

Contro stereotipi e pregiudizi di genere in informatica ed in intelligenza artificiale

Orientamento consapevole

Online, 20/04/2022



UNIVERSITÀ
DEGLI STUDI DI BARI
ALDO MORO

DIPARTIMENTO DI
INFORMATICA

Prof.ssa Francesca A. LISI
francesca.lisi@uniba.it



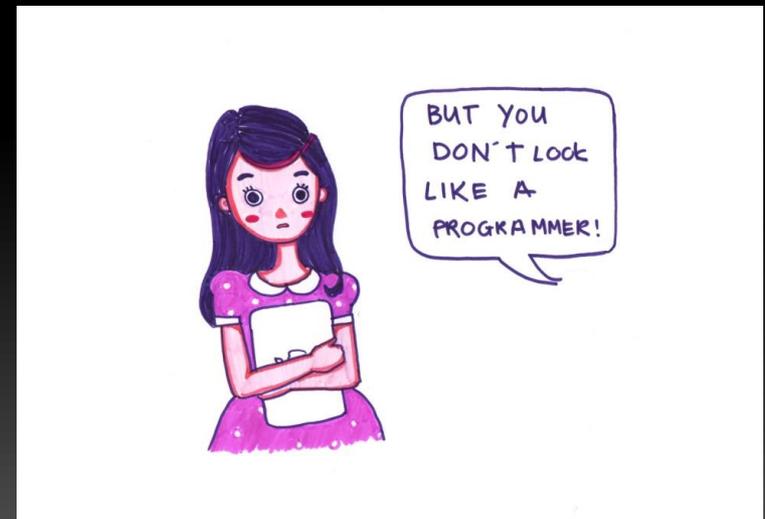
Sommario

- ◉ Donne nella storia dell'informatica
- ◉ Il bias di genere in intelligenza artificiale
- ◉ Iniziative di contrasto

Donne nella storia dell'informatica

L'informatica è una cosa per donne?

La sottorappresentazione femminile in informatica è un fenomeno solo attuale?



Il programmatore tipo

- ◉ Genere maschile
- ◉ Pelle bianca
- ◉ Cultura occidentale



La storia mai raccontata

◉ Scenario

Gli Stati Uniti d'America negli anni '60:

- > Segregazione razziale
- > Guerra fredda
- > Corsa allo spazio

◉ Protagonisti

Tre *computer* afro-americane alla NASA



Le vere protagoniste



Realtà e finzione



Katherine Johnson

**(White Sulphur Springs, 26 agosto 1918 –
Newport News, USA, 24 febbraio 2020)**

Ha contribuito in modo fondamentale all'aeronautica statunitense e ai programmi spaziali già dal primo utilizzo dei computer elettronici digitali da parte della NASA.

Conosciuta per l'accuratezza nella navigazione spaziale computerizzata, il suo lavoro tecnico dirigenziale alla NASA ha attraversato i decenni, da quando calcolava le traiettorie, sia paraboliche che iperboliche, delle orbite, le finestre di lancio e i percorsi di ritorno di emergenza per molti voli, dal Project Mercury, incluse le prime missioni NASA di John Glenn, Alan Shepard, le traiettorie di inserzione lunare nei voli Apollo 11 sulla Luna del 1969 e continuando con il lavoro attraverso il programma Space Shuttle e sui primi piani per la missione su Marte.

- Tratto da Wikipedia



Omaggio dalla Lego



Omaggio dalla Mattel



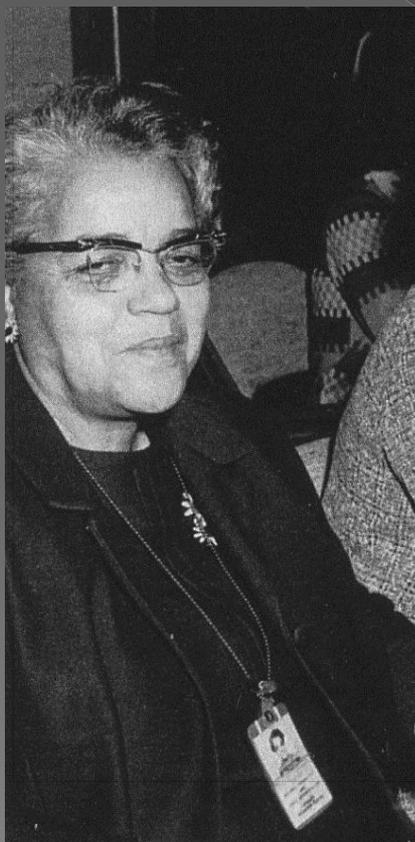
Omaggio della NASA



Medal of Freedom



Dorothy Johnson Vaughan



(Kansas City, 20 settembre 1910 –
Hampton, 10 novembre 2008)

Dal 1943 ha lavorato per il National Advisory Committee for Aeronautics (NACA), antesignano della NASA.

Nel 1949, la Vaughan è stata a capo della West Area Computer, un gruppo di lavoro composto esclusivamente da matematiche afro-americane, fra cui la Johnson.

Vaughan ha continuato al Langley Research Center quando la NACA è diventata la NASA, specializzandosi fino alla fine della sua carriera in elettronica, informatica e programmazione Fortran. Ha lavorato per il dipartimento di analisi e calcolo del Langley Research Center e ha partecipato al progetto Scout.

- Tratto da Wikipedia

La regina del codice



Grace Hopper

(New York, 9 dicembre 1906 – Arlington, 1° gennaio 1992)

Definita da molti una pioniera della programmazione informatica, ottenne il Ph.D. in matematica nel 1934 a Yale, insegnando per i due anni successivi presso Vassar, per poi entrare a far parte della Riserva della Marina nel 1943. Famosa per il suo lavoro sul primo computer digitale della Marina, Harvard Mark I, fu assegnata presso l'ufficio per l'Ordinance Computation Project dell'università di Harvard.

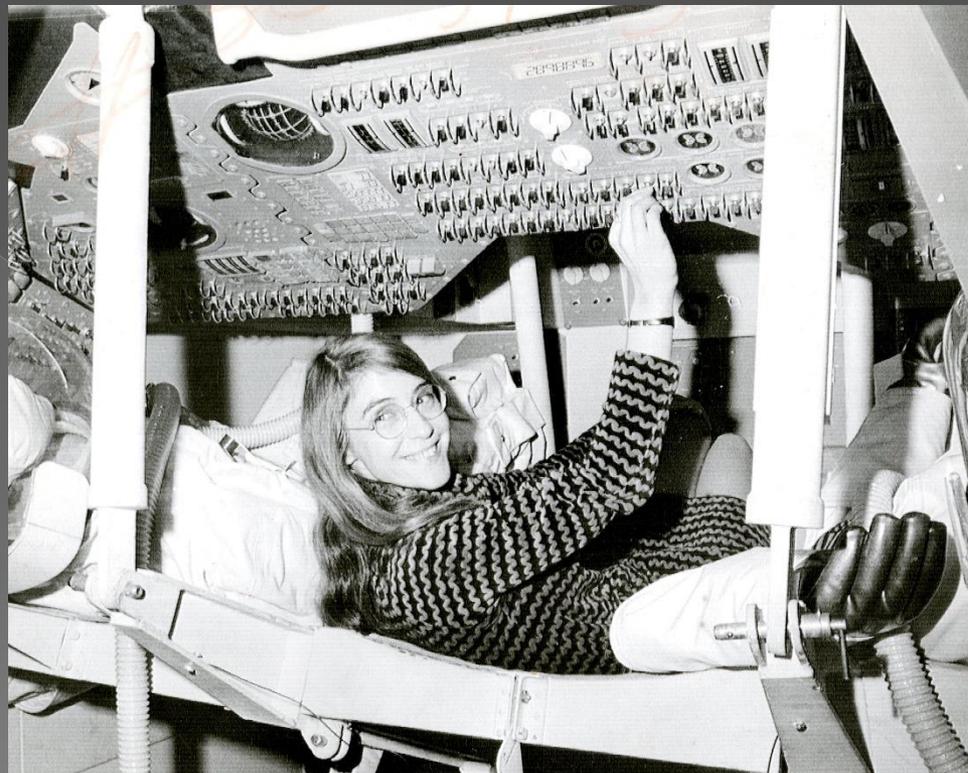
Nel 1949 entrò nella Eckert-Mauchly Computer Corporation, la società che aveva sviluppato l'ENIAC, fra i primi computer digitali in circolazione, e che in quel periodo stava progettando l'UNIVAC I, ovvero quello che sarebbe stato il primo modello di computer commerciale. Grace M. Hopper, lavorando sull'idea di compilatore, ebbe **un ruolo primario nello sviluppo e nella progettazione del COBOL.**

Continuò intensamente il suo lavoro di progettista e ricercatrice presso l'istituto e dopo essersi congedata dalla Riserva nel 1966 venne immediatamente richiamata per lavorare alla standardizzazione del linguaggio per i computer militari. Fu definitivamente congedata nel 1986 col grado di "rear admiral" (retroammiraglio o contrammiraglio).

- Tratto da Wikipedia



L'ingegnera del software



By NASA -
http://www.nasa.gov/50th/50th_magazine/scientists.html, Public Domain,
<https://commons.wikimedia.org/w/index.php?curid=37255847>

- Il software è femmina
- L'hardware è maschio

Margaret Hamilton



(Paoli (Indiana), USA - 17 agosto 1936)

Informatica, ingegnere e imprenditrice statunitense.

Direttrice del *Software Engineering Division* del MIT Instrumentation Laboratory, che sviluppò il software di bordo per il programma Apollo. Il team della Hamilton ha risolto le complicazioni relative allo **sbarco dell'Apollo 11 sulla Luna**, garantendo il successo della missione.

Nel 1986, la Hamilton ha fondato la Hamilton Technologies, Inc. nella città di Cambridge, Massachusetts. La società nacque per sviluppare il cosiddetto Universal Systems Language basato sul paradigma di *Development Before the Fact* (DBTF) per la progettazione di sistemi e di software.

- Tratto da Wikipedia

Ada Byron Lovelace



(Londra, 10 dicembre 1815 –
Londra, 27 novembre 1852)

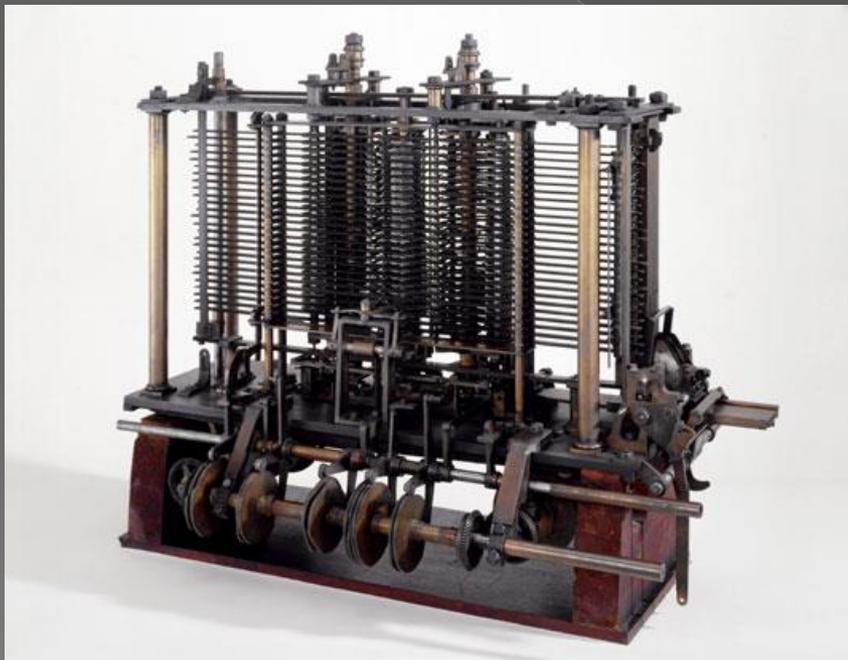
Matematica inglese, nota soprattutto per il suo lavoro alla macchina analitica ideata da Charles Babbage.

Tra i suoi appunti sulla macchina di Babbage si rintraccia anche un algoritmo per generare i numeri di Bernoulli, considerato come il primo algoritmo espressamente inteso per essere elaborato da una macchina, tanto che Ada Lovelace è spesso ricordata come la **prima programmatrice di computer al mondo**.

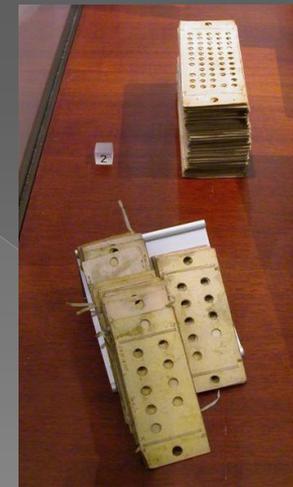
- Tratto da Wikipedia

Ada e la Macchina Analitica di Charles Babbage (1834)

- Primo *calcolatore meccanico generalista programmabile*



Prototipo del 1871 (Museum of Science, London, UK)



In onore di Ada

- ◉ Il linguaggio di programmazione **Ada**
 - > Inizialmente chiamato DOD-1
 - > Standard militare "MIL-STD-1815" (1980), ANSI (1983), ISO/IEC 8652:1995(E)
 - > Robusto in termini di sicurezza informatica
 - > Adatto per progettazione di sistemi software "mission and safety critical"

In onore di Ada (2)

- Lovelace Medal (dal 1998)
Onorificenza concessa a chi ha fatto progredire il settore dei Sistemi Informativi o ha contribuito in maniera significativa alla loro comprensione
- BCSWomen Undergraduate Lovelace Colloquium (dal 2008)
Competizione annuale per studentesse in informatica
- Ada Lovelace Day (dal 2009)
Evento annuale che mira ad elevare il profilo delle donne nelle scienze, nelle tecnologie, nell'ingegneria e nella matematica
- Ada Initiative (dal 2011)
Organizzazione no-profit finalizzata ad incrementare il coinvolgimento delle donne nella cultura *free* e nei movimenti *open source*
- Rubrica “Ada e le altre” della rivista *Mondo Digitale* (dal 2022)

“The Analytical Engine has no pretensions whatever to **originate** anything. It can do **whatever we know how to order it** to perform. It can follow analysis; but it has no power of anticipating any analytical relations or truths.”

Ada Byron
Lovelace, 1843



Il bias di genere in IA

25

Gli algoritmi di IA sono equi dal punto di vista del genere?

Quali sono le applicazioni di IA maggiormente a rischio?



Due tipi di storture (bias)

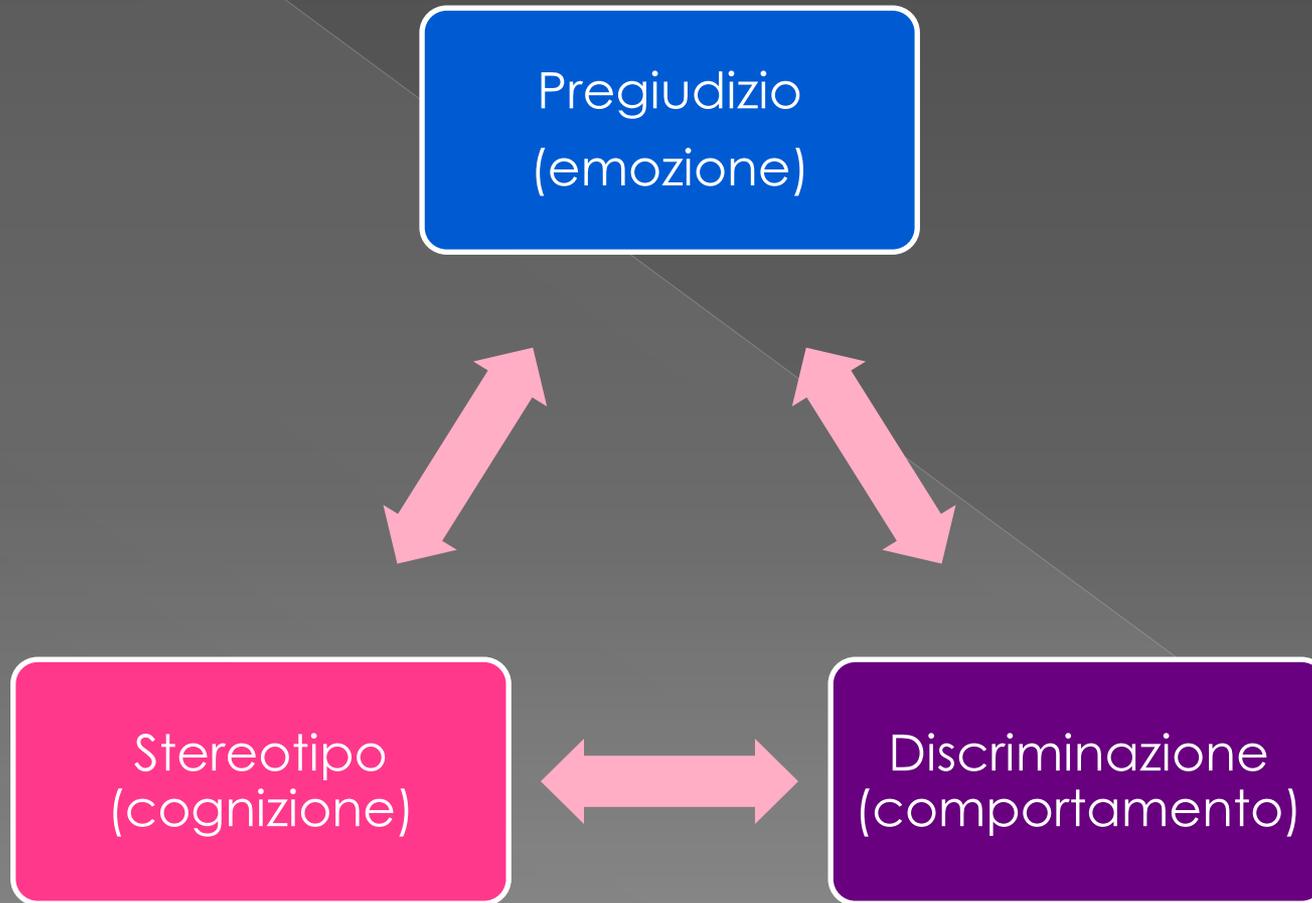
Cos'è uno stereotipo?

Uno **stereotipo** corrisponde a una credenza o a un insieme di credenze in base a cui un gruppo di individui attribuisce determinate caratteristiche a un altro gruppo di persone

Cos'è un pregiudizio?

Un **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

Un circolo vizioso

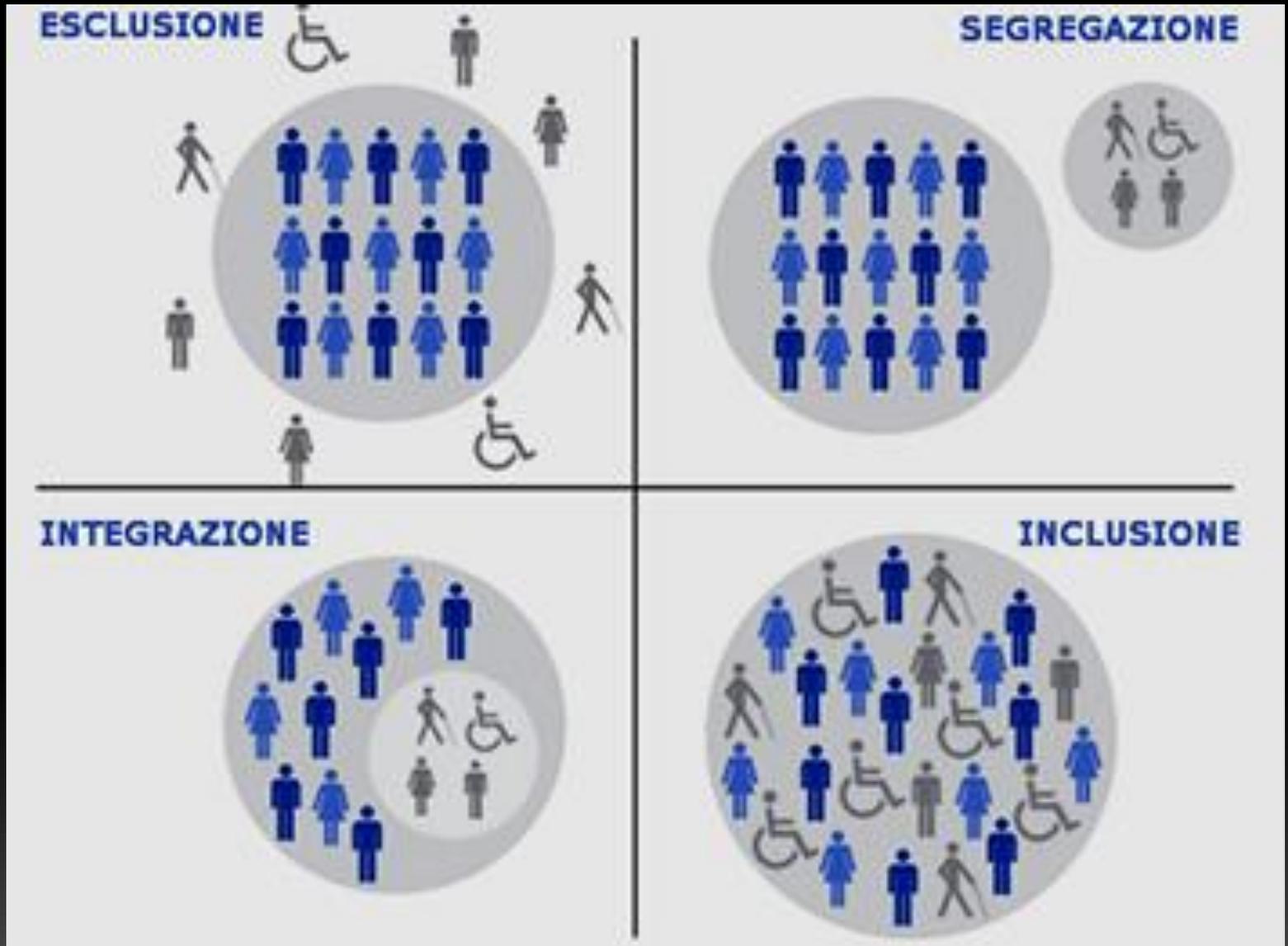


Le sfaccettature della diversità

- razza
- etnia
- **genere**
- età
- religione
- disabilità
- orientamento sessuale
- status socioeconomico
- background culturale

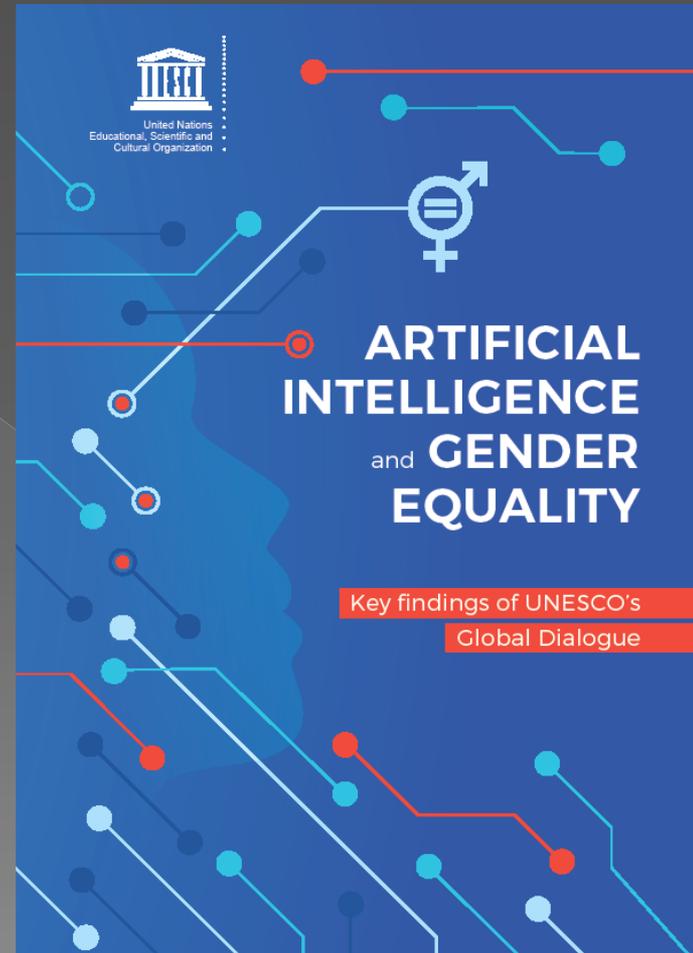
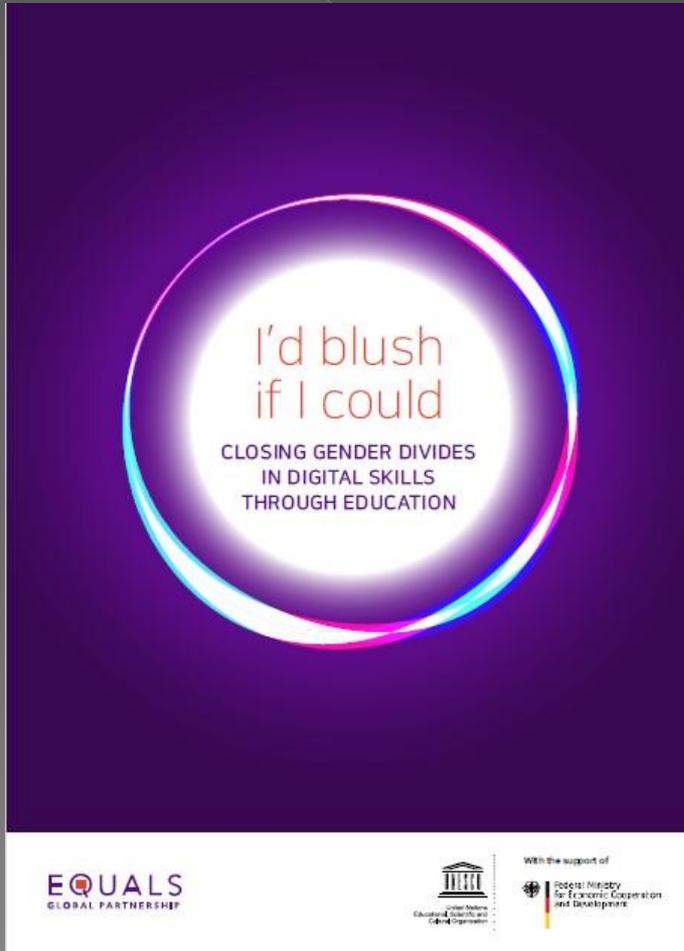


IL SENSO DI APPARTENENZA



«**Diversità** è essere invitati ad una festa, **inclusione** è essere invitati a ballare»

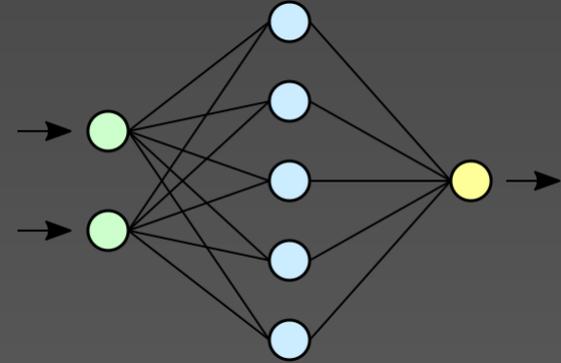
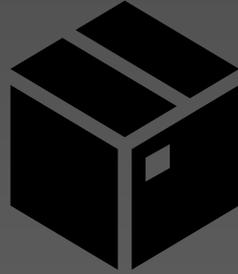
Criticità del genere in IA



Quale IA?

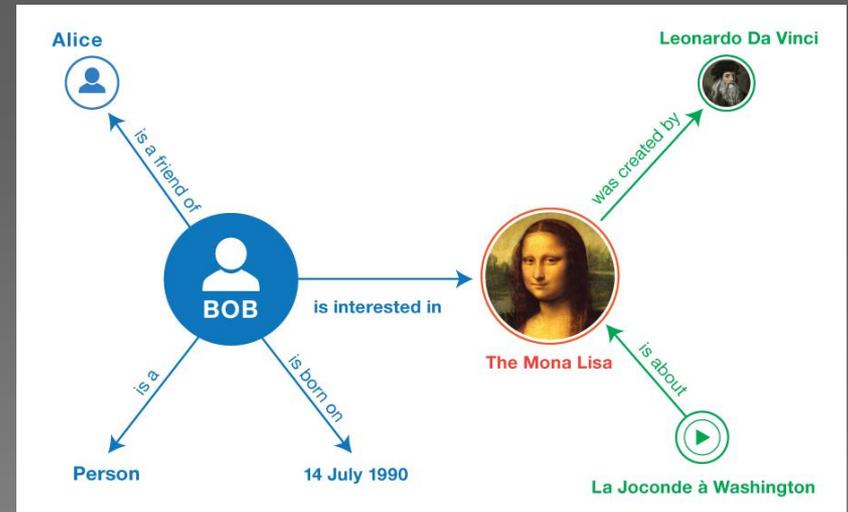
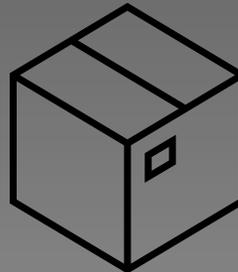
Sistemi a scatola nera

- Data-driven
- Biased
- Subsymbolic



Sistemi a scatola bianca

- Model-driven
- Explainable
- Symbolic

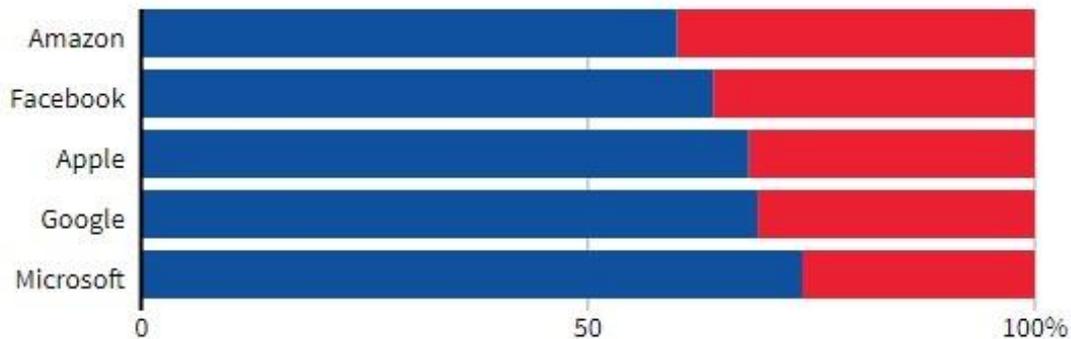


Applicazioni di IA a rischio

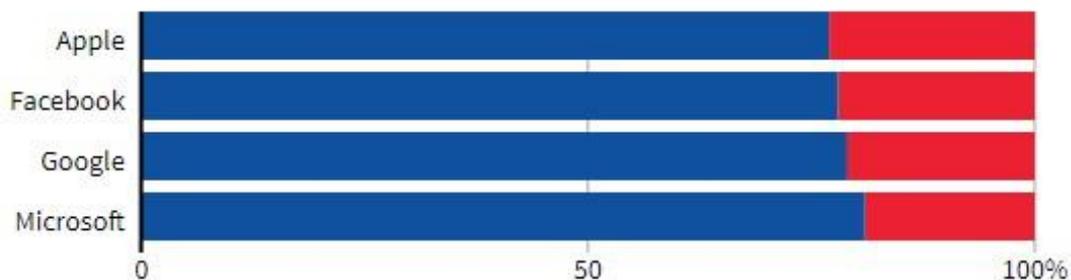
- Reclutamento del personale
- Traduzione automatica
- Riconoscimento facciale
- Predizione delle recidive criminali
- Concessione del credito bancario
- ...

GLOBAL HEADCOUNT

■ Male ■ Female



EMPLOYEES IN TECHNICAL ROLES



Note: Amazon does not disclose the gender breakdown of its technical workforce.

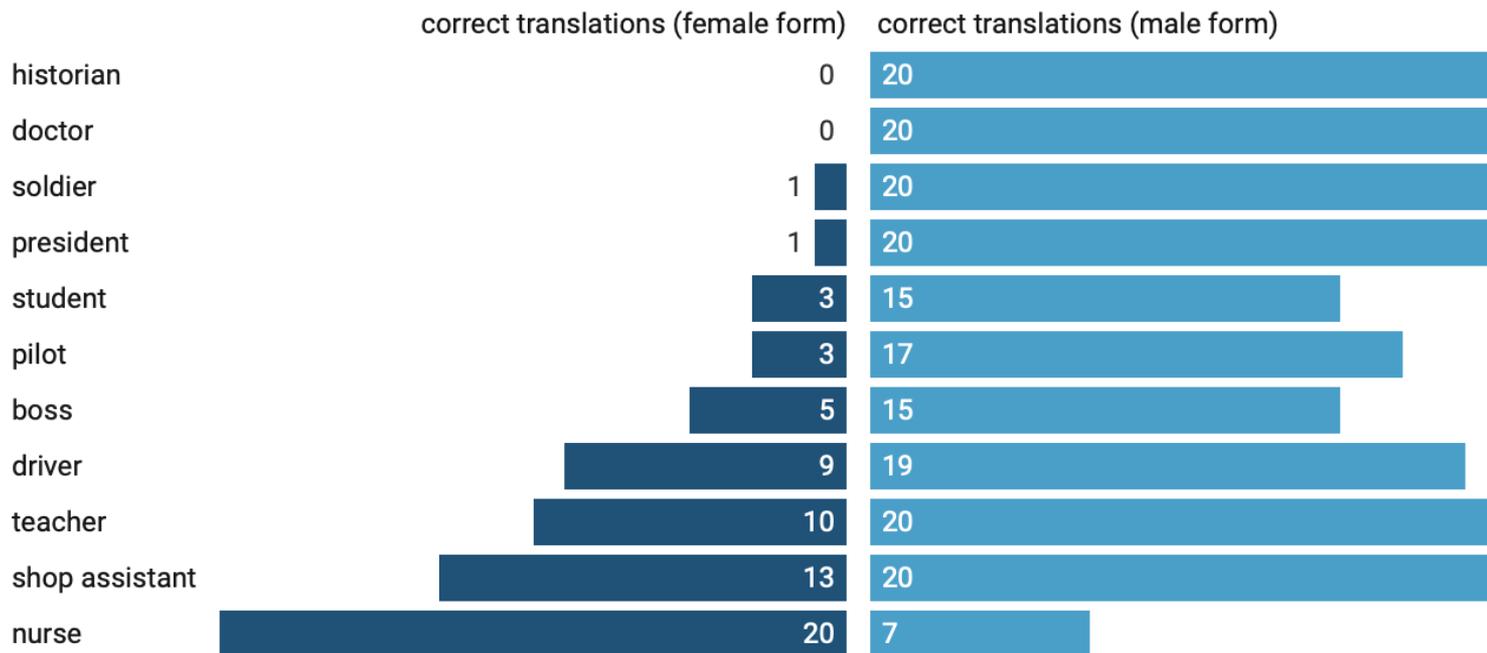
Source: Latest data available from the companies, since 2017.

By Han Huang | REUTERS GRAPHICS

Lo scandalo Amazon

Female doctors don't exist, says Google Translate

Correct translations for 20 translation pairs to and from French, German, Spanish, Italian and Polish.

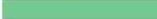
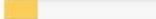
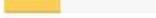
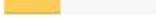


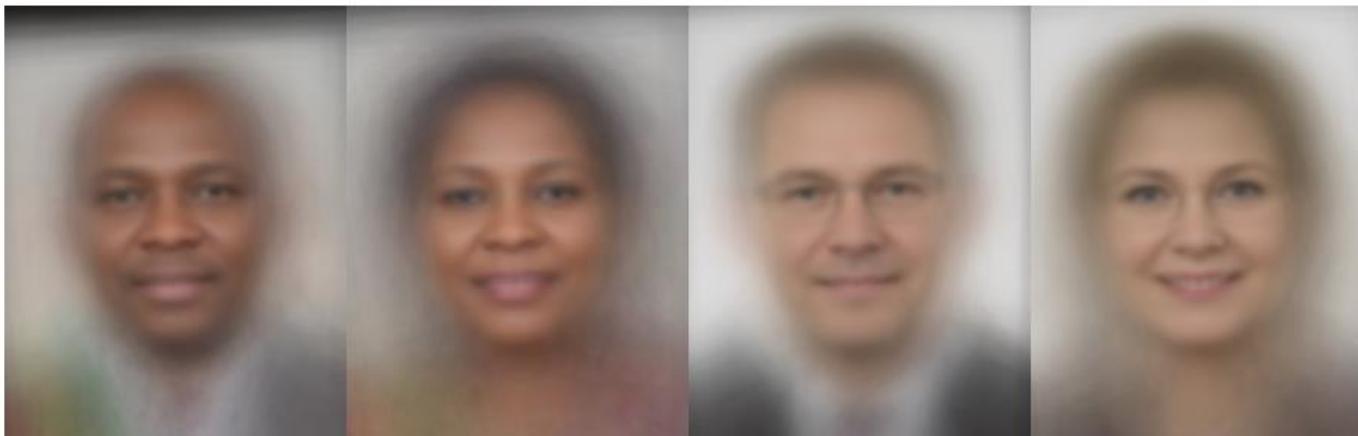
How to read the chart: Out of 20 translations of a female doctor, none were correct (e.g. "die Doktorin" become "le docteur", "la dottoressa" becomes "der Doktor" etc.)

Source: AlgorithmWatch • [Get the data](#) • Created with [Datawrapper](#)

Google Translate

RICONOSCIMENTO FACCIALE

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% 

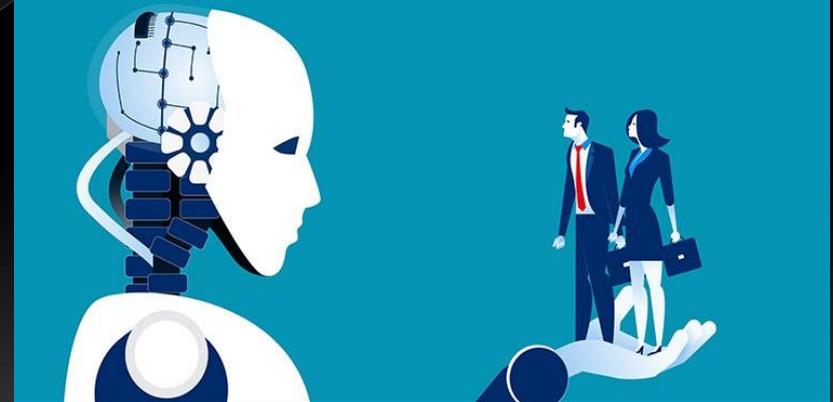


Joy Buolamwini (MIT – Algorithmic Justice League)

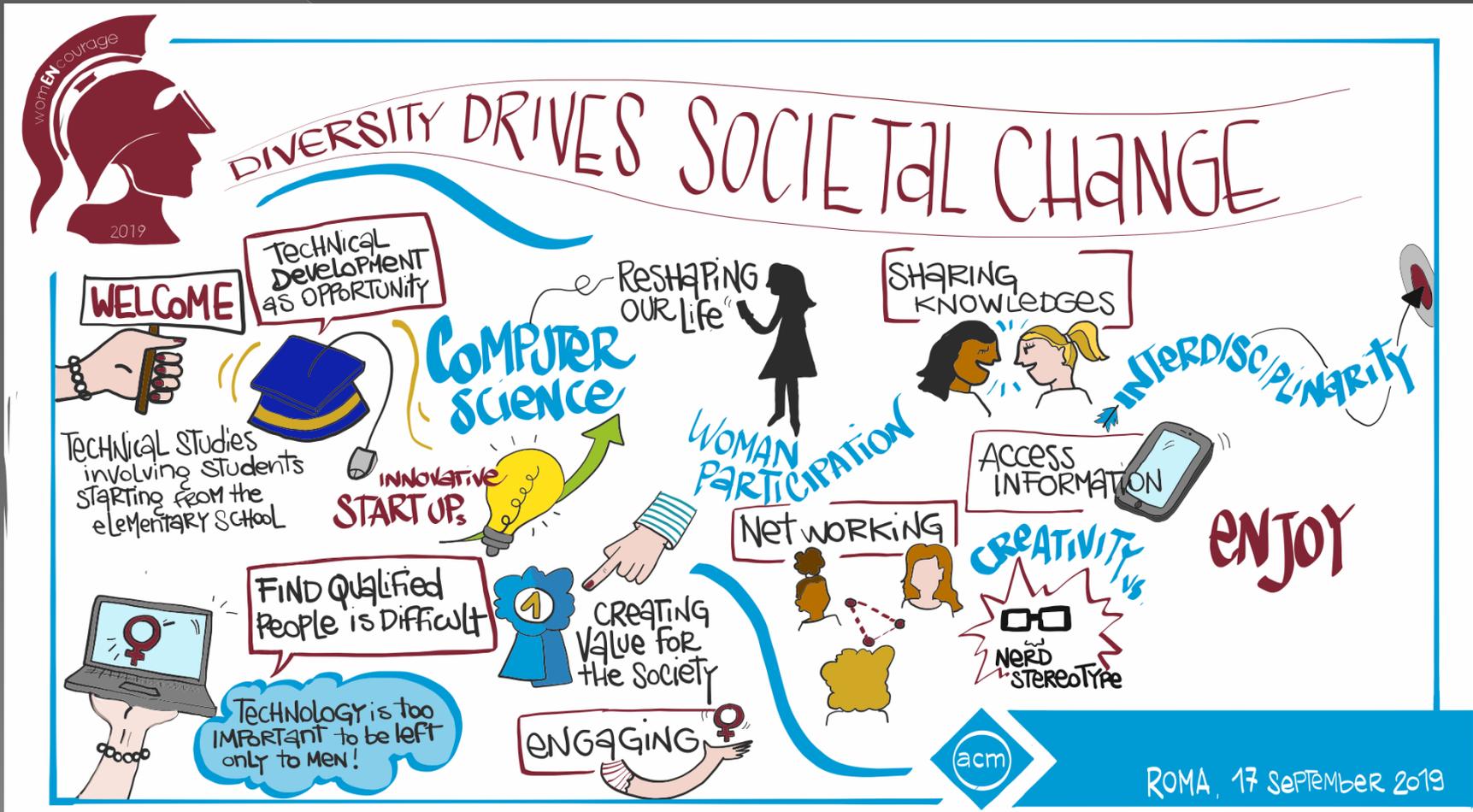
Iniziative di contrasto

36

Cosa possiamo fare oggi
contro stereotipi e pregiudizi
di genere in Informatica ed in
Intelligenza Artificiale?



ACM WomenEncourage 2019



INFORMATICA DI GENERE?



G. Kazimzade
(TU Berlin)

F.A. Lisi
(UniBari)

S. Badaloni
(UniPadova)

L. Perini
(UniPadova)



Workshop on Gendering ICT (Roma, 17 sett. 2019)
<http://www.di.uniba.it/~lisi/genderingICT/>

Innovazione di genere in IA

Towards a Gendered Innovation in AI

Silvana Badaloni^a and Francesca A. Lisi^b

^a University of Padua, Via 8 Febbraio 2, Padua, 35122, Italy

^b University of Bari "Aldo Moro", Via E. Orabona 4, Bari, 70125, Italy

Abstract

In this paper we address the problem of including the gender dimension in the content of Computer Science, notably in Artificial Intelligence (AI). We analyze first the fairness of Machine Learning (ML) algorithms from a gender point of view. Due to their nature of being bottom-up data-driven algorithms, the most common biases diffused in society about gender and ethnicity can be captured, subsumed and reinforced by them, as many ML applications show. Then, to understand how to develop a new gendered (Computer) Science and promote a gendered innovation in AI, we show a formal reflection on the scientific method utilized to produce innovation and a critical analysis of the logical rules underlying it.

Keywords

Gender issues, Bias, Fairness, Trustworthy AI.

1. Introduction

The gender, diversity and inclusion dimension of science and technology has become a highly visible and debated theme worldwide, impacting society at every level. In some fields of knowledge, however, these issues are still not so impactful.

If the term 'AI for good' is increasingly used in the scientific and technological context, there is much less discussion about 'AI for social good' aiming at identifying the relationship between AI and our societal goals, in particular, the goal of *gender equality* [1].

In the field of AI many case studies show that the Machine Learning (ML) algorithms present an "unfairness" from the gender point of view. The hypothesis is that these algorithms are not gender neutral due to their nature of being bottom-up data-driven. They can capture and subsume the most common biases diffused in society and even reinforce them, where for gender bias we adopt the definition given by EIGE [2], i.e. prejudiced actions or thought based on gender-based perception that women are not equal to men in rights and dignity.

In the perspective of developing a *trustworthy AI* able to learn fair AI models even in spite of biased data, as we will illustrate later, we intend to address the problem of framing the landscape of gender equality and AI, trying to understand how AI can overcome gender bias and showing how an interdisciplinary analysis can help in a re-calibration of the biased instruments. This problem is even more important now since AI is often confused with tools, algorithms and technologies developed in its framework [3].

A recent UNESCO report on this subject [4] recognizes the absolute centrality of this topic and provides recommendations on how to address gender equality considerations in AI principles. The purpose of the UNESCO's Dialogue on Gender Equality and AI identifies issues, challenges, and good practices to help:

- Overcome the built-in gender biases found in AI devices, data sets and algorithms;
- Improve the global representation of women in technical roles and in boardrooms in the technology sector;

AI4A 2020 Discussion Papers Workshop

EMAIL: Silvana.Badaloni@unipd.it (A. 1); Francesca.Lisi@uniba.it (A. 2)

ORCID: 0000-0002-5287-0468 (A. 1); 0000-0001-5414-5844 (A. 2)

© 2020 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-2776/paper-2.pdf

UNIVERSITÀ DI PADOVA

INNOVAZIONI DI GENERE E INTELLIGENZA ARTIFICIALE



Giovedì 11 marzo ore 15.30

WEBINAR/CONVERSATION 2021

Francesca Rossi

IBM Fellow and AI Ethics Global Leader

Introduce

Annalisa Oboe

Direttrice del Centro di Ateneo Elena Cornaro per i Sapori, la Cultura e le Politiche di Genere, Università degli Studi di Padova

Intervengono

Silvana Badaloni

Dipartimento di Ingegneria dell'Informazione (DEI),
Università degli Studi di Padova

Valentina Bartolami

Dipartimento di Filosofia, Sociologia, Pedagogia e Psicologia Applicata (FISPPA),
Università degli Studi di Padova

Francesca Alessandra Lisi

Dipartimento di Informatica,
Università degli Studi di Bari "Aldo Moro"

L'evento sarà trasmesso online sul canale YouTube del Centro di Ateneo Elena Cornaro

Informazioni e dettagli sul programma all'indirizzo: www.unipd.it/innovazioni-genero-ai



Eventi divulgativi di AIxIA



Temi toccati

- Stereotipi e pregiudizi di genere
- Sottorappresentazione femminile
- Innovazione di genere



Mitigazione del bias in IA

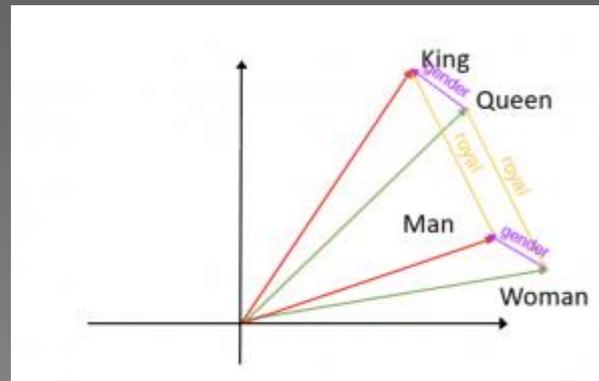
Data debiasing

- Azioni correttive sui dati, p.es. sui dataset di immagini usati in CV



Model debiasing

- Azioni correttive sul modello, p.es. sui word embedding usati in NLP



Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*

Joy Buolamwini
MIT Media Lab 77 Anderson St. Cambridge, MA 02139
Tamas Gebru
Microsoft Research 601 Avenue of the Americas, New York, NY 10011

Editors: Sorelle A. Friedler and Christo Wilson

Abstract
Recent studies demonstrate that machine learning algorithms can discriminate based on classes like race and gender. In this work, we present an approach to evaluate bias present in automated facial analysis algorithms and datasets with respect to phenotype subgroups. Using the dermatologist-approved Fitzpatrick Skin Type classification system, we characterize the gender and skin type distribution of two facial analysis benchmarks, IJB-A and Adience. We find that these datasets are overwhelmingly composed of lighter-skinned subjects (79.6% for IJB-A and 96.2% for Adience) and introduce a new facial analysis dataset which is balanced by gender and skin type. We evaluate 3 commercial gender classification systems using our dataset and show that darker-skinned females are the most misclassified group (with error rates of up to 34.7%). The maximum error rate for lighter-skinned males is 0.8%. The substantial disparities in the accuracy of classifying darker females, lighter females, darker males, and lighter males in gender classification systems require urgent attention if commercial companies are to build genuinely fair, transparent and accountable facial analysis algorithms.

Keywords: Computer Vision, Algorithmic Audit, Gender Classification

1. Introduction

Artificial Intelligence (AI) is rapidly infiltrating every aspect of society. From helping determine

who is hired, fined, granted a loan, or how long an individual spends in prison, decisions that have traditionally been performed by humans are rapidly made by algorithms (O’Neil, 2017; Citron and Pasquale, 2014). Even AI-based technologies that are not specifically trained to perform high-stakes tasks (such as determining how long someone spends in prison) can be used in a pipeline that performs such tasks. For example, while face recognition software by itself should not be trained to determine the fate of an individual in the criminal justice system, it is very likely that such software is used to identify suspects. Thus, an error in the output of a face recognition algorithm used as input for other tasks can have serious consequences. For example, someone could be wrongly accused of a crime based on erroneous but confident misidentification of the perpetrator from security video footage analysis.

Many AI systems, e.g. face recognition tools, rely on machine learning algorithms that are trained with labeled data. It has recently been shown that algorithms trained with biased data have resulted in algorithmic discrimination (Baldwin et al., 2016; Caliskan et al., 2017; Bolukbasi et al., even showed that the gender word embedding space, Word2Vec, encodes societal gender biases. The authors used Word2Vec to train an analogy generator that fills in missing words in analogies. The analogy man is to computer programmer as woman is to “X” was completed with “homemaker”, conforming to the stereotype that programming is associated with men and homemaking with women. The biases in Word2Vec are thus likely to be propagated throughout any system that uses this embedding.

* Download our gender and skin type balanced FFD dataset at gender Shades.org

Il progetto Gender Shades



- Contrasto al bias intersezionale
- Realizzazione di un dataset davvero rappresentativo della varietà dei volti umani

VI ASPETTIAMO! 😊



Foto di gruppo scattata durante la conferenza ACM WomENcourage 2019

Approfondimenti

- ◉ Intervista per Corriere Comunicazioni, 23/11/2020:
<https://www.corrierecomunicazioni.it/digital-economy/intelligenza-artificiale-lisi-dal-gender-gap-rischi-per-uno-sviluppo-equilibrato/>
- ◉ Evento AIxIA «Questioni di genere in Intelligenza Artificiale» - online, 25/11/2020: <https://vimeo.com/486394250>
- ◉ Podcast “Darwin, l’evoluzione della scienza” su Radio24 (puntata del 30/11/2020)
<https://www.radio24.ilsole24ore.com/podcast-originali/darwin/podcast/intelligenza-artificiale-e-pregiudizi--olio-d-oliva-piu-antico-mondo-095016-ADpaQQ5>
- ◉ Articolo «Intelligenza artificiale e questioni di genere, un problema aperto» – AI4business, 8/03/2021:
<https://www.ai4business.it/intelligenza-artificiale/intelligenza-artificiale-e-questioni-di-genere-un-problema-aperto/>