Artificial Intelligence (AI) has become a buzzword in modern society, being heralded as a revolutionary technology that will free humans from mundane responsibilities that should be embraced wholeheartedly, but also as threat to civilisation as it is known, with fears a general super intelligence could reign supreme, superseding humans.

In truth, neither of these scenarios represent AI’s current capabilities and offer a very black or white picture (Schwarz, 2019). However, there are multiple moral and ethical concerns arising as machine learning is ingrained deeper into all facets of modern life. This brief research aims to provide an insight into just two examples of the dilemmas faced both by private institutions and at national level.

Internet usage in China has quickly grown to 854.59 million individual users as of June 2019 (China Internet Network Information Center (CNNIC), 2019) and is widely considered to be at the forefront of becoming a digital society. The state announced in 2014 it would launch a social credit system (Creemers, 2014), similar to a traditional credit score, expanded to monitor all aspects of citizens lives, with “untrustworthiness” potentially resulting in reduced travel, employment and financial prospects, whilst positive behaviour is rewarded (Kostka, 2019).

One of the larger firms, Sesame credit which operates what is soon to be a mandatory credit score service, uses AI to rate individuals on a scale of 350-950 using big data from “a thousand variables across five data sets” (Campbell, n.d.) such as age, gender, “growth potential based on educational and professional history”, compliance on social media and the scores of people with which they communicate online (Reis & Press, 2019).

The use of AI and big data to automatically assign scores to individuals allows an alarming degree of automated control over a population. It is believed Facial Recognition Technology (FRT) is already in use in combination with over 200 million surveillance cameras (Donnelly, 2021) to allow the system to also identify and track individuals offline. Given AI’s history of making gender and racially biased predictions (Noor, 2020) in scenarios where it was not intentional, giving it direct control of an individual’s livelihood where a bias, political or otherwise may be implemented poses alarming questions if used by a regime with an agenda,

\*\*\*\*

FRT specifically is one of the most recognisable and favoured biometric methods, it has for example been used throughout COVID-19 in South Korea to track contacts and has aided global efforts to combat the virus. Its use resulted in lower incidences curves and lower mortality rates than countries without such tracking (Whitelaw et al., 2020).

It can however pose extreme ethical issues. Clearview AI is a relatively new company founded in 2017 which scrapes the web for faces matching that of the input. The program uses a neural net to convert images into vectors and store them in “neighbourhoods” based on similarities, when a photo is uploaded to be searched, vectors that are stored in the same neighbourhood are returned (Hill, 2020). This technology was quickly marketed toward law enforcement worldwide, offering a database of over 3 billion faces, it is used by over 600 law enforcement agencies in the US alone.

Concerningly, a private company is able to offer such data to anyone of their choosing, having already faced security breaches (Anon, 2020). The technology has been used in the US to arrest and charge an individual for committing a crime, typical police software would not have identified as they were in no government databases, yet due to social media content they were traced within 20 minutes (Hill, 2020).

Twitter and Youtube among others have sent cease-and-desist letters regarding data scraping but the practice is passionately defended by the company due to its ability to combat many crimes, including “terrorism and child exploitation” (Rezende, 2020). Australia and the UK have begun legal proceedings against the practice on the grounds of data protection, despite being used in both countries previously (Anon, 2021).

Despite the controversary, the software is still growing in popularity and raises the question of the need for global discussion on data’s use in AI. As the reader of this report there is a considerable chance that your face is accessible through the Clearview AI database, it must be discussed whether this is an acceptable price to pay to a private company’s algorithm for the national security it may provide.

References

Abd Rahman, R., Ramli, R., Jamari, Z. & Ku-Mahamud, K. (2016) *Evolutionary Algorithm with Roulette-Tournament Selection for Solving Aquaculture Diet Formulation*. [online]. Hindawi Publishing Corporation. [Accessed 18 December 2021].

Anon (2020) Controversial firm Clearview gets hacked. 2020 (3), pp. 12-12. [Accessed 1 January 2022].

Campbell, C. (n.d.) *How China Is Using Big Data to Create a Social Credit Score*. Available from: https://time.com/collection/davos-2019/5502592/china-social-credit-score/ [Accessed 1 January 2022].

China Internet Network Information Center (CNNIC) (2019) *Statistical Report on Internet Development in China*. [online]. pp. 13. Available from: https://www.cnnic.com.cn/IDR/ReportDownloads/201911/P020191112539794960687.pdf [Accessed 16 December 2021].

Creemers, R. (2014) *Planning Outline for the Construction of a Social Credit System (2014-2020)*. Available from: https://chinacopyrightandmedia.wordpress.com/2014/06/14/planning-outline-for-the-construction-of-a-social-credit-system-2014-2020/ [Accessed 20 December 2021].

Donnelly, D. (2021) *China Social Credit System [Punishments & Rewards] in 2021*. Available from: https://nhglobalpartners.com/china-social-credit-system-explained/ [Accessed 24 December 2021].

Hill, K. (2020) *The Secretive Company That Might End Privacy as We Know It (Published 2020)*. Available from: https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html [Accessed 1 January 2022].

Houser, K. (2021) Artificial Intelligence and the Struggle Between Good and Evil. *Washburn law journal*. 60 (3), pp. 475-478. [Accessed 14 December 2021].

Kostka, G. (2019) China’s social credit systems and public opinion: Explaining high levels of approval. *New Media & Society*. 21 (7), pp. 1565-1593. [Accessed 20 December 2021].

Noor, P. (2020) Can we trust AI not to further embed racial bias and prejudice?. *BMJ*. [Accessed 1 January 2022].

Reis, A. & Press, L. (2019) *Sesame Credit and the Social Compliance Gamification in China*. [online]. pp. 270-275. Available from: https://www.sbgames.org/sbgames2019/files/papers/ArtesDesignFull/196937.pdf [Accessed 1 January 2022].

Rezende, I. (2020) Facial recognition in police hands: Assessing the ‘Clearview case’ from a European perspective. *New Journal of European Criminal Law*. 11 (3), pp. 375-389. [Accessed 27 December 2021].

Schwarz, E. (2019) Günther Anders in Silicon Valley: Artificial intelligence and moral atrophy. *Thesis Eleven*. 153 (1), pp. 94-112. [Accessed 11 December 2021].

Surjanovic, S. & Bingham, D. (2013) *Styblinski-Tang Function*. Available from: https://www.sfu.ca/~ssurjano/stybtang.html [Accessed 19 December 2021].

Anon (2021) *The UK’s Information Commissioner’s Office and the Office of the Australian Information Commissioner conclude joint investigation into Clearview AI Inc.*. Available from: https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2021/11/clearview-statement/ [Accessed 23 December 2021].

Whitelaw, S., Mamas, M., Topol, E. & Van Spall, H. (2020) Applications of digital technology in COVID-19 pandemic planning and response. *The Lancet Digital Health*. 2 (8), pp. e435-e440. [Accessed 1 January 2022].