Compulsory HLT

Practical: Artificial Intelligence (AI)

1. 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.

AI brings unprecedented opportunities to businesses, but also incredible responsibility. The output from AI systems has a real bearing on people’s lives, raising considerable questions around AI ethics, data governance, trust, and legality. The more decisions a business puts into the hands of AI, the more they accept significant risks, such as reputational, employment/HR, data privacy, health and safety issues. However, according to an Accenture global research study, 88% of respondents do not have confidence in AI-based decisions.

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

Some of the real life examples, where the AI failed in the past are: Amazon’s Recruiting Tool, Microsoft Tay, Inverness Caledonian Thistle F.C. Ball Tracking System, Uber Self Driving Car Fatality, Face ID Hacked Using a 3D Printed Mask, Tesla Cars, [facial-recognition](https://www.analyticsinsight.net/is-facial-recognition-stooping-down-with-unauthorized-surveillance/) technologies, Data limitation in Excel(Public Health England (PHE)) and much more examples. Tweet chat bot is one of the cases where AI been used maliciously by users.

1. 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).

The UK GDPR gives people the right not to be subject to solely automated decisions, including profiling, which have a legal or similarly significant effect on them.  These provisions restrict when you can carry out this type of processing and give individuals specific rights in those cases. Implications of when AI fails can be breach of Equality Act 2010, data and protection law, wrong recognition, wrong given data, wrong advices in healthcare or other sectors. Some effects with example: automatic refusal of an online credit application; or e-recruiting practices without human intervention. In extreme cases, it might exclude or discriminate against individuals. Decisions that might have little impact generally could have a significant effect for more vulnerable individuals, such as children.

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

Accountability and transparency are important elements of preventing faults. Sharing practices, giving right data,

Ensuring the right technical guardrails, creating quality assurance and governance to create traceability and auditability for AI systems. This is another important part of every organisation’s toolkit to allow operational and responsible AI to scale.

Investing more in their own AI education and training so that all stakeholders – both internal and external – are informed of AI capabilities as well as the pitfalls.