**HOME LEARNING TensorFlow**

**RESPONSIBLE ARTIFICIAL INTELEGENCE (AI)**

* Find out what Responsible AI is?
* Find instances where AI has failed? Or been used maliciously or incorrectly.
* 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).
* What should organisations do to ensure that they are being responsible with AI and the wider use of data in general?
* Maximum 500 words **(Total Count = 480)**

Artificial intelligence (AI) is intelligence demonstrated by machines, as opposed to the natural intelligence displayed by animals including humans. However, many concerns have risen such as unfair decisions, workers being replaced, and a lack of privacy and security are valid. To make things worse, many of these issues are unique to AI. This means existing guidelines and laws are not suitable to address them. This is where Responsible AI comes in. It aims to address these issues and create accountability for AI systems. 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. The four key principles of responsible AI are as follows:

* transparency and interpretability
* human-centeredness
* fairness,
* privacy and security

The AI hype machine is real and big-name companies have fallen prey to AI failures. These experiences have been alarming to consumers, embarrassing to companies involved, and an important reality-check for us all.

There are many cases known cases of AI goof-ups and here are a few from only 2020;

### **The inaccuracy of YouTube’s AI moderators (COVID-19 related clips)**

### **Fumbling of facial recognition technology (COVID-19 masks)**

### **The A-Level fiasco (marking students for cancelled exams)**

### **The most devastating software mistake (COVID-19 inaccurate forecast of deaths)**

### **The law in place that deals when there are implications of AI failure. The UK GDPR, article 22(1),** 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. The UK GDPR applies to all automated individual decision-making and profiling. Article 22 of the UK GDPR has additional rules to protect individuals if you are carrying out solely automated decision-making that has legal or similarly significant effects on them. You can only carry out this type of decision-making where the decision is:

* + necessary for the entry into or performance of a contract; or
  + authorised by domestic law applicable to the controller; or
  + based on the individual’s explicit consent.

You must identify whether any of your processing falls under Article 22 and, if so, make sure that you:

* + give individuals information about the processing;
  + introduce simple ways for them to request human intervention or challenge a decision;
  + carry out regular checks to make sure that your systems are working as intended.

Consequently, it is imperative that organisations ensure that they are being responsible with AI and the wider use of data in general. The internal governance of organisations need to establish in-house guidelines such as having an objective review panel that is diverse and has the specialised expertise to understand the possible consequences of AI infused systems. A key success factor is leadership support and the power to hold leadership accountable.

In order for the development of responsible AI that operational this must be implemented and reviewed frequently and modified as appropriate.