

## TT – Home Learning Task 9

- Find out what Responsible AI is?

Responsible AI is the practice of designing, developing, and deploying AI with good intention to empower employees and businesses, and fairly impact customers and society, allowing companies to engender trust and scale AI with confidence.

Principles include:

- Fairness and inclusivity – so that no bias towards any groups of people are incorporated into the AI
- Transparency – in the methods that went into the development of the AI, so that it can be reproduced
- Privacy and security – to protect the individuals whose data has been used in the development of the AI
- Comprehensiveness – comprehensive AI has clearly defined testing and governance criteria to prevent machine learning from being hacked easily.
- explainable AI is programmed to describe its purpose, rationale and decision-making process in a way that can be understood by the average end user.
- Ethical AI initiatives have processes in place to seek out and eliminate bias in machine learning models.
- Efficient AI is able to run continually and respond quickly to changes in the operational environment.

Responsible AI is the right thing to do. It helps ensure that any AI system will be efficient, comply with laws and regulations, operate based on ethical standards and prevent potential for reputational and financial damage down the road.

Recommended practices:

1. Use a human-centered design approach
  2. Identify multiple metrics to assess training and monitoring
  3. When possible, directly examine your raw data
  4. Understand the limitations of your dataset and model
  5. Test
  6. Continue to monitor and update the system after deployment
- Find instances where AI has failed? Or been used maliciously or incorrectly.

In healthcare, where the data used has not been stored correctly

- 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).

With the healthcare example, the implications to this are privacy, security, possibly safety and inclusivity

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

Organisations should ensure that throughout the whole process, they follow and have check points to check that they are following the principles and guidelines of responsible AI. This should also be clearly documented. There should also be an independent reviewer coming in and making sure that organisations are doing this correctly and are able to give advice on how to improve their processes.