

IISL COLLOQUIUM ON THE LAW OF OUTER SPACE (E7)
Young Scholars Session with Keynote Lecture (1)

Author: Mr. Thomas Graham
Swinburne University of Technology, Australia, thomasgraham@swin.edu.au

Dr. Kathiravan Thangavel
Royal Melbourne Institute of Technology (RMIT), Australia, kathiravan.thangavel@student.rmit.edu.au

ARTIFICIAL INTELLIGENCE IN SPACE: AN ANALYSIS OF RESPONSIBLE AI PRINCIPLES FOR
THE SPACE DOMAIN**Abstract**

Advances in Artificial Intelligence (AI) technologies are enabling a plethora of new applications across many industries. There are already a multitude of applications for AI systems in the space industry, but as the AI and Space industries continue to rapidly grow in size and value, further use-cases will become apparent and proliferate to all corners of space operations and data analytics. Such space-based AI systems will bring many economic, scientific and environmental benefits, however they could also enable harms to individuals, organisations, and the environment if there are not developed properly. Potential breaches privacy through AI-assisted analysis of Earth observation imagery and collisions between objects in orbit due to malfunctioning automated maneuvering systems are examples of the harms that could eventuate in the event of poorly designed AI systems being deployed in the space-sector. Responsible AI practices are needed to ensure such risks do not eventuate. ‘Responsible’ (or ‘ethical’) AI has emerged as a discipline designed to guide responsible AI development through the seeking to maximise the benefits of AI systems for individuals and society while mitigating against any potential harms that they may cause. Commonly accepted Responsible AI principles include: accountability, contestability, fairness, security, privacy, transparency, explainability, and reliability, with notions of ‘do-no-harm’ and generating net-benefits for society and the environment sometimes also being included. These principles of Responsible AI are generalizable and industry agnostic, however they should be carefully considered in the context of the unique physical, economic, political, and technological characteristics of the space domain before being adopted wholesale by the space industry. While concepts such as security and reliability can be readily applied to applications of AI systems in the space domain, other ideals such as contestability, fairness, and explainability may not be as relevant to the use cases found within the space industry. This paper reviews a range of ‘Responsible’, ‘Ethical’, and ‘Trusted’ AI literature setting out guiding principles and governance mechanisms for AI systems and then critically examines the applicability and appropriateness of widely accepted Responsible AI principles in the context of the space industry. This serves as a first step towards creating a standardized regulatory framework for the responsible development of space-based AI systems and preventing harms associated with such systems occurring.