



Funded by
the European Union

MONIKA HANYCH

algoritmická diskriminace



ERA, 19 MARCH 2026

algoritmická diskriminace

Jednoho dne se psycholožka Miriam rozhodla na policii nahlásit svého manžela, který jí **opakovaně vyhrožoval násilím**.

Pustila policistům nahrávky, kde jí doslova vyhrožoval, že „jí vezme to, co je jí na světě nejdražší“, a popsala, kdy se muž projevoval násilnicky vůči jejich dětem.

Policisté se Miriam dotázali řadu otázek a **odpovědi vložili do programu VioGén**, softwaru, který odhaduje míru rizika recidivy genderově podmíněného násilí. Program vyhodnotil nebezpečí dalšího incidentu jako nízké. Když Miriam žádala soudce, aby manželu zakázal stýkat se s dětmi, soudce mj. s poukazem na **vypočtené nízké riziko programem** návrhu nevyhověl.

O sedm měsíců později její manžel děti zavraždil a spáchal sebevraždu.



Co je algoritmická diskriminace?

přímé či nepřímé diskriminační jednání,
založené na automatizovaném
rozhodování skrze algoritmy

algoritmy = obecné označení souboru výpočetních
instrukcí, které ze vstupů určují výstupní hodnoty

Dělení algoritmického rozhodování



rule-based algorithms, pravidlové

- příklad: systém automatických pokut při vysoké rychlosti auta

machine learning algorithms, autonomní

- příklad: software odhalující spam v příchozích e-mailech

znaky algoritmické diskriminace

1

rychlost a škála
diskriminačních
účinků

2

skrytá reprodukce
lidského faktoru

3

zkreslenost
vstupních dat

4

falešné
kauzality

5

netransparentnost
používaných metod

6

intersekcionalita

znaky algoritmické diskriminace

1

rychlost a škála
diskriminačních
účinků

Příklady:

- profilování uchazečů k zaměstnání
- rozdělování sociálních dávek

Problémy:

- bez přihlédnutí ke konkrétním okolnostem a individualitě člověka
- chyba se aplikuje na celý vzorek
- zvýšení škodlivého potenciálu na více dat

znaky algoritmické diskriminace

2

skrytá reprodukce
lidského faktoru

Problémy:

- opakují podjatost v rozhodování (a zveličují jí)
- nejsou neutrální či apolitické
- ovlivněné vstupními daty i konkrétním designem vývojářů

Příklad:

- odmítnutí půjčky kvůli demografickým charakteristikám namísto reálné finanční situace

znaky algoritmické diskriminace

3

zkreslenost
vstupních dat

- kvalita vstupů ovlivňuje kvalitu výstupů
- nevyvážená, neúplná, zkreslená vstupní data
- užití datasetů k jinému účelu, než byly zamýšleny (*transfer learning*)
 - falešné představa modelu – falešné korelace a výsledky

znaky algoritmické diskriminace

- problém *proxy variables*
(náhradní – často opět diskriminační – ukazatele)
- strojové učení založeno na rozeznávání korelací, nikoliv na porozumění kauzalitám (nedokáží sdělit, proč rozhodly; na základě čeho publikovaly daný výsledek)

4

falešné
kauzality

znaky algoritmické diskriminace

- systémy se vyvíjejí v čase – diskriminace skupiny či člověka např. jen jednorázově či nahodile; těžko prokazatelná
- předstírání shody (*alignment faking*) a další problémy
- problém výpočetní neredukovatelnosti (*computational irreducibility*) u složitých operací – nelze vysvětlit, jak stroj dospěl k danému výsledku

5

netransparentnost
používaných metod

znaky algoritmické diskriminace

- nejen kumulace a kvantita, ale propojenost diskriminačních důvodů
- výrazně silněji dopadne na jednu kategorii lidí oproti jiným
- Příklad: nošení náboženských symbolů ve školách; pozůstalostní důchod pro registrované partnery až v určitém věku

6

intersekcionalita

problémy obětí algoritmické diskriminace

1. Rozeznat ji

- oběť se často ani nedozví, že rozhodnutí vzniklo na základě algoritmického rozhodování
- stroj nedokáže vysvětlit, na základě čeho k danému závěru dospěl (a často ani experti)

2. Prokázat ji

- není možné provést účinnou zpětnou analýzu strojové úvahy
- strojové uvažování se může měnit a vyvíjet v čase

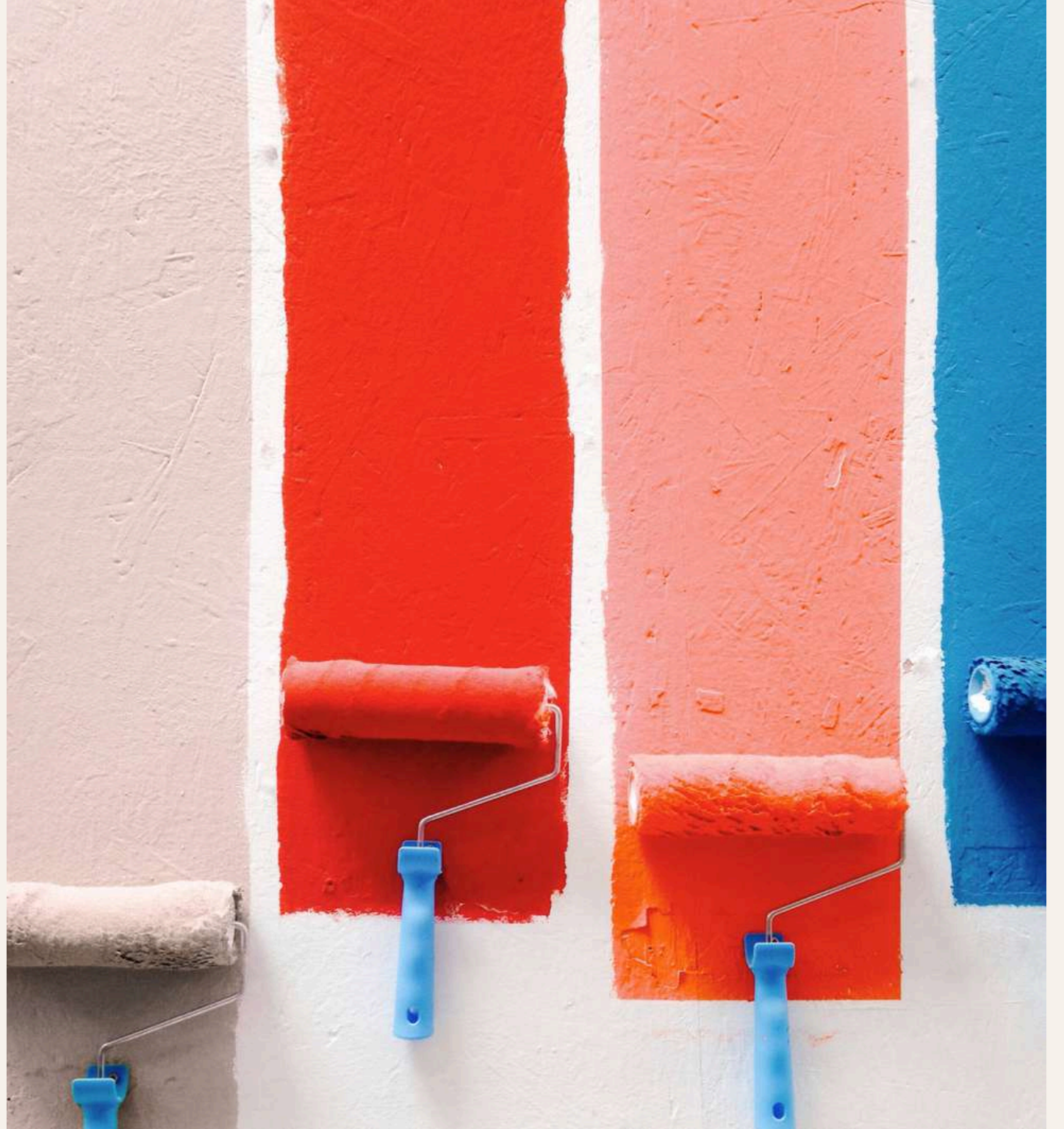
3. Identifikovat odpovědnost

- vývoj a trénink – supervize – implementace – užití systémů – konkrétní nasazení a výsledky
- mnoho zapojených osob v řetězci



regulatorní rámeček

stručný
přehled



Algoritmická diskriminace a regulatorní rámec

1

Evropská úmluva
o ochraně lidských práv
a Protokol č. 12

2

Mezinárodní pakt
o občanských
a politických právech

3

Listina základních
práv EU

4

EU nařízení o AI
(Akt o umělé inteligenci)

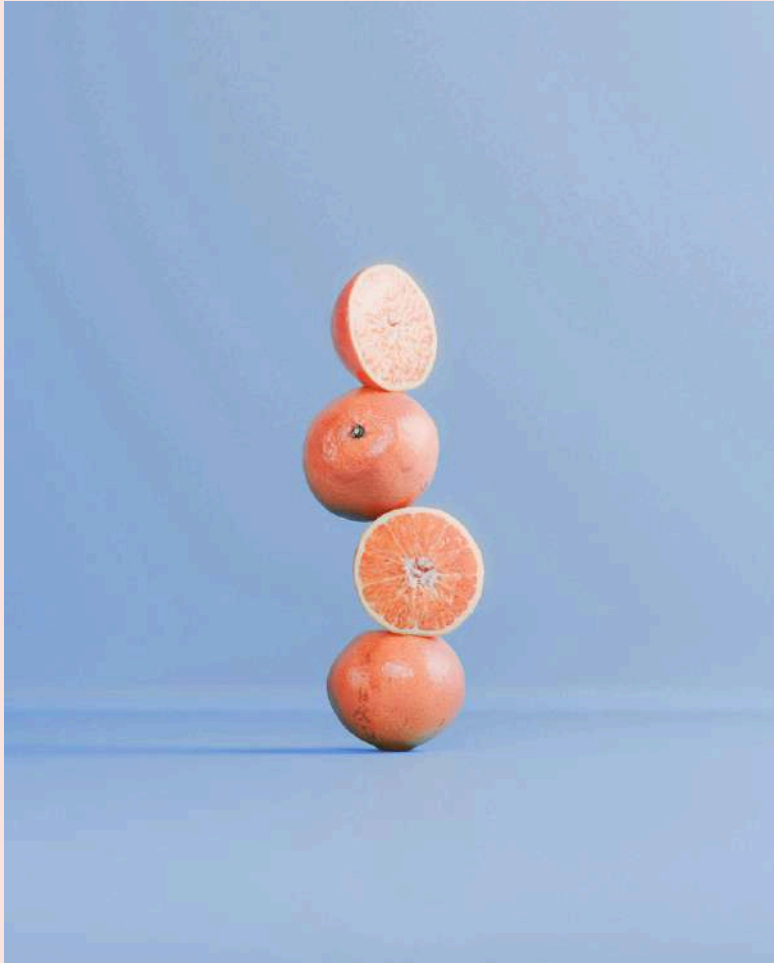
5

Rámcová úmluva
Rady Evropy o AI
a lidských právech

6

EU směrnice o
zástupných žalobách na
ochranu kolektivních
zájmů spotřebitelů

shrnutí a cesty



Shrnutí

- nový druh diskriminace, ale užití stávající regulace
- netransparentnost vývoje; zástupné ukazatele
- současné koncepty přímé a nepřímé diskriminace nepostačují pro řešení intersekcionality
- oběti: důkazní nouze, obchodní tajemství, black box, dlouhé řízení, slabé postavení

Cesty ven?

- kolektivní postupy obětí diskriminace
- zapojení orgánů – KVOP, ČOI, DIA, nevládky, odbory
- vhodná regulace, tlak na korporace

zdroje

- CATANZARO, M. In Spain, the VioGén Algorithm Attempts to Forecast Gender Violence [online]. *Algorithm Watch*. 27. 4. 2020. Dostupné z: <https://algorithmwatch.org/en/viogen-algorithm-gender-violence/>
- PIVODA, M. a HANYCH, M. Algoritmická diskriminace: neviditelné nerovnosti ve světě zkreslených dat. In *Mužské právo II*. Praha: Wolters Kluwer, 2025, s. 781-807. ISBN 978-80-286-0038-9.
- SURDEN, H. Artificial Intelligence and Law: An Overview. *Georgia State University Law Review*. 2019, roč. 35, č. 4, s. 1310–1321.
- CRIADO, N., SUCH, J. *Digital Discrimination*. In: YEUNG, K. a LODGE, M. (ed). *Algorithmic Regulation*. Oxford: Oxford University Press, 2019,
- GERARDS, J., XENIDIS, R. *Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law*. Luxembourg: Publications Office of the European Union, 2021.
- HADWICK, D., LAN, S. Lessons to be learned from the Dutch childcare allowance scandal: A comparative review of algorithmic governance by tax administrations in the Netherlands, France and Germany. *World Tax Journal*, 2021, roč. 13, č. 4.
- BUOLAMWINI, J., GEBR, T. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research*, 2018, roč. 81, s. 1–15.

děkuji za pozornost

MONIKA@HANYCH.COM

