

# SIT799 Human Aligned Artificial Intelligence

## Pass Task 2.1: Ethics in AI Quiz

### Overview

During week 2, you have been introduced to Ethics and Ethics in AI. As we have discussed AI can have controversial applications and thus may induce serious ethical issues in its employment. This quiz gives you a chance to demonstrate your understanding of what you have learned.

### Quiz

For each of the following scenarios, indicate the possible ethical issues that you may have to address. **Please justify your answers in no more than 5 sentences.**

1. You are asked to work on a project that aims to develop an AI system that will assist doctors during complex surgeries.  
A: Following points must be considered:
  - The AI system must have continuous human feedback to ensure patient safety
  - There should be accountability mechanism set up to handle potential mishaps during surgery
  - Patient's or family's consent is a must before using such a solution
  - All stakeholders involved must be briefed about the risks associated
  - The software must be peer reviewed and subject to rigorous testing with hard evaluation benchmarks before operating on humans.
2. You are asked to work on a project that aims to build a computer vision system for person identification.  
A: As person identification is a solution that deals with sensitive data, following points must be considered when building a person identification system:
  - Privacy of the public must be considered; such that unauthorised privacy violations do not occur.
  - Solution must be rigorously tested for bias against all groups of society; there should be no bias based on ethnicity, gender, etc.
  - Equal representation of subgroups must be present in data during model training.
  - Data collection should be consensual for training a person identification model.
  - Laws of the country must be adhered to.

3. You are asked to work on a project that aims to build automatic driving vehicles.

A: As automatic driving vehicles must take critical decision to ensure individual wellbeing, following points must be considered:

- System should prioritise pedestrian as well as passenger safety.
- System should be tested rigorously to follow traffic rules.
- There should be a human override option to let human driver take control in case of emergencies.
- System should be tested for response in emergencies before deploying into real world.
- The automatic driving vehicles must consider uncertainties caused by human drivers.

4. You are asked to work on a project that aims to build a speech recognition tool.

A: Speech recognition is a tool which can aid multiple users if deployed in an effective way. Some points to consider when designing such a solution are:

- Training data must be trained on a large sample size with equal representation for subgroups of different countries and regions to account for difference in pronunciation.
- System must be tested rigorously for performance on different language speaker and bias towards any specific linguistic region must be considered.
- System must also consider people with speech impediments like stammering or stuttering to serve this specific subgroup as well.