
Ethische und gesellschaftliche Perspektiven in der Informatik lehren – Ein Praxisbeispiel

4ING Workshop

12.10.22 | Catharina Rudschies | Universität Hamburg

Rise of the racist robots - how AI is learning all our worst impulses

There is a saying in computer science: garbage in, garbage out. When we feed machines data that reflects our prejudices, they mimic them - from antisemitic chatbots to racially biased software. Does a horrifying future await people forced to live at the mercy of algorithms?



▲ Current laws 'largely fail to address discrimination' when it comes to big data. Photograph: artpartner-images/Getty Images

In May last year, a stunning **report** claimed that a computer program used by a US court for risk assessment was biased against black prisoners. The program, Correctional Offender Management Profiling for Alternative Sanctions (Compas), was much more prone to mistakenly label black defendants as likely to reoffend - wrongly flagging

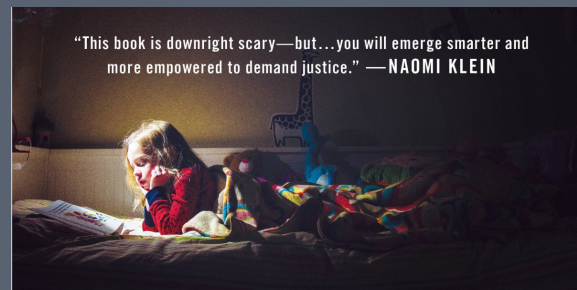


Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)

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



AUTOMATING INEQUALITY

HOW HIGH-TECH TOOLS PROFILE,
POLICE, AND PUNISH THE POOR





FEATURES

ARTIFICIAL INTELLIGENCE IS GOING TO SUPERCHARGE SURVEILLANCE

What happens when digital eyes get the brains to match?

By [James Vincent](#) | Jan 23, 2018, 10:54am EST

Illustrations by [Garret Beard](#)

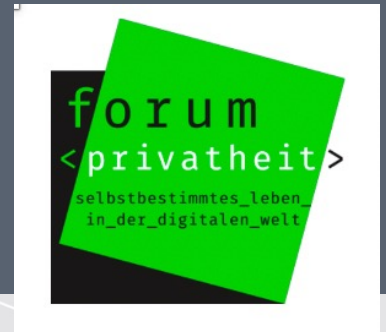
MIT
Technology
Review

We Need to Talk About the Power of AI to Manipulate Humans

Our tendency to become emotionally attached to chatbots could be exploited by companies seeking a profit.

by [Liesl Yearsley](#)

June 5, 2017



Policy Paper

RISIKEN KÜNSTLICHER INTELLIGENZ FÜR DIE MENSCHLICHE SELBSTBESTIMMUNG

Philosophie, Gesellschaft & IT (PGIT)

Modul: Vorlesung und begleitendes Seminar

- am Fachbereich Informatik der Universität Hamburg
- philosophische und sozialwissenschaftliche Perspektive auf Informationstechnologien, ihre Voraussetzungen und potentiellen Auswirkungen

PGIT - Ziele

- Sensibilisierung von Informatik-Studierenden für Risiken und Implikationen von IT
- Ursache von Problemen
- Ansätze, Probleme zu vermeiden

PGIT - Inhalte

- Computerethik
- Erkenntnis- und Wissenschaftstheorie
- Ethische Prinzipien wie Transparenz, Privatsphäre, Fairness, etc. im Kontext von Big Data
- Ansätze wie Values in Design / Value-Sensitive-Design
- Forschungs- und Datenethik
- Big Data und Politik

Für wen?

- Bachelor-Studierende des Fachbereichs Informatik verschiedener Studiengänge
 - B. Sc. Informatik, B. Sc. Mensch-Computer-Interaktion, B. Sc. Software-System-Entwicklung, B. Sc. Wirtschaftsinformatik, B. Sc. Computing in Science
- Bachelor- sowie Master-Studierende im Lehramtsstudium an der Universität Hamburg



PGIT – Entwicklung Didaktik

Präsenz-Format

- Vorlesung
- Referate und Diskussionen im Seminar
- Hausarbeit

2017 - 2019

Distanz-Format asynchron

- Aufgezeichnete Vorlesungen
- 1 Abgabe / Woche (bewertet)
- 1 Abschlussreflexion

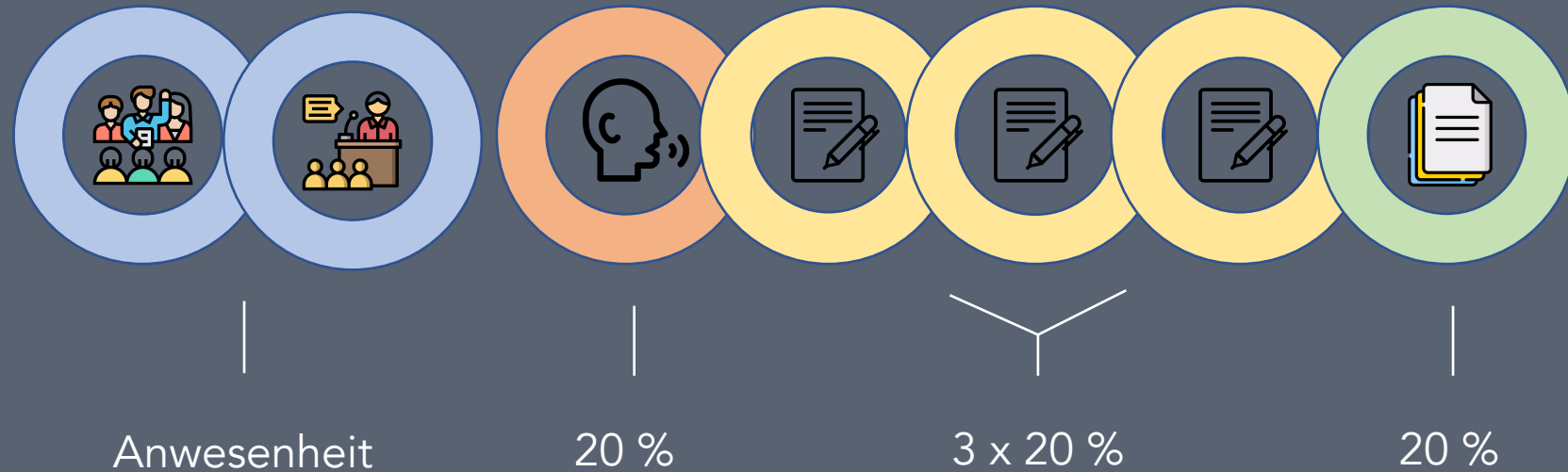
2020 (Corona-Lockdown)

Distanz-Format asynchron/synchron

- Aufgezeichnete Vorlesungen
- 1 Abgabe / Woche (bewertet)
- Alle 3-4 Wochen Online-Seminar
- 1 Abschlussreflexion

2021 (Corona-Lockdown)

PGIT – Entwicklung



2022

- Vorlesung und Seminar in Präsenz (wöchentlich)
- 1 Abgabe / 4 Wochen (bewertet) (insg. 3 Abgaben)
- Bewertung der mündlichen Mitarbeit
- 1 Abschlussreflexion

Studierenden-Feedback Lehrformate



Assignments:

- Anreiz, die Texte überhaupt zu lesen
- Anreiz, die Texte anders zu lesen/bearbeiten
 - Intensive Beschäftigung mit dem Text bis zum Verständnis / Extrahieren der wichtigen Elemente
- Regelmäßige Bearbeitung der Texte / geben Struktur über das Semester hinweg / kein Aufschieben
- Nachhaltiges Lernen
- Fragen in Assignments geben wichtige Hilfestellung beim Lesen / Bearbeiten

Studierenden-Feedback Lehrformate



Mündliche Benotung:

- Anreiz, sich intensiv vorzubereiten
- Erhöhte mündliche Beteiligung und intensivere Diskussionen
- Höhere Qualität der Diskussionen
- Durch gute und regelmäßige Vorbereitung Verknüpfungen zwischen Lerninhalten gemacht

Studierender PGIT: „...Antriebskraft für ein dynamisches Klima und einer größtenteils tollen Mitarbeit sowie Austausch untereinander...“

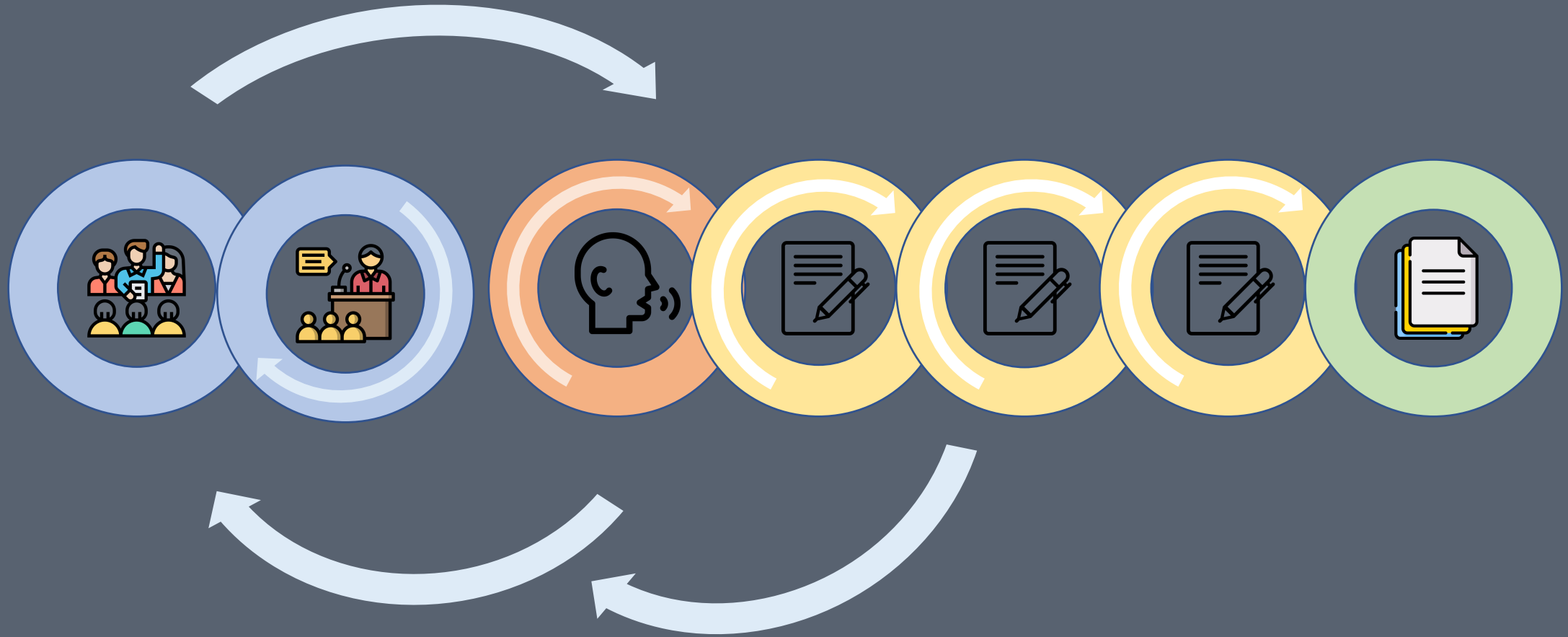
Studierenden-Feedback Lehrformate



Präsenzseminar:

- Vertieftes Verständnis und Erweiterung des Wissens
- Diskussion und Interaktion im Seminar hilfreich für Bearbeitung der Assignments
- Vortragen von eigenen Gedanken, Standpunkten und Wissen

Studierenden-Feedback Lehrformate



Studierenden-Feedback Inhalte



Erkenntnisse und Erlerntes:

- Ethik ist relevant/wichtig
- Interesse und Sensibilisierung für ethische Themen/Probleme
- Reflexionsfähigkeit und kritisches Hinterfragen
- Strukturierter Umgang mit Fachlektüre

Studierenden-Feedback Inhalte



Erkenntnisse für die Arbeit als Informatiker*in:

- Verantwortung der eigenen Arbeit, (Design-) Entscheidungen und deren Auswirkungen
- Selbstreflexion → Entscheidungen, eigenes Handeln und „Objektivität“ hinterfragen
- Neue/vermehrte Wahrnehmung von Risiken und Risikogruppen (z.B. vulnerable Gruppen)

PGIT - Entwicklung

Steigende Studierendenzahlen

Zunahme in den Anfragen für Bachelor- und Masterarbeiten

	PGIT Teilnehmende	EIT (englischsprachiges Parallelangebot)
SoSe 2017	61	
SoSe 2018	45	26
SoSe 2019	30	25
SoSe 2020	112	24
SoSe 2021	101	39
SoSe 2022	100 (ca. 150 Anmeldungen)	

PGIT – Ergebnisse

- Sensibilisierung von Informatik-Studierenden für Risiken und Implikationen
- Studierende wollen Probleme/Aspekte weiter verfolgen
- Studierende wollen Ansätze anwenden, um Probleme zu vermeiden

Danke!
