

Im Loop – Ungleichheiten in algorithmischen Entscheidungssystemen und ihren Daten

<https://www.pexels.com/de-de/foto/bildschirm-verschlüsselung-programmierung-web-design-2061168/>

Zentrale Fragen:

- Inwiefern sind AES an der Produktion von Wissen und an der Konstruktion von Wirklichkeit beteiligt?
- Inwiefern werden in diesen Prozessen gesellschaftliche Ungleichheiten (re-)produziert?

Inhalt: Im Loop...

1. Einleitung: Zentrale Fragen
2. Was sind algorithmische Entscheidungssysteme (AES)?
3. Ungleichheiten, Daten und AES: Problemfelder
 - 3.1 Repräsentativität von Daten
 - 3.2 Auslöschung von ‚Abweichungen‘ – Reproduktion der Norm
 - 3.3 Manifestation von gesellschaftlichen Ungleichheiten in Lebensrealitäten und Daten
 - 3.4 Datafizierung des Sozialen
4. Ungleichheiten und AES: weitere Problemfelder
5. Und nun?

2. Was sind algorithmische Entscheidungssysteme (AES)

?

„Buzzword“ Künstliche Intelligenz

- Was gilt als intelligent?
- meist sehr rationalistisch gedacht (z.B. Schachcomputer)
- Automatisierung kognitiver Prozesse

Algorithmus

- = mathematische Handlungsanweisung
- Legt Schritt für Schritt fest, wie eine Aufgabe bearbeitet werden soll
- Erzielt ausgehend von einem bestimmten Input, einem Datensatz, einen beachsichtigten Output
- Regelbasierter vs. lernende Algorithmus

Lernende Algorithmen (siehe auch: Machine Learning)

- Finden in Daten selbstständig Muster und Korrelationen
- können die erlernten Korrelationen auf neue Datensätze anwenden
- → Relevanz von ‚Trainingsdaten‘

vgl. Günther 2022, AlgorithmWatch 2022, Fry 2019, Bächle 2016

2. Was sind algorithmische Entscheidungssysteme (AES)?

Algorithmische/Automatisierte Entscheidungssysteme (AES)

- In AES werden Algorithmen automatisiert ausgeführt
- AES sprechen Empfehlungen für Entscheidungen aus, häufig in Form von Scores
- AES berechnen Wahrscheinlichkeiten
- Ziel: realweltliche Problemstellungen lösen, indem diese in computerles- und automatisierbaren ‚Code‘ übersetzt und formale Lösungswege definiert werden

- sind soziotechnische Systeme: Menschen bestimmen, welche Aufgaben AES wie lösen sollen + AES beinhalten eigene technische und epistemologische Affordanzen

- formen Gesellschaft und Subjekte und sind gleichzeitig durch diese geprägt

Vgl. Günther 2022, AlgorithmWatch 2022, Fry 2019, Bächle 2016

3. Ungleichheiten, Daten und AES: Problemfelder

- 3.1 Repräsentativität von Daten
- 3.2 Auslöschung von ‚Abweichungen‘ – Reproduktion der Norm
- 3.3 Manifestation von gesellschaftlichen Ungleichheiten in Lebensrealitäten und Daten
- 3.4 Datafizierung des Sozialen und Datenfundamentalismus

3.1 Repräsentativität von Daten

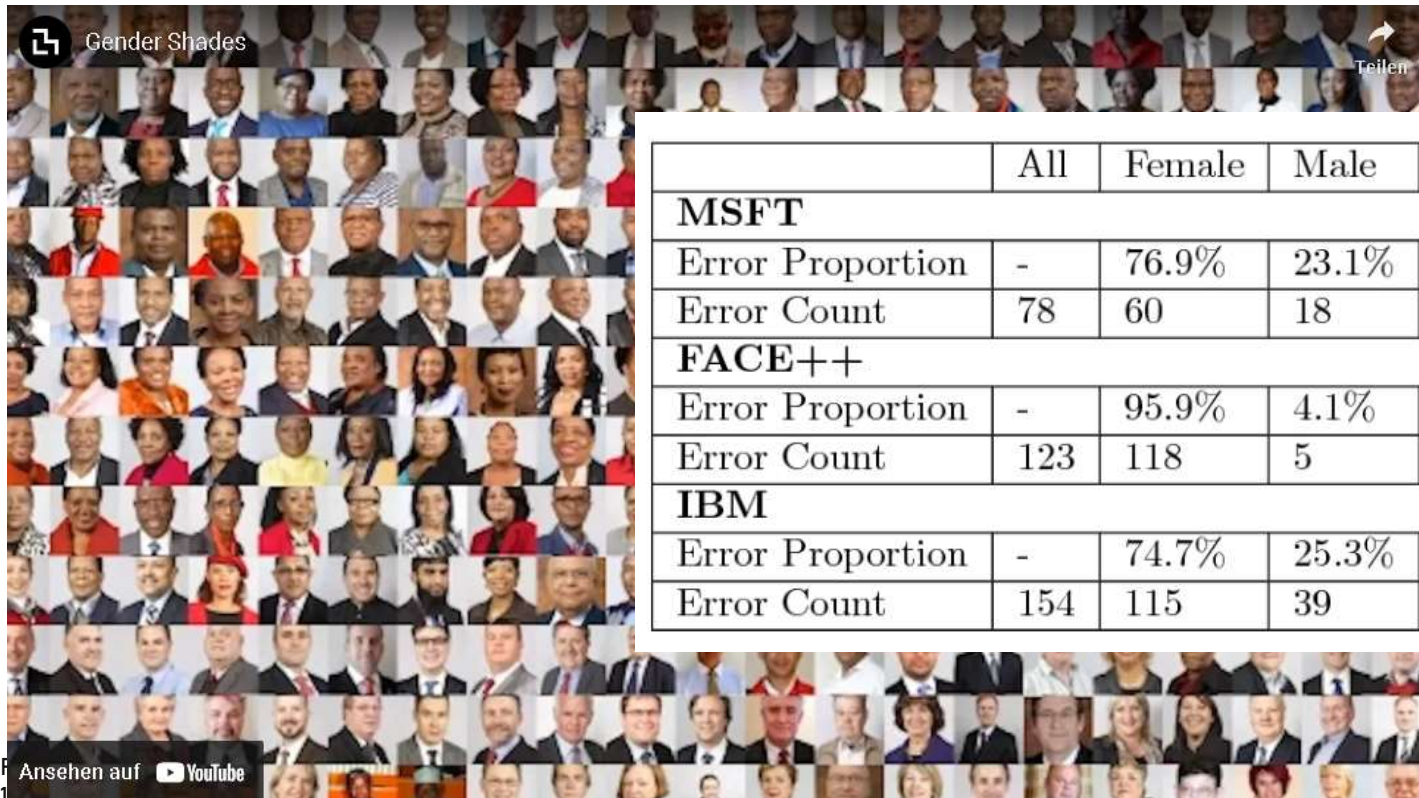
- Beispiel: Gesichtserkennung
- Studie ‚Gender Shades‘ von Joy Buolamwini & Timnit Gebru (2018)



https://de.wikipedia.org/wiki/Joy_Buolamwini



https://de.wikipedia.org/wiki/Timnit_Gebru



	All	Female	Male	Darker	Lighter	DF	DM	LF	LM
MSFT									
Error Proportion	-	76.9%	23.1%	93.6%	6.4%	70.5%	23.1%	6.4%	0.0%
Error Count	78	60	18	73	5	55	18	5	0
FACE++									
Error Proportion	-	95.9%	4.1%	74.0%	26.0%	72.4%	1.6%	23.6%	2.4%
Error Count	123	118	5	91	32	89	2	29	3
IBM									
Error Proportion	-	74.7%	25.3%	85.7%	14.3%	61.0%	24.7%	13.6%	0.6%
Error Count	154	115	39	132	22	94	38	21	1

3.2 Auslöschung von ‚Abweichungen‘ – Reproduktion der Norm

- Beispiel: selbstfahrende Autos und Menschen mit Behinderungen (Treviranius 2019)
- Beispiel: bildgenerierende Systeme (Ugless et al. 2023)

Stereotypes and Smut: The (Mis)representation of Non-cisgender Identities by Text-to-Image Models

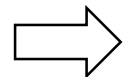
Eddie L. Ungless
School of Informatics
University of Edinburgh
Scotland

Björn Ross
School of Informatics
University of Edinburgh
Scotland

Anne Lauscher
Data Science Group
University of Hamburg
Germany



<https://www.pexels.com/de-de/foto/frau-bur-o-bewegung-geschafft-4064333/>

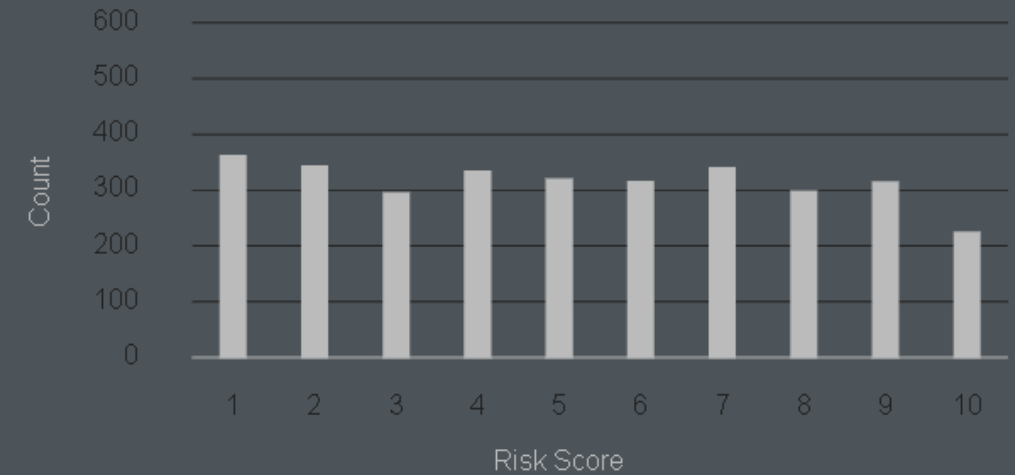


Welche Normen/Normalitäten werden mittels AES (re-)produziert?

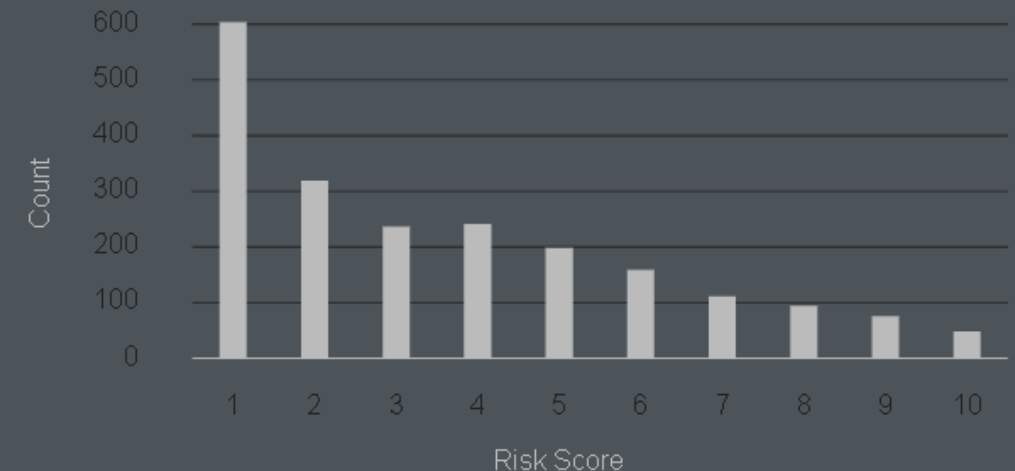
3.3 Manifestation von Ungleichheiten in Lebensrealitäten und Daten

- Beispiel Justizsystem: Vorhersage von erneuten Straftaten in den USA mittels AES (ProPublica 2016)
- Risiko-Scores für die Rückfälligkeit von Angeklagten bei Schwarzen Menschen höher
- Score berechnet sich aus 137 Fragen
 - z.B. *Was one of your parents ever send to jail?*
- Strukturelle Ungleichheiten bilden sich in Lebensrealitäten und Daten ab
- Selbstverstärkender Zirkel: Discrimination-Feedback-Loop (Whittaker et al. 2019)

Black Defendants' Risk Scores



White Defendants' Risk Scores



3.4 Datafizierung des Sozialen und Datenfundamentalismus

Zur Konstruktion von Daten

- Auf Grundlage welcher Daten wird ein AES trainiert?
 - Daten sind nie neutral
 - Daten sind immer ausschnitthaft
 - Datensätze werden produziert, kuratiert
 - Zeitpunkt, Kontext, Klassifizierungen
 - manuelles Labeling

Datenextraktion und Privacy

- (Zu) Wem gehören die extrahierten Daten?
 - Willkür in Datenextraktion vs. consentful-tech und privacy

vgl. Crawford 2022, Prietl 2019



3.4 Datafizierung des Sozialen und Datenfundamentalismus

- „Datenfundamentalismus“ (Crawford 2013): Daten als reiner Ausdruck von Realität, Wahrheit findet sich in Daten, Abkehr von theoretischem Verstehen
- „Technochauvinismus“ (Broussard 2019): computerisierte Lösungen sind immer besser
- Beispiel: Umgang mit Geschlecht in Data Analytics

```
array([[238, 233, 229],
       [238, 233, 229],
       [238, 233, 229],
       ...,
       [242, 237, 233],
       [242, 237, 233],
       [242, 237, 233]]],

[[238, 233, 229],
 [238, 233, 229],
 [238, 233, 229],
 ...,
 [242, 237, 233],
 [242, 237, 233],
 [242, 237, 233]]],

[[238, 233, 229],
 [238, 233, 229],
 [238, 233, 229],
 ...,
 [242, 237, 233],
 [242, 237, 233],
 [242, 237, 233]]],

...,
```

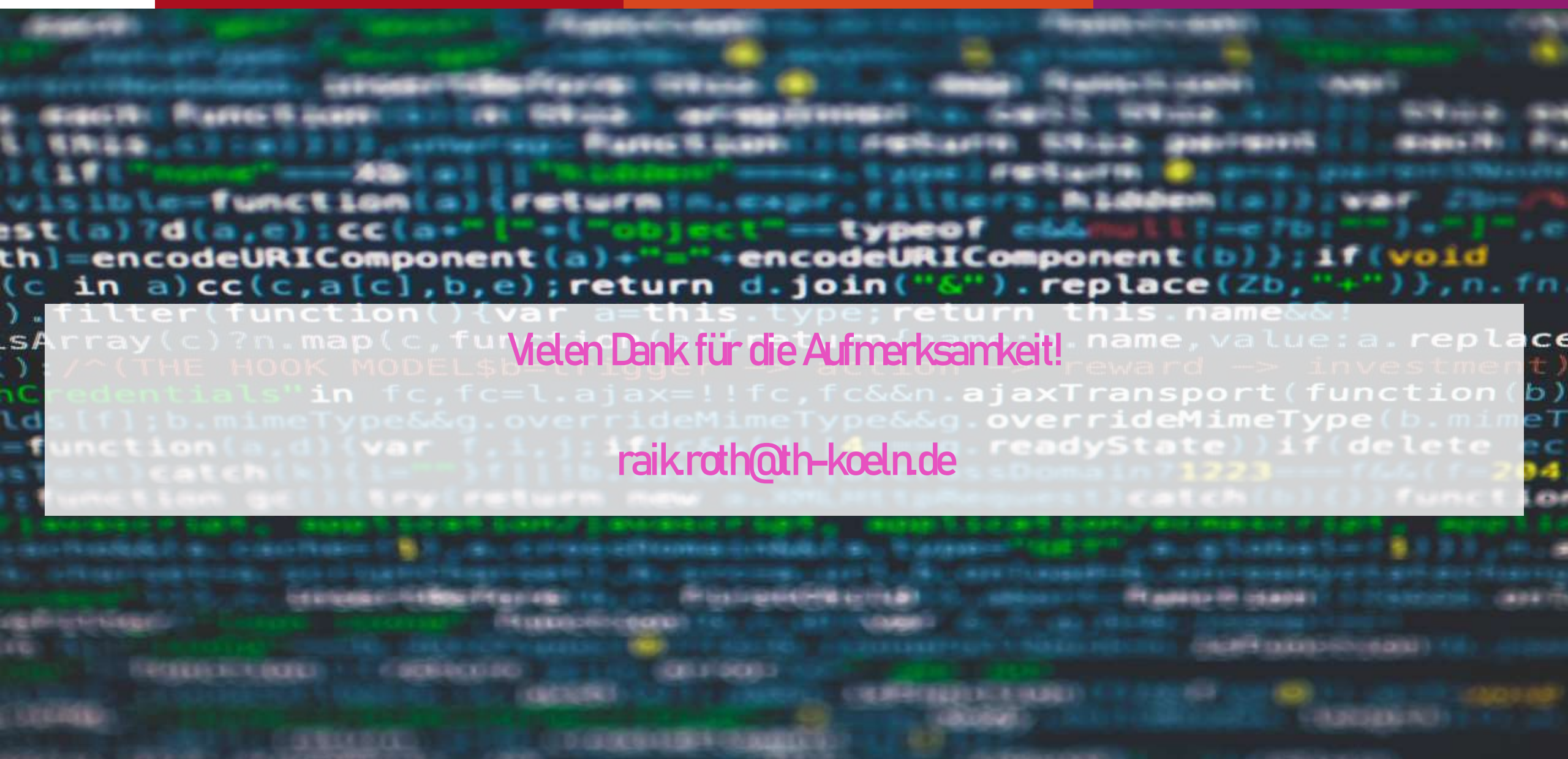
Quelle: Eigene Darstellung

Zur (Re-)Produktion von Ungleichheiten mittels AES: weitere Problemfelder

- Marktmacht, Big Tech, Surveillance- und Datenkapitalismus
 - Wer hat die Ressourcen zur Entwicklung aufwändiger Modelle? Wer hat Zugang zu welchen Daten?
 - Zu welchen Zwecken werden AES entwickelt und eingesetzt?
 - Wer bestimmt darüber? Wer profitiert (nicht)?
 - Datenextrahierendes Werbe-Überwachungssystem vs. Consentful Tech
- ‚Diversity Crisis‘ in Tech – personell und inhaltlich
 - Wer hat Zugang zur Berufen in Tech?
 - Wer ist an der Programmierung von AES beteiligt?
 - Welches Wissen ist zur Reproduktion von Ungleichheiten mittels AES bekannt?
- Ausbeutung natürlicher Ressourcen und Arbeitskraft
 - Wer leistet die schlecht bezahlte notwendige menschliche Arbeit?
 - Wie viele natürlichen Ressourcen werden in der Anwendung von AES verbraucht und auf wessen Kosten?

Und nun?

- Es braucht mehr Wissen und Reflexion darüber...
 - ...wo konkret AES überall wie zur Anwendung kommen (sollen)
 - ...wie gesellschaftliche Ungleichheiten und Herrschaftsverhältnisse unsere Wirklichkeiten prägen, wie sie in Technologien und in AES eingelagert sind und mittels dieser verstärkt werden
- Es braucht sozialwissenschaftliche Expertise in Entwicklung und Anwendung von AES
- Es braucht weniger Berührungängste bezüglich der Auseinandersetzung mit AES
- Es braucht politische, wertorientierte Antworten auf die Frage, wie und wozu wir ‚KI‘ (nicht) einsetzen wollen
- Es braucht digital data literacy und Medienpädagogik



Vielen Dank für die Aufmerksamkeit!

raik.roth@th-koeln.de

<https://www.pexels.com/de-de/foto/bildschirm-verschlüsselung-programmierung-web-design-2061168/>

Literatur

- AW AlgorithmWatch/Spielkamp, Matthias (2022): Automatisierte Entscheidungssysteme und Diskriminierung: Ursachen verstehen, Fälle erkennen, Betroffene unterstützen. Ein Ratgeber für Antidiskriminierungsstellen. Berlin. <https://algorithmwatch.org/de/autocheck/> [Zugriff 17.11.2023]
- Bächle, Thomas Christian (2016): *Digitales Wissen, Daten und Überwachung zur Einführung*. Hamburg: Junius Verlag.
- Criado-Perez, Caroline (2021): Unsichtbare Frauen: Wie eine von Daten beherrschte Welt die Hälfte der Bevölkerung ignoriert. Bundeszentrale für Politische Bildung.
- Broussard, Meredith (2019): *Artificial Unintelligence. How Computers Misunderstand the World*. MIT Press.
- Crawford, Kate (2021): *Atlas of AI. Power, Politics and Planetary Costs of Artificial Intelligence*. New Haven/London: Yale University Press.
- Crawford, Kate (2013): The Hidden Bias in Big Data. Harvard Business Review, 01.04.2013. <https://hbr.org/2013/04/the-hidden-biases-in-big-data>. [Zugriff 17.11.2023]
- Fry, Hannah (2019): *Hello world: was Algorithmen können und wie sie unser Leben verändern*. München: C.H. Beck.
- Günther, Elisabeth (2022): Topic Modeling. Algorithmische Themenkonzepte in Gegenstand und Methodik der Kommunikationswissenschaft. Herbert von Halem Verlag.
- Haraway, Donna (2017 [1995]): Situiertes Wissen. Die Wissenschaftsfrage im Feminismus und das Privileg einer partialen Perspektive. In: Bauer, Susanne/Heinemann, Torsten/Lemke, Thomas (Hrsg.) *Science and Technology Studies. Klassische Positionen und aktuelle Perspektiven*. Suhrkamp, 369–403.
- Houben, Daniel & Prietl, Bianca (Hrsg.). (2018): *Datengesellschaft. Einsichten in die Datafizierung. des Sozialen*. Bielefeld: transcript.
- Myers West, Sarah/Whittaker, Meredith/Crawford, Kate (2019): *Discriminating Systems. Gender, Race, and Power in AI*. AI Now Institut. <https://ainowinstitute.org/discriminatingystems.html> [Zugriff 17.11.2023]
- Whittaker, Meredith u.a. (2019): *Disability, Bias, and AI. Report*. AI Now Institute. <https://ainowinstitute.org/publication/disabilitybiasai-2019> [Zugriff 17.11.2023]
- O’Neil, Cathy (2018 [2016]). *Angriffe der Algorithmen*. Bonn: Bundeszentrale für politische Bildung.
- Reigeluth, Tyler. 2015. "Warum ›Daten‹ nicht genügen. Digitale Spuren als Kontrolle des Selbst und als Selbstkontrolle." *Zeitschrift für Medienwissenschaft*, Heft 13: Überwachung und Kontrolle, Jg. 7 (2): 21–34. Aufgerufen am 09.08.2021,
- Prietl, Bianca (2023) "Das Geschlecht der Datafizierung: MachtWissen im digitalen Zeitalter", *Journal Frauen- und Geschlechterforschung NRW*, 51, 67–72, https://www.netzwerk-fgf.nrw.de/fileadmin/media/media-fgf/download/publikationen/netzwerk_fgf_journal_51_f_web_230427.pdf. [Zugriff 17.11.2023]
- Prietl, B. (2019): "Die Versprechen von Big Data im Spiegel feministischer Rationalitätskritik", *GENDER Heft 3/2019*, 11–25, <https://doi.org/10.3224/gender.v11i3.02>.
- Treviranius, Jutta (2019): The Value of Being Different. W4A '19: Proceedings of the 16th International Web for All Conference, May 2019, 1–7, <https://dl.acm.org/doi/10.1145/3315002.3332429> [Zugriff 17.11.2023]
- Vipra, Jai/Myers West, Sarah (2023): *Computational Power and AI*. AI Now Institute. https://ainowinstitute.org/wp-content/uploads/2023/09/AI-Now_Computational-Power-an-AI.pdf [Zugriff 17.11.2023]
- Zweig, Katharina (2019): *Ein Algorithmus hat kein Taktgefühl*. Heyne.

Links

- AI Intersectional Toolkit: https://intersectionalai.miraheze.org/wiki/Intersectional_AI_Toolkit
- AI Now Institute: <https://ainowinstitute.org>
- AlgorithmWatch: <https://algorithmwatch.org/de/>
- Big Data Literacy Network: <https://www.bigdataliteracy.net> (inkl. sehr gutem und ausführlichem Glossar)
- Data for Black Lives: <https://d4bl.org/>
- LA Times (2021): <https://www.latimes.com/business/technology/story/2021-12-10/uber-transgender-drivers-blocked-accounts-rejected-ids>
- ProPublica (2016): Machine Bias. There's software used across the country to predict future criminals. And it's biased against blacks. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Queer in AI: <https://www.queerintai.com/>
- The Radical AI Podcast: <https://www.radicalai.org/> (inkl. Transkripten zu Interviews mit AI Scholars)