Your task is to:

**• Find out what Responsible AI is?**

Responsible AI is the notion of ethical and accountability which has been adapted by stakeholders (Government, industry etc) Making AI systems transparent, fair, secure, and inclusive are core elements of widely asserted responsible AI frameworks.

**• Find instances where AI has failed? Or been used maliciously or incorrectly.**

1. **Tesla cars crash due to autopilot feature** - Tesla Model S crashed north of Houston, killing two people. The car had missed a slight curve in the road, leading it to ram into a tree. it is believed that Tesla’s Autopilot or Full Self Driving (FSD) system was engaged during the crash. Tesla’s AI-based Autopilot feature can control steering, acceleration, and in some cases, braking. According to Musk, the AI is designed to learn from drivers’ actions over time.
2. **AI camera mistakes linesman’s head for a ball** - an AI-powered camera designed to automatically track the ball at a soccer game ended up tracking the bald head of a linesman instead. The camera kept on mistaking the ball for the bald head on the side-lines, denying viewers of the real action while focusing on the linesman instead.
3. **False facial recognition match leads to Black man’s arrest** - Black man living in Paterson, New Jersey, was accused of shoplifting and trying to hit a police officer with a car in Woodbridge, New Jersey. Although he was 30 miles away at the time of the incident, the police identified him using facial recognition software. He was arrested following a “high profile comparison” from a facial recognition scan of a fake ID left at the crime scene. Facial recognition technology, which uses machine learning algorithms to identify a person based on their facial features, is known to have many flaws. In fact, a 2019 study found that facial recognition algorithms are “far less accurate” in identifying Black and Asian faces.

**• Implications of when AI fails. There is a specific article in the GDPR Law that covers this, especially with automated decision making. (opt in and out options).**

Automated decision making – i.e. profiling cannot be used to make decisions about a person.

Article 22(1) of the UK GDPR limits the circumstances in which you can make **solely automated decisions**, including those based on profiling, that have a **legal or similarly significant effect on individuals.**

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

**•Maximum 500 words.**

Deploying AI requires careful management to prevent unintentional but significant damage, not only to brand reputation but, more important, to workers, individuals, and society as a whole.

Organisations should be responsible when building and applying AI by taking care to confirm that AI outputs are fair, that new levels of personalisation do not translate into discrimination, that data acquisition and use do not occur at the expense of consumer privacy, and that their organisations balance system performance with transparency into how AI systems make their predictions.

Organisations should:

* provide the best data through using informed algorithms, if there are inaccuracies then there wasn’t sufficient data in the first place
* Provide the proper oversight - Establish a system of governance with clear owners and stakeholders for all AI projects. Define which decisions you’ll automate with AI and which ones will require human input. Assign responsibility for all parts of the process with accountability for AI errors, and set clear boundaries for AI system development. This includes monitoring and auditing algorithms regularly to ensure bias is not creeping in and the models are still operating as intended.
* Consider ramifications of new technologies - In order for individuals to enforce policies, the technology must allow humans to make adjustments. Humans must be able to select and adjust the training data, control the data sources and choose how the data is transformed. Likewise, AI technologies should support robust governance, including data access and the ability to guide the algorithms when they are incorrect or operating outside of ethically defined boundaries.