# Are *non-diverse* AI research and development teams risking bias in innovated algorithms and artefacts?

Dr Huma Shah

Director of Science, EU Horizon2020 CSI-COP research and innovation project

Associate Member, Data Science Research Centre

Assistant Professor 'Trust in AI', School of Computing, Electronics and Mathematics

Faculty of Engineering, Environment and Computing

AI4EU Cafe: 5 February 2020

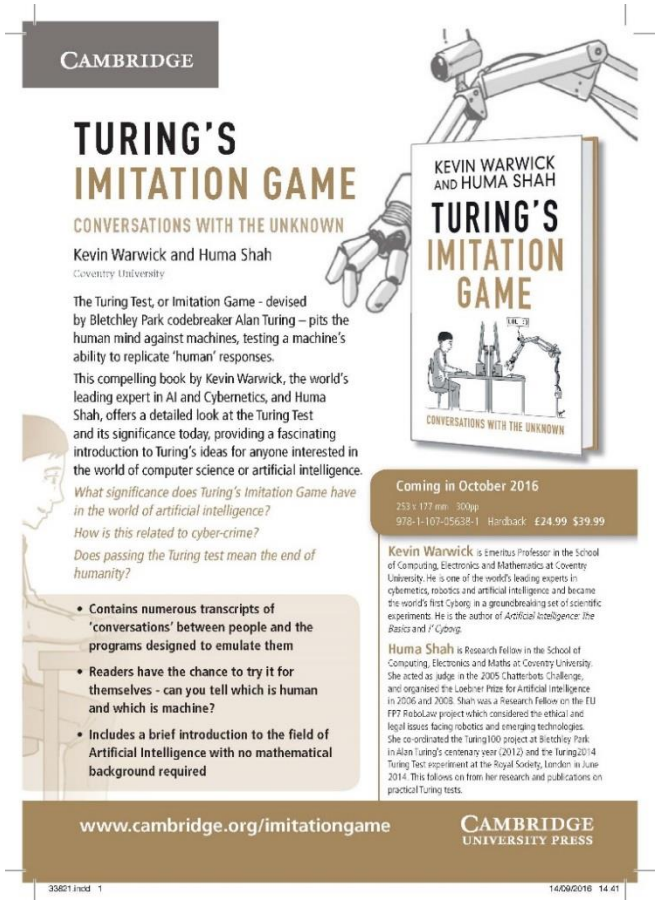
# My Activities include….

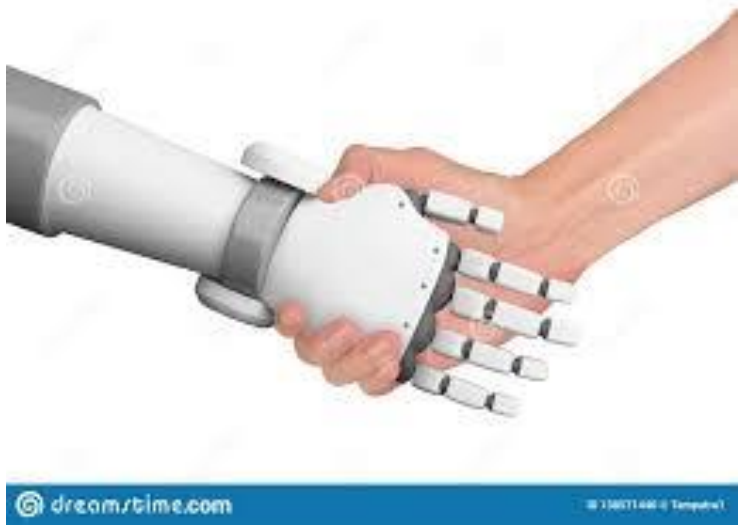
- **Science Director** (CoI), **EU** Horizon2020 **SwafS15**-2019 ***CSI-COP***. Coventry University coordinating 30month €1.99m, 11-partner international consortium. CORDIS: https://cordis.europa.eu/project/id/873169
- **Member of the EU AI Alliance**
- Member of the **European Platform for Women Scientists** (EPWS)
- West London **STEM Ambassador**

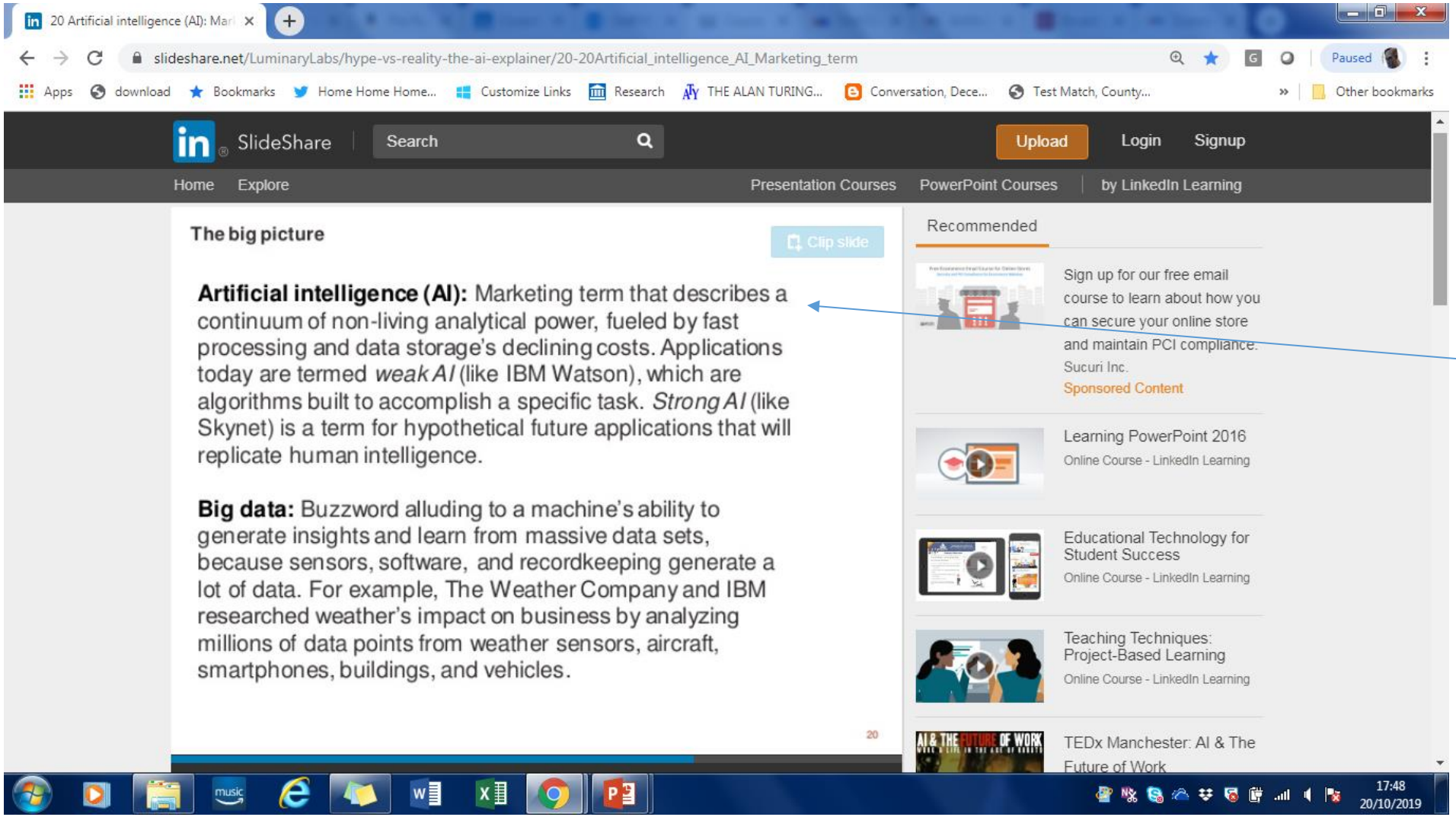
**Book:**

**published by**

**Cambridge University Press, 2016**




dreamstime.com




AI is not a marketing term!!!
It is a scientific field

https://www.slideshare.net/LuminaryLabs/hype-vs-reality-the-ai-explainer/20-20Artificial_intelligence_AI_Marketing_term

Promising magic by "unlocking data".

https://twitter.com/msuvarna/status/1176936163581386752

Clever AI?

Ad on Metropolitan Line tube,
Thursday October 17, 2019

# Artificial Intelligence meets Natural Stupidity

Article by **Drew McDermott**

published in

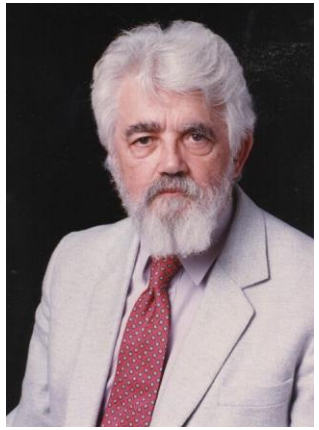
*ACM SIGART Bulletin* Issue 57, April 1976: pages 4-9

# Seriously though….

- The *I* in *AI* represents intelligence

# Origins of Artificial Intelligence (AI)

- Credit where it's due! AI coined by computer scientist **John McCarthy** in **1955 -** field emerged from a scientific gathering in 1956 at a now famous Dartmouth Conference in the US.



- Followed on from **Alan Turing'**s scholarship on machine intelligence during and after WWII: ***imitation game***

# Alan Turing

- Turing's 1948 essay *Intelligent Machinery*:
  - "…first manifesto of artificial intelligence"

Copeland, B. J. (2004).

*The Essential Turing: The Ideas That Gave Birth to the Computer Age.*

Clarendon Press: Oxford, UK. p. 486

# Most visual representation of an ‘artificial intellect’ would be an embodied, *thinking* robot

# Seeking Artificial Intelligence

KYOTO – A 400-year-old temple is attempting to hot-wire interest in Buddhism with a robotic priest it believes will change the face of the religion

"This robot will never die; it will just keep updating itself and evolving,"

"**With AI we hope it [robot] will grow in wisdom to help people overcome even the most difficult troubles. It's changing Buddhism**"

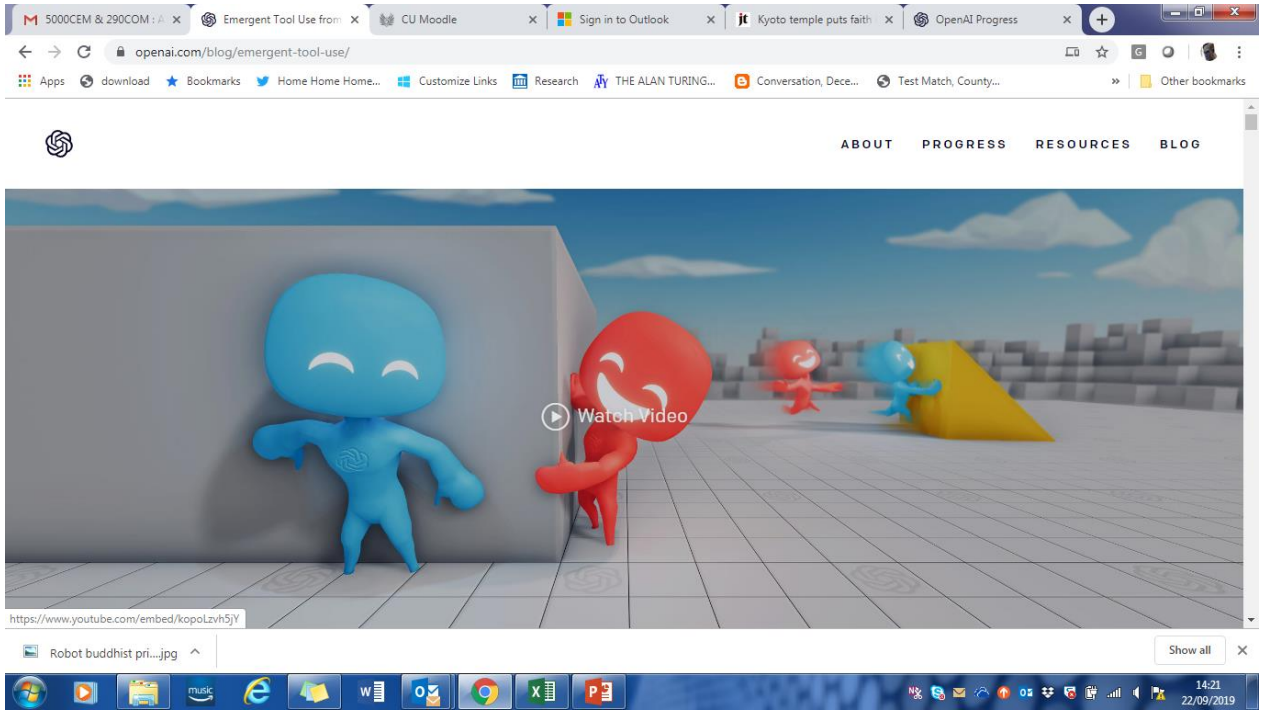
JAPAN TIMES: https://www.japantimes.co.jp/news/2019/08/15/business/tech/kyoto-temple-robot-priest/#.XYdvI1VKjIU

# Smart machines and algorithms

**IBM Watson 2011**

**Google DeepMind's AlphaGo 2016**

# AI ‘Playing hide and seek’



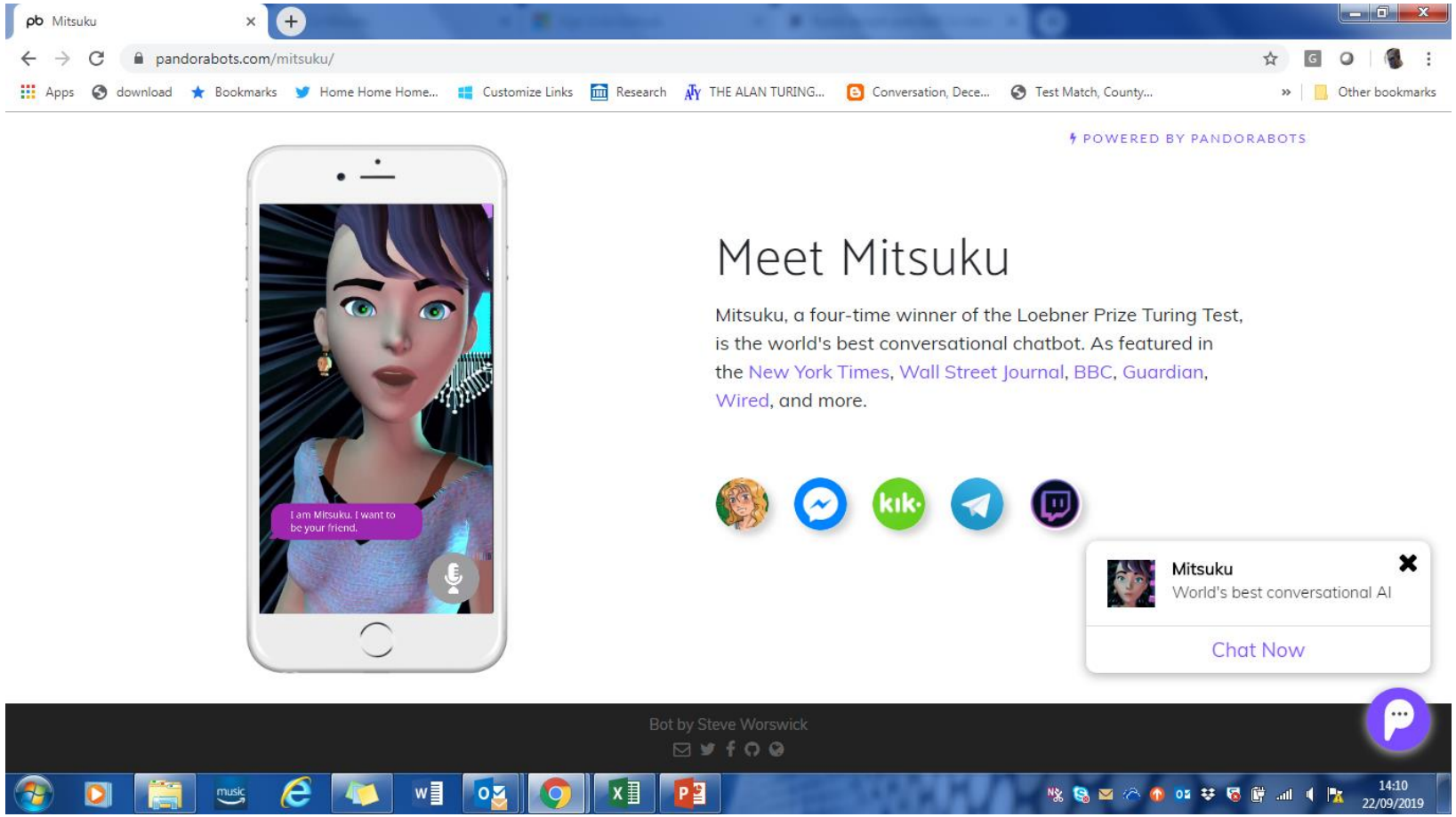

## OpenAI’s project

***Emergent Tool Use from Multi-Agent Interaction (2019)***

Agents observed “**progressively more complex tool use** while playing a simple game of hide-and-seek. …in new simulated hide-and-seek environment, **agents build a series of six distinct strategies and counterstrategies** …**some of which we did not know our environment supported**.”

https://openai.com/blog/emergent-tool-use/

# Four Loebner Prize-winning Chatbot




Steve Worswick's **Mitsuku** chatbot wins
2019 Loebner Prize

AISB: https://aisb.org.uk/new_site/?p=350

# But what is *Intelligence*?

- Turing (1948):
  - "..idea of intelligence is emotional, rather than mathematical"

# Turing's view

- Turing (1948):
  - **"The extent to which we regard something as behaving in an intelligent manner is determined as much by our own state of mind and training as by the properties of the object under consideration"**

# AI: Simulating the human mind

- "Programming a machine to think….. process should bear a close relation to that of *teaching*"

Turing, 1951 Lecture **Can Digital Computers Think?** in

Copeland, B. J. (2004). *The Essential Turing: The Ideas That Gave Birth to the Computer Age.*

Clarendon Press: Oxford, UK. p. 486

# As Turing said……

- ….. the pursuit of building thinking machines would:

  "help us greatly in finding out how we think ourselves"

Turing, 1951 Lecture **Can Digital Computers Think?** in

Copeland, B. J. (2004). *The Essential Turing: The Ideas That Gave Birth to the Computer Age.*

Clarendon Press: Oxford, UK. p. 486

# Intelligence testing and scientific method

- H.J Eysenck (1985):
  - "Intelligence tests are not based on any very sound scientific principles"
  - "..**not a great deal of agreement among experts regarding the nature of intelligence**"

(p. 8 in 1985 reprint of book ***Know your own IQ***, Penguin books).

# Information processing view

- Joseph **Fagan**: information processing as an “access to opportunity” to enter the world of knowledge.

# Intelligence - Fagan

- Fagan (2000: p. 168):
  - "… consider the implications for society of defining intelligence as processing …controversy surrounding the term *intelligence* has arisen and continues because **intelligence has historically been defined as *how much one knows* rather than how well one processes**.

# Intelligence - Fagan

- Fagan (2000: p. 168):
  - **“IQ scores, by convention, are based on how much one knows relative to one's age peers.”**

# Information processing & knowledge

- **Fagan:**
  - "…**intelligence is processing** and that processing can be **measured by performance on certain elementary cognitive tasks**."
  - "An **IQ score**, on the other hand, **depends** not only on processing ability but **on *what one has been taught***."

# Information processing & knowledge

- **Fagan:**
  - **"…information is processed, the mind changes."**
  - **"That change is called *knowledge*".**

# How do we convey our knowledge?

# Communication

- "Language is a set of vocal or written signs and symbols that permits a social group to communicate, and facilitates thinking and actions of individuals" (Fischler and Firschein, 1987: p. 157)

# Communication contd.

- “Language, in the full sense of the term, is species-specific to [hu]man…”

# Turing on language

- Turing's 1948 essay *Intelligent Machinery*:
  - "… learning of languages an impressive human activity"

# Turing test for intelligent machinery

- Turing explored intelligent machinery through ability of a machine ***to think***

# Turing test for intelligent machinery

- Turing: Thinking is *buzzing in the head*

# Avoid definitions

- Turing did not define intelligence or machine intelligence (please see Plato's *Laches* dialogue, or read W.V.O. Quine)

# Avoid definitions

- Turing proposed practical way to assess intelligence through Imitation Game, popularly known as the Turing test.

# Learning - Imitation

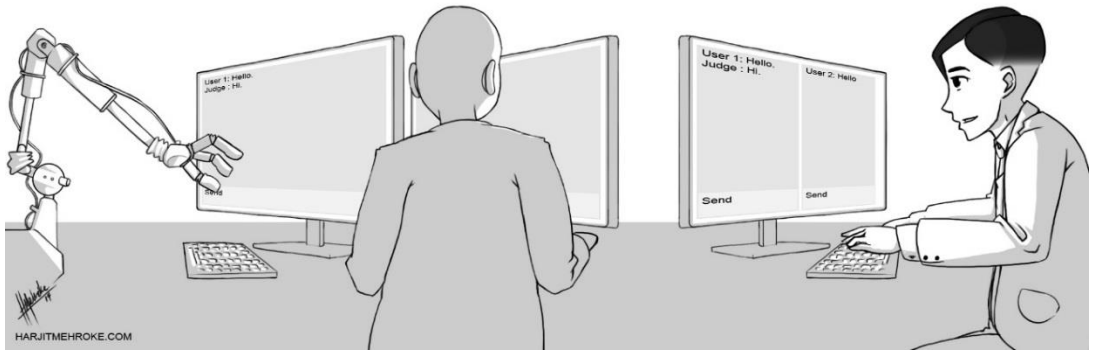
- Learning by imitation – look at how children learn

# Learning - Imitation

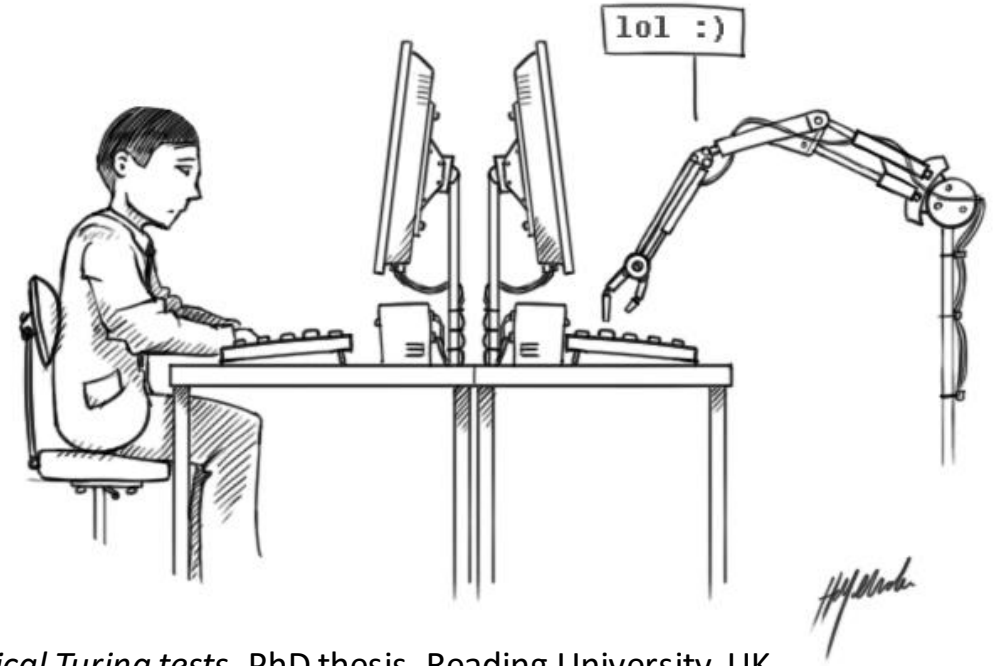
- Turing test measures a machine's intellectual capacity/ to answer any question in a "satisfactory" and "sustained" way (1950) by imitating the sorts of answers a human would give to any questions

# Turing's two tests for machine thinking

**Simultaneous Comparison**




**Viva Voce**




Shah, H. 2011. *Deception-detection and machine intelligence in Practical Turing tests.* PhD thesis, Reading University, UK
PDF available for download from **Academia.edu** and **ResearchGate**

# Current state of the Art: Turing test1

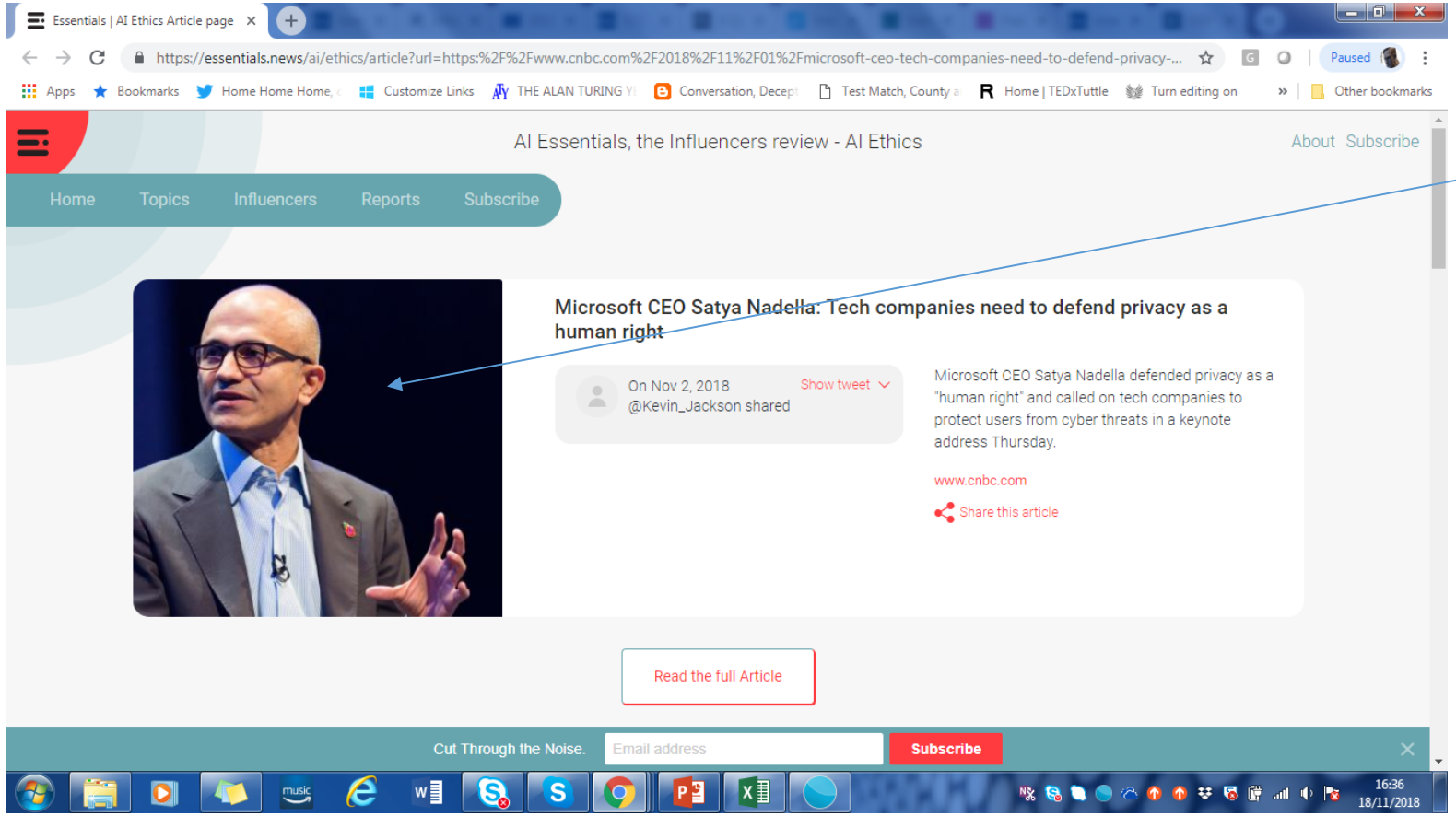
# Turing test2

# Issues in Nascent AI technologies

# Privacy

- **Privacy**:
  - **data protection concerns** arising from artificial intelligence 'approaches' (machine learning applied to big data sets)

# Privacy is a human right




**Microsoft CEO**
Satya Nadella:
- "Tech companies need to defend privacy as a human right"

# Privacy as a fundamental right

- "**Privacy** is a fundamental right:
  - **Essential to autonomy**
  - **Protection of human dignity**
  - Serving as the foundation upon which many other human rights are built."

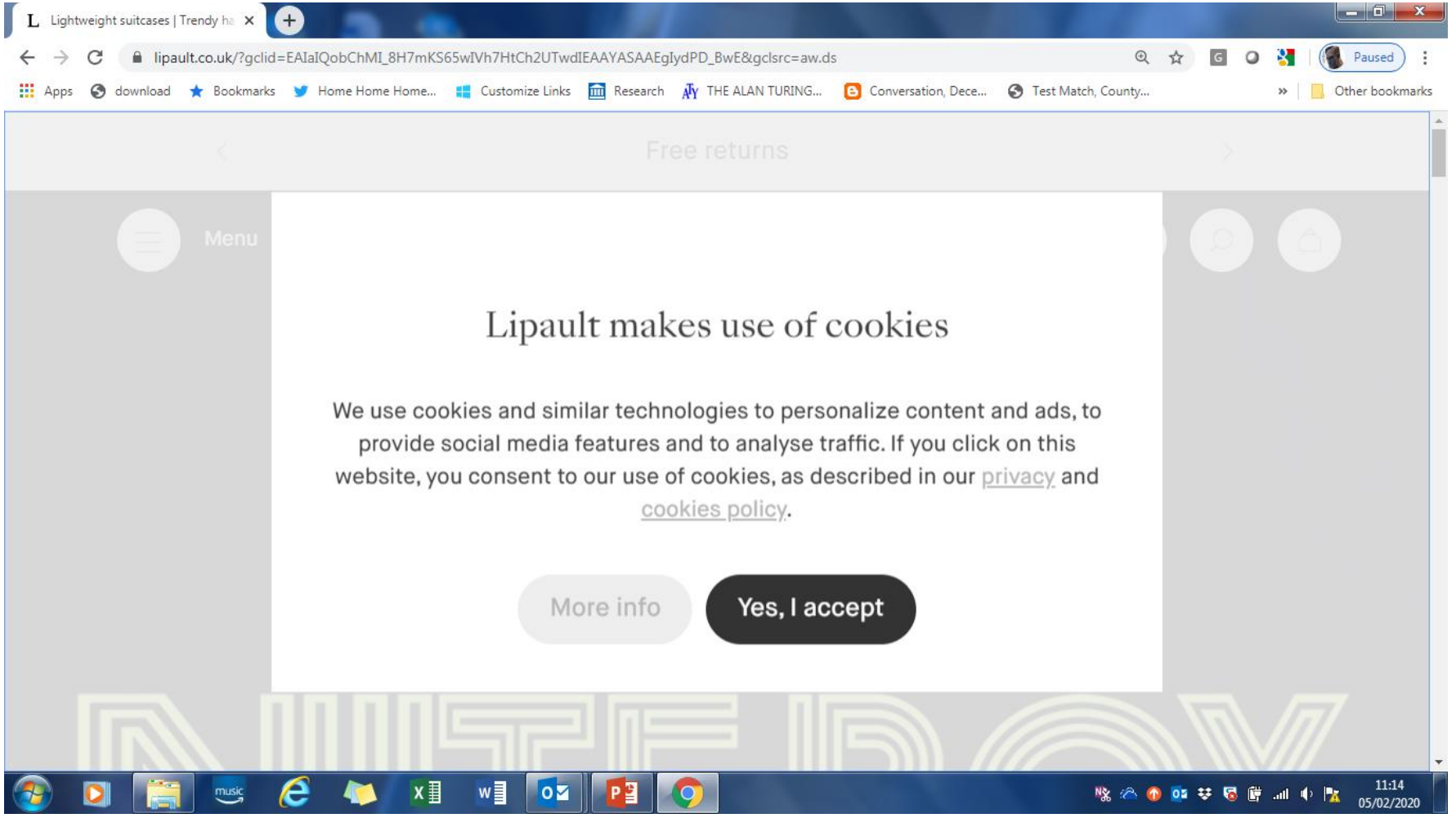
https://privacyinternational.org/explainer/56/what-privacy

# Sharing personal data

**A. Weigend**. **2017**. ***Data for the People: How to Make Our Post-Privacy Economy Work for You***. Basic Books: New York, USA (pages 2-3):

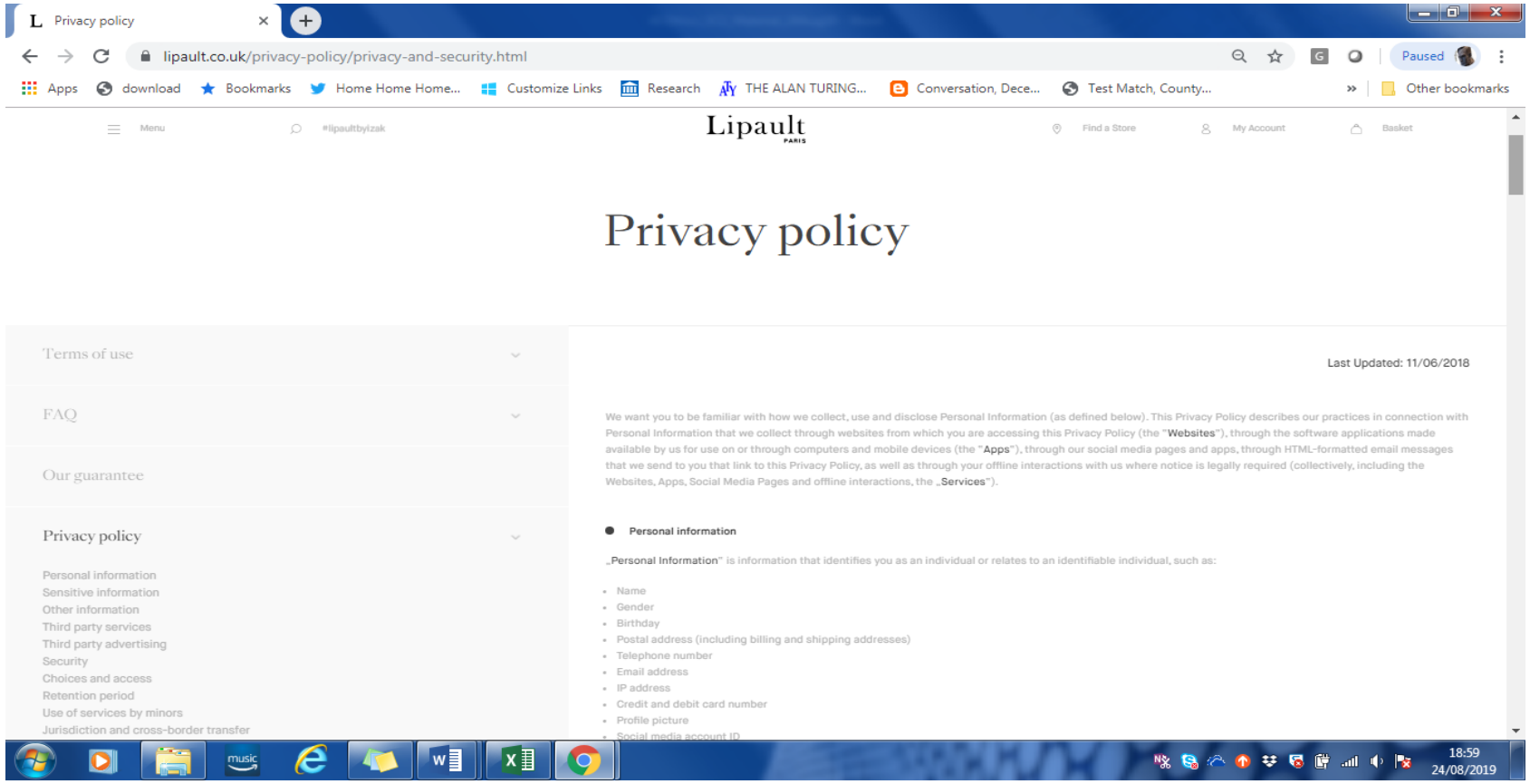
- "Every day more than a billion people create and share social data [online]…."
- "Some of these data are shared knowingly and willingly…"
- **"..part and parcel of the convenience of using the Internet and mobile devices**."

# Tracked without *willing* **consent**




**Lipault** luggage

# Lipault's Privacy Policy




AI4EU Cafe: 5 February 2020

# Online trading and collecting data

- **Lipault**'s online privacy information reveals: "We and our service providers **collect Personal Information** in a variety of ways"

- Lipault defines **personal information**: "..information *that identifies you as an individual or relates to an identifiable individual*, such as:
  - Name
  - Gender
  - Birthday

# Lipault's data sharing

- Lipault **personal information** sharing "..information *that identifies you as an individual or relates to an identifiable individual*, such as:
  - Postal address (including billing and shipping addresses)
  - Telephone number
  - Email address
  - IP address
  - Credit and debit card number
  - Profile picture
  - Social media account ID
  - Travel information, including trip flight information

# Websites and Cookies

- Cookies:
  - **"… are small text files that websites place on your device as you are browsing"**
  - **"… can store a wealth of data, enough to potentially identify you without your consent."**
  - **"…primary tool that advertisers use to track your online activity so that they can target you with highly specific ads"**

From here: https://gdpr.eu/cookies/

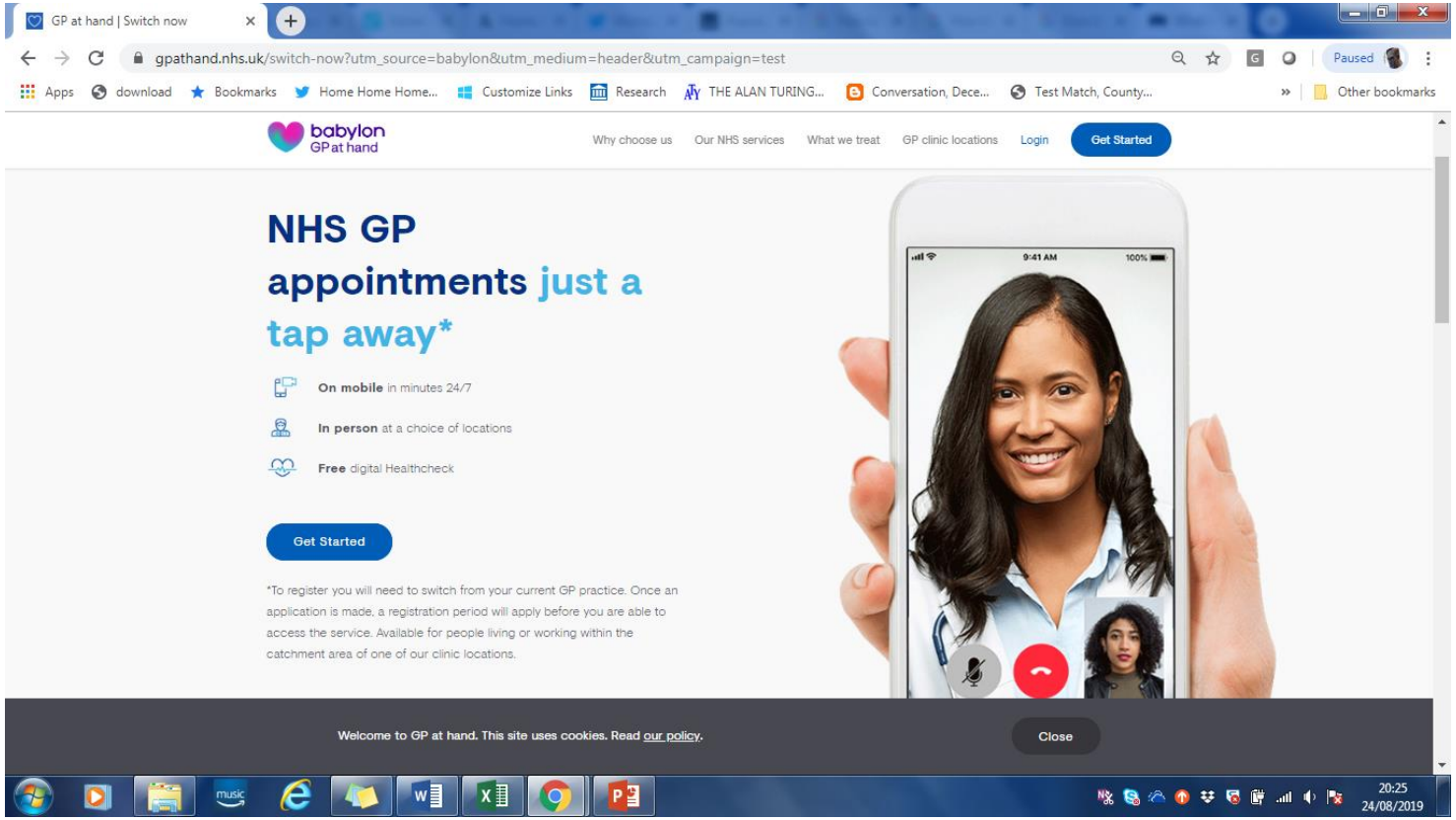
# Third-party access to personal information

- **Lipault**'s online practice of sharing your personal information with third-parties: "**We disclose Personal Information**":
  - "To our affiliates for the purposes described in this Privacy Policy."
    - You can consult the list and location of our affiliates at www.aboutbags.com . Affiliates as used in this document means the entities controlled by, under common control with or controlling us.

# Third-party access to Lipault customers' personal information

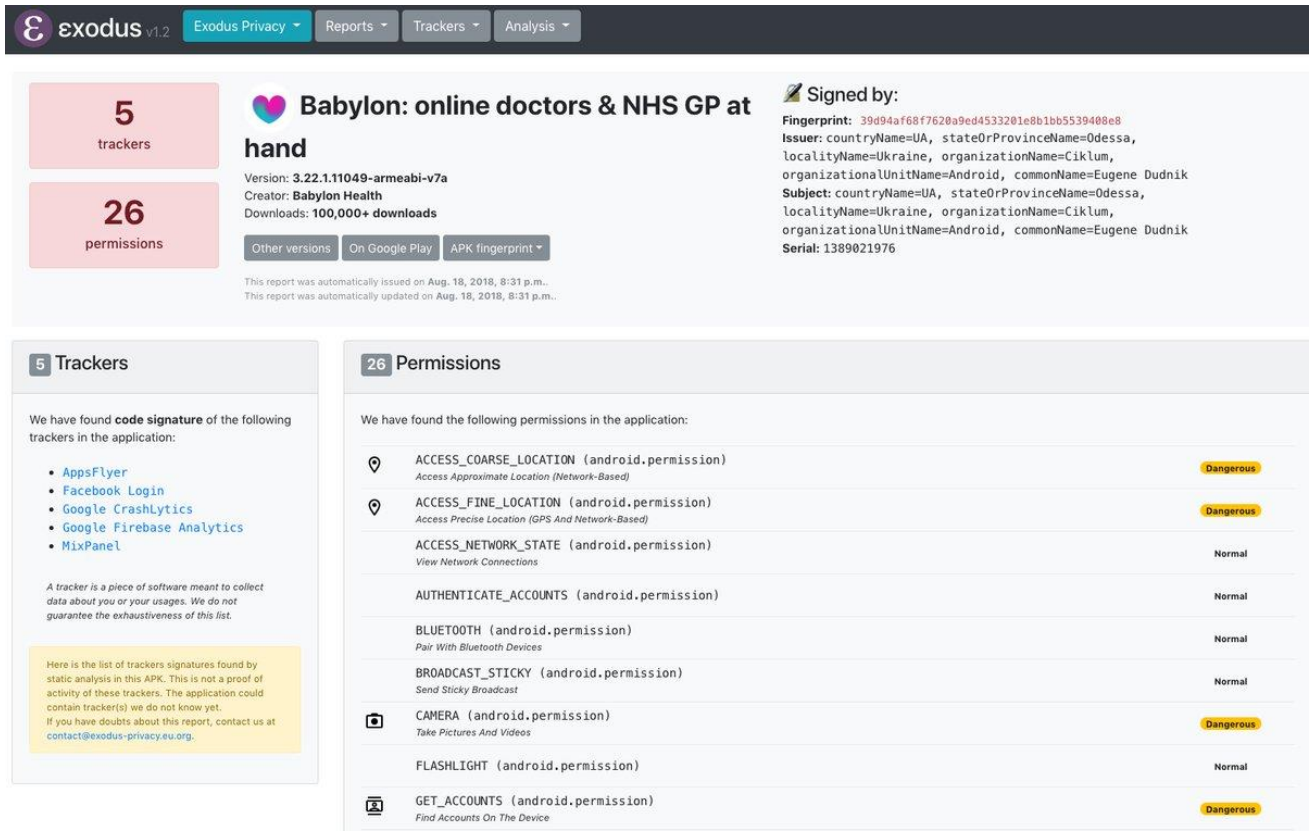
- "To our third party service providers, to facilitate services they provide to us [Lipault]".
  - These can include providers of services such as website hosting, data analysis, payment processing, order fulfilment, information technology and related infrastructure provision, customer service, email delivery, auditing, and other services.
- "To third parties, to permit them to send you marketing communications, consistent with your choices".
- "To third-party sponsors of sweepstakes, contests, and similar promotions"

# Privacy and Smart Phone Apps




***Babylon*** health app:
https://www.babylonhealth.com/

# Embedded Trackers in Smart Phone Apps




**Babylon** health **app** has these **embedded trackers**:

**Facebook Login**
AppsFyer
Mixpanel
Google Firebase
Google Crashlytics

# Sensitive Personal Information and Privacy

- Menstruation data is being monetised:
  - "Out of **36 apps** tested by ***Privacy International** **61% automatically transfer data to Facebook the moment a user opens the app**."
  - "This **happens whether the user has a Facebook account or not, and whether they are logged into Facebook or not**"
  - "**Menstruation apps are not just concerned with …. menstruation cycles**"
  - **"Menstruapps – How to Turn Your Period Into Money (For Others)**, they collect information about your health, your sexual life, your mood and more .."

*** https://www.privacyinternational.org/long-read/3196/no-bodys-business-mine-how-menstruation-apps-are-sharing-your-data**

# General Data Protection Regulations (*GDPR*)

- **EU GDPR:**
  - "Stronger rules on data protection mean **people have more control over their personal data and businesses benefit from a level playing field.**" https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en

# GDPR and consent

- From the UK Information Commissioner Office (ICO):
  - The GDPR sets a high standard for **consent**.
  - Consent means offering individuals real choice and control.
  - Genuine consent should put individuals in charge, build trust and engagement, and enhance your reputation
  - Consent requires a positive opt-in.
  - Explicit consent requires a very clear and specific statement of consent.

From here:

https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/consent/

# General Data Protection Regulations (*GDPR*)

- "[UK] Information commissioner Elizabeth Denham's **investigation into the [mis]use of data in elections and campaigns, finds:**
  - "**a disturbing disregard for voters' personal privacy**" https://ico.org.uk/media/action-weve-taken/2260271/investigation-into-the-use-of-data-analytics-in-political-campaigns-final-20181105.pdf

# General Data Protection Regulations (*GDPR*)

- In **October 2018**, UK Information Commissioner's Office (**ICO**) **fined Facebook £500,000**, the maximum under applicable data protection laws, for its **part in passing data on 87 million users to Cambridge Analytica**, which used it **during the 2016 US presidential campaign**.

https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2018/10/facebook-issued-with-maximum-500-000-fine/

# General Data Protection Regulations (*GDPR*)

- In **July 2019** UK's ICO announced **intent to fine British Airways £183.39m** for infringement of the GDPR following British Airways' data breach:
  - "involved user traffic to the British Airways website being diverted to a fraudulent site. Through this false site, customer details were harvested by the attackers. Personal data of approximately 500,000 customers were compromised in this incident, which is believed to have begun in June 2018." https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2019/07/ico-announces-intention-to-fine-british-airways/

# BIAS

# Bias

- **Bias**:
  - how some algorithms discriminate against certain sections of society

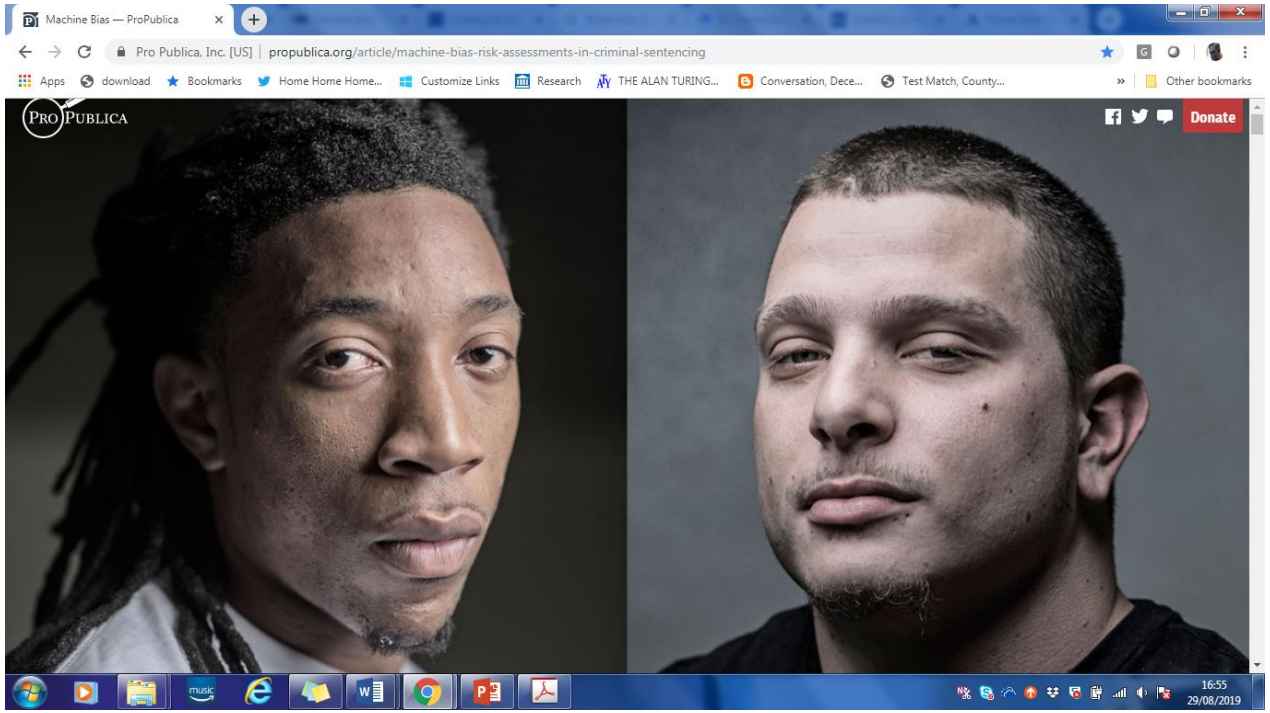
# Using 'AI' in recruitment: *Self-fulfilling prophecy*?

- Biased algorithms – Amazon's Recruitment model
  - "Amazon's computer models were trained to vet applicants by observing patterns in resumes submitted to the company over a 10-year period. Most came from men, a reflection of male dominance across the tech industry."
  - "Amazon's system taught itself that male candidates were preferable. It penalized resumes that included the word "women's," as in "women's chess club captain." And it downgraded graduates of two all-women's colleges…"

From Jeffrey Dastin, 2018 in **Reuters**:

https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

# Bias in AI technologies



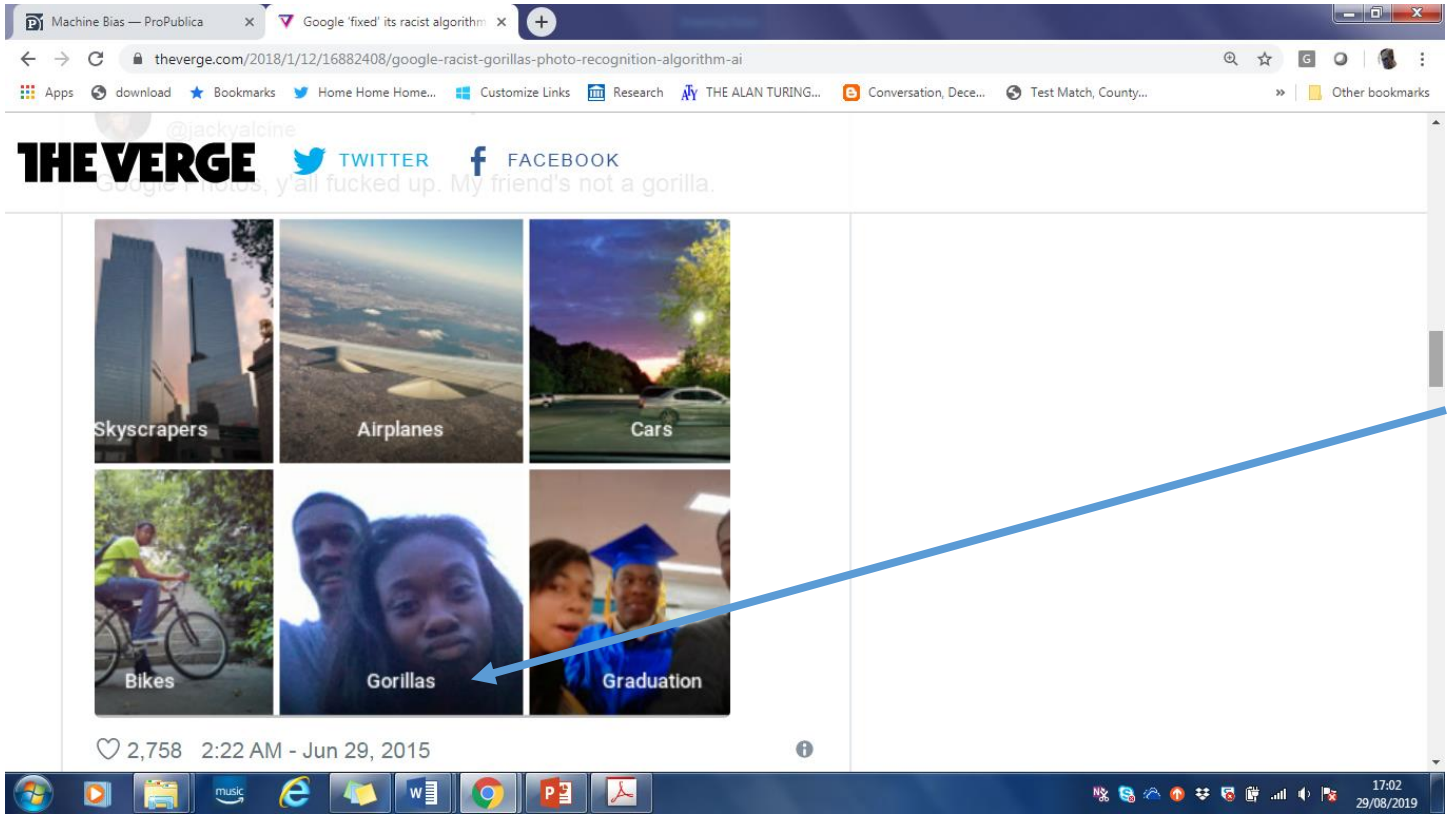

https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

**Anguin** et al., 2016:

-"There's **software** across the country [US] to **predict future criminals**. And it's biased against blacks."

-"…**the algorithm was somewhat more accurate than a coin flip**"

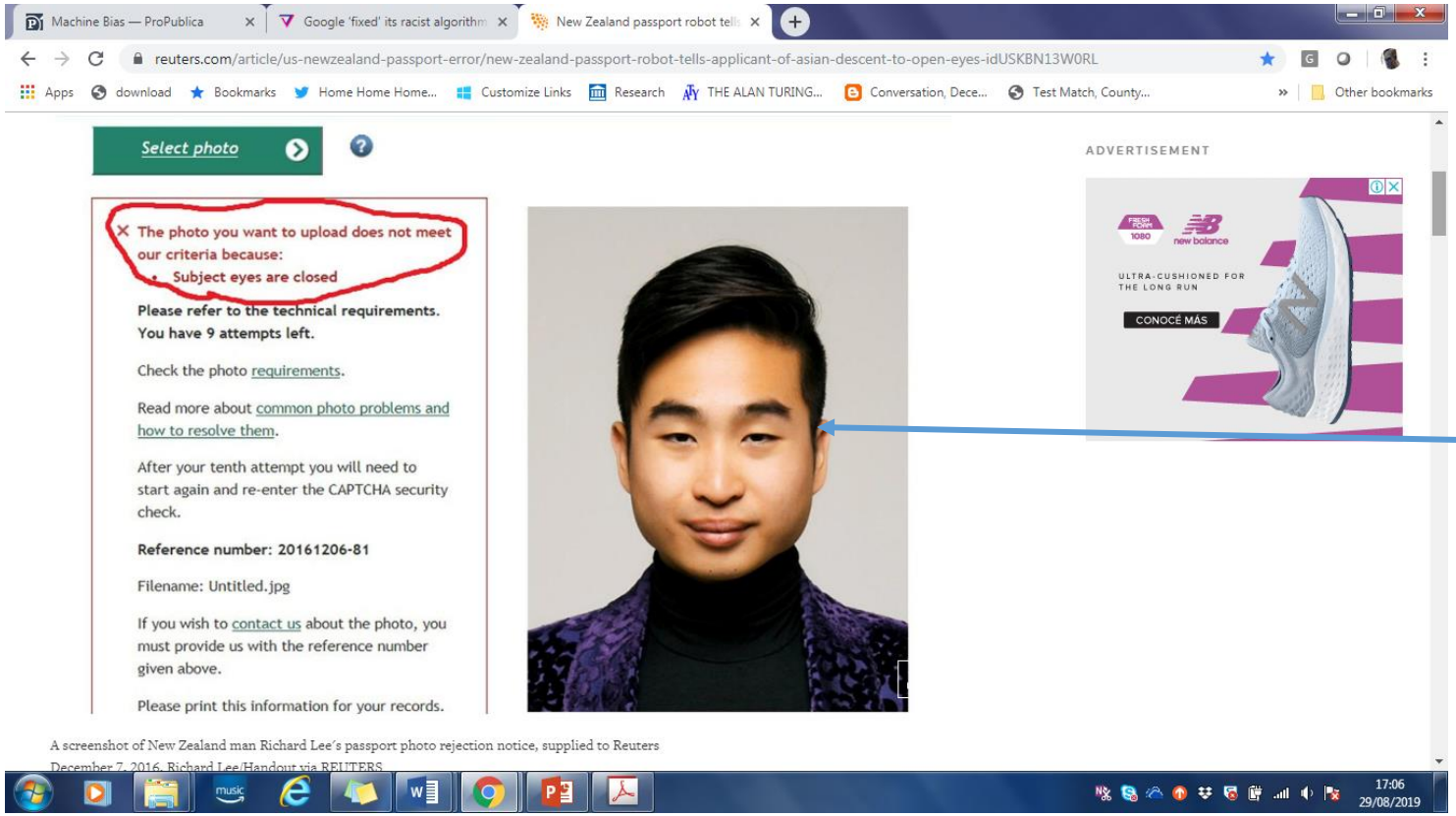
# Google's algorithm and bias




"Google 'fixed' its racist algorithm by **removing gorillas** from its **image-labeling tech**"

https://www.theverge.com/2018/1/12/16882408/google-racist-gorillas-photo-recognition-algorithm-ai

# New Zealand's passport renewal system




A screenshot of New Zealand man Richard Lee's passport photo rejection notice, supplied to Reuters
December 7, 2016. Richard Lee/Handout via REUTERS

"New Zealand passport robot tells applicant of Asian descent to open eyes"

https://www.reuters.com/article/us-newzealand-passport-error/new-zealand-passport-robot-tells-applicant-of-asian-descent-to-open-eyes-idUSKBN13W0RL

# Live Facial Recognition

- “Over the past few years six [UK] forces, including the Metropolitan police, have trialled facial recognition, with spectacularly poor results.”
- “After surveilling hundreds of thousands of people, the Met has made a mere handful of arrests using this technology in the past four years.”
- “Indeed, 81% of the “suspects” caught by the cameras were simply innocent bystanders not on the police’s watchlists. Two of the trials even had an error rate of 100%.”

David Davis is Conservative MP for Haltemprice and Howden:

https://www.theguardian.com/commentisfree/2019/sep/20/facial-recognition-technology-privacy

# Twitter @CatHallam1: April 6, 2019

- Permission granted October 21, 2019 by Cat Hallam to showcase her experience from April 2019:
  - "….thank you for asking in advance. Certainly, happy to see the information used in an educational context. Hope it helps the students & contextualises the subject matter."

**Cat Hallam**: “Today, I simply wanted to renew my passport online.”

### 3. Photo quality and format

Your photo must:

- be in colour, with no effects or filters
- not be blurred or have 'red eye'
- be unedited – you can't 'correct' your passport photo

**Our automated check suggests**

- it looks like your eyes are closed
- we can't find the outline of your head

**You can use the photo**

- if you have a medical or religious reason why it doesn't meet the rules
- if you think it meets all the rules

If your photo doesn't meet the rules you may have problems with your application or when you travel.

BETA This is a new service – your feedback will help us to improve it.

‹ Back

## Check your photo

**Your photo**

### 1. Background and lighting

Your photo must have:

- a plain light-coloured background – without texture or pattern
- balanced light – no shadows on your face or behind you
- no objects behind you

### 2. Your appearance

Make sure:

- the photo is a good likeness taken in the last month
- your whole face is visible with your eyes open
- you have a plain expression – no smile and mouth closed
- there are no reflections or glare (if you have to wear glasses)
- you're not wearing headwear (unless for religious or medical reasons)

### 3. Photo quality and format

Your photo must:

- be in colour, with no effects or filters
- not be blurred or have 'red eye'
- be unedited – you can't 'correct' your passport photo

**Our automated check suggests**

- we can't find the outline of your head

**You can use the photo**

- if you have a medical or religious reason why it doesn't meet the rules
- if you think it meets all the rules

If your photo doesn't meet the rules you may have problems with your application or when you travel.

**Do you want to submit this photo?**

- Yes, I want to submit it
- No, I want to get another photo

Continue

https://twitter.com/CatHallam1/status/1114590857397788673

GOV.UK Apply for a passport

BETA This is a new service – your feedback will help us to improve it.

‹ Back

## Check your photo

### 1. Background and lighting

Your photo must have:

- a plain light-coloured background – without texture or pattern
- balanced light – no shadows on your face or behind you
- no objects behind you

### 2. Your appearance

Make sure:

- the photo is a good likeness taken in the last month
- your whole face is visible with your eyes open
- you have a plain expression – no smile and mouth closed
- there are no reflections or glare (if you have to wear glasses)
- you're not wearing headwear (unless for religious or medical reasons)

### 3. Photo quality and format

Your photo must:

- be in colour, with no effects or filters
- not be blurred or have 'red eye'
- be unedited – you can't 'correct' your passport photo

Your photo

**Our automated check suggests**

- it looks like your eyes are closed
- we can't find the outline of your head

**You can use the photo**

- if you have a medical or religious reason why it doesn't meet the rules
- if you think it meets all the rules

If your photo doesn't meet the rules you may have problems with your application or when you travel.

**Do you want to submit this photo?**

Yes, I want to submit it

No, I want to get another photo

Continue

## You need another photo

This photo has failed our automated checks for the following reasons:

Your pho

**It looks like your eyes are closed**

You must keep your head straight and in the centre of the photo.

You can still submit your photo if you believe it should be accepted (eg there's a medical reason why you can't open your eyes).

**Your photo quality is too low**

You need to use a photo that's between 50kb and 10m

**It looks like your mouth is open**

Sometimes this is caused by facial hair or shadows near your mouth. Please try different lighting.

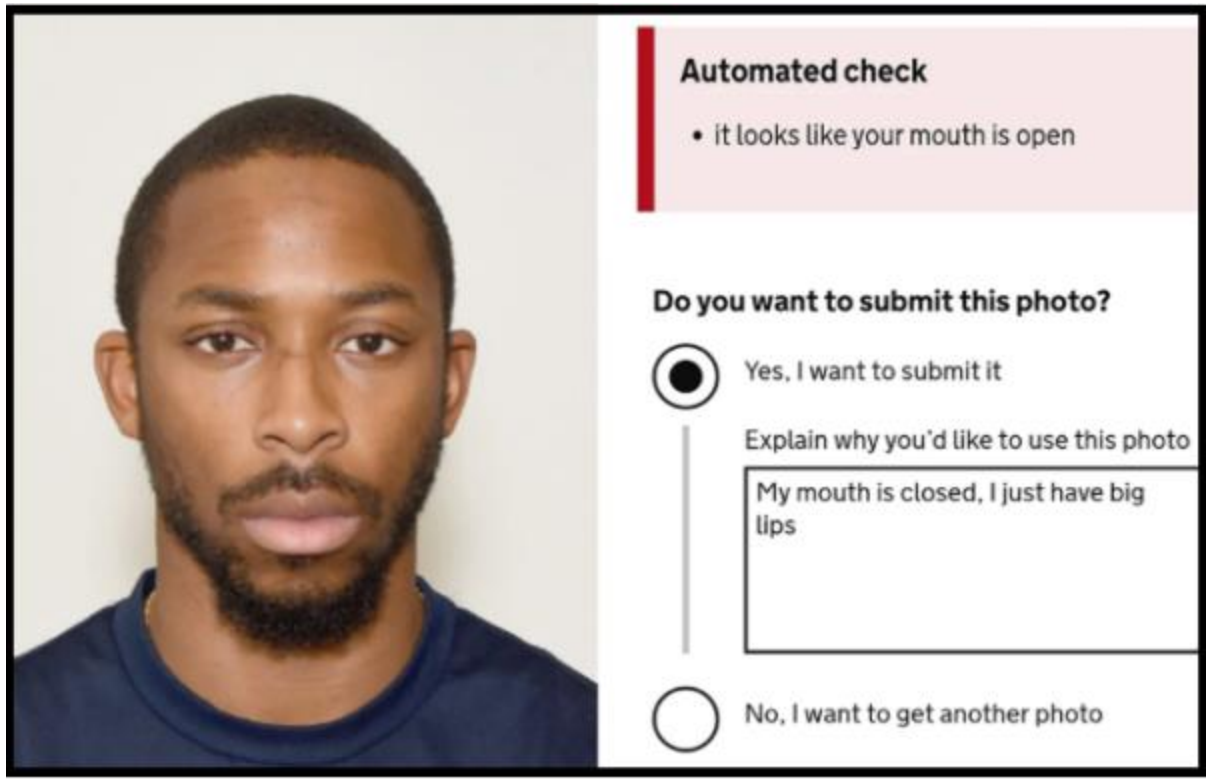
You can still submit your photo if you believe it should be accepted (eg there's a medical reason why your mouth appears to be open).

https://twitter.com/CatHallam1/status/1114590857397788673

# Twitter @CatHallam1: April 6, 2019

- “After numerous attempts and changing my clothes several times, this example illustrates why I regularly present on Artificial Intelligence/Machine Learning **bias**, equality, diversity and inclusion. #passport ”

# “Black man is stunned after passport photo checker mistakes his lips for an open mouth as Home Office facial recognition system is accused of bias”




https://www.msn.com/en-gb/money/technology/black-man-is-stunned-after-passport-photo-checker-mistakes-his-lips-for-an-open-mouth-as-home-office-facial-recognition-system-is-accused-of-bias/ar-AAHx5on?li=AABMOzg

# 2019: UK Home Office use of Facial Recognition

- “The UK government went ahead with a face-detection system for its passport photo checking service, despite knowing the technology failed to work well for people in some ethnic minorities.”
- “… was aware of problems with its website’s passport photo checking service, but decided to use it regardless”
- “User research was carried out with a wide range of ethnic groups and did identify that people with very light or very dark skin found it difficult to provide an acceptable passport photograph”

New Scientist: October 9, 2019:

https://www.newscientist.com/article/2219284-uk-launched-passport-photo-checker-it-knew-would-fail-with-dark-skin/#ixzz62XKSWWcF

# 2020: Met Police to roll out live facial recognition

- **BBC, 30 January 2020:**
  - **“The Metropolitan Police has announced it will use live facial recognition cameras operationally for the first time on London streets.”**
  - “The cameras will be in use for five to six hours at a time, with bespoke lists of suspects wanted for serious and violent crimes drawn up each time.”
  - “An independent review of its use has shown that most matches produced by the cameras are false alarms.”
  - “Privacy campaigners said it was a "serious threat to civil liberties".
  - “Following earlier pilots in London and deployments by South Wales Police, the cameras are due to be put into action within a month.”
  - “Police say they will warn local communities and consult with them in advance.”

From here: https://www.bbc.co.uk/news/uk-51237665

# Purpose of Today

- How diverse do AI teams need to be to deliver *AI for majority good*?
- Diversity could lead to new biases
- New documentary *Coded Bias* presented at the 2020 Sundance Film Festival:

https://www.wmm.com/sponsored-project/codedbias/

-

# References

**Copeland**, B. J. (2004). *The Essential Turing: The Ideas That Gave Birth to the Computer Age.* Clarendon Press: Oxford, UK

**Eysenck**, H. J. 1985. ***Know your own IQ***. Penguin Books: Middlesex, UK

**Fagan**, J. F. 2000. **A Theory of Intelligence as Processing: Implications for Society**. *Psychology, Public Policy and Law*. Vol 6(1), 168-179

**Fischler** M. A. and **Firschein**, O.I. 1987. ***Intelligence: the Eye, the brain, and the computer***. Addison-Wesley: USA

**Gardner**, H., Kornhaber, M.L. and Wake, W.K. 1996. ***Intelligence: Multiple Perspectives***. Harcourt Brace College Publishers: Texas: USA

**Turing**, A. M. 1948. 'Intelligent Machinery' in S.B. Cooper and J. van Leeuwen (Eds) *Alan Turing: His Work and Impact*, 2013 (Elsevier), 501-516

**Turing**, A. M. 1950. Computing Machinery and Intelligence', in *Mind* Vol 59 (236), 433-460

**Turing**, A.M. 1951. Can Digital Computers Think? Lecture in (Ed) B.J. Copeland). *The Essential Turing: The Ideas that Gave Birth to the Computer Age.* Clarendon Press: Oxford, UK, 2004, pp 482-486