**AI quiz: discussion**

## Question 2: Hacking

|  |  |  |
| --- | --- | --- |
|  |  | **Potential for AI systems to be hijacked and used in harmful ways**  An AI should benefit society throughout its lifecycle. In planning and development, the impacts of the AI need to be future-focused, forecast and then considered and addressed.  It is vital that teams creating the AI system have a mechanism to raise any concerns. In some cases the individual raising the concern may be negatively impacted for speaking up. A person facing this dilemma has conflicting values; loyalty and courage.  This scenario covers human, social and environmental wellbeing. |
| Discuss their responses using the following question prompts:   * What if there is a possibility an AI can be used in a way that may be harmful to society? * Whose role is it to stand up and call out these potential threats? * What do you think the term ‘AI for good’ means? | |

## Question 3: Self-driving car and Question 4: Responsibility

|  |  |  |
| --- | --- | --- |
|  |  | **An AI can make mistakes**  AI systems make predictions based on their training. These predictions are accompanied by a level of confidence. There is always a chance of error. In scenarios involving a self-driving car an unpredictable event may result in an accident – even a loss of life. What happens in these situations? How does the AI make the ‘best’ decision? Who is responsible when something goes wrong?  This scenario covers human, social and environmental wellbeing; reliability and safety; accountability; and contestability. |
| * How does the AI decide what to do in dangerous and life threatening situations? * Who is ultimately responsible when something goes wrong? * What can someone do if they are impacted badly by an AI system? Can they seek compensation? | |

## Question 5: Marketing

|  |  |  |
| --- | --- | --- |
|  |  | **AI can impact negatively on users**  Recommended systems in simple terms, find good matches and suggest them to a user. The algorithm though, could be deliberately designed with a bias towards a particular product or point of view.  In this example, the company could train the AI to include a certain parameter which skews the expected outcome.  As a result the AI would recommend products that may not necessarily be accurately matched to the user’s profile and online behaviours. The company however would benefit in a monetary sense.  This scenario raises the issues of fairness, honesty and human-centred values. |
| Discuss their responses using the following question prompts:   * How would you feel if a site recommends products that are skewed towards a particular company or point of view? * What is at the heart of this dilemma?   This can also happen in news feeds where you are only shown stories based on a particular point of view. | |

## Question 6: Automation

|  |  |  |
| --- | --- | --- |
|  |  | **AI may result in job losses**  Should we integrate AI into industry if it means many people will lose their jobs? Society benefits when workers are employed. AI systems don’t need to be paid or contribute back into the economy. Where do the cost savings go? Do the wealthy get wealthier?  This scenario covers human, social and environmental wellbeing and human-centred values such as inequality. |
| Discuss their responses using the following question prompts:   * What are the benefits of automation? * What are some of the negative points? * What is the relationship between automation and human wellbeing? | |