



Promoting gender equality, advancing the economy

26th March 2019 | JUMP Forum Brussels

Bias and Diversity in Machine Intelligence

Dr Allison Gardner



jump.brussels

WWW.JUMP.EU.COM



#JUMP2019
@JumpEquality



Allison Gardner

FOUNDER WOMEN LEADING IN AI
TEACHING FELLOW KEELE UNIVERSITY



@allisoncgardner

Bias and Diversity in Artificial Intelligence

Two sides of the same coin

Tech
Gadgets

Amazon ditched AI recruiting tool that favored men for technical jobs

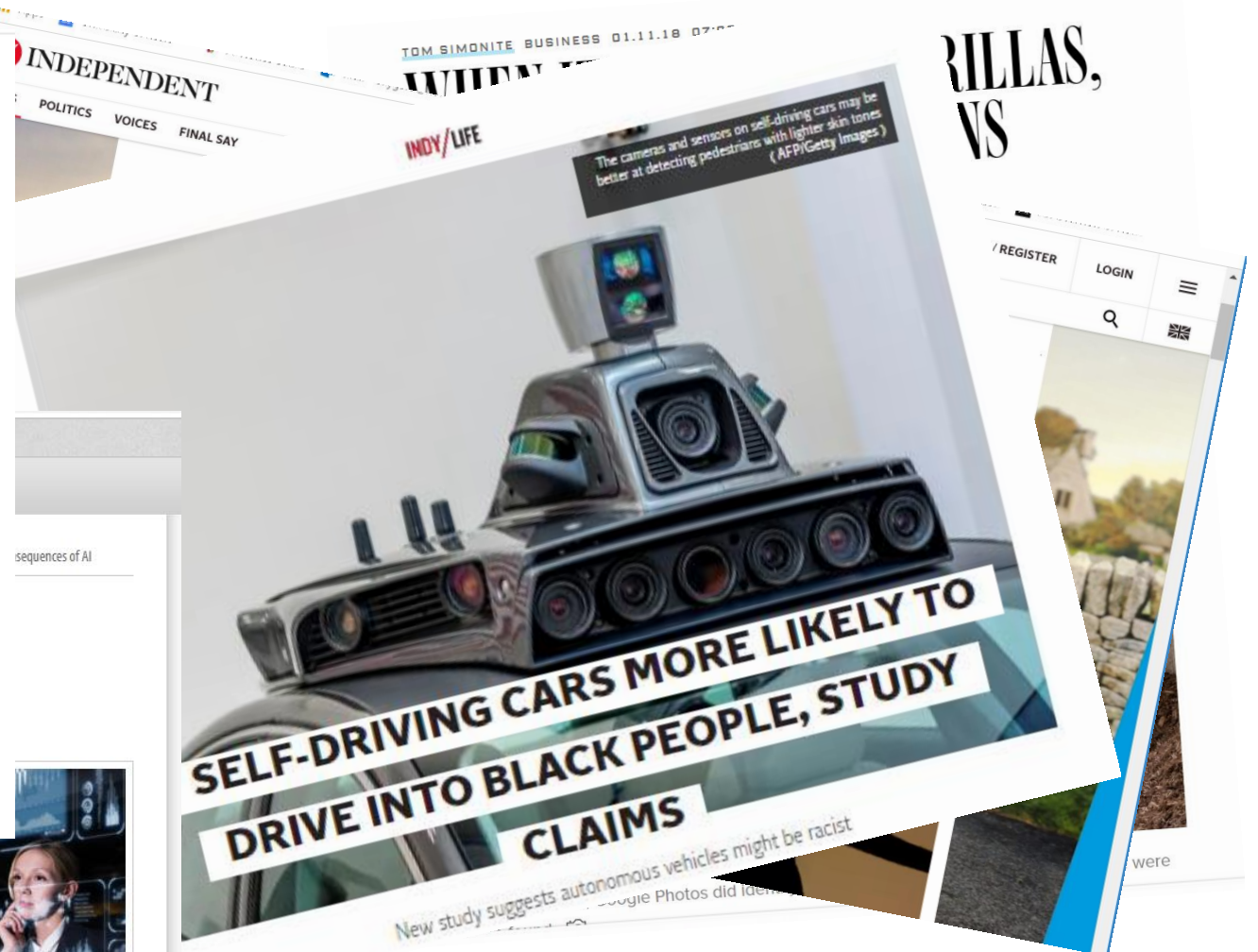
Specialists had been building computer programs since 2014 to review résumés in an effort to automate the search process



▲ Amazon's automated hiring tool was found to be inadequate after penalizing the résumés of female candidates. Photograph: Brian Snyder/Reuters

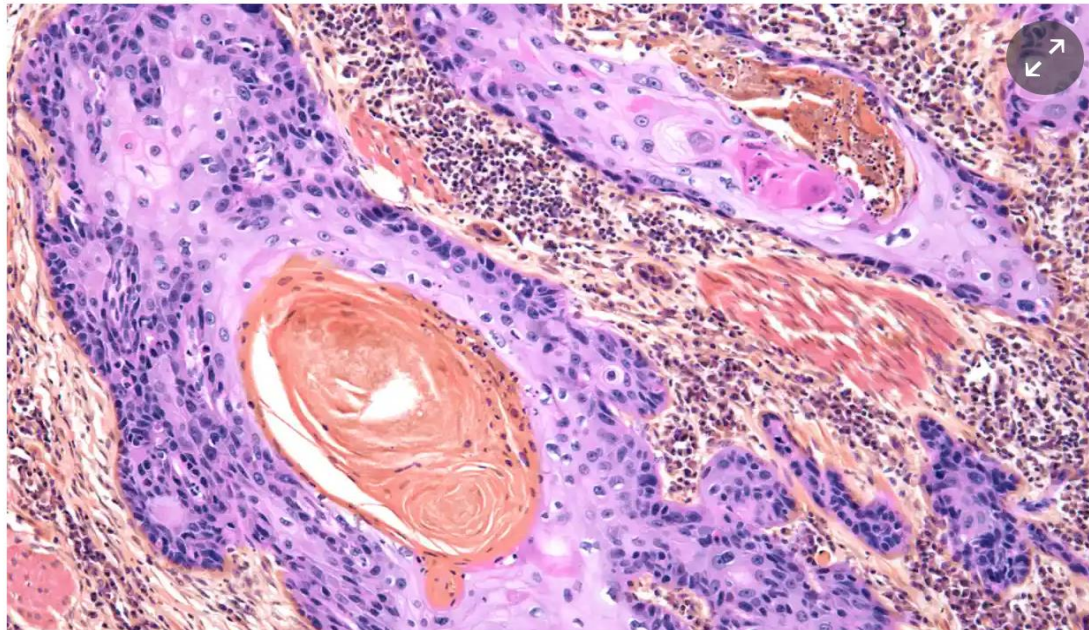
Amazon's machine-learning specialists uncovered a big problem: their new recruiting engine did not like women.

show men highly paid executive jobs, and the criminal recidivism model had racial bias. But what about the day when the online lending platform uses big data to determine a person's credit score and systematically rejects more loan applications from women or racial minorities than white men?



AI cancer detectors

Researchers suggest artificial intelligence is now better and faster at detecting cancer than clinicians



▲ Squamous carcinoma. An AI system distinguishes dangerous skin lesions from benign ones. Photograph: Steve Gschmeissner/Getty Images/Science P

More accurate than dermatologists

95% correct predictions to
87.5% by dermatologists

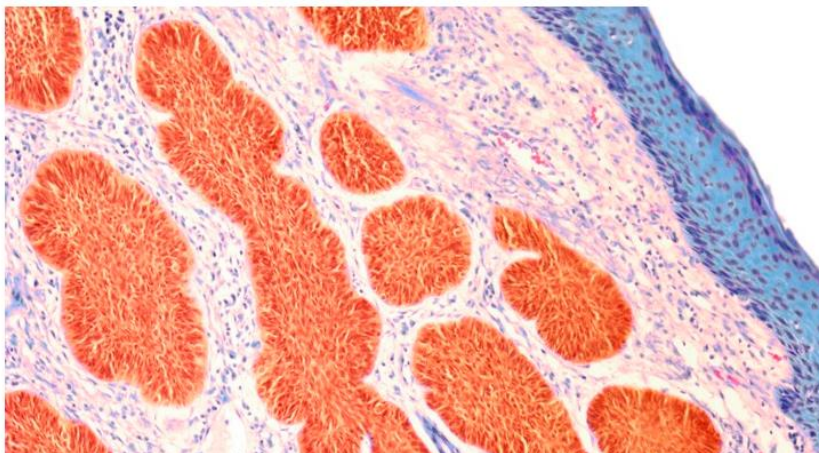
<https://www.theguardian.com/technology/2018/jun/10/artificial-intelligence-cancer-detectors-the-five>

HEALTH

AI-Driven Dermatology Could Leave Dark-Skinned Patients Behind

Machine learning has the potential to save thousands of people from skin cancer each year—while putting others at greater risk.

ANGELA LASHBROOK AUG 16, 2018



SECURE RX
SUPPLY WITH
BLOCKCHAIN

<https://www.theatlantic.com/health/archive/2018/08/machine-learning-dermatology-skin-color/567619/>

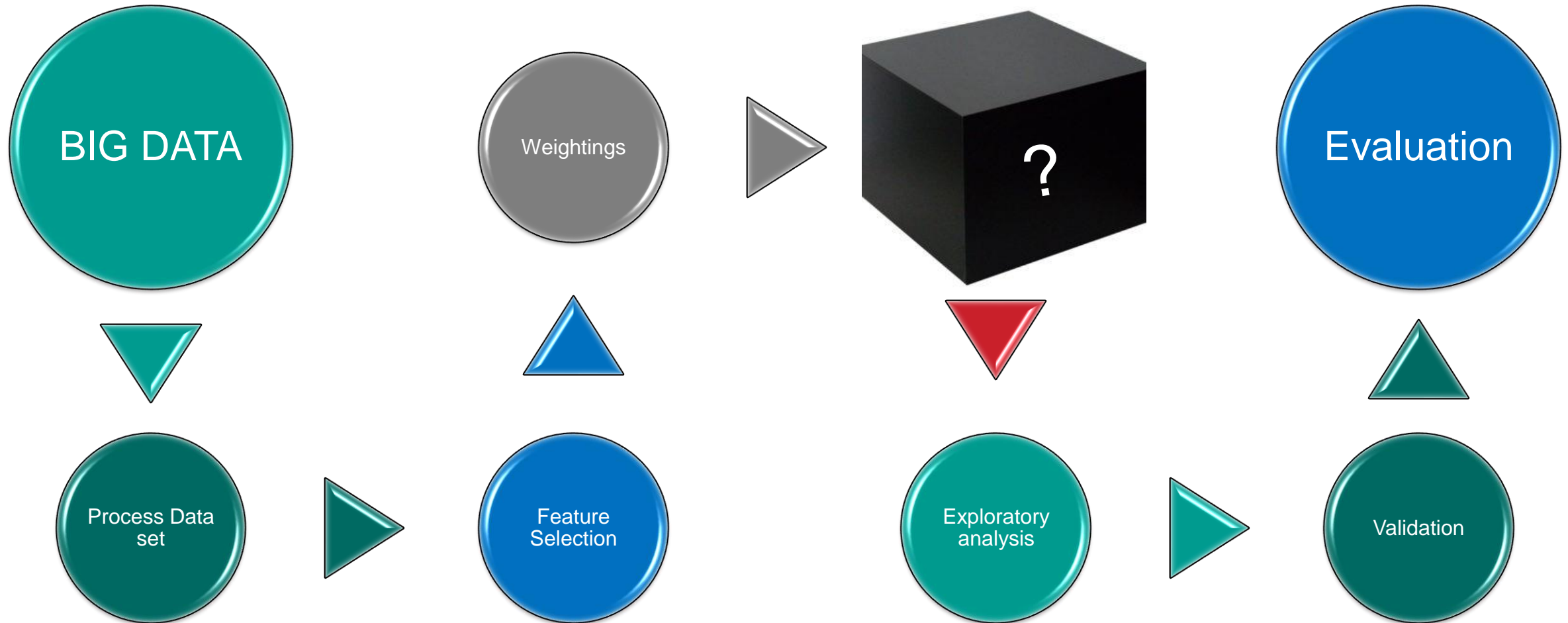
- Buolamwini and Gebru looked at three sex-recognition systems (IBM, Microsoft and Face++).
- Tested 1,270 photographs of international parliamentarians
- found that all classified lighter faces more accurately than darker ones.
- and classified males more accurately than females.
- **Eg. IBM's algorithm error rate:**
light male faces wrong - 0.3%
dark female faces - 34.7%



*Joy Buolamwini, Timnit Gebru ;
Proceedings of the 1st Conference on Fairness, Accountability and
Transparency, PMLR 81:77-91, 2018. available at
<http://proceedings.mlr.press/v81/buolamwini18a.html>*

**Some datasets are as much as
70-80% white and 70-80% male**

<https://www.perpetuallineup.org/>



- **C**orrectional **O**ffender **M**anagement **P**rofilng for **A**lternative **S**anctions
- Even when controlling for features such as prior crimes, future recidivism, age and gender black defendants were misclassified as higher risk at almost twice the rate of white defendants
- NOTE: It is disputed and it **depends on the definition of fairness** you use



Dressel and Farid (2018) have shown that:

- COMPAS provides **no greater accuracy than a simple linear classifier with two features**, age and previous convictions (COMPAS has 137 features).
- provided **no better prediction accuracy than a group of untrained people**.
- In all cases the accuracy level of COMPAS, the logistic regression classifier and the untrained individuals averaged at approximately **65%**.
- Also found **that even with race removed** COMPAS still showed bias:

	Incorrectly predicted to reoffend %	Incorrectly predicted to not reoffend %
Black	44.9	28
White	23.5	47.7

<http://advances.sciencemag.org/content/4/1/eaao5580.full>

- Explainability is IMPORTANT so don't accept the black box argument (AI can cheat) eg.
- An AI designed to detect hip fractures was using the imaging device model and the patient's age to make its prediction, rather than focusing on the fracture itself.
- The model was a good proxy for how sick someone was, because it could reveal if the scanner used was a portable version and therefore if the person was too sick to visit the hospital.



Lack of diversity and inclusivity at all stages of development : from identification of the problem to implementation and evaluation of the system results in **unconscious bias**.

This results in **machine bias** and produces **discrimination**.

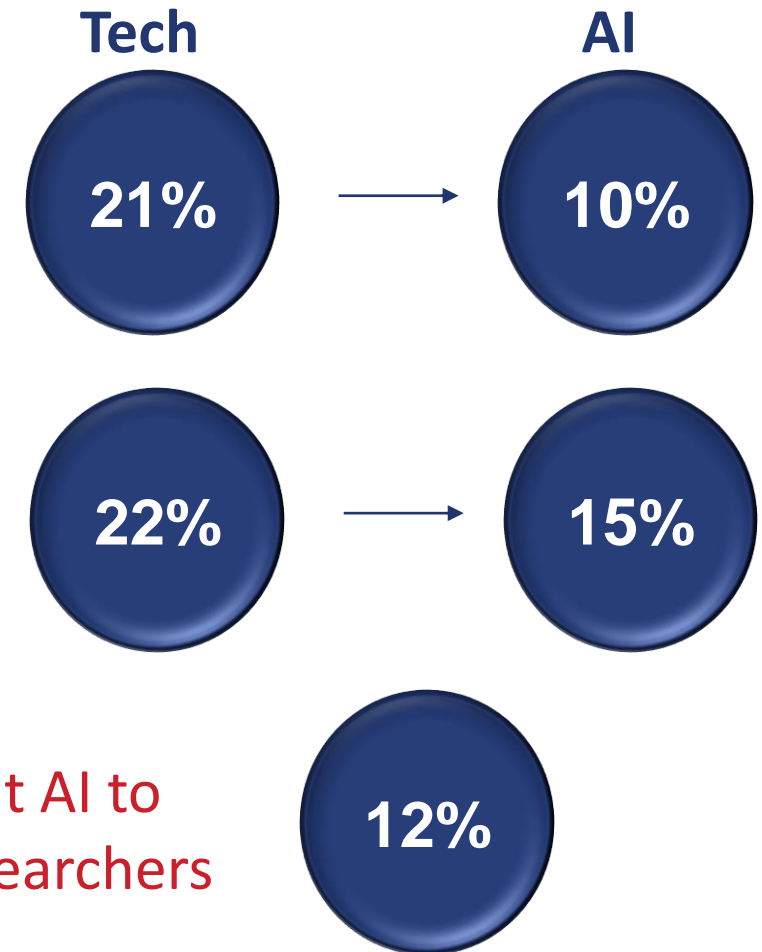
This risks exacerbating societal biases and **embedding inequality in our systems**



- **Google:** Google's AI research pages earlier this month, they listed 641 people working on "machine intelligence"

Facebook : Estimate of tech workers from reports and AI workers from pages for the company's AI research group listed 115.

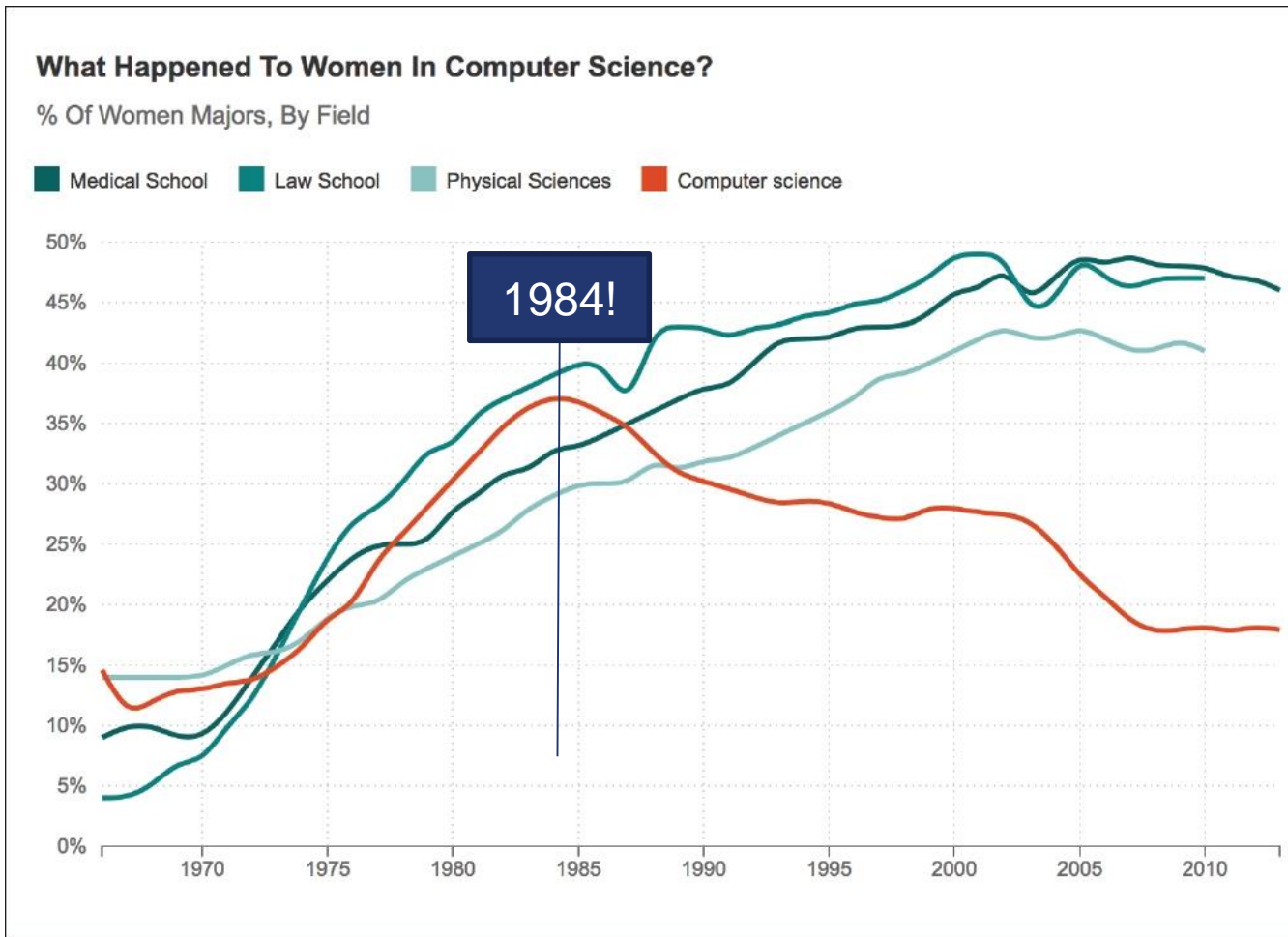
Overall: WIRED worked with Montreal startup Element AI to estimate the diversity of leading machine learning researchers from 3 top AI 2017 conferences:



The Gender Imbalance in AI Research Across 23 Countries

ELEMENT^{AI}

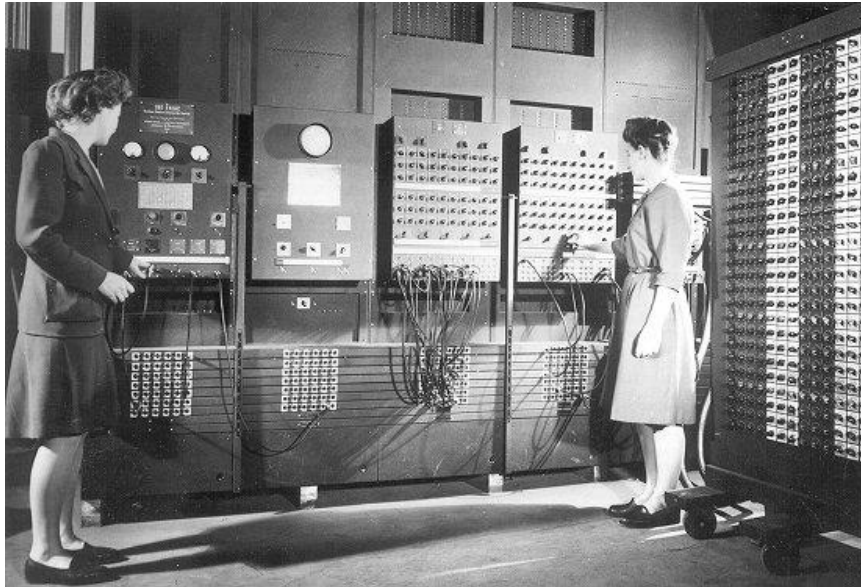
*Among 4000 researchers who have been published at the leading conferences NIPS, ICML or ICLR in 2017



NOTE: This is country dependent:

For the year 2011, women in the U.S. made up only 18% of undergrads in Computer Science and Engineering, whilst in India that number was 42%.

In 2005 in India, looking at CS alone, women earned 55% of the Bachelor's of Science degrees.

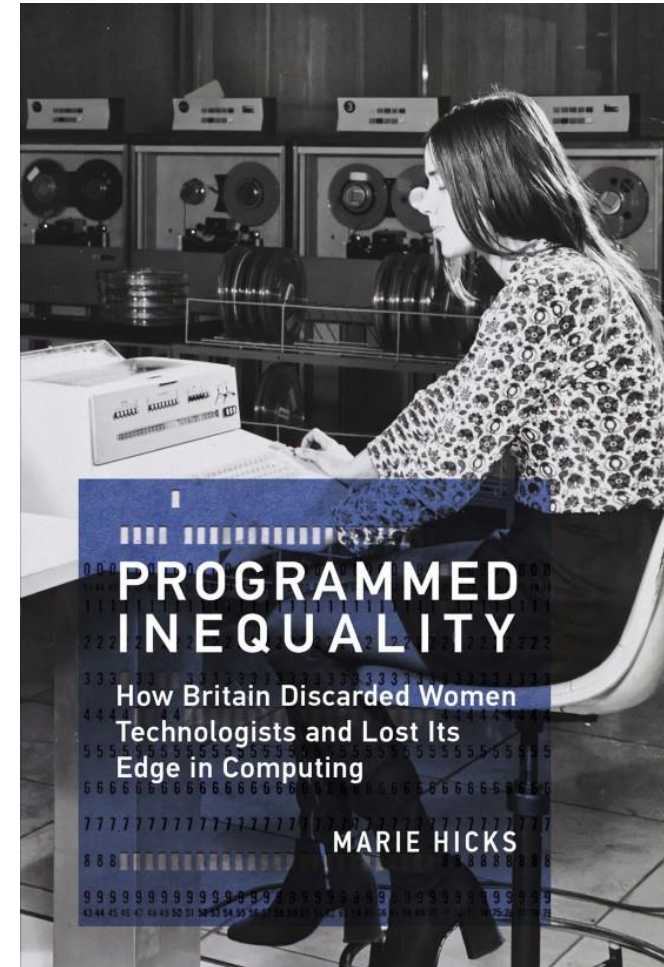


Programmers of ENIAC first electronic computer



- Programming was viewed as a feminized job and classed as a lower clerical grade (machine grade)
- In the 1950 - 1980's proportion of programmers who were female varied between 30-50%.

- **1970s and early 1980s:**
The Silicon Valley Gold Rush and UK 1960s White Heat policy
- Respect and need for programmers rose - and so did the pay.
- Men started becoming more interested in the field.
- To boost the field's prestige, men began creating professional associations, pushing educational requirements for programming careers, and discouraging the hiring of women.



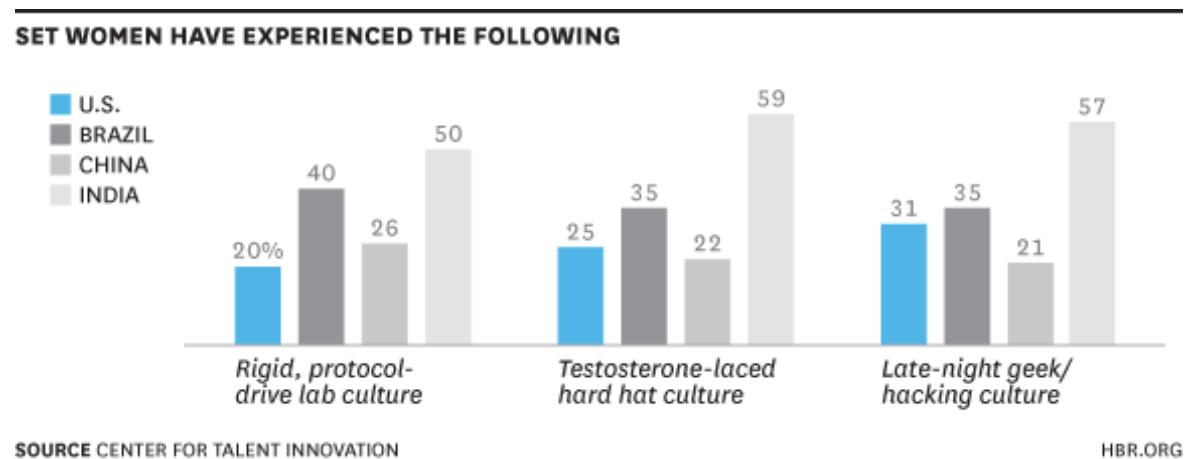
- **Aptitude tests** used to rank job candidates in programming played a role in masculinizing the field biased for male skills.
(Answers to the tests were circulated across male-only groups like fraternities)
- Companies used **personality profiles** to screen for applicants with anti-social tendencies, “a *stereotypically male trait*”, after two psychologists observed programmers shared the “striking characteristic” of “their disinterest in people.”



NES (third generation console, released in 1983).

<http://thewisdomdaily.com/women-used-to-dominate-tech-until-men-pushed-them-out/>

- Tech’s “geek workplace culture” compared to a super-competitive fraternity of arrogant nerds.



<https://hbr.org/2014/03/whats-holding-women-back-in-science-and-technology-industries>

Women in SET fields are 45% more likely than males to leave the industry within the year.

<https://www.talentinnovation.org/>



Push for Gender Equality in Tech? Some Men Say It's Gone Too Far

After revelations of harassment and bias in Silicon Valley, a backlash is growing against the women in tech movement.

<https://www.nytimes.com/2017/09/23/technology/silicon-valley-men-backlash-gender-scandals.html>

Change the culture

How code is taught in schools

How we recruit university students

How we recruit into jobs

Change the workplace culture

Change marketing and image of tech

Cut the b###sh#t

Regulate the algorithms:

Introduce regulation

GDPR

Algorithmic Impact Assessments

Algorithm Audits

Certification and an AI Fairness Stamp

KEY ELEMENTS OF A PUBLIC AGENCY ALGORITHMIC IMPACT ASSESSMENT

1. Agencies should conduct a self-assessment of existing and proposed automated decision systems, evaluating potential impacts on fairness, justice, bias, or other concerns across affected communities;
2. Agencies should develop meaningful external researcher review processes to discover, measure, or track impacts over time;
3. Agencies should provide notice to the public disclosing their definition of “automated decision system,” existing and proposed systems, and any related self-assessments and researcher review processes before the system has been acquired;
4. Agencies should solicit public comments to clarify concerns and answer outstanding questions; and
5. Governments should provide enhanced due process mechanisms for affected individuals or communities to challenge inadequate assessments or unfair, biased, or otherwise harmful system uses that agencies have failed to mitigate or correct.

<https://ainowinstitute.org/aiareport2018.pdf>

AIA

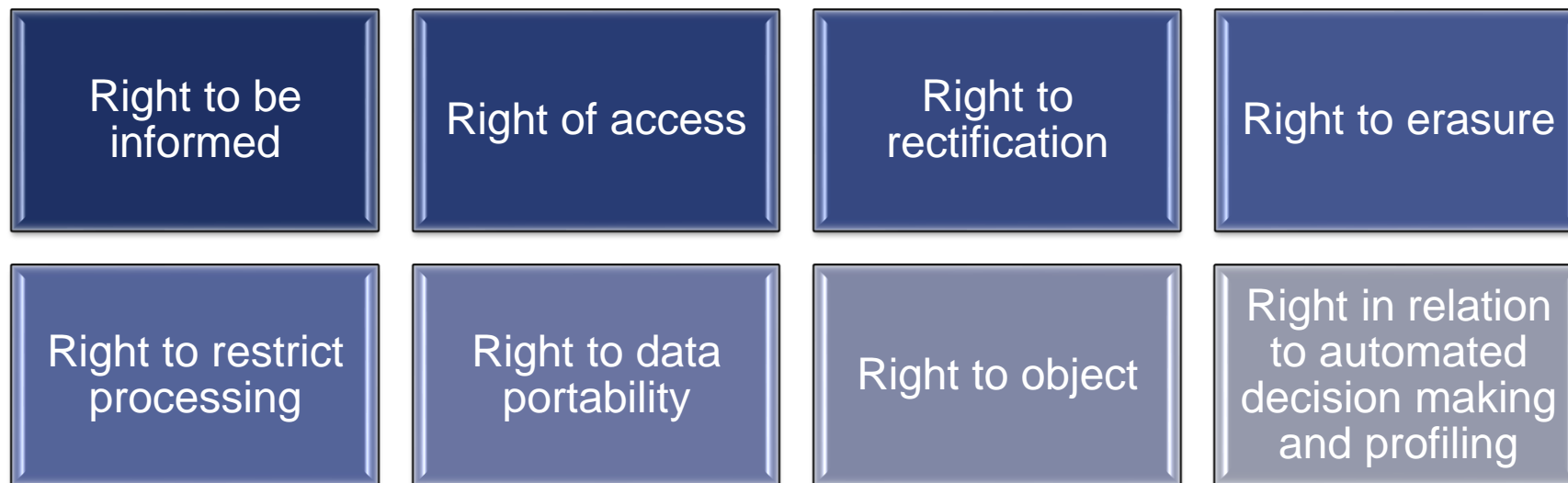
AI auditing

Certification/trustmarking

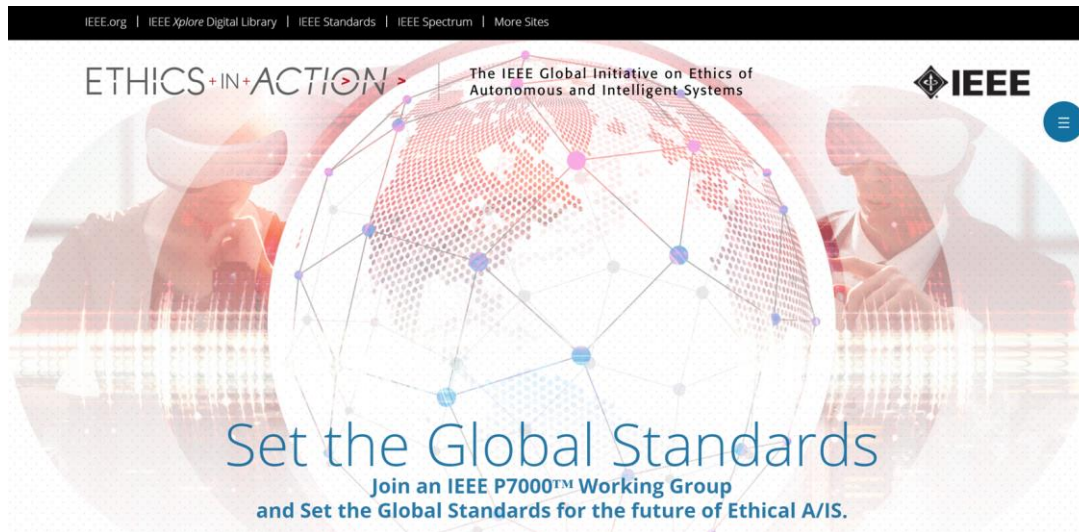
Citizen panels

Regulation/legislation

- Article 22: Right not to be subject to a solely automated decision
- Article 42: Certification
- Recital 71 Right to an explanation



Data Subject Rights



- Build in a requirement for diverse teams and ethics panels for AI certification
- Fairness certification and trustmarks will give a competitive edge
- Once there is a profit incentive we will see a sustainable change



**WOMEN
LEADING
IN AI**

JUMP

Promoting gender equality, advancing the economy

26th March 2019 | JUMP Forum Brussels

Thank you for your attention



jump.brussels



#JUMP2019
@JumpEquality

WWW.JUMP.EU.COM