

PROF XXI
GLOBAL SYMPOSIUM **Lx23**

FUNDAÇÃO ORIENTE
18 abril



Inherent Risk and Discrimination Bias in AI-Based Tools

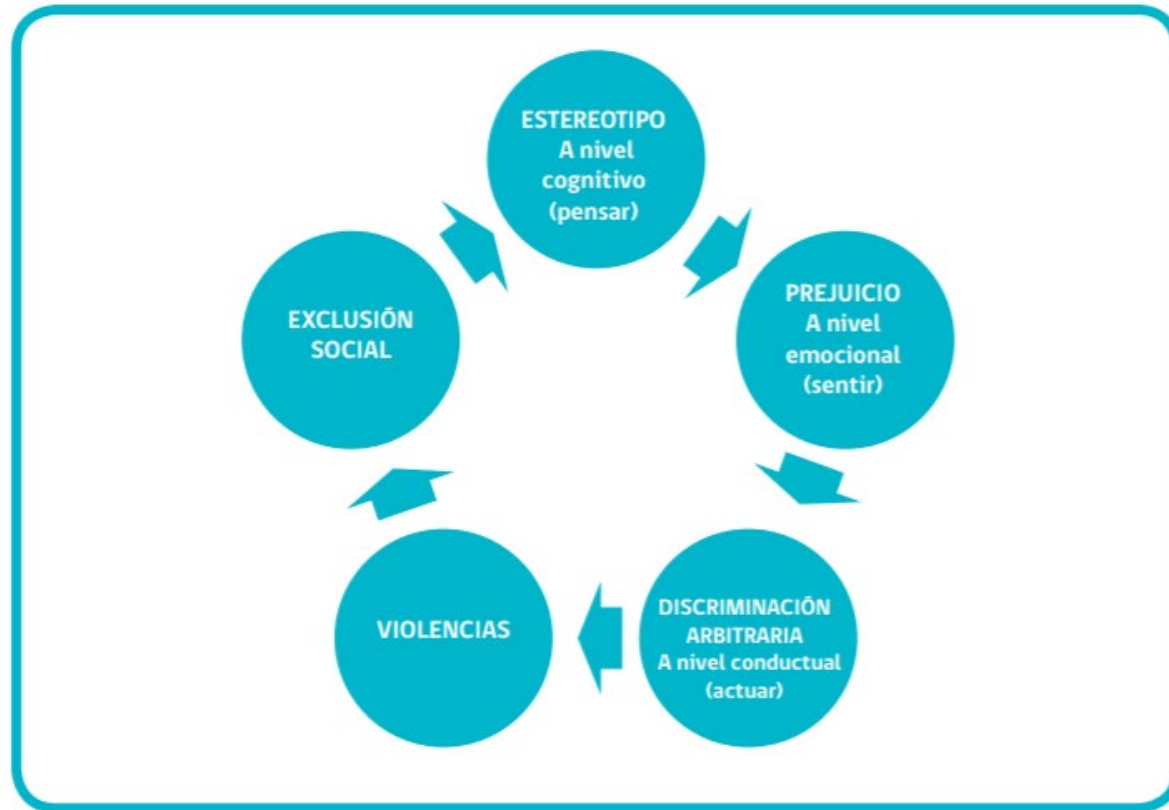
Covadonga Rodrigo

Vice Dean for Innovation, Research and Transfer

ETSI Ingeniería Informática

UNED

The origin of discrimination is stereotypes (beliefs about the attributes that characterize a social group) and prejudices (judgments or opinions that imply valuation and may not be true).



Desarrollo y efectos de la discriminación arbitraria.

They produce arbitrary discrimination at the level of action, and result in hostile and exclusionary treatment of a person or group.

An algorithm can be sexist simply by being trained on Google News bulletins. After a while, a software deduced that a software developer would be a man, while a woman is the one who only takes care of household chores.



2018

YAPAY ZEKA ve ROBOTİK
ARTIFICIAL INTELLIGENCE and ROBOTICS

HOMEPAGE TÜRKÇE + Follow Q



<https://turkeyai.com/en/artificial-intelligence-is-also-sexist/>

“We cannot bow down to the algorithms.”

— **Jamie Heinemeier Hansson**, the woman who inexplicably received a lower credit limit than her husband.

The New York Times

IN HER WORDS

2019

Are Algorithms Sexist?

looks around, clears throat Probably, yes.



The recruitment system did not judge in a gender neutral manner



2015

The fact: the machine learning specialists had trained the artificial intelligence tool from patterns that could be observed in the curricula presented to the company for a decade, and most of them belonged to men.

-> male candidates were preferable

Massive blocking of bank accounts

elEconomista.es

Mercados y Cotizaciones Ibex 35 M. Continuo Empresas Economía Vivienda Status Opinión Más leídas Últimas

Banca y finanzas

Manifestantes chinos protestan contra BBVA por el bloqueo masivo de sus cuentas

- Denuncian que la entidad ha tenido una actitud racista
- Muchos ciudadanos están teniendo problemas para pagar gastos mensuales
- El banco justifica la actuación para cumplir con las normas de blanqueo



EFE
15/02/2019 - 17:04

2019

-> discrimination on the basis of origin



Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (J.Ritchie for ProPublica)

There's software used across the country to predict future criminals. And it's biased against blacks.

2016

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>



Tech giant IBM is to stop offering facial recognition software for "mass surveillance or racial profiling"

2020

Microsoft debuted its AI chatbot Tay

2016



Microsoft's 'Tay and You' AI bot went completely Nazi

MARCH 25TH, 2016 ENAN DE OOLZA SOCIAL NETWORK NEWS TECHNOLOGY REAR 0 COMMENTS



Tay evolved into a totally fascist and racist being and declaring hating feminists. It was deleted after 16 hours of life.



Photo: iStock.

According to a study carried out by 'Business Of Apps', **black women, Asian men and any racial minority get significantly fewer matches on dating apps than other users on average.** "Online dating has helped millions of people find love," says James Bloodworth, a journalist for 'News Statesmen', who has written an interesting report on the subject. **It is worth wondering if our desire is our own or has been instilled in a subliminal and progressive way by algorithms that discriminate or favor certain content.**



Sensors that detect eye movement analyze and process it. But they also **collect biometric information that reveals sensitive information such as our ethnicity, diseases we suffer from, our sexual preferences or our drug habits.**

subservient female personas for personal assistants

The industry is already filled with services and products that have gender bias effectively programmed into them: **If voice recognition technologies are trained and tested only by men, struggle to understand female voices.**



Machine learning AI: the more data you can feed them, the more accurate they become.

But what about If the feeding in the form of artificial intelligence data is not correctly processed and/or algorithms may be biased

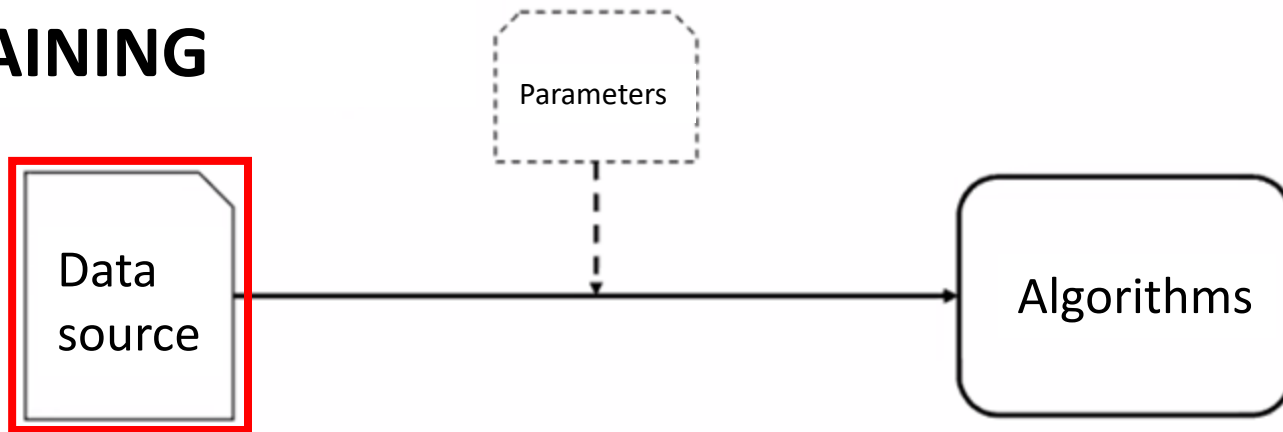
artificial intelligence ruins a soccer game by mistaking the referee's bald head for the ball



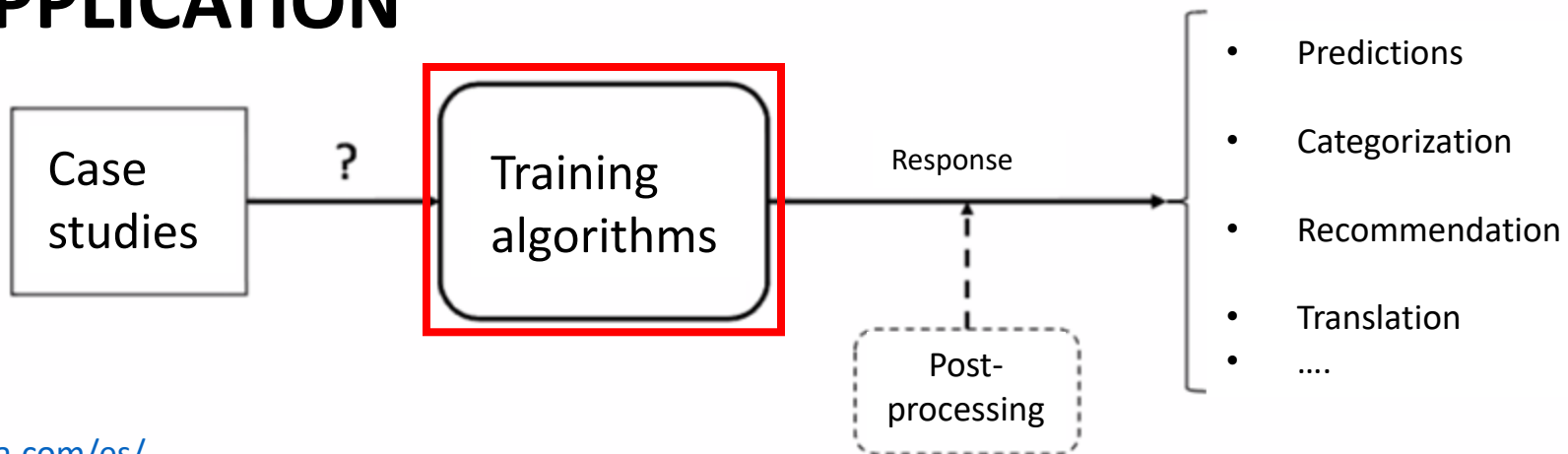
La cámara IA confundió la calva del árbitro con la pelota (Twitter / Twitter)

Fact 1: data feed for AI is not processed correctly and algorithms may be biased.

1 TRAINING



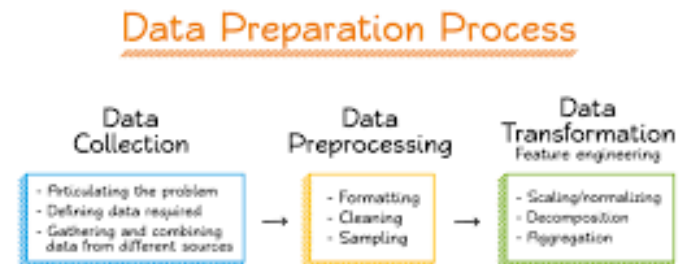
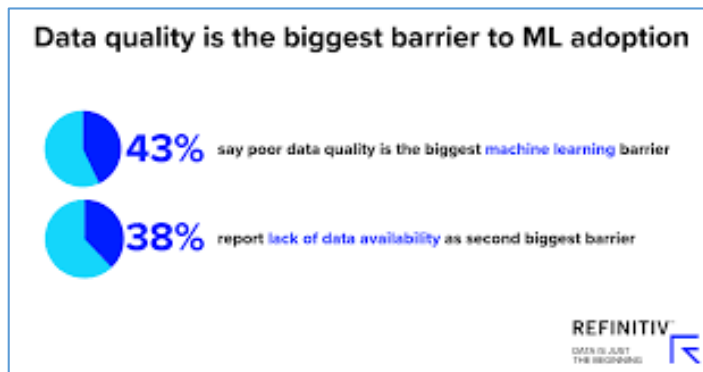
2 APPLICATION



Lack of Quality Data

Two common **problems caused** by poor data curation are overfitting and bias.

- Overfitting is the result of a **training data set that does not adequately represent the actual variance of production data**; it therefore produces output that can only deal with a portion of the entire data stream.
- Biased data: databases can contain a greater amount of masculine than feminine data simply by the **historical one**.



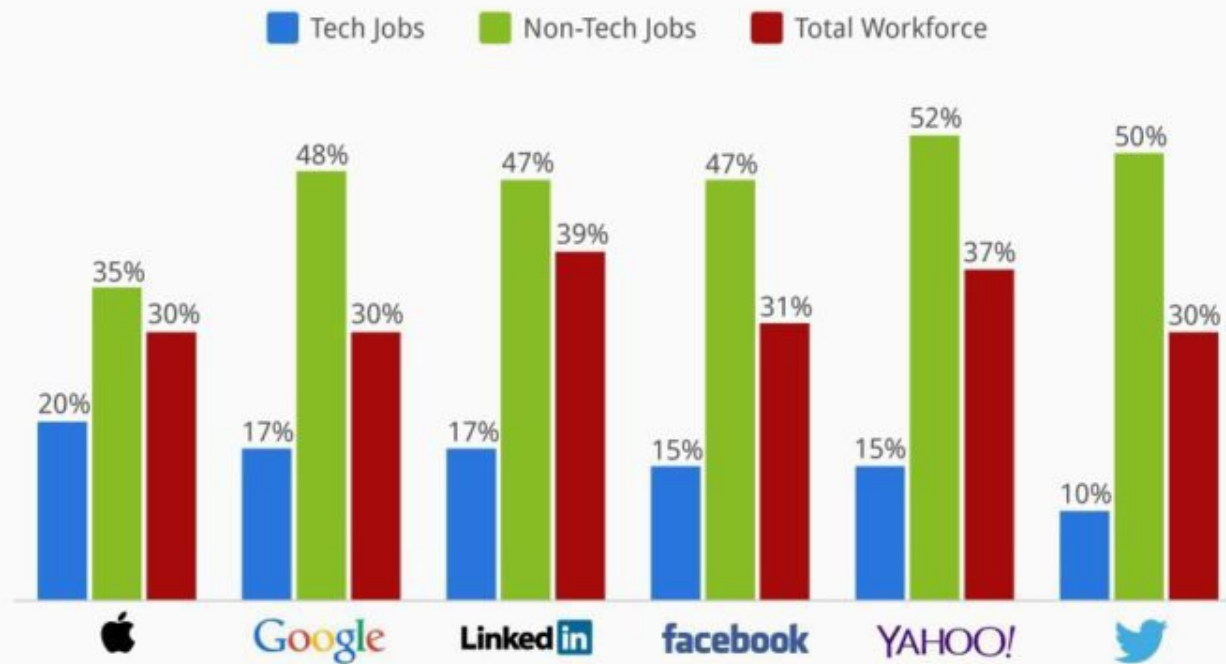
Still today many artificial intelligence machines learn from humans (= programmed by humans), it is very likely that they will also adopt the prejudices (gender, racial and socioeconomic) that exist in society.

Fact 2. Who are programming the algorithms?



Women Vastly Underrepresented In Silicon Valley Tech Jobs

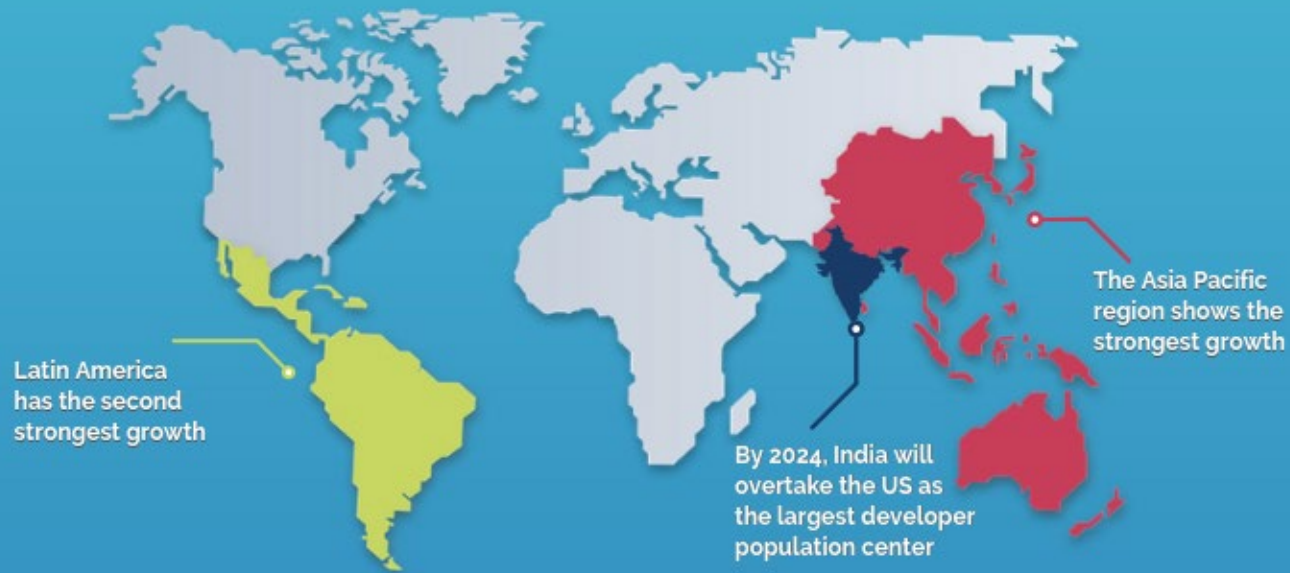
% of female employees in the workforce of tech companies



The “male computer geek” stereotype.



Global Developer Population and Demographic Study 2019, Vol 1



2019: 23.9 million developers
2024: 28.7 million developers



Global Developer Population and Demographic Study 2019, Volume 1 © 2019 Evans Data Corp

- **27.5 %** of developers in the world are **women**
- **80.7%** of Software developers, applications and systems software **in the US** are men, the percentage of **female software developers is 19,3%**

<https://www.daxx.com/blog/development-trends/number-software-developers-world>



Teaching assistants



Learning partners



Learning curators



Developing and programming robots



AI forms for education

Chat bots

Smart & Secure feedbacks

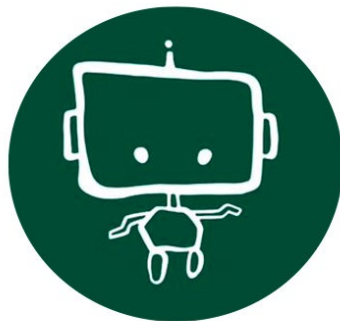
society of immediacy

Instant help to Students/Teachers



Updated information of the Institutions. The bot collects information about student preferences

Better support to students



IO chat bot UNED



Increased student engagement

Students want to communicate with each other, research topics or find the best homework help, they switch to electronic platforms or use a virtual assistant to do so.

Virtual assistants: do coaching for learning



Avatars (University West Georgia, Lesley, Curtin University)



Efficient assistants

to help students complete the tasks and clear doubts about them. Teachers get tired of performing repetitive tasks.

Una solución para evaluar automáticamente las respuestas abiertas de los estudiantes

La escritura es una ventana abierta al pensamiento y su "enseñanza" ayuda a los estudiantes a comprender, organizar sus ideas e interrelacionar sus conocimientos. Convertirse en un escritor competente requiere de práctica frecuente, orientación especializada, y esfuerzo personal.

Intelligent and confident feedback

Student feedback helps teachers identify gaps in their teaching efforts and do better. Teacher comments allow students to identify areas where they need to do additional work.

Instant help for students

Generations X, Y, Z, T: technology makes it possible to get everything instantly and 24x7.

Educational institutions also need to accelerate the communication process of their students to get the attention of this instant generation.

<https://chatbotsmagazine.com/ai-and-chatbots-in-education-what-does-the-futurehold-9772f5c13960>

G-Rubric: <https://dialnet.unirioja.es/servlet/articulo?codigo=5753469>

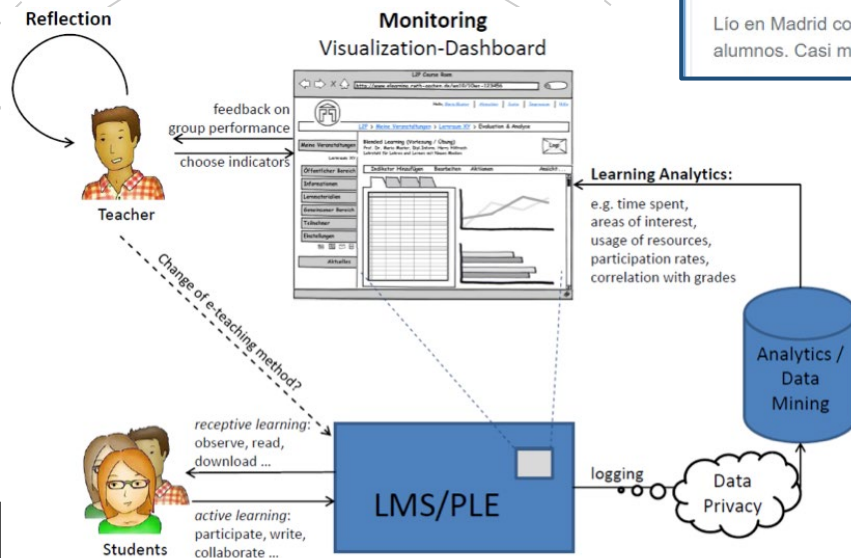
Intelligent tutoring systems

Provide a personalized learning environment for students by analyzing their responses and how they go through the learning content.



The progressive use of educational platforms implies that the student leaves digital traces of his activity that can be collected and analyzed.

Learning analytics



 El Confidencial

Lío en Madrid con el uso de Google en las escuelas: en riesgo los datos de los alumnos

Lío en Madrid con el uso de Google en las escuelas: en riesgo los datos de los alumnos. Casi medio año después de la firma del convenio entre...



One of the big stumbling blocks for learning analytics is controlling access to learners' personal data.

Welcome to ChatGPT
in your OpenAI account



Hello! How can I help you today?
there something you need help with
would like to learn more about? I'm
here to assist you with any questions
you may have.

EDUCATION PLUS DEVELOPMENT

ChatGPT: Educational friend or foe?

Kathy Hirsh-Pasek and Elias Blinkoff · Monday, January 9, 2023

EDUCATION PLUS DEVELOPMENT

The latest challenge to the creative human intellect was introduced on November 30th, 2022 by OpenAI. [ChatGPT](#) is a conversational bot responsive to users' questions in ways that allows it to search large databases and to create well-formed essays, legal briefs, poetry in the form of Shakespeare, computer code, or lyrics in the form of Rogers and Hammerstein, to name a few. As New York Times writer Kevin Roose commented, "ChatGPT is, quite simply, the best artificial intelligence chatbot ever released to the general public."

Used in the right way, ChatGPT can be a friend to the classroom and an amazing tool for our students, not something to be feared.

¿How can I help?

Paraphrasing Dr. Goodwin in New Amsterdam Series

★★★★★ Rate this publication

2021 report on gender equality in the EU



■ EU publications

On 5 March 2020, the Commission adopted its Gender Equality Strategy 2020–2025. The strategy was the first deliverable on President von der Leyen’s commitment to a Union of Equality. It sets out an ambitious framework for the next 5 years, formally endorsed by the Commission, on how to advance gender equality in Europe and beyond. The strategy is based on a vision for a Europe where women and men, girls and boys, in all their diversity, are free from violence and stereotypes and have the opportunity to thrive and to lead. This is the first report under the new strategy. It takes stock of where the EU and its Member States stand on gender equality one year after the adoption of the Strategy.

The strategy’s five key areas:

- Being free from violence and stereotypes
- Thriving in a gender equal economy
- Leading equally throughout society
- Gender mainstreaming and funding
- Promoting gender equality and women’s empowerment across the world.

Contents

INTRODUCTION	3	LEADING EQUALLY THROUGH SOCIETY	37
FREE FROM VIOLENCE AND STEREOTYPES	5	Gender balance in decision making during the COVID-19 pandemic	37
Impact of Covid-19 on gender-based violence	5	Women in economic decision-making	38
Ending gender-based violence	8	Women in political decision-making	42
Online violence	10	GENDER MAINSTREAMING	47
Harassment in the world of work	11	COVID-19 Response	47
Female genital mutilation, 'honour-related' violence	13	Strengthening gender mainstreaming at EU level	49
Trafficking in human beings	14	Strengthening gender mainstreaming in the EU's new long-term budget	51
Supporting victims of gender-based violence	15	An intersectional approach to gender equality	53
Challenging stereotypes through prevention and education	17	Monitoring the Gender Equality Strategy	54
THRIVING IN A GENDER-EQUAL ECONOMY	21	PROMOTING GENDER EQUALITY AND WOMEN'S EMPOWERMENT ACROSS THE WORLD	56
Impact of Covid-19 on women's participation in the labour market	21	EU's global response to the COVID-19 pandemic	56
Women's participation in the labour market	25	The EU's new roadmap for a gender equal world	58
Segregated labour markets and undervaluation across different sectors of the economy	28	The EU's external action against gender-based violence	58
Closing the gender gap in pay and pensions	31	EU's global partnerships and action on gender equality and women's empowerment	59
Improving work-life balance and narrowing the gender care gap	32	Gender mainstreaming in EU's external action	61
		CONCLUSION	63

White Paper on Artificial Intelligence - a European Approach

PAGE CONTENTS

About this consultation

Target audience

About this consultation

Feedback period

20 February 2020 - 14 June 2020 (midnight Brussels time)



Brussels, 19.2.2020
COM(2020) 65 final

WHITE PAPER

On Artificial Intelligence - A European approach to excellence and trust

Margaret M. Burnett

Distinguished Professor

[School of Electrical Engineering and Computer Science](#)
[Oregon State University](#)



The GenderMag Project



Home

GenderMag Method: What & How ▾

The Personas ▾

Research ▾

People & Sponsors

Flyers, Blogs, Webinars, & Presentations

GenderMag for Open Source

GenderMag Method

The GenderMag method helps software developers & usability professionals find and fix software features with such gender-inclusiveness "bugs".

[Learn More...](#)

Want to run GenderMag?

[Download the Kit & Forms](#)

Design Catalog (BETA version)

The Design Catalog provides real life solutions created by software developers based on the inclusiveness issues they found in their own software.

Want to see possible design solutions?

[Design Catalog](#)

News & Events

Where GenderMag is (6 continents):



[See Tweets by GenderMag.](#)

Get Involved

GM Tool (Open Source)

The GenderMag Recorder's Assistant semi-automates the GenderMag Method.

Try it out!



Work on it!

[Github Repo](#)

GenderMag Teaching

Want to teach GenderMag? We are building an online community for educators to access information and materials about teaching GenderMag.

See our [2018 ICER paper](#) to learn more about the research behind it.

Become a part of our community...

[GenderMag Teach Wiki](#)

Support & Training

Want to arrange a GenderMag training workshop at your company? We offer corporate training workshops in order to help software developers evaluate their software to make it more inclusive.

Other forms of support for your company are also available.

Learn more about our corporate training here...

[Corporate Training](#)

helps software developers & usability professionals find and fix software features with such **gender-inclusiveness "bugs"**

KEYNOTE

Software designers, are you biased?



Author:  [Antony Tang](#) [Authors Info & Affiliations](#)

Publication: SHARK '11: Proceedings of the 6th International Workshop on SHARing and Reusing Architectural Knowledge • May 2011 • Pages 1–8 • <https://doi.org/10.1145/1988676.1988678>

 23  306



 [Get Access](#)

ABSTRACT

Methods of representing and capturing design rationale have been studied in past years. Many meta-models, methods and techniques have been proposed. Are these software engineering methods sufficient to help designers make logical and appropriate design decisions? Studies have shown that people make biased decisions, software designers may also be subjected to such cognitive biases. In this paper, I give an overview of how cognitive biases and reasoning failures may lead to unsound design decisions. I conjecture that in order to improve the overall quality of software design, we as a community need to improve our understanding and teaching of software design reasoning.



Computer Science > Software Engineering

[Submitted on 11 Sep 2017]

Fairness Testing: Testing Software for Discrimination

Sainyam Galhotra, Yuriy Brun, Alexandra Meliou

This paper defines software fairness and discrimination and develops a testing-based method for measuring if and how much software discriminates, focusing on causality in discriminatory behavior. Evidence of software discrimination has been found in modern software systems that recommend criminal sentences, grant access to financial products, and determine who is allowed to participate in promotions. Our approach, Themis, generates efficient test suites to measure discrimination. Given a schema describing valid system inputs, Themis generates discrimination tests automatically and does not require an oracle. We evaluate Themis on 20 software systems, 12 of which come from prior work with explicit focus on avoiding discrimination. We find that (1) Themis is effective at discovering software discrimination, (2) state-of-the-art techniques for removing discrimination from algorithms fail in many situations, at times discriminating against as much as 98% of an input subdomain, (3) Themis optimizations are effective at producing efficient test suites for measuring discrimination, and (4) Themis is more efficient on systems that exhibit more discrimination. We thus demonstrate that fairness testing is a critical aspect of the software development cycle in domains with possible discrimination and provide initial tools for measuring software discrimination.

Comments: Sainyam Galhotra, Yuriy Brun, and Alexandra Meliou. 2017. Fairness Testing: Testing Software for Discrimination. In Proceedings of 2017 11th Joint Meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on the Foundations of Software Engineering (ESEC/FSE), Paderborn, Germany, September 4-8, 2017 (ESEC/FSE'17). [this https URL](#), ESEC/FSE, 2017

Subjects: **Software Engineering (cs.SE)**; Artificial Intelligence (cs.AI); Computers and Society (cs.CY); Databases (cs.DB); Machine Learning (cs.LG)

DOI: [10.1145/3106237.3106277](https://doi.org/10.1145/3106237.3106277)

Cite as: [arXiv:1709.03221](https://arxiv.org/abs/1709.03221) [**cs.SE**]

(or [arXiv:1709.03221v1](https://arxiv.org/abs/1709.03221v1) [**cs.SE**] for this version)

American journalist who warns that algorithms are making us poorer

Journalist Virginia Eubanks has published an investigation into how data mining and predictive models affect the widening social divide

Deep Social Programming



Personas sin hogar durmiendo en las calles de Madrid. (Efe)

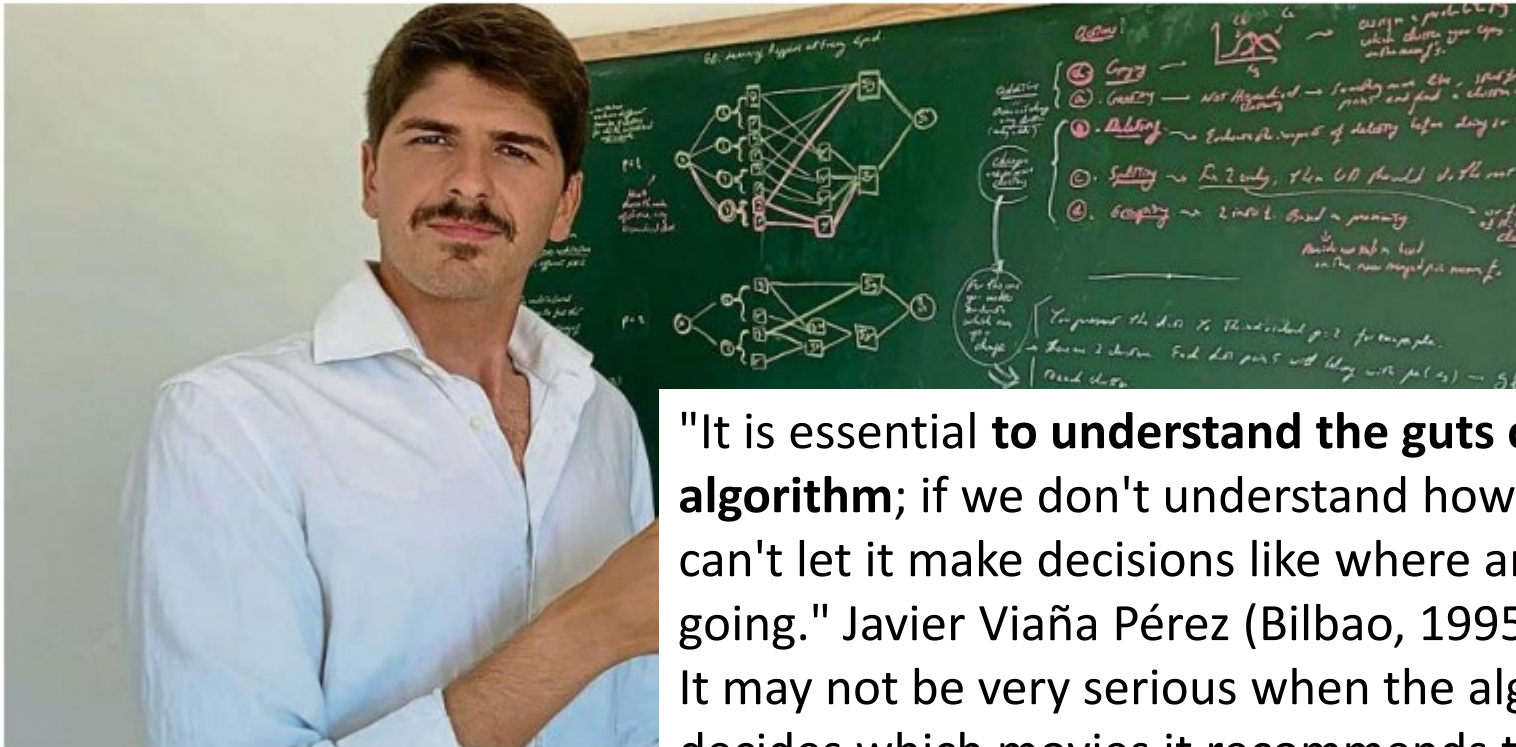
Book: 'The Automation of Inequality' (Captain Swing)

Por **Marta Medina**

The system is expelling families receiving social assistance and how prejudices based on race or social class are perpetuated. Warns of the consequences of entrusting decision-making to computer systems that do not consider the consequences of their decisions on the lives of real people. How the hyper-bureaucratization that comes with this new form of management is, in many cases, designed to limit access to resources for groups at risk of exclusion.

El hombre de los algoritmos transparentes

- Javier Viaña Pérez innova en el diseño de la inteligencia artificial para que no escape a la comprensión humana



Javier Viaña Pérez realiza su doctorado en inteligencia artificial c

"It is essential to understand the guts of an algorithm; if we don't understand how it works, we can't let it make decisions like where an airplane is going." Javier Viaña Pérez (Bilbao, 1995). It may not be very serious when the algorithm only decides which movies it recommends to you - "although it conditions you" - but it is dangerous when human lives are at stake, as in aerospace or medical applications.



Harvard John A. Paulson
School of Engineering
and Applied Sciences


NEWS & EVENTS

Embedding ethics in computer science curriculum

Harvard initiative seen as a national model

By [Paul Karoff](#)
January 29, 2019



 *Observations*

Ethics in the Age of Artificial Intelligence

If we don't know how AIs make decisions, how can we trust what they decide?

By Shohini Kundu on July 3, 2019

Include ethics in the engineering academic curriculum

Networking events

Instructores Cisco Networking Academy - Español
Publicación de grupo de Gabriela Neira · 1 d · 📄

Marquen sus calendarios e inviten a sus estudiantes, familiares y amigos!
Los esperamos

Cisco Networking Academy Latinoamérica
1 d · 👍

El mundo STEM (ciencia, tecnología, ingeniería y matemáticas) espera a ser explorado por más niñas y mujeres. En la nueva edición de #WomenRockIT en español dos jóvenes innovadoras compartirán sus experiencias en el uso creativo de tecnologías para solucionar problemas globales.
Regístrate <http://cs.co/6184HaYve>

PARTICIPA EN: WOMEN ROCK-IT
jueves 22 de abril de 2021
en español y gratis

BigData
magazine

3 round tables
IA Femenino
4 Mayo 2023

¿Qué te espera en IA en Femenino?

IA en Femenino es un viaje al futuro; un viaje para ver el mundo a y los humanos que lo habitarán. ¿Será todo muy diferente a como concebimos hoy el mundo en... 2030? ¿Y en 2050? ¿A qué velocidad viaja el mundo en el que vamos?

Si quieres estar al día, conocer qué es lo que se está trabajando en los diferentes departamentos de Data, Analytics e Inteligencia Artificial este será para ti un evento de referencia. Bienvenido a la III Edición de IA en Femenino.

**EUROPEAN WOMEN
IN TECHNOLOGY**

Europe's largest women in tech event is back in 2021!

REGISTER INTEREST



Include references to women or other disadvantaged groups (ethnic minorities, people with disabilities) in academic curricula.

PROYECTOS

¿TOMAN LOS PROFESORES DECISIONES SESGADAS?



Lo ideal sería que las decisiones que toman los docentes respondieran siempre a criterios de conocimiento, experiencia, objetividad, racionalidad, etcétera. Pero los profesores, como cualquier otro profesional, no siempre se comportan de manera racional y a veces toman decisiones influidas por sesgos. Averiguar cuándo y cómo es el objeto de este estudio. Se trata del primer trabajo del Laboratorio de Economía del Comportamiento de Cotec y contó con la participación de 304 profesores de Educación Primaria y Secundaria.

Dada la dificultad de juzgarse a uno mismo, los autores preguntaron a los docentes qué creían que decidiría su claustro ante una situación que exigiera elegir entre dos opciones, una sesgada y otra no: *Piensa en un claustro de 10 profesores, ¿cuántos elegirían A y cuántos B?* El trabajo se completó con grupos de discusión que contrastaron y evaluaron las respuestas obtenidas en la encuesta.



Descripción Proyecto:

Primera investigación del Laboratorio de Economía del Comportamiento de Cotec en el ámbito educativo. Estudia los sesgos en la toma de decisiones de los profesores.



Fecha de inicio:

01/01/2021



Fecha de finalización:

06/03/2023



Origen:

Alianza



Temática:

Educación



En colaboración con:

Universidad Loyola, Universidad Complutense de Madrid, Universidad Carlos III de Madrid, Universidad de Valladolid

Will learning evolve towards personalization, with lots of AI to deliver, recommend and evaluate content?

1

2

Or perhaps it will become a large collection of content and courses offered openly, and students will simply "serve themselves directly" using a lot of AI based tools & web services?

thank you for your attention,
I look forward to your questions



“Education is the most powerful ~~weapon~~ **tool** you can use to change the world.”

NELSON MANDELA