AI Assignment

What is responsible AI?

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 be trusted and scale AI with confidence.

AI brings a vast number of opportunities to businesses, but also incredible responsibility. Its direct impact on people's lives has raised ethical questions around AI, data governance, trust, and legality. In fact, Accenture's 2022 Tech Vision research found that only 35% of global consumers trust how AI is being implemented by organizations. And 77% think organizations must be held accountable for their misuse of AI.

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

Instagram's skin-showing Artificial Intelligence algorithm

In 2020, an investigation into the social media company Instagram found that one of its algorithms prioritized photos of men and women that show more skin. This can directly impact content creators and has an especially negative impact on many young people who use Instagram.

The investigation analysed 2,400 photos and found that a computer program recognized 21% of them as containing women in bikinis or underwear or bare-chested men. Not posting images showing body parts significantly decreases the organic reach on Instagram. This means that male and female content creators could face pressure to show skin to reach a larger audience.

Smart speakers listening to you

Smart speakers, such as Google Nest, Alexa and so on have been criticised for listening to private conversations. While these speakers typically only wake up and start listening after an active word, such as "Alexa," they are prone to mistakes. A study carried out in 2020 found that some software, such as Alexa, Google Assistant, Siri, and Cortana can be activated by mistake up to 19 times a day. Companies do this since it gives them access to people's private conversations, which is valuable data and can be used for marketing products.

In 2019 the Finnish newspaper Helsingin Sanomat interviewed four Finns who work for Google and other technology companies listening to private conversations of Google Home device users. It is worth noting that users of these devices do not know that their conversations are being shared and listened to by company workers. "Generally, you know that you are collecting data, but then when you listen to other people's conversations, it was somehow "sick" commented one worker. Having virtual assistants is good but they would have to work if the need to share private conversations with the servers of large technology companies.

Implications of when AI fails:

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.

An example of AI failing is when a company uses automated processes to pick who they should interview for a job. This can lead to discrimination and lead the public not to trust AI:

Example

As part of their recruitment process, an organisation decides to interview certain people based entirely on the results achieved in an online aptitude test. This decision has a significant effect since it determines whether or not someone can be considered for the job.