**Q: What is responsible AI?**

A: Responsible AI is a governance framework that documents how a specific organization is addressing the challenges around artificial intelligence (AI) from both an ethical and legal point of view. It also focuses on ensuring the use of AI technologies is transparent and used in a manner consistent with user expectations, organizational values and societal laws and norms. Responsible AI can help maintain user trust and individual privacy.

**Q: Where has AI failed, or been used maliciously or irresponsibly?**

1. Facial recognition: - In 2018 Chinese police reportedly admitted to wrongly accusing a woman of jaywalking, after a facial recognition system designed to catch jaywalkers identified her face on an advert on a passing bus. The algorithm’s lack of ability to distinguish between people and photos of people was identified as the issue.Similarly, also in 2018, Amazon’s facial recognition software falsely matched 28 congresspeople with mugshots of criminals. According to the American Civil Liberties Union (ACLU), nearly 40% of the matches were of people of colour, indicating that the technology may be racially biased.
2. IBM’s Waston - a supercomputer that defeated some of the world’s best-known quizzers in a game of televised Jeopardy - was updated in 2018 in an attempt to launch a medical AI system that would make suggestions for treating cancer patients. However, hospitals and oncologists identified major flaws, including Watson’s suggestion of putting a patient with excessive bleeding on a medication that would cause even more bleeding, possibly killing the patient. IBM has blamed its engineers, stating they programmed Watson with hypotheticals and fictional cases, rather than relying on actual patient data and historical medical charts.

**Q: What are the implications of AI failing? (GDPR law opt in and out options).**

These failures of AI highlight that using AI as the only decision maker can led to wrong decisions and predictions being made, and that such failures can have huge, unfair, negative impacts on and for the individuals/groups concerned. In recognition of this, Article 22 of GDPR law focuses on concerns around automated profiling and decision-making. It states that AI — including profiling — cannot be used as the sole decision-maker in choices that can have legal or similarly significant impacts on individuals’ rights, freedoms and interests. For instance, an AI model cannot be the only step for deciding whether a borrower is eligible to qualify for a loan. The goal is to prevent any unfair bias or discrimination from entering into a decision. In addition, GDPR also now requires organisations to gather consent from users to store their personal data for any use, including analytics and profiling. This is referred to as the right to opt-in and opt-out.

**Q: What should organisations do to ensure they are being responsible with AI and the wider use of data in general?**

Organisations need to ensure that AI outputs are fair, that outputs/prediction do not translate into discrimination, that data acquisition and its use are in line with consumer privacy and expectations, and that the use of data and how AI systems make their predictions is reasonably transparent.

This is covered by GDPR, which requires organisations and individuals collecting data to ensure that the data is used inline with stakeholder and/or participants expectations and understanding. This includes stating at the time of data collection what the data will be used for, how it will be used, if and how the data and any outcomes will be shared with others (including the wider public if applicable), making it clear that participants have control of their own data and can ask to withdrawn from the data collection now or in the future (unless collection is legally enforced), and explain who long the data will be held for. All of this must be expressed in plain language, along with the process to follow to opt out.

If data is to be used to make assumptions/predications of a particular population, the organisation also has a responsibility to ensure that the data collected reflects the real-world population (for example, that it includes a reasonable amount of data from any minority groups). Test should be run to test predictions/outcomes for any inaccuracies/biases, and any potential limitations should be clearly stated.