

COMP2400/6400: Intelligent Machines, Ethics and Law
SGTA Task
Week 5: AI and Warfare

Points of Reference

- Lecture in W4 on “Blame and Responsibility in AI Systems” held on 17 March 2025.
- Required Reading 1: Santoni de Sio, F., & Mecacci, G. (2021). Four responsibility gaps with Artificial Intelligence: Why they matter and how to address them. *Philosophy & Technology*, 34(4), 1057–1084. <https://doi.org/10.1007/s13347-021-00450-x>
- Required Reading 2: Hindriks, F., & Veluwenkamp, H. (2023). The risks of autonomous machines: From responsibility gaps to control gaps. *Synthese*, 201(1), 21. <https://doi.org/10.1007/s11229-022-04001-5>

Please reply to the following questions in your own words by critically and argumentatively reflecting on the content of the required readings:

1. Santoni de Sio and Mecacci (2021) identify four responsibility gaps: a culpability gap, a moral accountability gap, a public accountability gap, and an active responsibility gap. Out of these, choose two gaps that you find most relevant and interesting. How can these gaps be characterised? Do you agree or disagree with Santoni de Sio and Mecacci’s (2021) assessment of these gaps? Provide reasons for your evaluation.*
2. Hindriks and Veluwenkamp (2023) deny that the employment of AI systems can create responsibility gaps. Instead, they argue that the problems identified by previous research are better described as control gaps. Reconstruct and critically discuss their argument. Is their argument valid and plausible?*
3. Santoni de Sio and Mecacci (2021) and Hindriks and Veluwenkamp (2023) agree that the realisation of ‘meaningful human control’ is a feasible strategy for mitigating problems arising from human-AI interactions. Describe the concept of ‘meaningful human control’ and show how it can be applied to cases of either algorithmic decision-making or automated driving.

***Replies to asterisked (starred) questions are mandatory.**

Suggested word count: 150 words per question.

Restriction: You are **not allowed to use any AI based tools** (such as ChatGPT) to reply to these questions. This includes directly copying and paraphrasing artificially generated text. It also includes the integration of chatbot-generated revisions of text that you have drafted yourself. Presenting work created by AI technologies as your own is academically dishonest and any submission doing so will be referred to the Macquarie University Academic Integrity Disciplinary Committee. The submitted text must be your own academic work.