

Autonomous Decision Making System Case Study



El Equipo

Verity Miles (s3644459)

Simon Karumbi (s3455453)

Christopher LeMarshall (s3482127)



Pre-assessment screening of applications

- Standardised video interviews
- Conversational AI
- Assessments

Claim to decrease recruitment time by 90%

Has performed over 20 million interviews since 2004

Over 700 clients, including:

- The Foxtel group
- PWC
- RioTinto
- Telstra
- Vodafone
- BP

Fairness

Accountability

Transparency

Ethics

FATE - Fairness

Machine Bias

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

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

May 23, 2016

Image: Angwin et al 2016

Technology

Researchers Combat Gender and Racial Bias in Artificial Intelligence

Companies use AI to predict everything from the credit worthiness to preferred cancer treatment. The technology has blind spots that particularly affect women and minorities.

Image: Bass & Huet 2017

Facial Recognition Is Accurate, if You're a White Guy

Image: Lohr 2018

Color Matters in Computer Vision

Facial recognition algorithms made by Microsoft, IBM and Face++ were more likely to misidentify the gender of black women than white men.



Gender was misidentified in up to 1 percent of lighter-skinned males in a set of 385 photos.



Gender was misidentified in up to 7 percent of lighter-skinned females in a set of 290 photos.



Gender was misidentified in up to 12 percent of darker-skinned males in a set of 318 photos.



Gender was misidentified in 35 percent of darker-skinned females in a set of 271 photos.

Image: J. Buolamwini from Lohr 2018

FATE - Transparency



- Competency measures determined by organisational psychologists
- No understanding of what goes into these
- No access to training data or test scores

EPIC 2019

Verity Miles

Best Practices

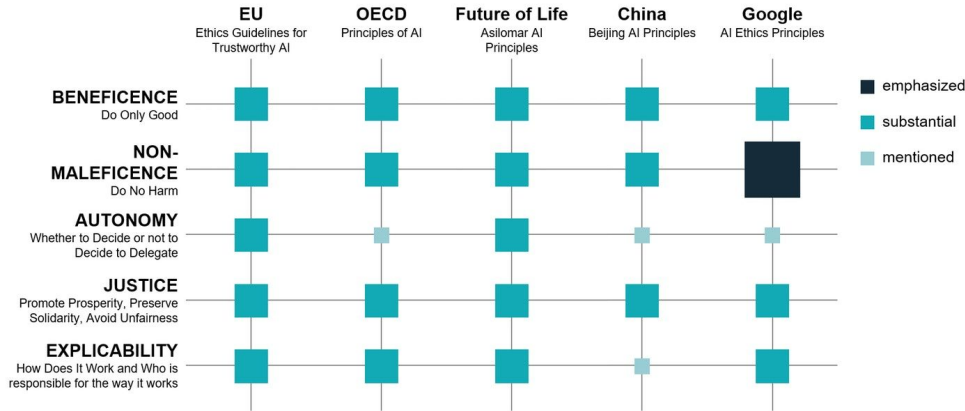


Figure 2: Five organizations in alignment with Floridi's AI principles.
The darker and larger the square, the more focus on the respective principle

Image: [CIONET](#) 2021

- OECD: Principles of AI
- EU: Ethics Guidelines for Trustworthy Artificial Intelligence
- US: Guidance for Regulation of Artificial Intelligence Applications

Legal Codes



Image: [ZDNet](#) 2021

- No formal legal codes
- Local Privacy Laws and Non-Discrimination Laws
- Interpretation of the GDPR Articles 13 and 15
- Google and Facebook self-regulate

HireVue's potential issues

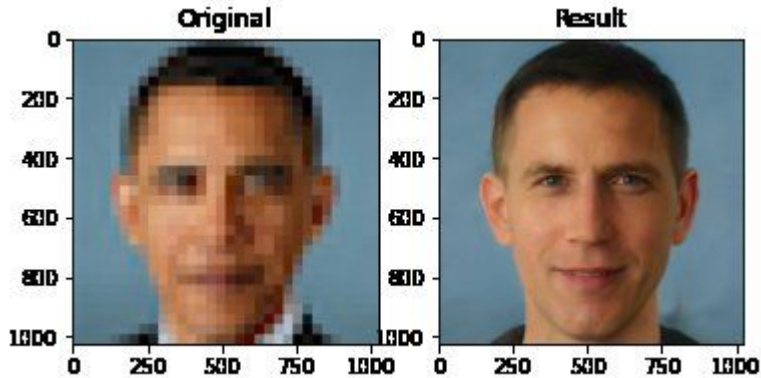


Image: [The Verge](#) 2020

- Fairness and non-discrimination
- Human oversight and determination
- Responsibility and accountability

Promoting Fairness & Ethics

- Ensure no difference between demographics
- Find and remove leaked protected characteristics
- Engage and protect minority groups



Image: Photo of a girl laying left hand on white digital robot by Andy Kelly. Copyright 2017 by Andy Kelly. Reuse allowed under the [Unsplash license](#).

Promoting Transparency & Accountability HireVue



Image: *Person Pointing Paper Line Graph* by Lukas. Copyright 2017 by Lukas. Reuse allowed under the [Pexels license](#).

- Provide candidates with specific, measurable feedback
- Offer independent review
- Make AI implementer accountable

Thank you



El Equipo

Verity Miles - s3644459@student.rmit.edu.au

Simon Karumbi - s3455453@student.rmit.edu.au

Christopher LeMarshall - s3482127@student.rmit.edu.au

References



- Angwin J, Larson J, Mattu S & Kirchner L (2016), *Machine Bias*, ProPublica, viewed 23 April 2021, url<<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>>
- Bass, D and Huet, E (2017), *Researchers combat gender and racial bias in artificial intelligence*, Bloomberg.com, viewed 23 April 2021, url<<https://www.bloomberg.com/news/articles/2017-12-04/researchers-combat-gender-and-racial-bias-in-artificial-intelligence>>
- Filipe, M (2019), *Artificial Intelligence and Data Ethics II*, CIONET, viewed 25 April 2021, url<<https://www.cionet.com/blog/2019/11/05/artificial-intelligence-and-data-ethics-ii>>
- EPIC (2019), Complaint and Request for Investigation, Injunction, and Other Relief, The Electronic Privacy Information Centre to Federal Trade Commission, Washington, viewed 22 April 2021, url< https://epic.org/privacy/ftc/hirevue/EPIC_FTC_HireVue_Complaint.pdf>
- Kahn, Jeremy (2021), *HireVue drops facial monitoring amid A.I. algorithm audit*, Fortune, url<<https://fortune.com/2021/01/19/hirevue-drops-facial-monitoring-amid-a-i-algorithm-audit/>>, viewed 22 April 2021
- Lohr, Steve (2018), *Facial recognition is accurate, if you're a white guy*, New York Times, viewed 22 April 2021, url<<https://www.nytimes.com/2018/02/09/technology/facial-recognition-race-artificial-intelligence.html>>
- Kelly, Andy (2017), *Photo of a girl laying left hand on white digital robot*, viewed 24 April 2021, url<https://unsplash.com/photos/0E_vhMVqL9g>
- Lukas (2017), *Person Pointing Paper Line Graph*, viewed 24 April 2021, url<<https://www.pexels.com/photo/person-pointing-paper-line-graph-590041>>
- Palmer, D. (2019), *What is GDPR? Everything you need to know about the new general data protection regulations*, ZDNet, viewed 25 April 2021, url<<https://www.zdnet.com/article/gdpr-an-executive-guide-to-what-you-need-to-know/>>
- Vincent, J (2020), *What a machine learning tool that turns Obama white can (and can't) tell us about AI bias*, The Verge, viewed 25 April 2021, uel<<https://www.theverge.com/21298762/face-depixelizer-ai-machine-learning-tool-pulse-stylegan-obama-bias>>