**Find out what Responsible AI is?**

AI uses machine learning models to make predictions based on data and in turn these are used to make decisions. AI can be used from detecting cancer to self-driving cars. There is limited human input into decisions made and this can cause issues. It is vital you can trust its outputs. Responsible AI is using AI ethically and transparently. This can be done using a framework to consider the performance and security of AI use. Responsible AI needs to consider the bias (and fairness), interpretability (and explanation), privacy (and security) and safety. It is possible models may discriminate against certain groups of people and we need to ensure to be able to evaluate the fairness of models. This can be more easily done with more interpretable models where we can get an explanation for how predictions/decisions are made. Responsible AI evaluates and considers what to do when models are biased/unfair, which is important on less interpretable models.

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

February 2019 – Facial recognition software used to incriminate a person who was innocent. A study found facial recognition software is far less accurate in identifying black and Asian faces.

2014 – Amazon used AI to review job applicants, however it was biased against women. The model was trained on data (CVs/resumes) that were mostly of men and the model taught itself to favour men.

Examples of malicious use can include the use of deepfakes whereby you can create a clone of someone and manipulate their words. This could be a threat to democracy and political stability.

**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).**

When AI uses or creates personal data then it falls under data protection laws which are made up of General Data Protection Regulation (GDPR) and Data Protection Act 2018 (DPA 2018). Whilst there is no direct reference to AI or associated technologies there is references to use of profiling and automated decision making (part of machine learning).

The right to be informed

Articles 13 and 14 of the GDPR give individuals the right to be informed of:

* the existence of solely automated decision-making producing legal or similarly significant effects;
* meaningful information about the logic involved; and
* the significance and envisaged consequences for the individual.

The right of access

Article 15 of the GDPR gives individuals the right of access to:

* information on the existence of solely automated decision-making producing legal or similarly significant effects;
* meaningful information about the logic involved; and
* the significance and envisaged consequences for the individual.

Recital 71 provides interpretative guidance on rights related to automated decision-making. It mainly relates to Article 22 rights, but also makes clear **that individuals have the right to obtain an explanation of a solely automated decision after it has been made**.

The right to object

Article 21 of the GDPR gives individuals the right to object to processing of their personal data, specifically including profiling, in certain circumstances.

There is an absolute right to object to profiling for direct marketing purposes.

Rights related to automated decision-making including profiling

Article 22 of the GDPR gives individuals the right not to be subject to a solely automated decision producing legal or similarly significant effects. There are some exceptions to this and in those cases it obliges organisations to:

* adopt suitable measures to safeguard individuals, including the right to obtain human intervention;
* express their view; and
* contest the decision.

Reference: <https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/explaining-decisions-made-with-artificial-intelligence/part-1-the-basics-of-explaining-ai/legal-framework/>

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

Organisations should ensure the following:

Identified the people who are in key roles across the decision-making pipeline and how they are responsible for contributing to an explanation of the AI system.

Ensured that different people along the decision-making pipeline are able to carry out their role in producing and delivering explanations, particularly those in AI development teams, those giving explanations to decision recipients, and their compliance teams.

If buying the AI system from a third party, they know that they have the primarily responsibility for ensuring that the AI system is capable of producing explanations.

Reference: <https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/explaining-decisions-made-with-artificial-intelligence/part-3-what-explaining-ai-means-for-your-organisation/organisational-roles-and-functions-for-explaining-ai/>