

LECTURER:

MARK GILLAN

EXERCISE:

:::AI ALIGNMENT

::: ::: :::AI

EXPLOSION

KNOWLEDGE AND REFLECTION - EXERCISE

■ Short-Answer Questions

- Answer ALL the questions in clear sentences / paragraphs.
- You may use examples where appropriate.
- Write your answers with use of headings for each question onto a Word document, saving the file as:
6 Exercise – Alignment and Explosion
- Questions now follow...



Questions

Question 1 – AI Alignment Problem (Description)

- Describe what is meant by the *AI alignment problem*.

Question 2 – Risks of the AI Alignment Problem

- Describe two risks posed by the AI alignment problem.

Question 3 – Opportunities of AI Alignment

- Describe two opportunities or benefits of solving the AI alignment problem.

Question 4 – AI Explosion (Description)

- Describe what is meant by a *potential AI explosion*.

Question 5 – Risks of a Potential AI Explosion

- Describe two risks associated with a potential AI explosion.

Question 6 – Opportunities of a Potential AI Explosion

- Describe two opportunities or benefits associated with rapid AI development.

Question 7 – Connection Between Alignment and Explosion

- Explain why AI alignment becomes more important if an AI explosion were to occur.