**Practical: Artificial Intelligence (AI) – Week 7 Home Learning – Samesh Gumber**

Responsible AI involves developing and using artificial intelligence with the right intentions to help businesses and society as a whole. Therefore, it should allow users to trust and use AI with confidence because if users do not trust AI they are not going to feel comfortable using it with concerns surrounding any information they are uncomfortable with sharing or think may make biased decisions. Artificial Intelligence involves using a machine learning model to automate a given task. Risks with AI arise as there is limited to no human input into the decisions made by the system, so Responsible AI allows us to design frameworks to define a better approach to using AI. The main goal with Responsible AI is to create systems that are interpretable, fair, safe and respectful of a user’s privacy. We can use Responsible AI in various aspects and an important area is privacy and data governance. To begin with we must ensure that the quality of data used is of high standard and mistake-free. If there are mistakes in the data used by AI then the system could potentially make incorrect decisions. We should also not use AI with personal sensitive data, for example, medical history.

An instance where AI had failed was with the ‘Uber Self Driving Car’, which was seen to go through red lights during real-world testing. In 2016, Uber conducted tests on its self-driving cars in San Francisco without approval from California state regulations. The resulting documents showed that the vehicle did not stop at six red lights in the city during their test ride. The caveat was that there was a driver behind the wheel to take over if something went wrong, but this is a scenario where the ramifications could be dangerous if something were to go wrong.

For instances where AI fails with data privacy when storing and using a data from a user, we can use regulations to try and minimise the potential risks of using AI. The General Data Protection Regulation (GDPR) means that when organisations are dealing with users’ (from the EU) personal data. Under the GDPR, the data collection, storage, and processing should be carried out only after consent from the user. So the user can ‘opt-on’ to allow consent, ‘opt-out’ to refuse consent. These two options can be implemented using checkboxes and toggle buttons.

When it comes to AI, it is expected that organisations self-regulate so they must create their own Responsible AI framework. The issue with this is that across different organisations there could be inconsistencies between one framework used by an organisation to another. So a potential solution to this would be for all organisations to use the same guidelines to ensure uniformity across an industry, which in turn simplifies things for the user. Also, it could be proposed that each company should have a review panel, that is diverse and that has the knowledge to understand the possible consequences of AI systems. Similarly, organisations should strive to increase investment into AI education and training so all stakeholders are aware of the capabilities of AI.