ARITIFICIAL INTELLIGENCE

The use of Artificial intelligence (A.I) comes with its inherent challenges and risks from a legal and ethical point of view (Gillis 2021). Responsible A.I is therefore a governance framework that allows organisations to navigate around these challenges and to mitigate the risks which arise from A.I usage (Arietta et al 2020).

Although valuable in its usage artificial intelligence has been known to fail (Peters et al 2020). In 2016, Microsoft released a chatter bot A.I on twitter which was known as Tay (Wikipedia 2021). The intention was to study natural languages and conversation to gain a deeper understanding (Pykes 2021). However, a vulnerability in Tay meant that trolls maliciously manipulated it to send sexist, racial and offensive tweets (Wikipedia 2021). Subsequently, Microsoft withdrew Tay within 16 hours of its launch (Pykes 2021).

Another example of A.I failure is the fatal accident involving Uber’s self-driving car and a pedestrian crossing the road in 2018 (Dialani 2021). It was reported that the A.I of the self-driving car could not identify a pedestrian as an object except they were close to a crosswalk (Pykes 2021)

A third example of A.I gone wrong involved the Amazon facial recognition A.I known as Rekognition (Dialani 2021). In this case the A.I wrongly matched faces of well-known sports athletes to mugshots of criminals (Dialani 2021).

Although the above examples highlight situations where artificial intelligence either failed or was maliciously used, according to Araujo et al (2020) advancement in technology has seen a greater dependence on artificial intelligence through automated decision making. However, according to Butterworth (2018) the UK government has put in place safeguards to protect the consumer in Article 22 of the Data Protection Act 2018 (GDPR). The article states that individuals reserve “the right not to be subject to a decision based solely on automated processing, including profiling” (ICO 2021).

This means that individuals can “challenge a decision or request an intervention” (Which 2021).

In conclusion, the major aim of A.I responsibility is to ensure that there is an accountability element to every data that is collected by organisations (Araujo et al 2020). Furthermore, a responsible A.I ensures that such data is interpreted in a manner that is safe, fair and respects service users’ privacy (Butterworth 2018). Therefore, a responsible A.I must underpin the following principles: accountability, fairness, interpretability, safety, and privacy (Sullivan 2020). Subsequently organisations must ensure A.I responsibility by adhering to principles such as these or principles closely related to these (Araujo 2020). For example, on its website, Microsoft’s customers are alerted to its principles of artificial responsibility which include fairness, transparency and explainability, human-centeredness, privacy, and security (Microsoft 2021).

On the other hand, when it comes to responsible data management, it is expected that organisations put systems in place that ensure transparency, fairness and which respect data that affects individuals (Internet Society 2019). Examples of this include and is not limited to staff training, ensuring systems are working properly as intended and quality control measures that ensure constant process improvement (Internet Society 2019, Which 2021).

References

Araujo, T., Helberger, N., Kruikemeier, S. and De Vreese, C.H., 2020. In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & SOCIETY*, *35*(3), pp.611-623.

Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R. and Chatila, R., 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities, and challenges toward responsible AI. Information Fusion, 58, pp.82-115.

Butterworth, M., 2018. The ICO and artificial intelligence: The role of fairness in the GDPR framework. *Computer Law & Security Review*, *34*(2), pp.257-268.

Dialani P. (2021) <https://www.analyticsinsight.net/famous-ai-gone-wrong-examples-in-the-real-world-we-need-to-know/>

Information Commissioner’s Office, 2021 <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/rights-related-to-automated-decision-making-including-profiling/>

Microsoft (2021) https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1%3aprimaryr6

Peters, D., Vold, K., Robinson, D. and Calvo, R.A., 2020. Responsible AI—two frameworks for ethical design practice. *IEEE Transactions on Technology and Society*, *1*(1), pp.34-47.

Pykes K, (2021) <https://towardsdatascience.com/5-ai-failures-you-probably-should-know-about-417ddebbc323>

Sullivan C (2020) https://towardsdatascience.com/what-is-responsible-ai-548743369729

Wikipedia (2021) <https://en.wikipedia.org/wiki/Tay_(bot)>

Which (2021) https://www.which.co.uk/consumer-rights/advice/can-i-stop-a-company-making-automated-decisions-about-me-aOxZw4t2tHtG