

# Il genere nell'intelligenza artificiale

**Silvana Badaloni**

*Dipartimento di Ingegneria dell'Informazione,*

*Centro di Ateneo Elena Cornaro*

*Università di Padova*

*[silvana.badaloni@unipd.it](mailto:silvana.badaloni@unipd.it)*



*S.Badaloni, Generi, Saperi e Giustizia Sociale, UNIPD, 10 maggio 2024*

---

## *Fix the numbers of women*

<http://www.unipd.it/bilancio-genere>

*The world need Science and Science needs Women.  
And good solutions need reliable statistics and indicators*



# Studentesse immatricolate Ingegneria

Ing. IT

- 26,3 % Proporzione di donne immatricolate nei corsi ingegneristici 21/22 (25,2 % 19/20);

Ing. UNIPD

- Valor medio 22,6 % aa 23/24
- Ingegneria per l'Ambiente e il Territorio 39,4 %
- Ingegneria Civile 25,9 %
- Ingegneria Meccanica 12,3 %

DEI

- aa 23/24 Valor medio 26.08 %
- Ingegneria biomedica 62,6 %
- Elettronica 6 % - Ingegneria Informatica 15,4 %

# Sguardo al settore ICT

## *Information and Communication Technologies*

- In 35 paesi europei meno di 1 su 5 laureati in computer science sono donne<sup>1</sup>
- Divario di genere in un settore in grande sviluppo
- Rapporto di Almalaurea 2022<sup>2</sup> (Consorzio interuniversitario) ci dice che:
  - i/le laureati/e STEM si posizionano meglio rispetto al totale in tutte le discipline
  - sono le laureate in Ingegneria Industriale e dell'Informazione quelle con le retribuzioni mensili più elevate (1.723 euro) sebbene inferiori a quelle dei colleghi maschi (1.872 euro) - gender pay-gap
  - si tratta di una evidenza che va divulgata in ambito scolastico per aiutare le giovani a fare scelte consapevoli

1. OECD Gender Data Portal, Where are tomorrow's female scientists?

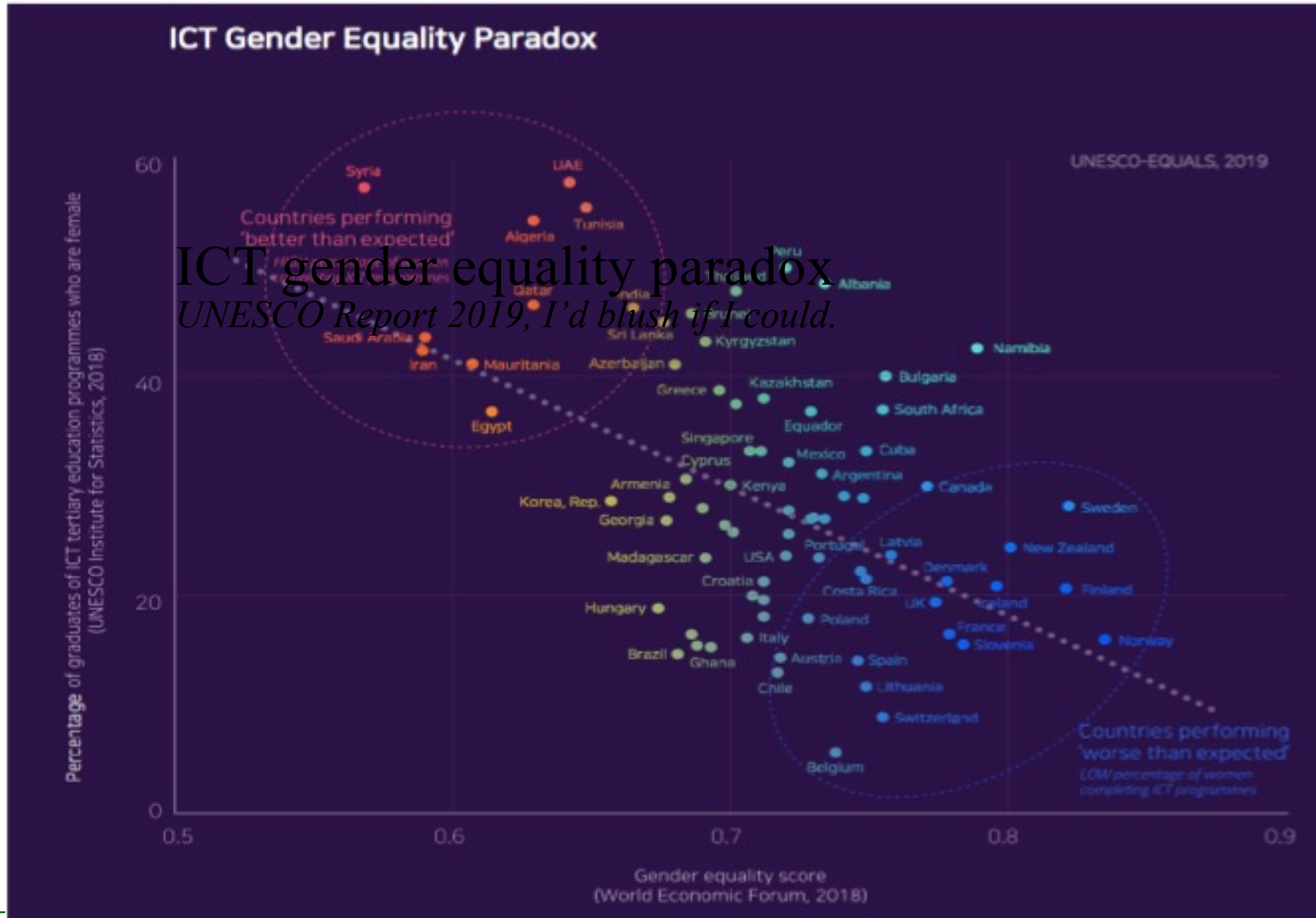
2. Rapporto Almalaurea 2022. Laureate e laureati: scelte, esperienze e realizzazioni professionali

[https://www.almalaurea.it/sites/almalaurea.it/files/convegni/gennaio2022/6\\_almalaurea\\_rapportocompleto\\_laureatelaureati.pdf](https://www.almalaurea.it/sites/almalaurea.it/files/convegni/gennaio2022/6_almalaurea_rapportocompleto_laureatelaureati.pdf)



# ICT gender equality paradox

*UNESCO Report 2019, I'd blush if I could.*

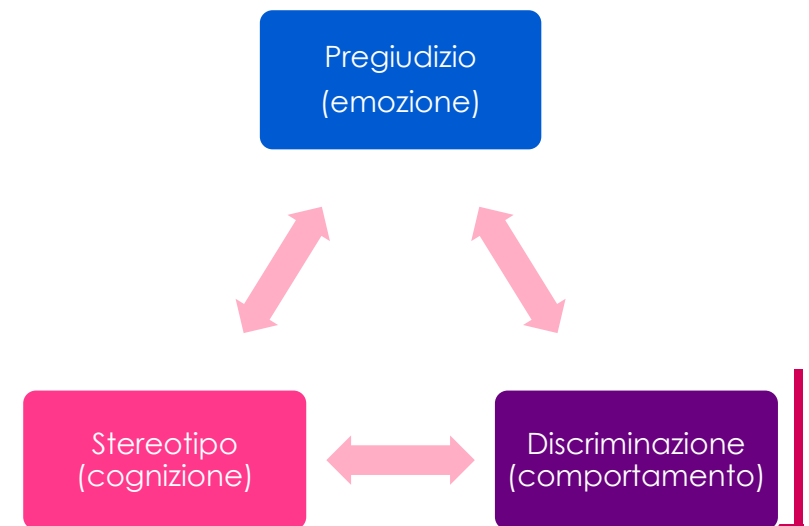


ICT gender equality paradox  
 UNESCO Report 2019, I'd blush if I could.

# Bias


- **Stereotipo o bias** è un sistema di conoscenza condivisa nella società, pro o contro qualcosa
- Gli stereotipi ci aiutano a confrontarci con un mondo diverso/sconosciuto, assumono allora lo status di giudizi che categorizzano fatti o persone: pregiudizi
- **Pregiudizio** è un'opinione preconcepita concepita non per conoscenza precisa e diretta del fatto o della persona, ma sulla base di voci e opinioni comuni
- **Circolo vizioso**

(source: F.A. Lisi)



# Ruolo degli stereotipi

- Ruolo degli stereotipi che condizionano la scelta delle giovani e dei giovani
- Stereotipo sulla Scienza: ce n'è uno che porta a considerare una **fatalità** il fatto che ci siano ancora poche donne nella Scienza
- Uno molto diffuso riguarda la presunta incapacità delle donne in ambito logico-matematico
- Non è vero che la presunta inferiorità femminile sia dovuta ad una diversità a livello cerebrale tra i due sessi: studi scientifici hanno mostrato che questa idea è falsa, ed è un pregiudizio<sup>1</sup>
- Il cervello non ha sesso (Antonella Viola)
- E' un problema culturale che va affrontato, a tutti i livelli in quanto il contesto familiare di provenienza, i modelli sociali proposti, l'attitudine delle scuole hanno una forte influenza sulle scelte formative delle giovani generazioni



*1 S.J.Ceci, W.M.Williams, S.M.Barnett. Women's underrepresentation in Science: sociocultural and biological considerations. Psychological Bulletin, 2009, vol 135, 2, 218-261*

# Quale ingegneria?

---

- Le ragazze, se scelgono ingegneria, sono selettive e scelgono corsi nuovi e interdisciplinari. I numeri delle giovani iscritte sembrano dipendere da quanto distante dalla società è percepito il sapere, la conoscenza nell'ambito dello specifico corso di ingegneria
- Le ragazze non si identificano nell'immagine stereotipata dell'ingegnere: rigido, pesante alleato delle macchine e della tecnologia
- Ingegneria/tecnologia in un contesto sociale/delle scienze della vita
- NERD





# Innovazioni di genere

## *Fix the knowledge*


---

- Genere e scienza non vuole dire solo aumentare il numero delle donne presenti per battere la sotto-rappresentazione delle donne attraverso adeguate politiche di genere
- *Fix the knowledge* vuol dire integrare la dimensione di genere nei contenuti scientifici dell'innovazione.
- Come possiamo ri-disegnare le teorie scientifiche tenendo conto della dimensione di genere?
- According to (Sanchez, 2013): *“There is a need to go beyond stereotypical feminization of products – so called “pinking” – as female preferences can be drivers for substantial innovation”*
- Non e' sufficiente applicare il cosiddetto “pinking” method – i.e. la *stereotipica femmininizzazione del prodotto*



# Come?

- Come possiamo dare una nuova interpretazione dei fatti rispetto a un universo proposto come neutro (ma neutro non è)?
- Come si può introdurre la dimensione di genere nei contenuti dell'innovazione scientifica in tutti i campi del sapere?
- Per produrre una scienza e una innovazione di genere, è necessario radicalmente cambiare la teoria e porre nuove domande scientifiche, riformulare le domande scientifiche, in qualunque campo del sapere
- Usando le regole dell'argomentazione logica abbiamo sviluppato una riflessione formale sul metodo scientifico e sulle modalità per produrre innovazione di genere



S. Badaloni, F. A. Lisi: Towards a Gendered Innovation in AI (short paper). In: G. Vizzari, M. Palmonari, A. Orlandini (Eds): Proceedings of the AIxIA 2020 Discussion Papers Workshop co-located with the the 19th International Conference of the Italian Association for Artificial Intelligence (AIxIA2020), Anywhere, November 27th, 2020

<http://ceur-ws.org/Vol-2776/paper-2.pdf>

# Gendered Innovation

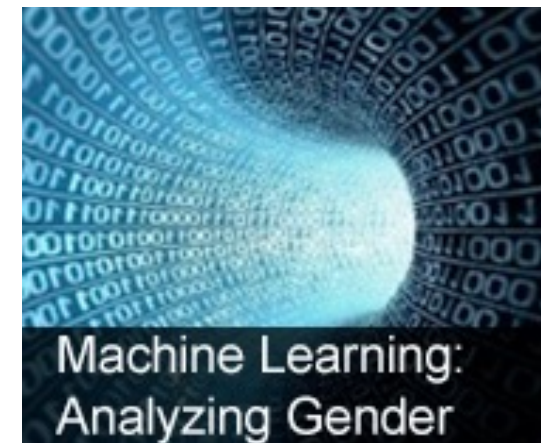
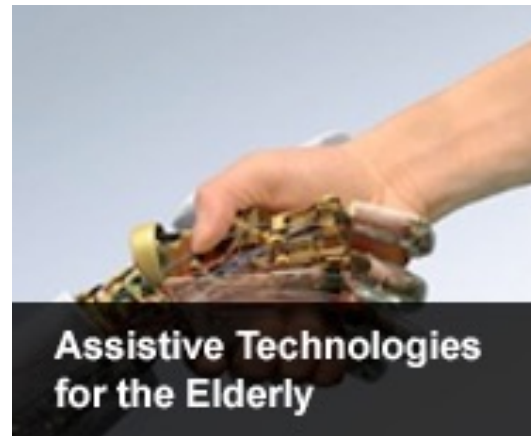
---



- Londa Schiebinger – Stanford University
- <http://genderedinnovations.stanford.edu/index.html>
- Perché questo approccio?
- Esempio iconico. Negli Stati Uniti, tra il 1997 e il 2000, sono stati ritirati 10 farmaci a causa degli effetti dannosi sulla salute (alcuni potenzialmente mortali) e 8 su 10 si sono rilevati più dannosi per la salute delle donne rispetto agli uomini
- L'approccio va rovesciato: l'innovazione di genere arricchisce la scienza e apre nuovi orizzonti. Solo così si può produrre una ricerca che sia gender-responsible, che permetta una crescita della qualità della vita in modo paritario per tutte e per tutti.

# Londa Schiebinger, Stanford University

## Case studies



<http://genderedinnovations.stanford.edu/index.html>

*S.Badaloni, Generi, Saperi e Giustizia Sociale, UNIPD, 10 maggio 2024*

# Heart diseases in Women

- La cardiopatia ischemica è il killer numero uno della popolazione statunitense ed europea, indipendentemente dal genere. Era stata considerata una malattia prevalentemente maschile, nelle donne spesso non era stata diagnosticata
- Gendered innovation:
  1. Ridefinire la patologia (diversa tra donne e uomini)
  2. Sviluppare nuove tecniche diagnostiche
  3. Capire le differenze dovute al sesso nei sintomi
  4. Fattori di rischio (fumo)
  5. ....

<https://genderedinnovations.stanford.edu/case-studies/heart.html>




# Algoritmocrazia

---

- E' iniziata l'era dell'algoritmo? può portare alla dittatura dell'algoritmo su tutta la conoscenza? in altri termini il calcolo algoritmico sviluppato per la soluzione dei problemi, può divenire un rischio per la conoscenza stessa per meccanismi impliciti di dominio e di ingiustizia?
- Algoritmi AI riflettono punto di vista del progettista
- Amplificano pregiudizi e stereotipi del gruppo dominante
- Etica del progettista precede etica delle macchine
- Ethics washing dei giganti dell'informatica
- Decliniamo la **Innovazione di genere nell'ambito dell'AI**: quali sono i problemi che il sempre maggiore diffondersi di tali applicazioni possono implicare a livello etico? Quale è il rapporto con le questioni di genere per poter parlare di gendered innovation di genere in questo ambito?

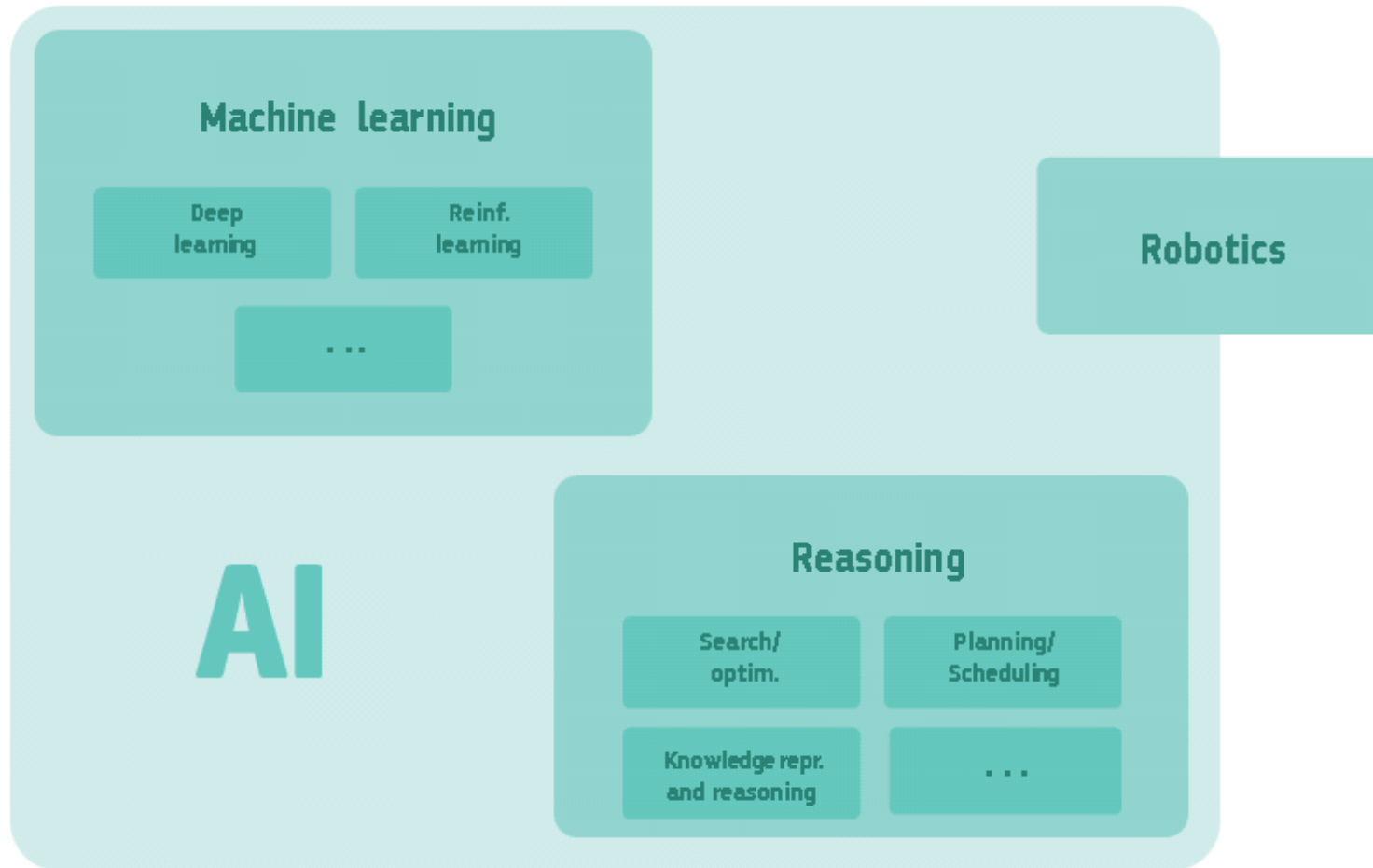


- L'AI è una disciplina scientifica che mira a definire e sviluppare programmi o macchine (software e/o hardware) che mostrano un comportamento che verrebbe definito intelligente se fosse esibito da un essere umano
- AI non è science fiction: è già nelle nostre vite ... nei sistemi di riconoscimento facciale, pagamento, traduzione, navigazione, etc.
- Come può una macchina essere considerata intelligente?
  - Cos'è l'intelligenza?
  - Approccio operativo del Test di Turing per definire l'intelligenza come imitazione basandosi sul concetto di imitazione: se una macchina fosse così abile da imitare l'uomo si potrebbe dire che pensa
- Come insegnare a una macchina a risolvere un problema
  1. **Ragionamento logico** – diciamo quali sono i passi che devono essere fatti per risolvere il problema
  2. **Machine learning** – forniamo esempi di soluzione di problemi e metodi di generalizzazione



Francesca Rossi. Il confine del futuro. Possiamo fidarci dell'Intelligenza Artificiale? Feltrinelli, 2019

# General architecture





# AI generativa

- AI Generativa: poggia il suo funzionamento sul **modello linguistico**. Nessuna semantica
- L'AI non è certo nata un anno fa, con il lancio di **ChatGPT**. Perché allora sembra che ci sia come un'era pre e post Gpt (novembre 2022)? Siamo davvero davanti a qualcosa che cambierà il mercato e tutti i soggetti?



# La nascita dell'AI

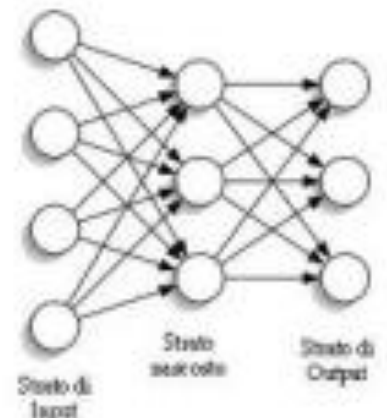
- Conferenza di Dartmouth 1956
  - John McCarthy ha coniato il termine **Artificial Intelligence**



Trenchard More, John McCarthy, Marvin Minsky, Oliver Selfridge, Ray Solomonoff

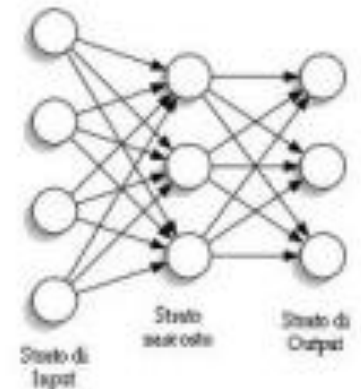
# Ragionamento vs ML

- L'avventura dell'AI è nata più di 60 anni fa
- **AI simbolica o logica** – esempio: trovare la strada migliore per andare da una città ad un'altra
- **AI Machine Learning**: queste tecniche permettono a un sistema AI di imparare come si può risolvere un problema che non può essere affrontato con un approccio simbolico.
- Il sistema – usualmente una rete neurale o sue modificazione evolute – viene addestrato con molti esempi di comportamento input-output in modo da poter generalizzare ed essere in grado di riconoscere e trattare input differenti da quelli del training set.
- Apprendono dagli esempi
- E' un sistema di classificazione di dati molto potente in contesti differenti.
- Da soluzione in moltissimi problema ed è all'origine di questo enorme successo dell'AI.



# Bias in AI

- Biased Machine Learning?
  - Sistemi di ML per la loro natura intrinseca, possono sussumere i bias diffusi nella società su genere ed etnia, rappresentati nei data-set usati per l'addestramento, e possono portare a decisioni "unfair"
  - "Fairness is the absence of any prejudice or favoritism toward an individual or a group based on their inherent or acquired characteristics"<sup>1</sup>
  - Consapevolezza come sviluppatori
- 1 N.Mehrabi, F.Morstatter, N.Saxena, K.Lerman, A.Galstyan. 2019. A survey on bias and fairness in Machine Learning. <https://arxiv.org/abs/1908.09635>



# Face recognition

**Joy Buolamwini - MIT**  
media lab researcher –  
*Facial Recognition Is  
Accurate, if You're a  
White Guy*



# FR System

Average Accuracy:

Lighter male 99 %

Lighter female 93 %




Darker male 88 %

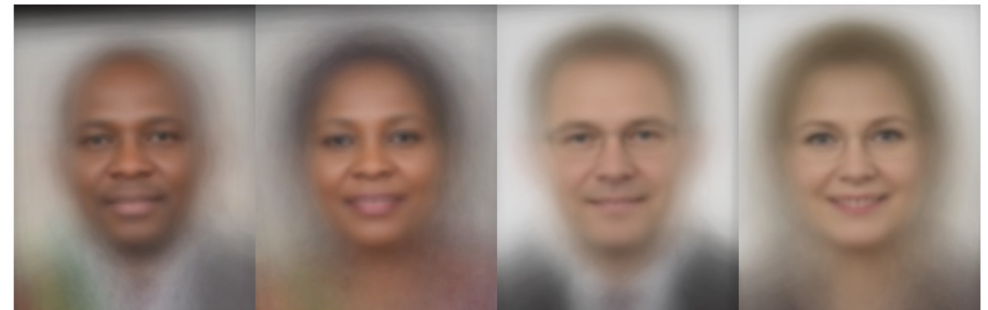
Darker female 65 %

- A.I. software is only as smart as the data used to train it. If there are many more white men than black women in the system, it will be worse at identifying the black women.

- *Coded Bias*

- **Timnit Gebru**, licenziata da Google perché denunciava i pregiudizi razziali dell'Intelligenza Artificiale

Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
 Microsoft	94.0%	79.2%	100%	98.3%	20.8%
 FACE++	99.3%	65.5%	99.2%	94.0%	33.8%
 IBM	88.0%	65.3%	99.7%	92.9%	34.4%



# Good news

---

- IBM ha deciso di non vendere più questo tool FR per scopi generali
- La chief executive Arvind Krishna ha detto che “vendors and users of AI systems have a shared responsibility to ensure that AI is tested for bias, particularly when used in law enforcement”.
- Amazon ha interrotto l’uso della tecnologia FR da parte della Polizia
- Nuove regolamentazioni in US



# Amazon Recruitment Tool

- Gli specialisti di machine-learning in Amazon hanno scoperto che “their new recruiting engine did not like women”
- Il sistema di assunzione di personale era basato su un sistema di ML
- Il sistema era stato addestrato sui cv delle persone assunte nei precedenti 10 anni. La maggior parte erano uomini quindi...
- “So the automatic recruitment tool preferred male candidates”
- E’ stato disattivato
- [Jeffrey Dastin](#). (2018). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters.





# Traduttori automatici

---

- “Ci sono una dottoressa e un infermiere”
- Italiano/Inglese (Inglese non distingue genere)
- → “*There are a doctor and a nurse.*”
- Inglese/Italiano
- → “*Ci sono un medico e un'infermiera.*”
- Non distinguendo l'inglese tra il genere, il sistema ha dovuto ‘indovinare’ il sesso del dottore e dell'infermiere



# Sistema Compas per il rischio di recidiva

<p><b>DYLAN FUGETT</b></p> <p>Prior Offense 1 attempted burglary</p> <p>Subsequent Offenses 3 drug possessions</p> <p><b>LOW RISK 3</b></p>	<p><b>BERNARD PARKER</b></p> <p>Prior Offense 1 resisting arrest without violence</p> <p>Subsequent Offenses None</p> <p><b>HIGH RISK 10</b></p>
---	--

<p><b>VERNON PRATER</b></p> <p>Prior Offenses 2 armed robberies, 1 attempted armed robbery</p> <p>Subsequent Offenses 1 grand theft</p> <p><b>LOW RISK 3</b></p>	<p><b>BRISHA BORDEN</b></p> <p>Prior Offenses 4 juvenile misdemeanors</p> <p>Subsequent Offenses None</p> <p><b>HIGH RISK 8</b></p>
--	---

**Two Drug Possession Arrests**

<p><b>DYLAN FUGETT</b></p> <p><b>LOW RISK 3</b></p>	<p><b>BERNARD PARKER</b></p> <p><b>HIGH RISK 10</b></p>
---	---

*Fugett was rated low risk after being arrested with cocaine and marijuana. He was arrested three times on drug charges after that.*

**Two Petty Theft Arrests**

<p><b>VERNON PRATER</b></p> <p><b>LOW RISK 3</b></p>	<p><b>BRISHA BORDEN</b></p> <p><b>HIGH RISK 8</b></p>
--	---

*Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not reoffend.*



# Bias detection and mitigation

---

- Nessuna blind application di tecniche ML
- Usando questi tools e algoritmi di grande successo, sviluppati nell'ambito del paradigma bottom-up del Machine Learning, dobbiamo analizzare se la conoscenza utilizzata per fare il training include bias su genere ed etnicità
- Il problema nasce fondamentalmente perchè si presta poca attenzione a come i dati sono raccolti, processati e organizzati
- Esistono metodi per fare il debiasing dei datasets
- Bias detection and mitigation




# ChatGPT

- **ChatGPT** è un modello di *chatbot* basato su apprendimento automatico sviluppato da OpenAI specializzato nella conversazione con un utente umano
- Poggia il suo funzionamento sul **modello linguistico**
- **GPT** sta per Generative Pre-trained Transformer

Da chatGPT3

- L'intelligenza artificiale (IA) è un ramo dell'informatica e della scienza dell'informazione che si occupa di creare macchine in grado di eseguire compiti che normalmente richiedono l'intelligenza umana. L'obiettivo principale dell'IA è quello di creare programmi o dispositivi che possono imparare, ragionare, risolvere problemi, comprendere il linguaggio naturale e riconoscere immagini e suoni, tra le altre cose.
- Ci sono diverse tecniche utilizzate nell'IA, tra cui il machine learning, la logica fuzzy, le reti neurali e l'apprendimento profondo. L'IA viene utilizzata in molti campi, tra cui la robotica, la medicina, la finanza, l'automazione industriale e i servizi di assistenza clienti.



L'IA continua a evolversi rapidamente e molti esperti ritengono che avrà un impatto significativo sulla società in futuro, portando a nuove opportunità e sfide in vari settori.

# Allucinazioni e plausibilità

- **Può soffrire di “allucinazioni”**. Ovvero «il modello può affermare con sicurezza cose/fatti che in realtà sono totalmente inventati»
- ‘Uno dei principali vantaggi della Chat GPT è la sua capacità di “imparare” dalle conversazioni che ha con gli utenti’
- Il sistema ha appreso a non contraddire l’utente, si mostra accondiscente con l’utente e da risposte **plausibili**. Ma la **plausibilità è diversa dalla verità**. Di qui nasce un forte pericolo. Il sistema può generare contenuti falsi, inappropriati e pericolosi. Tali contenuti poi vengono divulgati a un pubblico ampio creando false credenze. E’ necessario un controllo attento per evitare diffusione di falsità e possibilità di discriminazione a livello sociale.



# Questioni di genere?

- Interrogato su questioni di genere e se è vero che l'AI è ancora nelle mani degli uomini
- **«Più donne devono progettare i sistemi di AI, altrimenti si perpetuano i pregiudizi»**



# Possiamo fidarci?

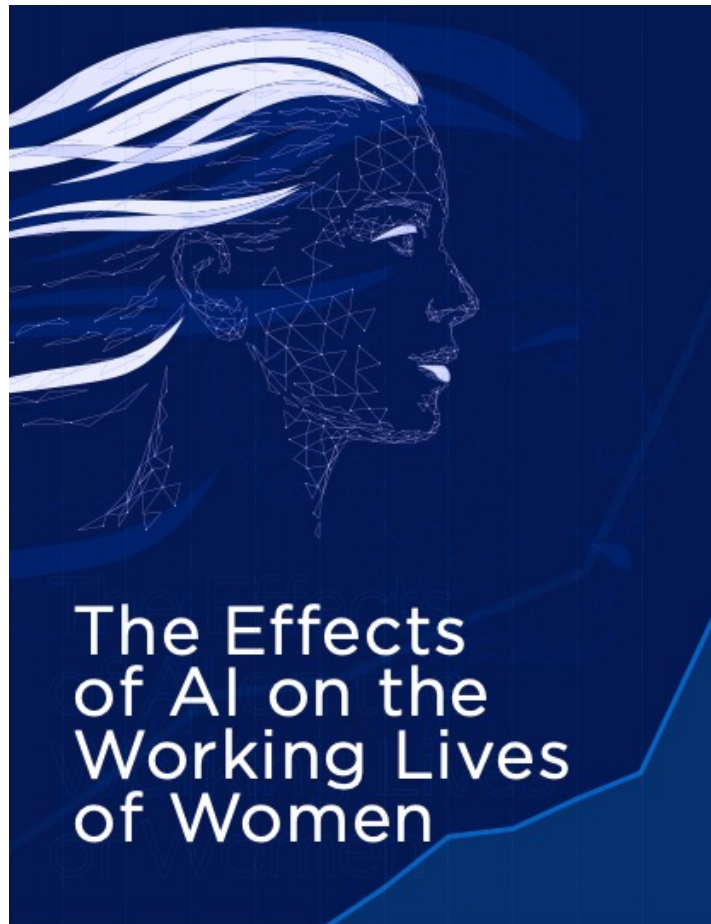
---

- Le tecniche di IA che vengono utilizzate da gran parte delle applicazioni sono degne della nostra fiducia (“trustworthy”)?
- Quali sono le caratteristiche/i requisiti che un sistema deve avere per poter essere degno della nostra fiducia?
  - Non possiamo permettere a una tecnologia e al suo uso di danneggiare valori umani fondamentali
  - AI trasparente, spiegabile, rispettosa della privacy e scevra da pregiudizi
- *Il vero scopo non è migliorare la tecnologia, ma migliorare noi stessi e la nostra società tramite l’uso della tecnologia*

*Francesca Rossi, IBM*



# Genere e tecnologie digitali



## The Effects of AI on the Working Lives of Women

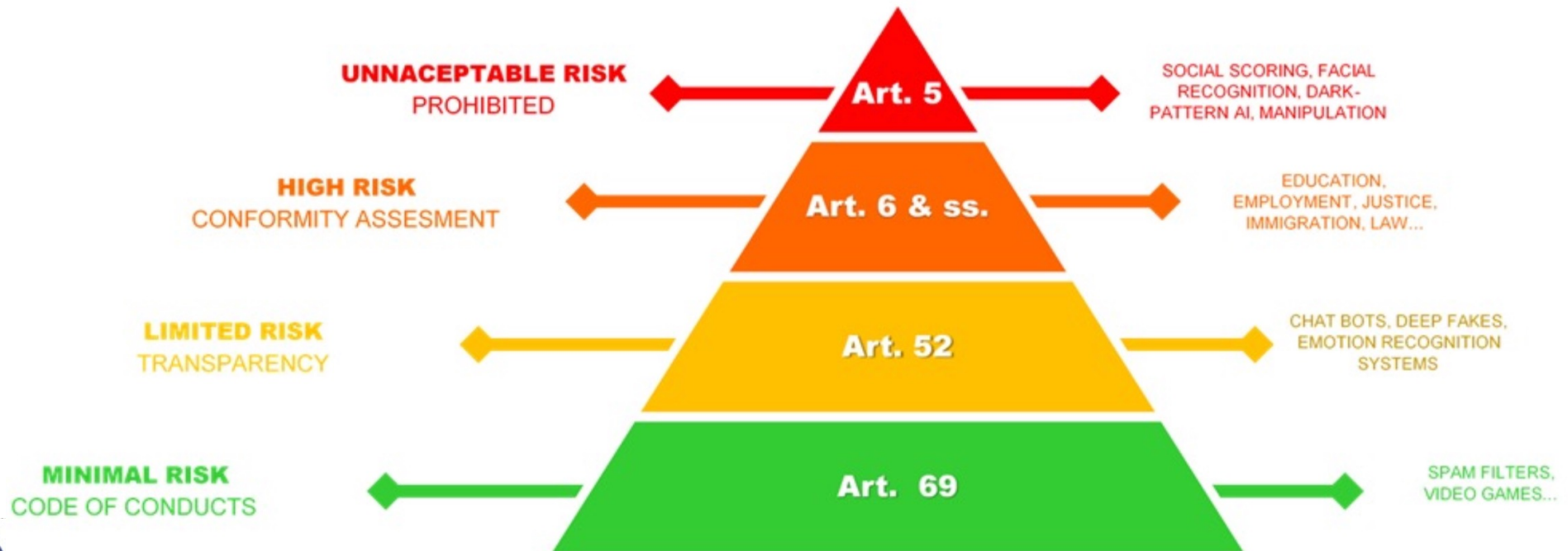


- UNESCO Report 2019, I'd blush if I could.  
<https://en.unesco.org/Id-blush-if-I-could>
- UNESCO Report 2020, Artificial Intelligence and Gender Equality.  
<https://en.unesco.org/AI-and-GE-2020>
- UNESCO, OECD and IDB (2022). The effects of Artificial Intelligence on the working lives of women.  
<https://oecd.ai/en/women-event-2022>



# AI-Act - A EU pyramid of ranking

## Levels of risk



# Etica dell'AI

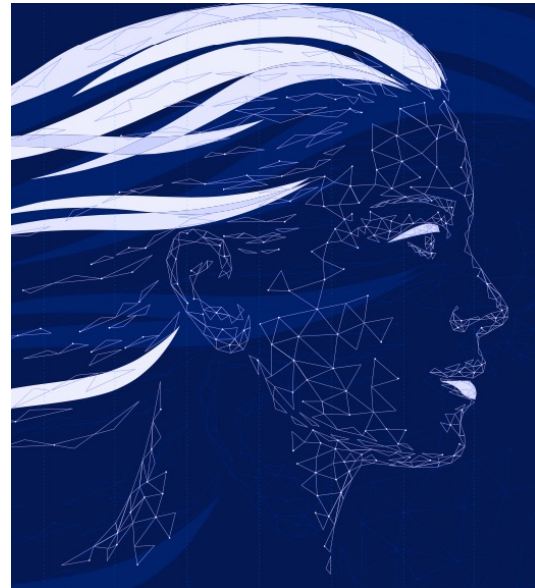
- Norme che regolamentano l'uso: AI-ACT per l'Europa
- Usare tecniche ML con consapevolezza
- Promuovere la carriera delle donne in AI: solo 16% di persone high-skilled AI sono donne
- AI pensa da uomo perchè è stata pensata quasi esclusivamente da uomini (bianchi)
- **3 prospettive**
  - 1 Sviluppare formazione e disseminare una cultura di genere
  - 2 Promuovere la carriera delle donne in AI: è essenziale per prevenire le distorsioni e migliorare l'efficacia dei sistemi basati sull'AI
  - 3 Debiasing i data-set, lavorare su dati e algoritmi

# Corso al DEI

- Saperi di genere ed etica nell'AI - 6 crediti liberi - a Ingegneria Biomedica ma aperto a tutti/e. General Course
- Docenti: S. Badaloni e Antonio Roda'

Un link al sito del corso e'

<https://www.dei.unipd.it/node/35894>



# Grazie!



- <https://www.youtube.com/watch?v=Lmq8hqlfKO0>
- <http://www.donnescienza.it>
- <http://www.epws.org>
- <http://eige.europa.eu>
- <http://genderedinnovations.stanford.edu/case-studies/>
- <http://www.unipd.it/bilancio-generi>

L'unico modo per arrivare a una visione più ampia è essere in un punto particolare  
(Donna Haraway, 1995)

S.Badaloni, Generi, Saperi e Giustizia Sociale, UNIPD, 10 maggio 2024

# Laboratorio

---

- Al sito [ustat.mur.gov.it](http://ustat.mur.gov.it) aprire il documento MONITORAGGIO DEL NUMERO DI IMMATRICOLATI e elaborare le statistiche di genere per le STEM.
- Tenere conto dei dati presentati dalle Prof.sse nelle slide durante le lezioni.
- Domanda 1. Quali sono le ragioni principali che NON spingono le giovani a iscriversi nelle STEM e in particolare nei Corsi di Laurea di Ingegneria Informatica e Informatica
- Domanda 2. Quali strategie potrebbero essere adottate per incentivare, in tempi brevi, la presenza delle giovani in tali percorsi universitari che portano sicuramente a un impiego interessante e ben remunerato.



# References

- S. Badaloni and L. Perini (Eds) A model for building a Gender Equality Index for academic Institutions. Padova University Press 2016 <http://www.padovauniversitypress.it/publications/9788869380983>
- S. Badaloni, F. A. Lisi (2020) Towards a Gendered Innovation in AI In: G. Vizzari, et al (Eds): Proceedings of the AIxIA 2020 Discussion Papers Workshop co-located with the the 19th International Conference of the Italian Association for Artificial Intelligence (AIxIA2020), Anywhere, November 27th, 2020 <http://ceur-ws.org/Vol-2776/paper-2.pdf>
- S.Badaloni, A.Rodà (2022). Gender Knowledge and Artificial Intelligence, *1st Workshop on Bias, Ethical AI, Explainability and the role of Logic and Logic Programming, BEWARE-22, co-located with AIxIA 2022, University of Udine, Udine, Italy.* <https://ceur-ws.org/Vol-3319/paper12.pdf>
- S. Lohr (2018) Facial Recognition Is Accurate, if You're a White Guy, The New York Times.
- N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, Aram Galstyan: A Survey on Bias and Fairness in Machine Learning. CoRR abs/1908.09635 (2019)
- F. Rossi. Il confine del futuro. Possiamo fidarci dell'Intelligenza Artificiale? Feltrinelli, 2019
- [European Commission 2020] European Commission (2020). Gendered Innovation 2: How Inclusive Analysis Contributes to Research and Innovation. ISBN 978-92-76-16416-62 [https://ec.europa.eu/info/news/gendered-innovations-2-2020-nov-24\\_en](https://ec.europa.eu/info/news/gendered-innovations-2-2020-nov-24_en)
- Ethics guidelines for trustworthy AI (2019) <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>
- I FOCUS DEL RAPPORTO ANVUR 2023 – Analisi di genere <https://www.anvur.it/wp-content/uploads/2024/01/Focus-equilibrio-di-genere-2023.pdf>

# Addio Giulia!

---



S.Badaloni, Generi, Saperi e Giustizia Sociale, UNIPD, 10 maggio 2024