

# MANIPULATION UND DISKRIMINIERUNG IM KONTEXT VON KÜNSTLICHER INTELLIGENZ



**FABIENNE GRAF, MLAW, LL.M. (DUKE UNIVERSITY)**

**AZENES DATA SCIENCE EVENT: KI IM GESCHÄFTSALLTAG**

# Übersicht

- I. Kontext
- II. Diskriminierung
- III. Manipulation
- IV. Bedeutung für den Geschäftsalltag
- V. Ausblick

# I. Kontext

<b>Node 1 (Sigmoid)</b> class = 1st: 1.690 class = 2nd: 2.374 class = 3rd: -0.947 class = crew: 1.125 age = adult: -2.100 age = child: 2.109 gender = male: 0.518 gender = female: -0.534 Bias: -2.113	age = child: -1.902 gender = male: 2.504 gender = female: -2.546 Bias: 0.811  <b>Node 4 (Sigmoid)</b> class = 1st: 1.040 class = 2nd: 1.941 class = 3rd: 4.920 class = crew: 0.309 age = adult: 0.796 age = child: -0.820 gender = male: 2.422 gender = female: -2.413 Bias: -4.122  <b>Node 5 (Sigmoid)</b> class = 1st: -1.015 class = 2nd: 3.839 class = 3rd: -1.629 class = crew: 3.143 age = adult: 1.828 age = child: -1.782 gender = male: -0.188 gender = female: 0.189 Bias: -2.219	<b>Node 6 (Sigmoid)</b> class = 1st: 0.845 class = 2nd: 3.572 class = 3rd: -0.611 class = crew: -0.521 age = adult: -0.322 age = child: 0.312 gender = male: 0.540 gender = female: -0.610 Bias: -1.616  <b>Class 'yes' (Sigmoid)</b> Node 1: 2.034 Node 2: -0.605 Node 3: -3.760 Node 4: -1.499 Node 5: -1.194 Node 6: -0.555 Threshold: 3.439  <b>Class 'no' (Sigmoid)</b> Node 1: -2.034 Node 2: 0.605 Node 3: 3.760 Node 4: 1.499 Node 5: 1.194 Node 6: 0.555 Threshold: -3.439
---	---	--

## Allgemein

- Dematerialisierung
- Virtualisierung
- Personalisierung

## «KI»?

- Trend/Hype
- Skalierung
- Rückkopplung

[KNN, nach Abraham Bernstein, Positionspapier Künstliche Intelligenz, präsentiert am 10. November 2021, Universität Zürich ]

# I. Kontext

## «Künstliche Intelligenz» vs. «algorithmische Systeme»

Kollektiver Begriff?

Wissensbasierte Systeme,  
statistische Methoden,  
maschinelles Lernen (KNN etc.)

 Funktion:

- Automat. Entscheidungen
- Empfehlungen
- Schlussfolgerungen
- Vorhersagen

Technologieneutraler Begriff?

Applikation von Technologien  
im **sozialen Kontext**

 Wirkung:

- Individuen
- Kollektiv/Gesellschaft

# I. Kontext

## ... anders: EU «AI Act» (Entwurf vom 21.4.2021)

**ANHANG I**  
**TECHNIKEN UND KONZEPTE DER KÜNSTLICHEN INTELLIZENZ**  
**gemäß Artikel 3 Absatz 1**

- a) Konzepte des maschinellen Lernens, mit beaufsichtigtem, unbeaufsichtigtem und bestärkendem Lernen unter Verwendung einer breiten Palette von Methoden, einschließlich des tiefen Lernens (*Deep Learning*);
- b) Logik- und wissensgestützte Konzepte, einschließlich Wissensrepräsentation, induktiver (logischer) Programmierung, Wissensgrundlagen, Inferenz- und Deduktionsmaschinen, (symbolischer) Schlussfolgerungs- und Expertensysteme;
- c) Statistische Ansätze, Bayessche Schätz-, Such- und Optimierungsmethoden.

**ANHANG III**  
**HOCHRISIKO-KI-SYSTEME GEMÄß ARTIKEL 6 ABSATZ 2**

Als Hochrisiko-KI-Systeme gemäß Artikel 6 Absatz 2 gelten die in folgenden Bereichen aufgeführten KI-Systeme:

1. Biometrische Identifizierung und Kategorisierung natürlicher Personen:
  - a) KI-Systeme, die bestimmungsgemäß für die biometrische Echtzeit-Fernidentifizierung und nachträgliche biometrische Fernidentifizierung natürlicher Personen verwendet werden sollen;

# I. Kontext

## Recht

- Regelung vs. Regulierung
  - EU: «AI Act»
  - Schweiz: Bundesverfassung? Bundesgesetz?
- (Rechts-)Politik
- Rechtsfortbildung: «Hinterherhinken»?
- Geschäftsalltag und Rechtsanwendung

# II. Diskriminierung

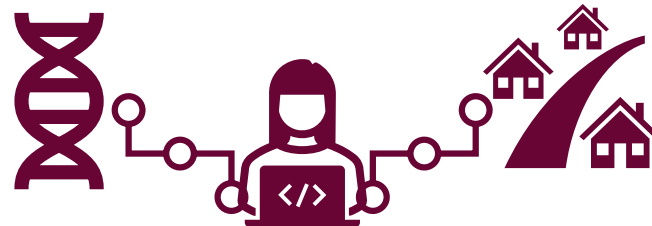
## Rechtsbegriff

- Qualifizierter Modus der Ungleichbehandlung:  
Abstellen auf Zugehörigkeit zu einer bestimmten Gruppe oder einem bestimmten Merkmal wie Geschlecht (einschließlich Geschlechtsidentität), sozialer oder ethnischer Herkunft, Sprache, Religion, politischer Überzeugung, Alter, Behinderung und sexueller Ausrichtung
- Art. 8 Abs. 2 BV (Diskriminierungsverbot)
- Art. 35 Abs. 3 BV (Drittwirkung von Grundrechten)

## II. Diskriminierung

### Fokus

- Direkte vs. indirekte Diskriminierung/Proxy-Diskriminierung?
- Bsp. HR-Applikation



### Relevanz

- Skalierung und Rückkopplung
- Prävention?



# II. Diskriminierung

## Rechtsdurchsetzung

- «Allg. Gleichbehandlungsgesetz»? DSG? ZGB?
- Lösungsansatz: Zwei-Stufen-Test
  1. Feststellung von Ungleichbehandlung  
z.B. Dokumentation; «human-in-the-loop»; externe Expertise  
⚡ Grundproblem: Feststellbarkeit und Korrelationen
  2. Qualifikation

[Test nach Florent Thouvenin et al, Workshop Report, Comprehensible Algorithms: A Legal Framework for the Use of Artificial Intelligence, Zürich/Basel 2023, S. 2]

# III. Manipulation

## Relevanz

- Beeinflussung von Menschen: Denken und Handeln
  - Sozialer Kontext
  - Abgrenzung: *Nudging, Dark Patterns* etc.
- Pro memoria: «KI» vs. «algorithmische Systeme»?

*Search Engines*

*Targeted Ads*

*Dating Apps*

*Social Media*

*Recommendations*

*Autofill*

*News Feeds*

*Personalized Prices*

*Infinite Scrolling*

# III. Manipulation

## Definitionsansatz

- Multidisziplinärer Begriff und Phänomen
- Willensbildung vs. Willensumsetzung
  - Ausgangslage Rechtsordnung
  - Durchsetzbarkeit
- Zum Nachteil?
- Rechtsgut: Autonomie

[siehe Fabienne Graf/Liliane Obrecht/Soraya Weiner, Erste Erkenntnisse zu Transparenz, Diskriminierung und Manipulation - Rechtliche Rahmenbedingungen für Künstliche Intelligenz in der Schweiz", in: Jusletter 2022]

# III. Manipulation

## Gegenstand des Rechts

- Technologieneutralität und Rechtsetzung

- Generalklauseln

### **Art. 28 Abs. 1 Schweizerisches Zivilgesetzbuch (ZGB)**

«Wer in seiner Persönlichkeit widerrechtlich verletzt wird, kann zu seinem Schutz gegen jeden, der an der Verletzung mitwirkt, das Gericht anrufen.»

### **Art. 2 Bundesgesetz gegen den unlauteren Wettbewerb (UWG)**

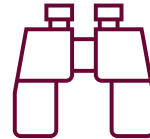
«Unlauter und widerrechtlich ist jedes täuschende oder in anderer Weise gegen den Grundsatz von Treu und Glauben verstossende Verhalten oder Geschäftsgebaren, welches das Verhältnis zwischen Mitbewerbern oder zwischen Anbietern und Abnehmern beeinflusst.»

## IV. Bedeutung für den Geschäftsalltag



### Ziele

- Innovationsförderung
- Verbreitungsförderung
- Ermöglichung und Anwendungsförderung
- Technikanwendung
  - Techniksicherheit
  - Technikhaftung



### Perspektiven

- Fokus Datenschutz?
- Verträge
- IP-Strategien  
(Werke, Designs, Marken und Erfindungen)

# V. Ausblick

## Herausforderungen

- Rechtsrahmen: Verhältnis zu EU und US?
- Transparenz (Politik)
- Mensch-Maschine-Interaktion: kritisches Denken, Bildung
- Entscheidungsfindung: menschliche/algorithmische/hybride?
- Rechtsdurchsetzung: individuelle/kollektive?

# V. Ausblick



## Comprehensible Algorithms: A Legal Framework for the Use of Artificial Intelligence

### Workshop Report

by FLORENT THOUVENIN, STEPHANIE VOLZ, FABIENNE GRAF and SORAYA WEINER (all University of Zurich), NADJA BRAUN BINDER and LILIANE OBRECHT (both University of Basel), as well as invited experts ANNIKA BAUMANN (Weizenbaum-Institut), YOCHANAN BIGMAN (Hebrew University), ANDREA BONEZZI (New York University), MIREILLE HILDEBRANDT (Radboud University), SANDRA MATZ (Columbia University), ELENA MUGELLINI (University of Applied Sciences of Western Switzerland, Fribourg), MASSIMILIANO OSTINELLI (Florida Atlantic University) and MATTHIAS SPIELKAMP (Private Researcher, AlgorithmWatch).

From 30 March to 1 April 2023, a group of experts in the fields of law, business administration, busi-

ness administration, law, and technology, a term that encompasses a range of technologies that automate decisions, recommendations, conclusions, or predictions. AI includes knowledge-based systems, statistical methods, and machine learning approaches (e.g., neural networks). The high-performance use of these technologies is mainly based on the combination of a large number of mathematical optimizations, extracting structures from significant amounts of data, using large computing capacities.

To avoid misleading associations, the participants agreed not to use the term AI in this workshop summary but rather speak of “algorithmic systems.” This term does not refer to specific current or future technologies but to the application of technologies in a social context. The need for legal assessment arises when technologies affect individuals and/or society. The term “algorithmic systems” also covers applications with similar effects as what is now referred to as AI that are based on technologies not usually qualified as AI, as well as technologies yet to be developed.



# HERZLICHEN DANK! FRAGEN?



**BIANCHISCHWALD LLC**  
bianchischwald.ch

5, rue Jacques-Balmat  
P.O. Box 5839  
CH-1211 Geneva 11  
**P** +41 58 220 36 00  
**F** +41 58 220 36 01

Genferstrasse 24  
P.O. Box 1534  
CH-8027 Zurich  
**P** +41 58 220 37 00  
**F** +41 58 220 37 01

12, avenue des Toises  
P.O. Box 5410  
CH-1002 Lausanne  
**P** +41 58 220 36 70  
**F** +41 58 220 36 71

Elfenstrasse 19  
P.O. Box 133  
CH-3000 Berne 15  
**P** +41 58 220 37 70  
**F** +41 58 220 37 71