Artificial Intelligence and Gender Equality

Costanza Nardocci

5 May, 2026

University of Milan, School of Law, Department of Italian and Supranational Public Law

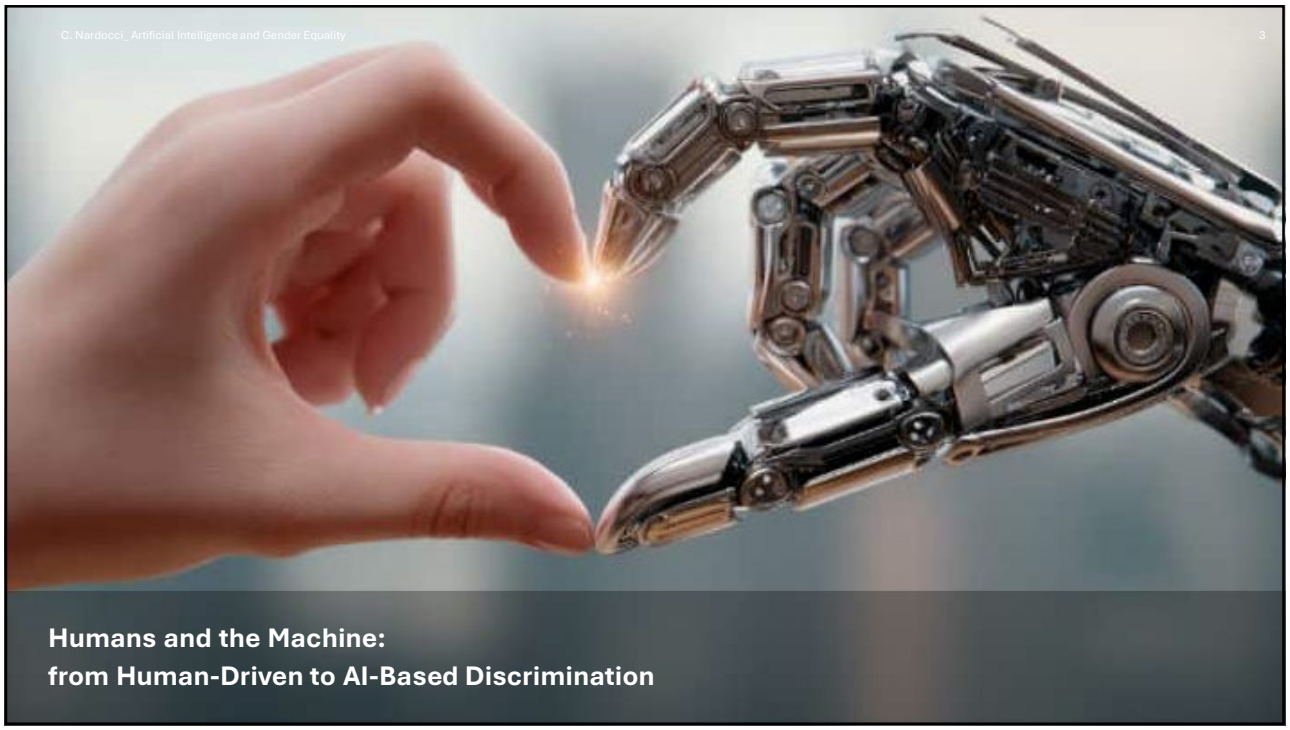
Funded by the European Union

1

Overview

- Humans and the Machine: from Human-Driven to AI-Based Discrimination
- When AI Meets Gender (In)Equality: Challenges and Cases to Watch

2



Humans and the Machine: from Human-Driven to AI-Based Discrimination

3

What Differentiate Humans from the Machines? & Why Does it Matter for Discriminatory AI?

- Metacognition
- Rationality
- Ethics and Morals

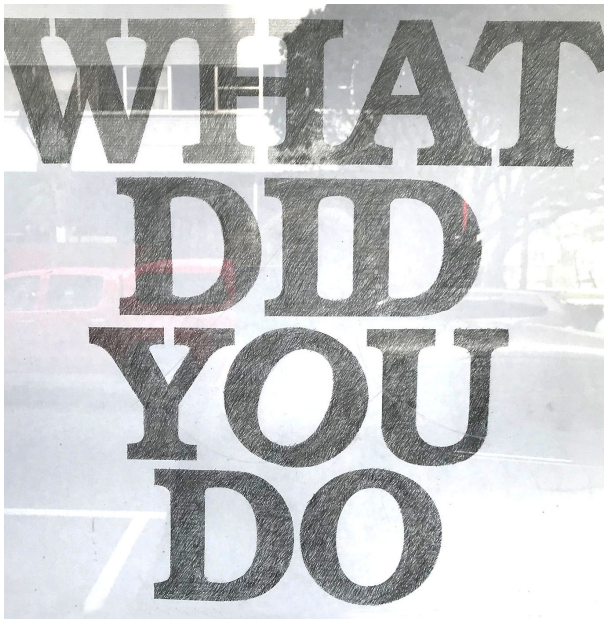
4

Human-Driven *versus* AI-Based Discrimination

- Agent(s)
- Conduct(s)
- Victim(s)
- Protected grounds (Proxies)
- Consequences on points of law (effects and harms)

The Agent(s)

- The humans: the programmer, the manufacturer, the deployer, the end-user(s)
- The machine: AI technologies (especially, machine learning and deep learning systems)



The Conduct(s)

- Mixed agency
- Intertwined
- Multiple
- Directly or indirectly linked with the discriminatory effect(s)
- Opaque

The Victim(s)

- Individuals (as in human-driven discrimination)
- Social groups
  - ❖ already marginalized communities (racial or religious minorities)
  - ❖ "new" minorities (AI-targeted minorities or algorithmic minorities) and «new» group rights

The Protected Grounds (Proxies)

- Proxies, not much traditional factors of discrimination
- Correlations: proxies are predicative of the individual affiliation to a protected group
- Intersectionality: among proxies, and among proxies and protected grounds

Consequences on Points of Law

- Direct discrimination? NO
- Indirect discrimination? NOT MUCH
- Discrimination by Association? MAYBE
- AI-based discrimination? YES

Direct Discrimination (& Disparate Treatment)
Why not …

- Lack of an explicit references to one or more protected grounds
- Challenges in tracing back the casual link between the conduct and the effect, linked to the black box theory and the opacity of AI systems
- Only under the disparate treatment theory: lack of intentionality on the side of the agent(s)

Indirect Discrimination (& Disparate Impact)
Why not much…

- The disparate impact in discriminatory AI is not based on protected characteristics, but, rather, on proxies
- The degree of the impact, that should be "disparate" or disproportionate, is hard to prove
- Difficulties in finding the comparator
- Challenges in proving the harm suffered by the victim(s)

Discrimination by Association Maybe...

- Based on correlations between proxies and protected grounds
- The victim is not necessarily a member of the protected group, but he/she is associated to it, or he/she perceived himself/herself as such

AI-Based Discrimination... YES

- Based on proxies more than factors of discrimination
- Multiple rather than single agency (and correlated responsibilities)
- Causality often hidden due to the opacity of AI systems

Challenges Ahead and Questions Left Unanswered

Regulatory gaps:

- Lacking legislative efforts to grasp the specifics of AI-based discrimination (e.g. AI Act, CoE Framework Convention on AI)

Liability gaps:

- Who might be responsible for discriminatory AI?
- The manufacturer? The programmer? The deployer? The end users? All from their own "part"?
- Which type of responsability?

Gaps in jurisprudence:

- Poor case-law on AI technologies and discrimination

Examples of The (Few) Cases... Some Speak of «Discrimination» and Some of «Other» (New?) Rights

Domestic Courts in Comparative Perspective:

- Italy → Tribunal of Bologna, «Deliveroo»
- United States → «Compas» (AI to assess likelihood of recidivism), «Workday» (AI and discrimination in the workplace)
- Canada → *Ewert v. Canada* (AI as *risk assessment tools*)
- Finland → National Non-Discrimination and Equality Tribunal, Discriminatory AI in credit granting procedures
- United Kingdom → *R. (Bridges) v. Chief Constable of South Wales Police* (2020)

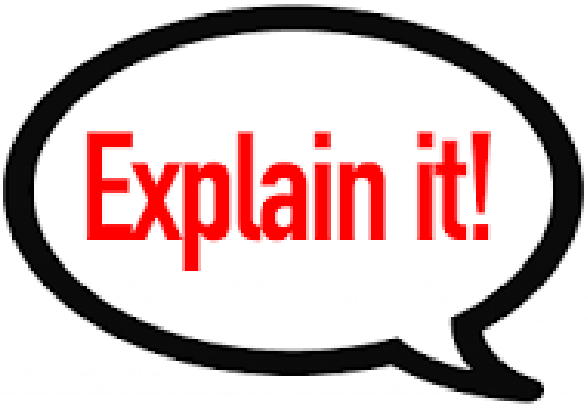
N.B. *R. (Thompson and Carlo) v. The Commissioner of Police of the Metropolis* (Pending)

- Australia → *Tickle v. Giggle* (gender identity and AI)

Supranational Courts:

- European Court of Human Rights → *Glukhin v. Russia* (Facial recognition system)

Coordinates to Interpret the Judgements (and their Outcomes)



- Missing references to the specifics of AI-based discrimination
- Very few mentions of the proxy
- Tendency to rely on anti-discrimination laws, stretched to to adapt them to "artificial" discrimination(See, Italy, «Deliveroo» case)
- Focus on other rights rather than on the discriminatory features of the case (See *R. (Bridges) v. Chief Constable of South Wales Police*)

Tackling AI-based Discrimination?
One Proposal

- Discrimination by Association (focus on correlations rather than on causation)
- Open(-ended?) list of factors of discrimination, including proxies
- Hypothetical *Tertium Comparationis*
- Effects instead of conducts
- Joint liability
- Collective fairness *versus* Individual fairness

What the Law Does Not Say... «Procedural» Remedies at the times of AI-based Discrimination

«AI-Asociated» Rights:

- Right to know
- Right to an explaination
- Human oversight

& the Others:

- Inversion of the burden of the proof
- Recognition and strengthening of the *locus standi* of Equality Bodies and NGOs (See Article 77 EU AI Act)

When AI Meets Gender (In)Equality: Challenges and Cases to Watch

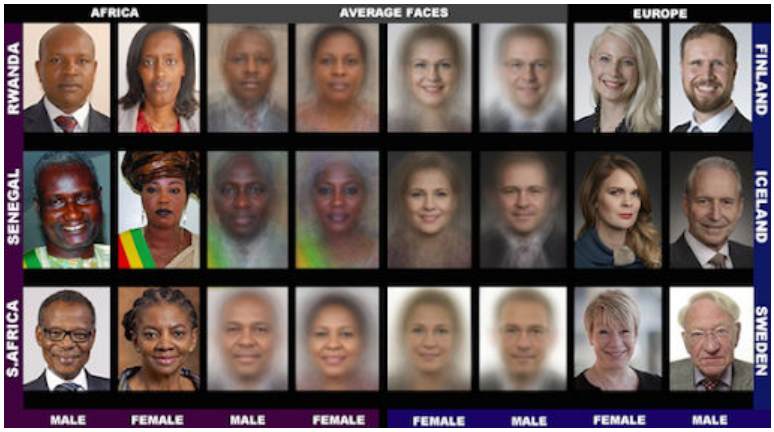
A Conflicting Relationship and Its Origins

- Few women enrolled in STEM disciplines
→ BUT increase in gender equality education is not itself increasing workforce participation

- Few women working in the Tech Industries
→ PLUS women's representation declines when approaching managerial roles

- Bias in the data-set
→ MEANING underepresentation of women and lack of intersectionality (See Gender Shades Project)

- Non-Inclusive Language



… and Some Explanations

- Lack of diversity in the dataset
- Male over-representation
- Gender stereotypes embedded in data, feature selections and correlations
- Predominance of male-centered proxies

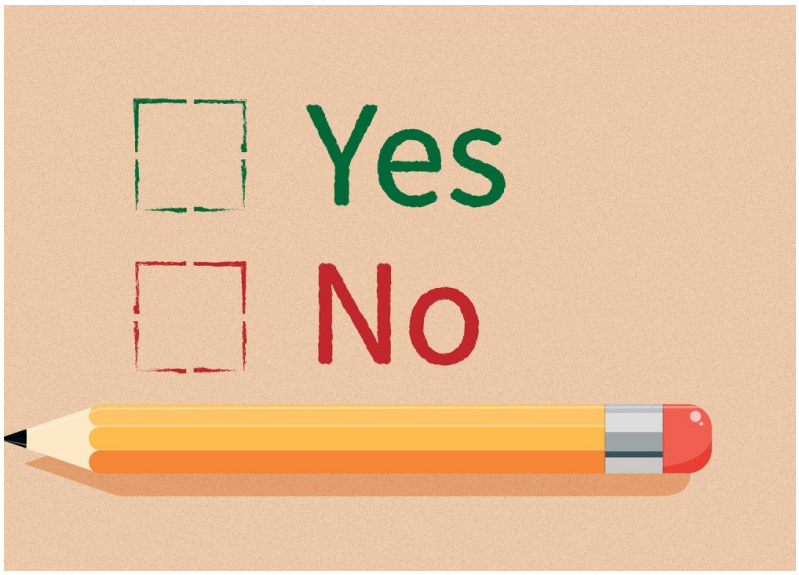
Are European AI Regulations Gender-Sensitive?

- The EU AI ACT features poor references to the gender dimension featuring AI systems → Article 68 «Scientific panel of independent experts» and Article 95 «Codes of conduct for voluntary application of specific requirements»

N.B. «Sex» and «Gender», including «Gender Equality» (Recitals Nos. 27 and 48) are mentioned mostly in the Preamble / No article prohibiting inequalities and discrimination

- The COE FRAMEWORK CONVENTION ON HUMAN RIGHTS DEMOCRACY AND THE RULE OF LAW → mentions «gender» twice, especially, under Article 10 «Equality and non-discrimination»

N.B. «Sex» is never mentioned!



Maybe Not (the EU AI Act, at least)

- ARTICLE 5 (Prohibited AI practices) → prohibits the use of biometric AI systems, but it does not explicitly mention «gender» among the protected features

- ARTICLE 6 (Classification rules for high-risk AI systems) and Article 43 (conformity assessment) → lack of an intersectional approach to evaluate the compliance of AI systems with EU law

- ARTICLE 27 (FRIA) and Article 9 → Lack of a Gender Impact Assessment alongside the gender-neutral risk assessment provided under EU law

Beyond the Law, but Talking about Laws: Areas of (Potential) Concern

- WORKPLACE
- LANGUAGE (and translations)
- VIOLENCE AGAINST WOMEN AND GIRLS

Weren't You Hired? Maybe, It's AI Related

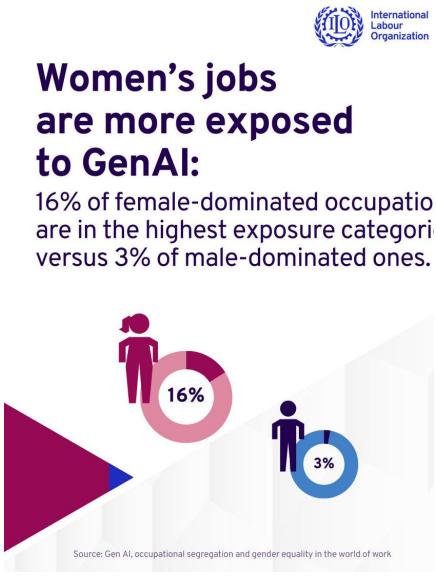
- Amazon's AI Recruiting Tool (Scrapped) → the AI tool was trained on resumes submitted by men, with the consequence that resumes containing the word "women" were penalized.
- Google Ads → gender discrimination in advertising publication: jobs with a higher salary are shown in a lower percentage to a female audience
- Video Interview AI Analysis Tools → in 2025 a complaint was brought against AI tools like HireVue, Retorio, and myInterview in that the video hiring software were allegedly discriminatory even based on gender
- Large language models (LLMs) → suspected of gender (and, also, racial bias), according to the findings of a 2024 research study of the University of Washington DC (US)

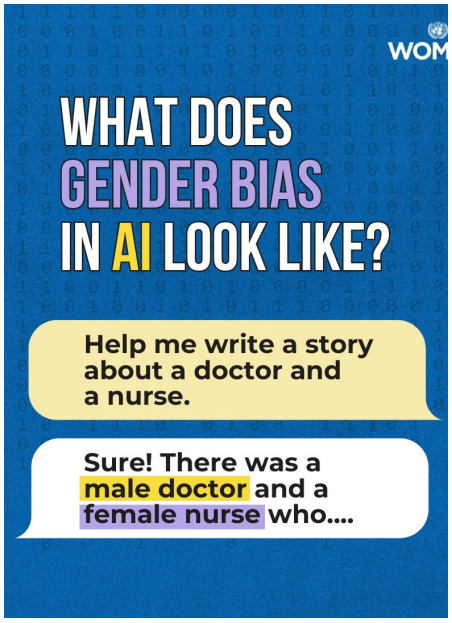
More about Gender Equality and AI-Issues in the Workplace

- Occupational segregation
- Women underrepresentation in AI-related occupations → «AI systems trained on incomplete or biased data can reproduce and amplify gender inequalities, with consequences for recruitment, participation, promotion, pay, workplace monitoring and access to skills and resources»
- Job losses → «Women are disproportionately concentrated in less senior, routine and codifiable roles, and less likely to perform abstract or high-autonomy tasks where AI acts as a complement rather than a substitute»

N.B. «Minority women», belonging to ethnic, racial or religious minority communities, are predominantly exposed

Source: ILO Report, 5 March 2026, https://www.ilo.org/publications/gen-ai-occupational-segregation-and-gender-equality-world-work





Were You Referred to as "Man" in a Text? Maybe It's AI (Again) !

- Some types AI technologies use words and language to perform their tasks (LLM AI systems)
- Non-gender-inclusive language negatively impacts on the fair functioning of LLM systems
- Despite global efforts in reviewing natural languages to embrace gender-sensitive approaches, languages continue to resort to the masculine (as neutral masculine)
- Gender bias in natural languages impact on AI causing risks of discriminatory LLM AI technologies

One Last Note: AI and the New Frontiers of Violence Against Women and Girls

- AI as an additional risk for women exposed to violence

→ AI is boosting the development of new or alternative forms of violence against women

- AI as a (potential) tool to tackle violence against women

When AI Exacerbates Women's Vulnerability and Creates (also) «New» Forms of Violence against Women

TECHNOLOGY-FACILITATED VIOLENCE AGAINST WOMEN AND GIRLS (TF VAWG) → image-based abuse; doxing and violent threats

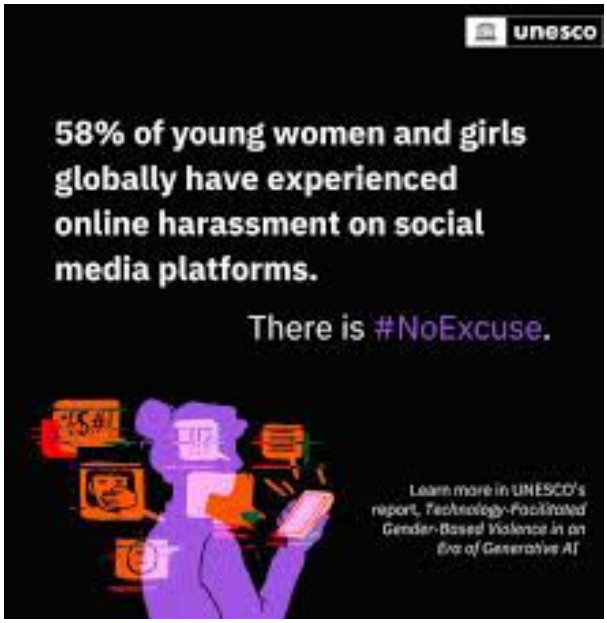
PORNOGRAPHY GENERATED BY AI → leading to isolation, school bullying and harassment (Report of the Special Rapporteur on violence against women and girls, its causes and consequences, Reem Alsalem, August 2024)

STRUCTURAL AND PERSISTEND GENDER-BASED VIOLENCE ASSOCIATED TO AI → bias reinforcement and misogyny amplification

Recent European Initiatives on Forms of Violence against Women AI-related

- The EU →
  - ❖ FRA Report 2026 and the Gender Equality Strategy 2026-2030
  - ❖ EU Parliament from the LIBE Committee are now pushing for a ban on so-called "nudifier" systems
- The CoE → 2 Recommendations (March 2026)
  - ❖ Equality and artificial intelligence
  - ❖ Accountability for technology-facilitated violence against women and girls

N.B. First international legal standard focused on this topic



Cases to Watch: The EU and States' Fight against AI and Gender-Based Violence

- (SPAIN) VioGén → AI system to assess and tackle gender-based violence risks (studies demonstrated its poor effectiveness)
- (EU) SHILED → focused on stalking, sexual harassment, dating violence and online grooming, the project introduces tech tools to empower young people, educators and youth workers
- (UNITED KINGDOM) DASH (Domestic Abuse, Stalking and Honour Based Violence) → risk assessment checklist to support police's evaluation of the situation
- (CANADA) SARA (Spouse Abuse Risk Assessment) → similar to VioGén

At the (Good) Intersection of AI and Gender-Based Violence

- LEGAL AND POLICY FRAMEWORKS → «ensure that technology-facilitated violence against women and girls is covered and addressed by criminal law»
- ACCOUNTABILITY → Looking at the roots….«addressing its root causes, including sexism, gender stereotypes, historical and structural inequalities, systemic discrimination, entrenched patriarchal or misogynistic norms and practices, and the persistence of a culture that disregards the need for consent in sexual interactions»
- INTERSECTIONALITY → «human rights-based, intersectional, inclusive, gender-transformative, victim-centred, trauma-informed, age-appropriate and child-friendly approach»
- ACCESS TO JUSTICE → and «prevent secondary victimisation»

THANK YOU!

Artificial Intelligence and Gender Equality

Costanza Nardocci

costanza.nardocci@unimi.it