



Die Macht der Algorithmen

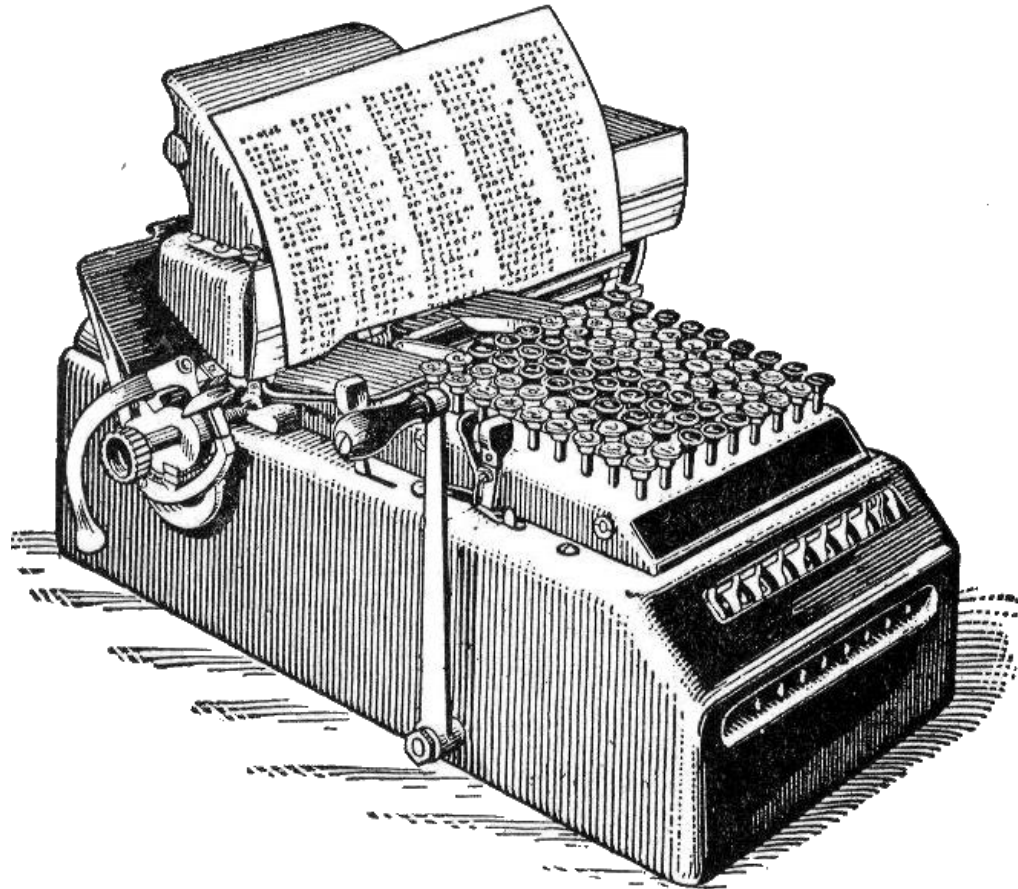
Prof. Dr. Katharina A. Zweig
TU Kaiserslautern

Jahrestagung des Ethikrates 2017
21.6.2017



ALGORITHM
WATCH

Wer entscheidet heute?



Kernaussagen



- Algorithmische Entscheidungssysteme bieten die Chance, bessere Entscheidungen zu treffen. (im Folgenden: ADM Systeme)
- Algorithmische Entscheidungssysteme und ihre Einbettung in gesellschaftliche Prozesse sind so komplex, dass es auf vielen Ebenen zu Fehlern kommen kann.
- Algorithmische Entscheidungssysteme sind inspizierbar und teilweise kontrollierbar.
- Wir brauchen eine demokratisch legitimierte, staatsferne Überprüfungsstelle.
- Dazu müssen wir als Gesellschaft die kleinste Menge an ADM Systemen definieren und weitere wichtige Eckpunkte festlegen.¹



¹ ADM-Manifest von Algorithm Watch, veröffentlicht 2016, <https://algorithmwatch.org/de/das-adm-manifest-the-adm-manifesto/>

Der Wunsch nach objektiven Entscheidungen



- Die USA haben die zweithöchste Gefangenquote (666 von 10.000 Einwohnern¹).
- Dabei stellen Afro-Amerikaner 40% der Inhaftierten, obwohl sie nur 13% der US-amerikanischen Bevölkerung ausmachen.²
- Soweit dieses Phänomen auf latentem Rassismus beruht, könnte man ihm mit ADM Systemen begegnen.



¹ http://www.prisonstudies.org/highest-to-lowest/prison_population_rate?field_region_taxonomy_tid=All, International Center for Prison Studies, basierend auf Daten im Frühjahr 2017, am 18.6.2017 abgefragt.

² https://en.wikipedia.org/wiki/Incarceration_in_the_United_States#Ethnicity

Wie lernen Computer?



Was hat eine Chance, von meinem Kind gegessen zu werden?



Wie lernen Computer



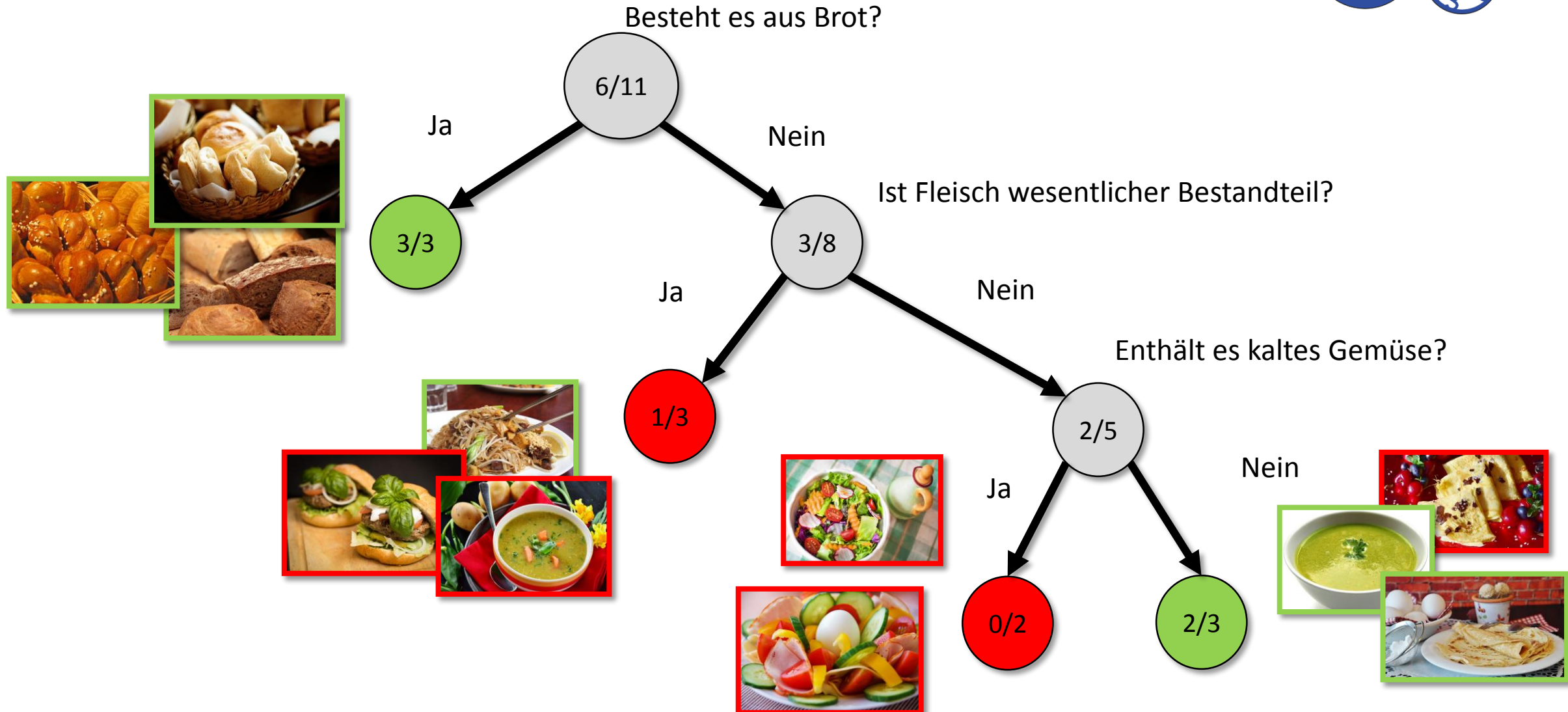
- Datenpunkte mit Erfahrungswerten
 - Z.B. Garmethode (gebacken, gebraten)
 - Art der Speise (Suppe, Gratin)
 - Zutaten
 - Das zu bewertende Verhalten:
 - Wurde gegessen oder wurde verweigert?
- Mensch und Computer versuchen in ähnlicher Weise, die wichtigsten Regeln zu extrahieren.



Daten



Entscheidungsbäume



Rückfälligkeitsvorhersage in den USA

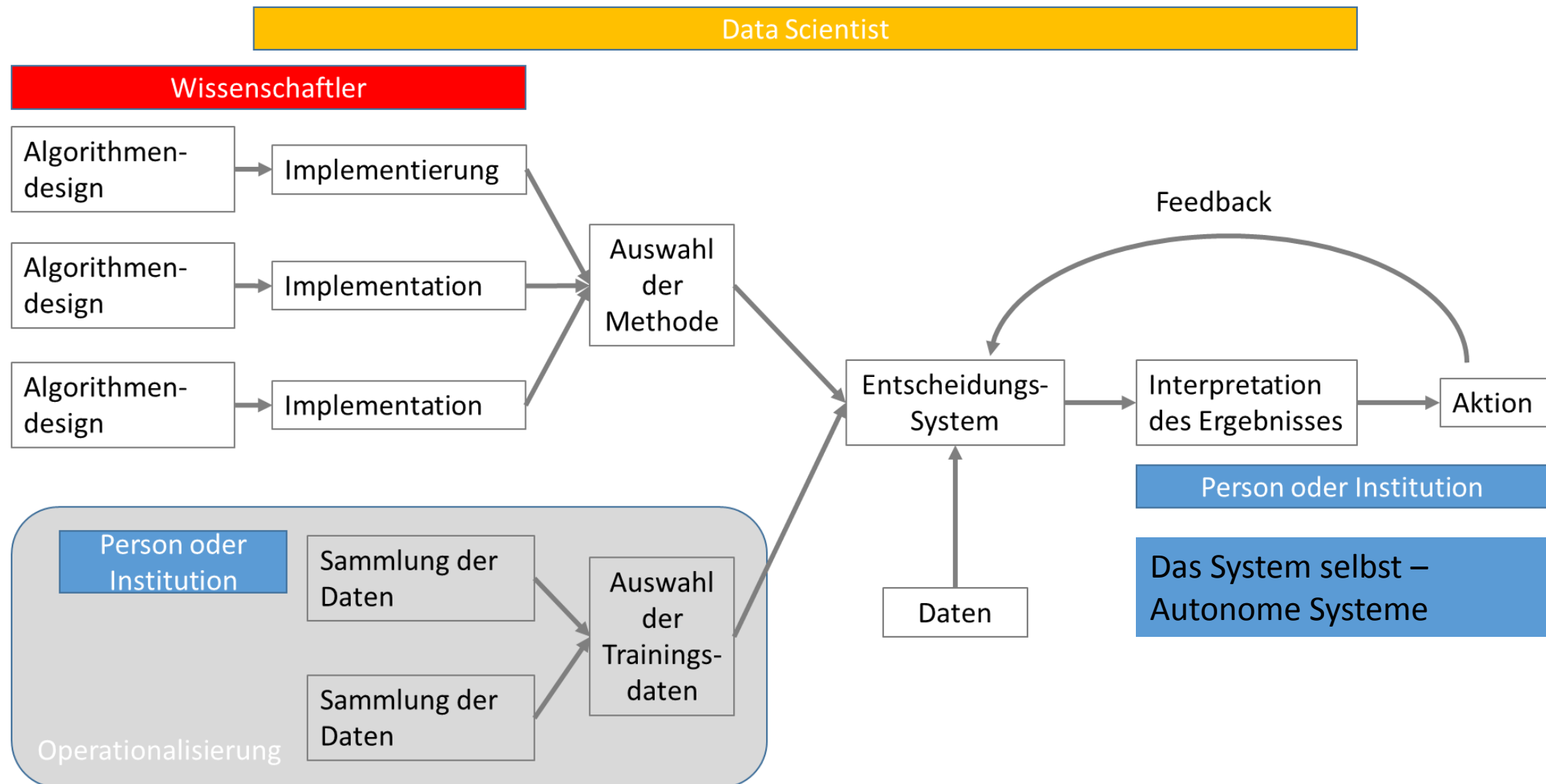


- Basierend auf allgemeinen Angaben (Alter, Vorstrafen, ...) und einem Fragebogen.
 - Auf Basis der Rückfälligkeitsraten von Kriminellen früherer Jahrgänge wurde das ADM System dann trainiert.
 - Neue Verdächtige werden mit einer Zahl hinsichtlich ihres Risikos bewertet.
 - Die obersten 30% werden als „Hochrisikogruppe“ bezeichnet.
 - Algorithmische Sippenhaftung;
 - Fehlerwahrscheinlichkeit von 50% bzw. 80% (!).¹
 - Ursprünglich für Vergabe knapper Ressourcen nach Haft/bei Bewährung, jetzt oft **vor** dem Schuldspruch.
- Frage 4: Würden Sie sagen, dass diese Person möglicherweise oder zugegebenermaßen ein Gang-Mitglied ist?
 - Fragen 33-36: War Ihr Vater, Mutter, Bruder, Lebenspartner(in) ihres Wissens jemals in Haft?

Fragen aus: <https://www.documentcloud.org/documents/2702103-Sample-Risk-Assessment-COMPAS-CORE.html>

¹ Eigene Forschung; vergleiche Angwin, J.; Larson, J.; Mattu, S. & Kirchner, L.: "Machine Bias - There's software used across the country to predict future criminals. And it's biased against blacks.", *ProPublica*, 2016

Wie subjektiv sind sie nun, die Algorithmen?



Was können wir also tun?



- Chancen eruieren, Probleme genau definieren, Qualitätsmaße für ADM Systeme entwickeln.
 - Dazu breite Diskussion notwendig.
- Entwicklung von ADM verbessern: interdisziplinäre Teams, Curriculum für Data Scientists, Ethikregeln
- Demokratisch legitimierte Prüfungsstelle aufbauen.

Weiterführende Literatur



- Katharina A. Zweig: „Was ist ein Algorithmus“, 1. Arbeitspapier von Algorithm Watch, veröffentlicht am 12.5.2016, <https://algorithmwatch.org/de/arbeitspapier-was-ist-ein-algorithmus/>
- Katharina A. Zweig: „Überprüfbarkeit von Algorithmen“, 2. Arbeitspapier von Algorithm Watch, veröffentlicht am 2.7.2016, <https://algorithmwatch.org/de/zweites-arbeitspapier-ueberpruefbarkeit-algorithmen/>
- Cathy O’Neil: „Weapons of Math Destruction“, Crown Publishing Group, New York, 2016
- Katharina A. Zweig: „Wo Maschinen irren können“, Studie der Bertelsmann Stiftung, erscheint im Sommer 2017
- Katharina A. Zweig, Tobias D. Krafft, Marc Hauer: „Dein Algorithmus – meine Meinung“, veröffentlicht im April 2017, Bayerische Landeszentrale für neue Medien, <https://www.blm.de/aktivitaeten/medienkompetenz/materialien/algorithmenbroschuere.cfm>

Twitter: @nettwwerkerin Email: zweig@cs.uni-kl.de