**ARTIFICIAL INTELLIGENCE (AI) HOME LEARNING TASK**

**1 - Find out what responsible AI is?**

Responsible AI is a framework that outlines how individual organisations address the challenges around artificial intelligence (AI), from both the legal and ethical perspective. They also need to clearly define who takes responsibility if problems arise. The four main principles which highlight responsible AI include fairness, transparency and explainability, human centredness (empathy), privacy and security.

**2 - Find instances where AI has failed. Or been used maliciously or incorrectly.**

In 2018 an Uber self-driving test vehicle hit and killed a woman. US safety investigators found that the vehicle had software problems which could have caused the accident.

Another example was in Japan 2015 when a revolutionary Henn-na hotel which opened its doors to the public with all its employees including the cleaners, porters and front desk being robots. Unfortunately, the robots continually broke down and were failing to adequately respond the customers’ request. The hotels ended up getting rid of their unreliable and costly robots in favour of human staff.

**3 –** **Implications where AI fails in regards to GDPR** law involves automated decision making which does not include any human involvement. Profiling is also an automated process of analysing an individual’s data by classifying them based on their characteristics such as interests, preferences or simply making predictions. Automated decision making can be useful for some organisations. However, if used irresponsibly can cause major risk for some individuals affected. According to Article 22(1), “The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.”

**4 –** **Organisations can show that they are being responsible with AI** by ensuring they have inbuilt mechanisms such as having a diverse review panel, that has the knowledge to understand the possible consequences of AI infused systems. Management should be held accountable for all the processes in the project, which can invoke trust in the users. Organisations should have a wide range of datasets to experiment with before implementation. They should examine the raw data where possible and understand the limitations of the data they are working with. They also have to continually test the data and statistics. Organisations need to monitor and update their systems to ensure that models of use remain current and relevant. Investing in more AI education and training ensure that stakeholders are kept informed of all the pros and cons and capabilities of AI.