**ARTIFICIAL INTELLIGENCE (AI)**

**What is Responsible AI?**

Responsible AI is a framework that governs and documents how an organisation addresses the challenges around AI from an ethical and legal point of view. It clarifies where responsibility lies if something goes wrong and is an important for responsible AI initiatives.

We need more governance of AI as often ***bias*** can be introduced by the data that’s used to train machine learning models which then creates unintended consequences. If the training data is biased, then decisions made by the programming will also be biased. The technology may then be misused accidently or intentionally.

**Examples of failed AI**

Microsoft’s AI chatbot failed and got corrupted on Twitter. The chatbot Tay could automatically reply to people and engage in casual and playful conversation. As more people talked with Tay, the chatbot would learn how to write more naturally and hold better conversations. Less than 24 hours after it was launched internet trolls had corrupted the chatbots personality by flooding the bot with racist, misogynist and anti-Semitic views.

Another example is Amazons AI for recruitment where the engineers inadvertently trained their model that white males were automatically better candidates for their software engineering roles as they based their data against their current engineering employees.

**GDPR and AI**

The GDPR (General Data Protection Regulation) has had an impact in creating a more regulated data market. There are 7 key principles:

* Lawfulness, fairness and transparency
* Purpose limitation
* Data minimisation
* Accuracy
* Storage limitation
* Integrity and confidentiality
* Accountability

As data is the key ingredient for AI applications this raises issues in policy related regulations. Under Article 22 the GDPR includes provisions that apply to all profiling and automated decision making. Such as the right to access and right to object. And in profiling and automated decision making tht include the processing of personal data, all GDPR provisions apply.

One major topic of GDPR in the context of AI is the ‘meaningful information about the logic involved’. It should be understood as information around the algorithmic method used rather than an explanation about the rationale of an automated decision. e.g., if a loan application is refused Article 22 may require the controller to provide information about the input data related to the individual and the general parameters set in the algorithm that enabled the automated decision. But Article 22 would not require an explanation around the code or how and why the decision was made.

It is also important to understand the value of considering data protection at an early stage, as AI relies on the data of people. So, it is paramount that explicit consent of processing personal data is obtained from individuals to whom the data belongs. People should be made aware of how their data is being used, and how the data is being held secure. One way of doing ensuring this is by being clear with people of their rights to opt in or opt out of how their information is used or shared.