HLTW8Tracey

Find out what Responsible AI is?

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

• Implications of when AI fails. There is a specific article in the GDPR Law that covers this, especially with automated decision making.

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

To determine if an AI (Artificial Intelligence) system is responsible it should be able to quantity several different aspects of how the data is gathered, stored and interpreted.

Does the data cover a wide and diverse cross section of target area, the validity of the data collected, is it a truthful source? How the data is stored what sort of security is in place, particularly if the data is highly sensitive. How is the data being queried, is there a bias in the queries?

Most recently the area were AI seems to be failing consistently is with facial recognition. Where this system has shown a distinct bias for white males. So that for all other groups the accuracy of the system is not reliable. This has been apparent with people of darker skin tones. One example show as a black man moved away from his laptop the system just ignored his face all together and just showed a blank space.

One of the biggest fails would be with the Police force and racial profiling, because of the systems inability to identify facial differences for black people it creates a generic model which could fit a wide cross section of the black community making it totally unreliable. Case in point when a black teenage boy was stopped for something a man in his 30’s was accused of doing. When an actual comparison was done the two individuals had nothing in common other than they were both black.

Organizations should make sure their data is collected from as diverse a group as possible and have a team that come from a wide range of view points to ensure the queries not now skewed to favour one particular demographic over another.