Artificial Intelligence (AI) HL Practical

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 engender trust and scale AI with confidence.

Here is an example of when AI has failed, the time Microsoft’s AI Chatbot Tay was trolled. Microsoft stood out as truly newsworthy when they reported their new chatbot. Composing with the slang-loaded voice of a teen, Tay could naturally answer to individuals and take part in easy going and lively discussion on Twitter. In actuality, Tay was repeating such offensive statements that were basically said by other human users. Those users were purposely trying to provoke Tay. Since it was programmed to imitate the language patterns of 18-24 old millennials, it was building a conversation by processing phrases of human users and merging in with other data fed to the software. It was programmed to talk and engage with people, which can make her smarter each day. But this unfortunately added to the examples of AI gone wrong and hence Tay was taken offline within 16 hrs.

#### Another example of the wrong approach to AI was in 2016, Uber tested its self-driving cars in San Francisco without taking permissions and approvals from the State. That is ethically and legally not right. Moreover, the internal documents of Uber stated that the self-driving car crossed around 6 red lights in the city during testing putting people in real danger from the experiment.

The rise in AI technologies creates more urgency for organisations to understand the implications of AI empowered decision making and how to ensure AI is being used responsibly. Organisations must think of AI technology in a holistic way – understanding where AI sits in the value chain and creating the right structures to ensure long-term governance by: Establishing internal governance, for example by an objective review panel, that is diverse and that has the knowledge to understand the possible consequences of AI infused systems. A key success factor is leadership support and the power to hold leadership accountable. Ensuring the right technical guardrails, creating quality assurance and governance to create traceability and auditability for AI systems. This is an important part of every organisation’s toolkit to allow operational and responsible AI to scale. Investing more in their own AI education and training so that all stakeholders – both internal and external – are informed of AI capabilities as well as the pitfalls.