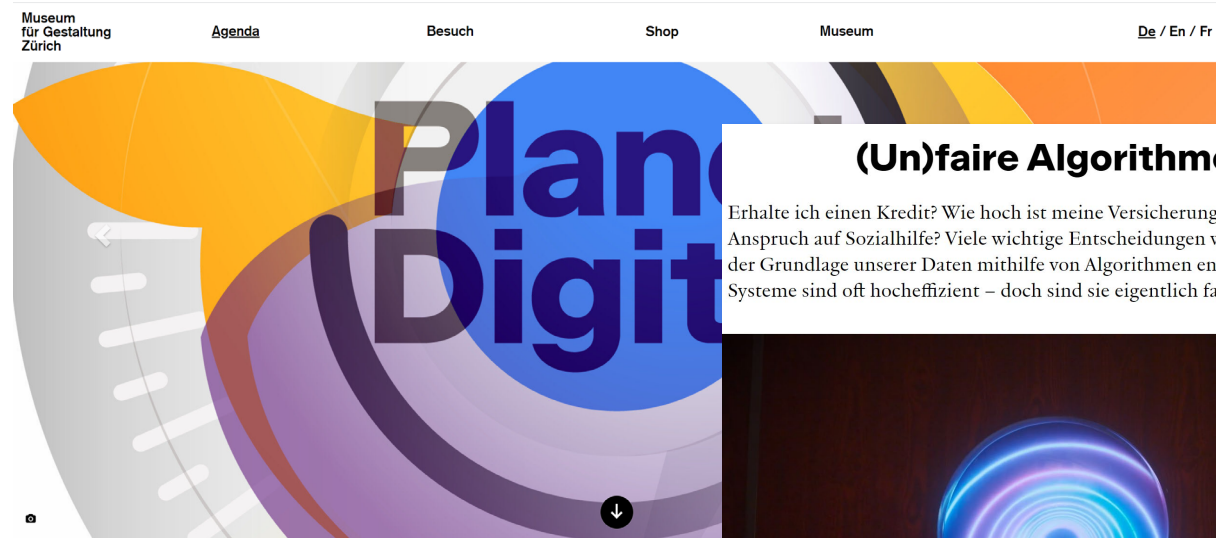


Christoph Heitz
Prof., Computer Science
ZHAW



Joachim Baumann
PhD Stud., Comp. Science
ZHAW

Nicht nur Forschung: Dialog mit Zivilgesellschaft



(Un)faire Algorithmen

Erhalte ich einen Kredit? Wie hoch ist meine Versicherungsprämie? Habe ich Anspruch auf Sozialhilfe? Viele wichtige Entscheidungen werden heute auf der Grundlage unserer Daten mithilfe von Algorithmen entschieden. Diese Systeme sind oft hocheffizient – doch sind sie eigentlich fair?

Können Algorithmen gerecht sein? Wird künstliche Intelligenz uns alle flüssig machen? Gibt es eine mediale Wahrheit? Und was haben Lava mit Datenschutz zu tun? Von seltenen Erden über gigantische Serverlandschaften und technische Gadgets bis zu sozialen Fragestellungen: Plan Digital nimmt unsere digitale Gegenwart und Zukunft unter die Lupe. Die Ausstellung von Universität Zürich und Museum für Gestaltung Zürich vom 2. Juni bis 1. Juli 2022.



Vier Kurzfilme zeigen in einer humorvollen Retroästhetik auf, wie Algorithmen unser Leben beeinflussen



Photo by Corinna Hertweck

Panel Discussion on (Un)fair Algorithms

June 2, 2022
Museum für Gestaltung
Zurich, Switzerland
In German

Einige Milestones

- **Forschung**
 - Publikationen, internationale Konferenzen, 2 best-paper awards
 - Gründung Konferenz «European Workshop on Algorithmic Fairness»
 - 2022: 40 Teilnehmende
 - 2023: 100-150 Teilnehmende
- **Forschung in die Praxis bringen**
 - Schweizweite Innovationsinitiative «Responsible AI» (2023)
 - Div. Workshops mit Firmen (z.B. Axa, UBS, ...)
- **Ausbildung :**
 - Bachelor Informatik/Data Science (FS 2023): Modul «Algorithmic Fairness»
 - Weiterbildung (2024): Modul «Umgang mit Daten» in CAS «Digitale Kompetenzen in der Sozialen Arbeit»
 - Weiterbildung (2024): Modul «Responsible AI» in CAS für Data Science