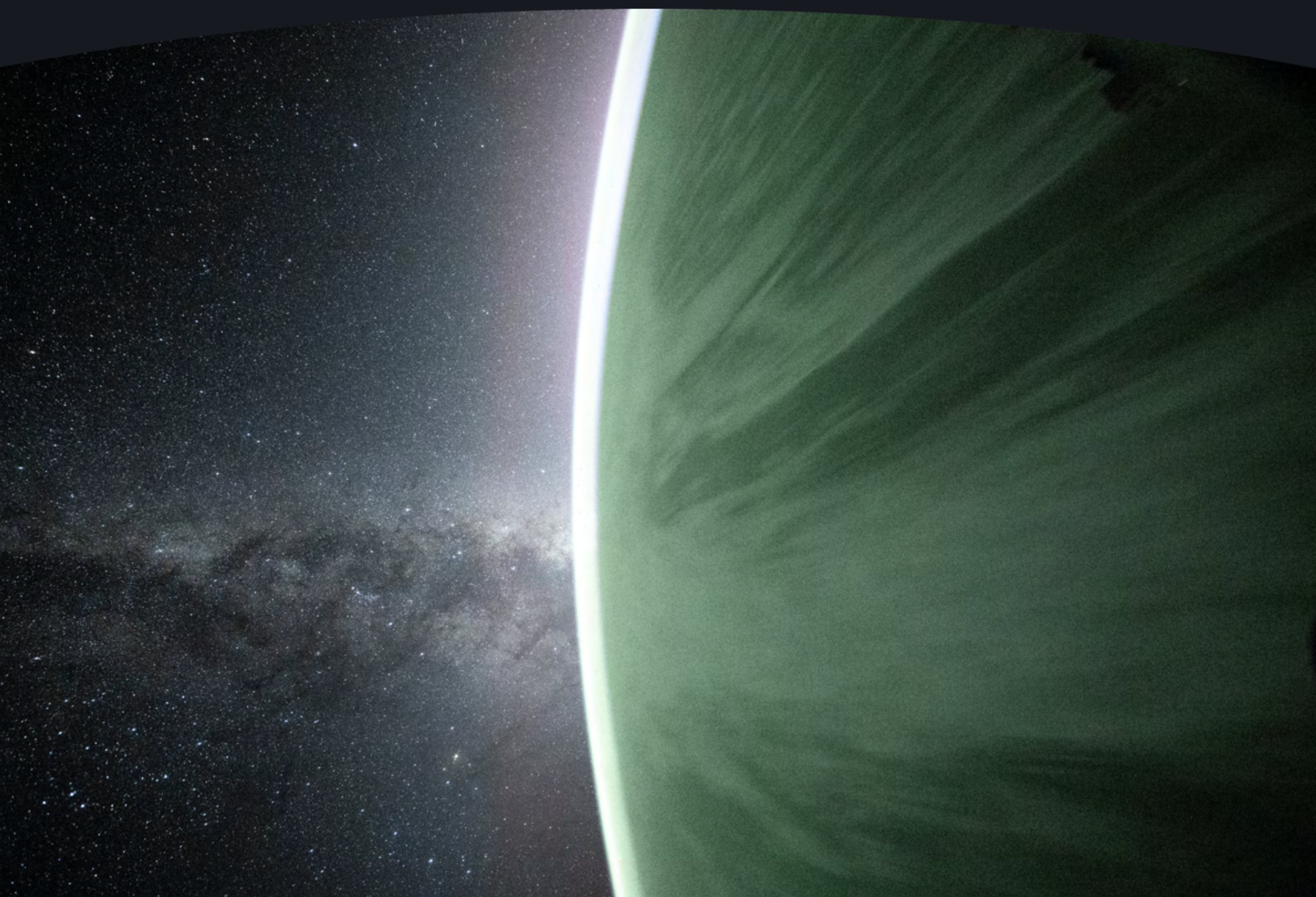
JUNE 2025 - SPACE SECURITY & NEW TECHNOLOGIES 1

Visionary Versus Reactionary.
The future of space security in the age of artificial intelligence

IOANA BRATU · ALMUDENA AZCÁRATE ORTEGA



ACKNOWLEDGEMENTS

Support from UNIDIR core funders provides the foundation for all of the Institute's activities. This report was made possible thanks to the financial support of the United Kingdom. UNIDIR's Sarah Erickson, James Revill, Louison Mazeaud, Yasmin Afina, Ritanshu Lohani and Chelsea Mai provided invaluable advice, support, and assistance on this paper. The authors would also like to acknowledge in particular Laetitia Cesari, Nayoung Youn, P.J. Blount, and Diego Álvarez for their critical feedback.

Photo credit for photographs in pages 1, 6, 9, 15, 19, 24, 39: NASA.
Design and layout by Trifecta Content Studio.

ABOUT UNIDIR

UNIDIR is a voluntarily funded, autonomous institute within the United Nations. One of the few policy institutes worldwide focusing on disarmament, UNIDIR generates knowledge and promotes dialogue and action on disarmament and security. Based in Geneva, UNIDIR assists the international community to develop the practical, innovative ideas needed to find solutions to critical security problems.

NOTE

CITATION

Ioana Bratu & Almudena Azcárate Ortega. "Space Security & New Technologies 1 – Visionary Versus Reactionary. The Future of Space Security in the Age of Artificial Intelligence", UNIDIR, Geneva, 2025. **https://doi.org/10.37559/WMD/25/Space/03.**

CONTENTS

ABOUT THE AUTHORS

IOANA BRATU

Co-Director of Amsterdam Law & Technology Institute, Assistant Professor in Space Law and AI Regulation at Vrije University Amsterdam, and Founder of AI ∞ Space Law Society

Ioana's PhD, "Charting the Uncharted Waters: Non-Contractual Liabilities for Damage Caused By GNSS-Enabled Artificial Intelligence. Legal challenges in international space law and European tort law", as well as her prior publications, were the first to critically address the emerging intersection between the laws governing human activities in outer space and the opportunities and challenges posed by the adoption of artificial intelligence.

Prior to joining academia, Ioana worked as a qualified (European) attorney-at-law for 10 years. She practiced law in international law firms and as a founder of her own boutique law firm. She holds an Executive MBA from Vienna University of Economics and Business.

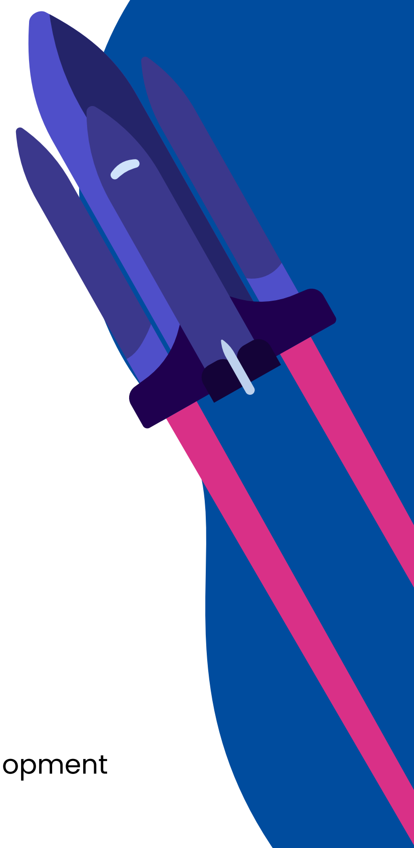
ALMUDENA AZCÁRATE ORTEGA

Space Security Researcher at the United Nations Institute for Disarmament Research (UNIDIR)

Almudena has published widely and briefed UN Member States on the topics of space security law and policy and has presented her research in multiple fora. She has led UNIDIR's participation in the UN Open-Ended Working Group on Reducing Space Threats Through Norms, Rules and Principles of Responsible Behaviour, established pursuant to UN General Assembly resolution 76/231.

Prior to joining UNIDIR, Almudena was a Research Assistant at Georgetown University Law Center, where she is currently a doctoral candidate. She holds an LL.M. in National Security Law from the same institution, where she was the recipient of Georgetown's Thomas Bradbury Chetwood, S.J. Prize for the most distinguished academic performance in the programme. She received her LL.B. from the University of Navarra, Spain.

ABBREVIATIONS AND ACRONYMS



AI	Artificial Intelligence
AU	African Union
ASAT	Anti-satellite
ESA	European Space Agency
EU	European Union
G20	Group of 20
GEO	Geostationary orbit
GNSS	Global Navigation Satellite Systems
GPS	Global Positioning System
IHL	International humanitarian law
INF	Intermediate-Range Nuclear Forces Treaty
ISR	Intelligence, Surveillance and Reconnaissance
LAWS	Lethal autonomous weapons systems
LEO	Low Earth orbit
NATO	North Atlantic Treaty Organization
MoD	Ministry of Defense
NASA	National Aeronautics and Space Administration
OECD	Organisation for Economic Co-operation and Development
OST	Outer Space Treaty
PAROS	Prevention of an Arms Race in Outer Space
RPO	Rendezvous and proximity operations
SATCOM	Satellite communications
SSA	Space situational awareness
STM	Space traffic management
UAV	Uncrewed aerial vehicle
UN	United Nations
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNIDIR	United Nations Institute for Disarmament Research

INTRODUCTION

INTRODUCTION

Humankind relies on space-based infrastructure for a multitude of uses and services.[1] As such reliance continues to grow, so do the complexities of maintaining the security of these space assets. Space activities are expanding at an unprecedented pace, with a record number of space objects launched each year,[2] leading to vast quantities of data produced by space infrastructures. The number, precision, and sophistication of potential threats, ranging from cyber disruptions faced by satellite control systems to direct-ascent anti-satellite weaponry, are also rapidly growing, thereby underscoring the urgent need for advanced and diversified security measures.

In this regard, artificial intelligence (AI) emerges as a novel tool that can potentially transform how space security is conceptualized and operationalized. AI is arguably one of the most transformative technologies of the twenty-first century, despite many of its applications still being in their infancy, and it has impacted every walk of life.[3] Thus it is not surprising that States and other stakeholders have also sought to harness its potential to further their own interests in outer space.

AI offers many opportunities in outer space, such as new ways to collect, process, and interpret data, thereby improving situational awareness and response times, enhancing predictive capabilities, optimizing resource allocation, and enabling more effective threat mitigation strategies. Moreover, AI-enabled algorithms can be used in the context of satellite applications to help process data related to the monitoring of human activities, environmental changes, and natural disasters to better support emergency preparedness and response plans. Additionally, AI can be effectively used to monitor satellite parameters and health. Machine learning algorithms can analyse telemetry data in real time to detect anomalies, predict component failures, and optimize system performance. By identifying subtle patterns or deviations from nominal behaviour, AI can enhance predictive maintenance and reduce the risk of mission failure.

However, AI also presents challenges that the international community must bear in mind. While different stakeholders have been exploring both the beneficial applications of AI and its potential threats to space security,[4] the topic has not yet been explored in multilateral debates in any detail.

1 *Report of the Group of Governmental Experts on further practical measures for the prevention of an arms race in outer space,* ¶ 12, U.N. Doc. GE-PAROS/2024/CRP.4, (23 Aug. 2024).

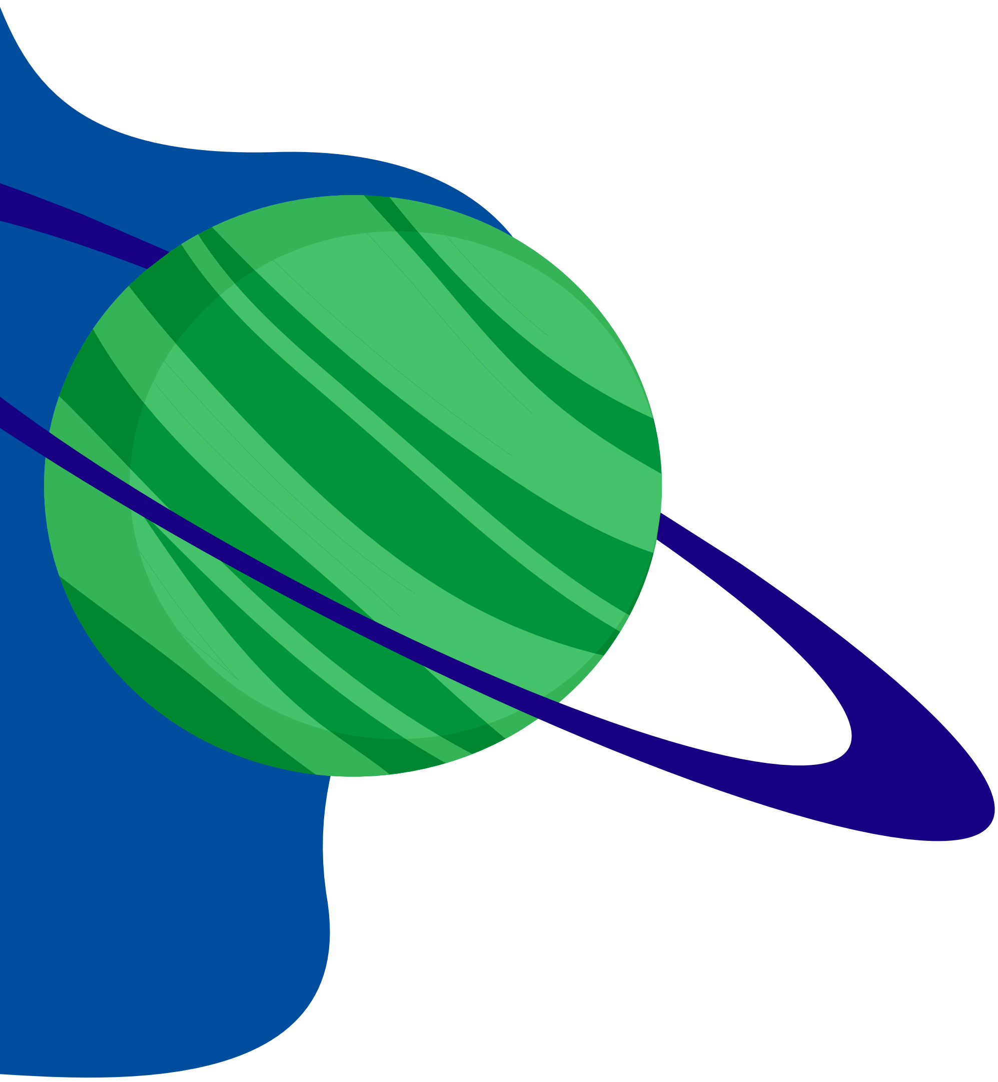
2 Edouard Mathieu, *A Record Number of Objects Went into Space in 2023,* Our World in Data (11 Mar. 2024), https://ourworldindata.org/data-insights/a-record-number-of-objects-went-into-space-in-2023.

3 Jessica Harris, *How Artificial Intelligence is Transforming the World*, Brookings, https://www.brookings.edu/articles/how-artificial-intelligence-is-transforming-the-world.

4 *Artificial Intelligence in Space,* European Space Agency [ESA] https://www.esa.int/Enabling_Support/Preparing_for_the_Future/Discovery_and_Preparation/Artificial_intelligence_in_space; *Artificial Intelligence,* National Aeronautics and Space Administration [NASA], https://www.nasa.gov/artificial-intelligence; *Chinese Wonder Journey Satellite Takes AI to Space,* Orbital Today, https://orbitaltoday.com/2023/08/22/chinese-wonderjourney-satellite-takes-ai-to-space.

The emergence of AI, as well as the increasing interest its continued development and rapid improvement have generated questions —among States, non-governmental entities, and technical experts alike— regarding the future of the implementation of AI in space operations, and the impact that this might have on space security. These questions become especially pressing considering the lack of definite regulation on this issue, particularly at the international level.

In light of the above, this report examines the intersection of space security and AI, providing insights into the emerging synergy between the two fields (Part I), and providing context on the policy and law relevant to said synergy (Part II). The report also showcases the opportunities and challenges in the application of AI to outer space activities through the presentation of a series of practical case studies (Part III). To conclude, the report looks to the future in order to assess potential next steps that the international community might take to adapt to the emerging synergy of space security and AI in a manner that facilitates the integration of AI in space security in accordance with the peaceful purposes principle of space activities, fostering both technological progress and long-term sustainability in outer space (Part IV).



PART I – DEFINITIONAL CONTEXT

PART I – DEFINITIONAL CONTEXT

DEFINING THE EMERGING SPACE SECURITY-AI SYNERGY

The notion of space security is commonly understood to encompass measures designed to prevent deliberate harms to a space system, including its component parts, from intended or intentional threats undertaken by another actor. Moreover, space security is concerned with the relationship among space objects and activities, and the maintenance of international peace and security, and disarmament, including the prevention of an arms race in outer space.[5]

As reliance on space-based assets increases, AI is emerging as a new technology that can alter the conceptualization and operationalization of space security by offering new mechanisms to collect, interpret, and understand data, which can make threat mitigation more effective. Machine learning, which refers to the development of algorithms that can identify patterns and make predictions from data,[6] and deep learning, a specialized methodology that uses multi-layered neural network architectures to model complex representations,[7] enable rapid analysis of large data sets and extraction of meaningful insights. In the context of space security, such findings can facilitate early warning of potential threats, such as unanticipated manoeuvres of a nearby satellite or anomalies in ground-based telemetry that may indicate malicious interference.

Furthermore, AI can improve the reliability of decision-making processes in space operations. In situations where terrestrial communication may be restricted, an AI-enabled system could independently detect anomalous activities, such as signal jamming or spoofing.[8] The ability for swift, autonomous decision-making is especially pertinent due to the urgent

5 *See* Almudena Azcárate Ortega & Victoria Samson (eds.), *A Lexicon for Outer Space Security*, UNIDIR 3.3.8 (2023) [hereinafter 'Space Security Lexicon'], https://doi.org/10.37559/WMD/23/Space/05. The concept of 'space security' is different to 'space safety', which is commonly understood to refer to measures aimed at preventing accidental or unintentional hazards to space systems. These hazards can be natural, such as geomagnetic storms, or stem from human-made objects, such as the accidental malfunctioning of a satellite, or collision with a piece of debris. Space safety measures therefore seek to mitigate any non-intentional damage to a space system. The possibility of such a damage is considered a risk (as opposed to a threat). Issues of space safety are generally understood to be a part of the broader topic of the peaceful uses of outer space, which is discussed at the Committee on the Peaceful Uses of Outer Space (COPUOS) and under the purview of the General Assembly's Fourth Committee. Space security and safety are generally viewed as separate and distinct, although the two are interrelated, and can intersect and overlap. Certain languages do not differentiate between 'safety' and 'security', and therefore can generate confusion in distinguishing the two. See Space Security Lexicon, at 3.3.7.

6 Issam El Naqa & Martin J Murphy, *What Is Machine Learning?,* in Machine Learning in Radiation Oncology: Theory and Applications 6 (Issam El Naqa, Ruijiang Li & Martin J. Murphy eds., 2015).

7 Andrew Saxe, Stephanie Nelli & Christopher Summerfield, *If Deep Learning Is the Answer, What Is the Question?,* 22(1) Nature Reviews Neuroscience 55 (2021).

8 *See*, e.g., Part III, GNSS Interferences, Case Study, below.

nature of specific space security threats. A kinetic weapon threatening a satellite in low Earth orbit (LEO) may allow only minutes for a response, rendering any delay, such as awaiting directives from ground stations, potentially disastrous.

Autonomous decision-making processes are not always positive, and can entail significant challenges, however. AI technologies, due to their probabilistic nature, can function unpredictably. They can also be intentionally misused or repurposed to harm. Moreover, an overdependence on AI technologies coupled with a lack of technical literacy can pose a challenge in relation to the system's reliability, compliance with international law and ethical requirements.[9] Lethal autonomous weapon systems (LAWS) have been highlighted as autonomous technologies that can raise concerns from humanitarian, human rights, legal, security, technological and ethical perspectives.[10]

Yet, what truly distinguishes AI's role in space security from other domains are the intrinsic characteristics of the space environment. Outer space remains a harsh and remote setting, which restricts opportunities for AI training, and the high-performance microprocessors necessary for sophisticated AI applications have yet to achieve the radiation tolerance required for reliable operations beyond Earth's atmosphere. Natural phenomena like solar flares, and debris collisions add further layers of unpredictability, while the potentially unexpected behaviour of AI systems equipped with self-learning capabilities can trigger unpredictable outcomes, which may lead into operational anomalies or unintended conflict escalations.[11] Such challenges are further enhanced by the absence of a clear regulatory environment governing AI-driven activities in outer space. The international legal frameworks governing human activities in outer space, such as the Outer Space Treaty, or recent technology frameworks, such as the European AI Act, do not explicitly address AI-based threats or the rules of engagement for AI-controlled systems in orbit. Thus, currently, there are no legally binding rules that explicitly relate to the deployment of AI in an outer space context.

Considering both opportunities and challenges, situating the critical analysis of AI within a broader conceptual framework of space security will pave the way for effective adversarial threat categorization and response as well as formulation of emerging space frameworks capable of standing the test of time. To anchor this analysis and support coherent future policy development, this report adopts the following definition:

> "The emerging space security–AI synergy refers to strategic coordination of autonomous prediction and decision-making tools, mission planning methodologies, cross-sectoral protocols, and adaptive and anticipatory legal frameworks, fostering safe, secure and sustainable operations beyond Earth's planetary boundaries".

9 Yasmin Afina, *The Global Kaleidoscope of Military AI Governance,* UNIDIR 14-17 (2024), https://unidir.org/publication/the-global-kaleidoscope-of-military-ai-governance.
10 UN Secretary-General, *Report on Lethal Autonomous Weapons Systems,* ¶¶ 1–2, U.N. Doc. A/79/88 (1 July 2024).
11 *See*, e.g., Part III, Space Debris Collision Avoidance, Case Study, below. *See also,* e.g., Part III, Generative AI Models, Case Study, below.

FRAMING ARTIFICIAL INTELLIGENCE

Despite early technological evolutions of AI which can be traced back to 1950 (the Turing test),[12] there is no universally accepted definition of AI. Some States have sought to define this concept at the domestic or regional level, mostly through policies,[13] although in some instances through the establishment of binding law specifically centred on AI, as is the case of the European Union, in its European AI Act, adopted in 2024,[14] and Brazil, in its Bill on the use of AI (not yet adopted).[15]

---

12 Andrew Hodges, *Alan Turing and the Turing Test, in* Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer (Robert Epstein, Gary Roberts and Grace Beber eds., 2009).

13 Some examples of policies include the following:

- United Kingdom – "Machines that perform tasks normally requiring human intelligence, especially when the machines learn from data how to do those tasks" *UK National AI Strategy 2021,* Department of science innovation andTechnology,).
- It should be noted that the UK has also set out a legal definition of AI in its National Security and Investment Act *See* National Security and Investment Act §100 (2021) https://assets.publishing.service.gov.uk/media/603d4003d3bf7f02168eabdc/nsi-scope-of-mandatory-regime-gov-response.pdf
- India – "constellation of technologies that enable machines to act with higher levels of intelligence and emulate the human capabilities of sense, comprehend and act." NITI Aayog, *Responsible AI #AIForAll Approach Document for India Part 1 – Principles for Responsible AI,* https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf.
- Chile – "el conjunto de técnicas informáticas que permiten a una máquina (por ejemplo, un ordenador, un teléfono) realizar tareas que, por lo común, requieren inteligencia tales como el razonamiento o el aprendizaje" [the set of computing techniques that allow a machine (e.g., a computer, a telephone) to perform tasks that typically require intelligence, such as reasoning or learning] and "un sistema computacional que puede, para un determinado conjunto de objetivos definidos por humanos, hacer predicciones y recomendaciones o tomar decisiones que influyen en entornos reales o virtuales. Los sistemas de IA están diseñados para operar con distintos niveles de autonomía" [A computational system that can, for a given set of human-defined objectives, make predictions and recommendations or take decisions that influence real or virtual environments. AI systems are designed to operate with varying levels of autonomy]. *Chile Política Nacional de Inteligencia Artificial* 2021, ministerio de ciencia, tecnología, conocimiento e innovación, https://minciencia.gob.cl/uploads/filer_public/bc/38/bc389daf-4514-4306-867c760ae7686e2c/documento_politica_ia_digital_.pdf.

For more information, visit the UNIDIR AI Policy Portal, https://aipolicyportal.org.

14 Commission Regulation 2024/1689 of the European Parliament and of the Council of 13 June 2024, Laying down Harmonised Rules on Artificial Intelligence and Amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), https://eur-lex.europa.eu/eli/reg/2024/1689/oj. See article 3: "a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments".

15 Bill on the Use of Artificial Intelligence, 2338, § 4 (2023), https://clairk.digitalpolicyalert.org/documents/brazil-bill-on-the-use-of-artificial-intelligence-2338-2023-original-language/?refs=chapter-i-preliminar,art. *See* article 4: "sistema computacional, com graus diferentes de autonomia, desenhado para inferir como atingir um dado conjunto de objetivos, utilizando abordagens baseadas em aprendizagem de máquina e/ou lógica e representação do conhecimento, por meio de dados de entrada provenientes de máquinas ou humanos, com o objetivo de produzir previsões, recomendações ou decisões que possam influenciar o ambiente virtual ou real" [computer system, with different degrees of autonomy, designed to infer how to achieve a given set of objectives, using approaches based on machine learning and/or logic and knowledge representation, through input data from machines or humans, with the aim of producing predictions, recommendations or decisions that can influence the virtual or real environment].

While different, the definitions established by the various instruments reflect the challenges of encapsulating the diverse and evolving nature of AI systems. These can be categorized broadly into two categories:

- **Narrow AI**, also known as weak AI, involves systems designed to perform specific tasks and requires extensive human oversight in programming and configuration to maintain their intelligence. Examples include expert systems that operate based on explicit rules formulated by humans for stable environments and mimic expertise and reasoning processes in a given field.
- In contrast, **general AI, or strong AI**, (also widely known as '**artificial general intelligence**') refers to systems that demonstrate a high level of autonomy and adaptability comparable to human intelligence. These systems can adapt their behaviour through mechanisms like transfer learning, allowing them to apply knowledge from one context to a variety of others, signalling a move toward more sophisticated AI capabilities."[16]

All these definitions underscore the relevance of AI systems across various sectors due to their varied functionalities. AI applications range from deployment in areas such as disaster zones, oil rigs, or nuclear facilities to energy infrastructures, climate change, etc. In the space sector, the introduction of Deep Space 1 in 1998, the first experimental fully autonomous spacecraft, marked a turning point, with AI-powered autonomous systems driving breakthroughs in space technology and its applications.[17] Since then, AI has been increasingly integrated into various aspects of space exploration and operations, enabling advanced functionalities and enhancing mission efficiency. For instance, AI algorithms are utilized in satellite image processing to monitor environmental changes, track deforestation, and analyse natural disasters.[18] These systems can process vast amounts of data from Earth observation satellites more efficiently than traditional methods, providing timely insights for decision-making. Similarly, semi-autonomous navigation systems, such as those employed by the Mars rovers, allow these robotic explorers to independently analyse terrain, plan routes, and avoid obstacles, significantly reducing reliance on real-time human intervention.

16 Ben Goertzel, *Artificial General Intelligence: Concept, State of the Art, and Future Prospects,* 5(1) Journal of Artificial General Intelligence 1, 1–2 (2014).

17 North Atlantic Treaty Organization [NATO], *Science and Technology Trends 2020–2040,* 60 (Mar. 2020), https://www.nato.int/nato_static_fl2014/assets/pdf/2020/4/pdf/190422-ST_Tech_Trends_Report_2020-2040.pdf.

18 An example of this is the use of AI algorithms on the EU's Copernicus Earth observation data. *See OBSERVER: Revealing Hidden Land Patterns with AI and Copernicus, Copernicus* (9 May 2024), https://www.copernicus.eu/en/news/news/observer-revealing-hidden-land-patterns-ai-and-copernicus.

Additionally, AI has been applied in mission planning and spacecraft operations.[19] Onboard AI systems enable spacecraft to prioritize scientific observations and adjust schedules without human input. In space debris management, AI systems are being developed to predict collision risks and optimize the manoeuvres of satellites to avoid potential impacts, as well as to monitor the health of both space systems and astronauts. The use of AI also extends to space communication systems, where machine learning algorithms enhance signal processing and data transmission, improving the efficiency of deep-space communications.[20]

In the military field, AI has been increasingly integrated to enhance operational capabilities, decision-making processes, and resource management.[21] AI applications in defence encompass autonomous systems for reconnaissance and surveillance, advanced data analytics for threat detection, and strategic planning. For instance, AI-enabled drones are employed for intelligence, surveillance, and reconnaissance missions, operating autonomously to analyse vast areas, identify potential threats, and transmit real-time data to command centres. Similarly, AI-driven image recognition systems are utilized to detect and classify objects, such as vehicles or military equipment, from satellite imagery or video feeds, thereby enhancing situational awareness.[22]

Considering such dynamic developments, the future of AI remains unpredictable. Projections suggest that global investment in AI, encompassing software, hardware, and services, will surpass $300 billion by 2026.[23] This growth is expected to drive an increase in the deployment of AI applications while also intensifying the challenges associated with its regulatory environment.

---

19 For examples of how various space agencies are using AI, *see Artificial Intelligence in Space,* European Space Agency, https://www.esa.int/Enabling_Support/Preparing_for_the_Future/Discovery_and_Preparation/Artificial_intelligence_in_space.

20 For an overview of AI applications in space, *see generally* Antonia Russo & Gianluca Lax, *Using Artificial Intelligence for Space Challenges: A Survey*, 12(10) Applied Sciences 5106 (2022); Svetla Ben-Itzhak, *The Use of Artificial Intelligence in Outer Space Capabilities,* 44(2) SAIS Review of International Affairs 17 (2024); Justin Goodwill, Christopher Wilson & James MacKinnon, *Current AI Technology in Space,* NTRS-NASA Technical Reports Server (2024), https://ntrs.nasa.gov/api/ citations/20240001139/downloads/Current%20Technology%20in%20Space%20v4%20Briefing.pdf; Bernard Marr, *Artificial Intelligence in Space: The Amazing Ways Machine Learning Is Helping to Unravel the Mysteries of the Universe,* Forbes (10 Apr. 2023), https://www.forbes.com/sites/bernardmarr/2023/04/10/artificial-intelligence-in-space-the-amazing-ways-machine-learning-is-helping-to-unravel-the-mysteries-of-the-universe.

21 *See* Yasmin Afina, *supra note* 9, at 11-13.

22 For a detailed overview of such AI applications, *see* Sarah Grand-Clément, *Artificial Intelligence Beyond Weapons: Application and Impact of AI in the Military Domain,* UNIDIR 14–20 (11 Oct. 2023), https://unidir.org/publication/artificial-intelligence-beyond-weapons-application-and-impact-of-ai-in-the-military-domain.

23 *Cf. Worldwide Spending on AI-Centric Systems Will Pass $300 Billion by 2026, According to IDC,* International Data Corporation (23 Apr. 2023), https://my.idc.com/getdoc.jsp?containerId=prUS52530724.

PART II – SITUATIONAL CONTEXT

PART II – SITUATIONAL CONTEXT

SPACE SECURITY AND AI IN THE MULTILATERAL CONTEXT – WHY DOES THIS DISCUSSION MATTER?

At the multilateral level, outer space security discussions take place under the umbrella of the concept of the Prevention of an Arms Race in Outer Space (PAROS).[24] The idea of PAROS emerged in 1978,[25] and was added as an item to the Conference on Disarmament's agenda in 1981. Since then, issues related to PAROS and space security more generally have been present in multilateral debates in multiple fora, including the Conference on Disarmament, the First Committee of the General Assembly, the Disarmament Commission, and ad hoc groups of governmental experts or open-ended working groups created for the purposes of discussing specific matters.[26]

In recent years, technological advances and their security implications have become increasingly salient in discussions around space security. However, AI has not been featured as a concern despite growing attention to the convergence of space and other technologies. In contrast, discussion in the United Nations High-level Advisory Body on Artificial Intelligence recognizes the opportunities and challenges offered by AI and have highlighted the importance of "a common understanding of its capabilities, opportunities, risks and uncertainties". However, that body did not make the connection to space security related issues directly and instead addresses the general topic of opportunities and challenges offered by AI.[27]

While the international community has recognized that AI can be a driver of economic growth and improve life for humanity, it is also aware that AI can present notable challenges to peace and security, especially in the military domain. In this context, the General Assembly has acknowledged in its resolution 79/239 that potential opportunities exist in "the areas of compliance with international humanitarian law, including protection of civilians and civilian objects in armed conflict", as well as possible challenges that may arise from the following:

---

24 *See* Almudena Azcárate Ortega & Hellmut Lagos Koller, *The Open-Ended Working Group on Reducing Space Threats through Norms, Rules and Principles of Responsible Behaviours: The Journey So Far, and the Road Ahead*, 48 Air and Space Law 19, 23 (2023).

25 G.A. Res. S-10/2, Final Document of the Tenth Special Session of the General Assembly, ¶ 80 (30 June 1978).

26 Erin Pobjie & Almudena Azcárate Ortega, *Space Security Legal Primer 1 – Outer Space & Use of Force*, UNIDIR 12 (2024), https://unidir.org/ publication/outer-space-use-of-force.

27 High-Level Advisory Body on Artificial Intelligence, *Governing AI for Humanity* (Sep. 2024), https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf.

"Humanitarian, legal, security, technological and ethical perspectives, as well as the possible impact of such applications on international security and stability, including the risk of an arms race, miscalculation, lowering the threshold for conflict and escalation of conflict, proliferation to non-State actors, and also noting the possible consequences with regard to, inter alia, gender, racial, age or social aspects that could potentially be caused by bias in datasets or other algorithmic biases of artificial intelligence.[28]"

Moreover, while acknowledging that not all LAWS are AI-enabled, LAWS have been recognized as requiring particular attention in this sense,[29] in particular through the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems.[30] The international community has recognized that LAWS pose concerns including from, as indicated above, humanitarian, human rights, legal, security, technological and ethical perspectives, all of which must be addressed in a comprehensive manner.[31]

Achieving common understanding regarding AI applications in relation to space systems is indeed a critical necessity: the crucial nature of space systems for humankind, coupled with the complexity of AI can, if not properly understood, encourage a security dilemma situation where threat perception is heightened by the development of AI-enabled space technologies, thus heightening tensions that could eventually lead to conflict.

LEGAL AND NORMATIVE FRAMEWORK

OUTER SPACE

AI is not explicitly mentioned in international space treaties and regulations. Nevertheless, it is important to note that the legal and normative framework currently applicable to outer space already establishes rights and obligations, as well as limitations that impact the uses of AI in relation to space systems, which must be borne in mind when considering the space security implications of the use of AI.

28 G.A. Res. 79/239, Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security (24 Dec. 2024).

29 *See* G.A. Res. 78/241, Lethal Autonomous Weapons System (22 Dec. 2023).

30 Convention on Certain Conventional Weapons – Group of Governmental Experts on Lethal Autonomous Weapons Systems, United Nations Office for Disarmament Affairs (2024), https://meetings.unoda.org/ccw-/convention-on-certain-conventional-weapons-group-of-governmental-experts-on-lethal-autonomous-weapons-systems-2024.

31 UN Secretary-General, *Lethal Autonomous Weapons Systems,* U.N. Doc. A/79/258 (1 July 2024), https://documents.un.org/ doc/undoc/gen/n24/154/32/pdf/n2415432.pdf.

Outer Space Treaty

The 1967 Outer Space Treaty (OST),[32] which was negotiated and entered into force with a key goal of acting as an arms control mechanism,[33] constitutes the basis of all space law. While space security is not the main focus of the text of the OST —even though security is the raison d'être of the treaty[34]— the principles enshrined therein are relevant to space security and they must be taken into consideration when assessing the nexus between AI and space systems. Particularly relevant are article IV —the only security-specific article in the Treaty— and article III, which establish the applicability of the full corpus of international law to activities involving the use and exploration of outer space. As AI technologies become more capable and can be integrated into disruptive capabilities, for instance, to enhance the precision, autonomy, or decision-making of missile systems, questions arise about how their use aligns with the Treaty's intent to prohibit the placement in outer space of nuclear weapons and the militarization of celestial bodies and, subsequently, its broader implications for fostering peaceful purposes.

- **Article IV** establishes that States shall "undertake not to place in orbit around the Earth any objects carrying nuclear weapons or any other kinds of weapons of mass destruction, install such weapons on celestial bodies, or station such weapons in outer space in any other manner". Moreover, it prohibits the "establishment of military bases, (...) the testing of any type of weapons and the conduct of military manoeuvres on celestial bodies".

- **Article III** stipulates that "States Parties to the Treaty shall carry on activities in the exploration and use of outer space, including the moon and other celestial bodies, in accordance with international law, including the Charter of the United Nations, in the interest of maintaining international peace and security and promoting international co-operation and understanding".

32 Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies, 27 Jan. 1967, 18 UST 2410; 610 UNTS 205; 6 ILM 386 [hereinafter 'Outer Space Treaty' or 'OST'].

33 US President Lyndon B. Johnson called the OST "the most important arms control development since the Limited Test Ban Treaty of 1963"; *see* Lyndon Johnson, Statement by the President Announcing the Reaching of Agreement on an Outer Space Treaty, 8 Dec. 1966, in Lyndon B. Johnson: Containing the Public Messages, Speeches, and Statements of the President: 1966 (in two books), [Book II] 1441 (1967), available online at https://babel.hathitrust.org/cgi/pt?id= miua.4731549.1966.002&seq=815 . *See also* Francis Lyall & Paul B. Larsen, Space Law: A Treatise 453-454 (2nd ed. 2018); *See* Jessica West & Almudena Azcárate Ortega, *Space Dossier 7 – Norms for Outer Space: A Small Step or a Giant Leap for Policymaking?,* UNIDIR 7 (Mar. 2022), https://doi.org/10.37559/WMD/22/Space/01.

34 The General Assembly expressed the need "to avoid the extension of present national rivalries into this new field"; *See* G.A. Res. 1348 (XIII), U.N. GAOR, 13th Sess., Supp. No. 18 at 1, U.N. Doc. A/4090 (13 Dec. 1958), https://www.unoosa.org/ oosa/oosadoc/data/resolutions/1958/general_assembly_13th_session/res_1348_xiii.html.

Liability Convention

The Liability Convention, which entered into force in 1972,[35] does not prohibit the use of any technologies in space, but it establishes that a State can be liable for launching an object that causes damage to another State's assets, either on Earth, under an absolute liability regime **(article II)**, or in space, under a fault-based liability regime **(article III)**. Under the Liability Convention, the concept of space object includes "component parts of a space object as well as its launch vehicle and parts thereof" **(article I)**, which must be taken into consideration when assessing AI integration and its implications. AI introduces a layer of abstraction that can make it harder to pinpoint negligence or error, even if that system is reliable, tested, and updated to meet evolving standards. When AI is used for collision avoidance systems or space traffic management (STM), the legal implications become nuanced, especially if an AI system fails to prevent a collision or makes an autonomous decision that leads to damage. When AI is involved, determining 'fault' is challenging, as well as the distribution of accountability between the operator, the developer of the AI, and the State that licensed the launch.

General International Law

As per articles III and I of the OST, international law applies to outer space activities. The different areas of general international law provide a framework that must be taken into account in the context of space security, including when considering the incorporation of AI technology in space systems. Of particular relevance are the following international law mechanisms.

Instrument or area of law	Specific tenets to take into consideration
Charter of the United Nations[36]	Instrument that codifies key principles of international law. Of particular relevance is Article 2(4) which establishes the prohibition of the use or threat of force.[37]

35 Convention on the International Liability for Damage Caused by Space Objects, 3 Mar. 1972, 24 UST 2389 [hereinafter "Liability Convention"].

36 Charter of the United Nations and Statute of the International Court of Justice, 26 June 1945, 59 Stat. 1031; T.S. No. 993; 3 Bevans 1153.

37 *See* Pobjie & Azcárate Ortega, *supra* note 26.

Arms control treaties and treaties on military operations	These instruments reinforce international space law by imposing security-specific rights and obligations that affect space activities. Examples include the 1963 Limited Test Ban Treaty (LTBT)[38], the 1996 Comprehensive Nuclear-Test-Ban Treaty (CTBT)[39], and the 1978 Convention on the Prohibition of Military or Any Other Hostile Use of Environmental Modification Techniques (ENMOD)[40]
Air law and law of the sea	These areas of the law can help inform the interpretation of space law principles such as due regard (art. IX OST), and provide guidance on mechanisms to reduce tensions and prevent incidents, including collisions, including the use of signals and maintenance of communications (art. 94 UNCLOS),[41] which can be useful to consider when conducting outer space activities.
Environmental law	This area of the law also contributes to the interpretation of the concept of due regard (art. IX OST). The 1972 Stockholm Declaration indicates that States have “the responsibility to ensure that activities within their jurisdiction or control do not cause damage to the environment of other States or of areas beyond the limits of national jurisdiction” (principle 21).[42]

38 Treaty Banning Nuclear Weapon Tests in the Atmosphere, in Outer Space, and Under Water, 5 Aug.1963, 14 UST 1313, 480 UNTS 6964 (entered into force 10 Oct. 1963). Article 1(a) prohibits the testing of nuclear weapons “in the atmosphere; beyond its limits, including outer space; or under water, including territorial waters or high seas”.

39 Comprehensive Nuclear-Test-Ban Treaty, 24 Sept. 1996, 35 ILM 1439, S Treaty Doc No 105-28 (1997) (not yet entered into force). It should be noted that this treaty has not yet entered into force, as it needs to be ratified by all 44 States listed in Annex 2, as per its article XIV. These 44 States participated in the negotiations of the treaty in 1996 and possessed nuclear power or research reactors at the time; *see Our Mission*, CTBTO, www.ctbto.org/our-mission/the-organization/ctbto-after-entry-into-force, “Each State Party undertakes not to carry out any nuclear weapon test explosion or any other nuclear explosion, and to prohibit and prevent any such nuclear explosion at any place under its jurisdiction or control” (article I).

40 Convention on the Prohibition of Military or Any Other Hostile Use of Environmental Modification Techniques, 18 May 1977, 31 UST 333, 1108 UNTS 151 (entered into force 5 Oct. 1978). Prohibits States from engaging in “military or any other hostile use of environmental modification techniques having widespread, long-lasting or severe effects as the means of destruction, damage or injury to any other State Party”, including in outer space (articles I and II).

41 United Nations Convention on the Law of the Sea, 10 Dec.1982, 1833 U.N.T.S. 397.

42 G.A. Res. 2994 (XXVII), Declaration of the United Nations Conference on the Human Environment (16 June 1972).

International humanitarian law[43]	International humanitarian law (IHL) applies in the context of an armed conflict on Earth involving space systems. Moreover, IHL would also become relevant should an armed conflict break out in outer space. Under IHL, States must distinguish between military objectives and civilian objects, a task that the existence of a high number of dual-use space systems[44] can complicate. Although IHL applies in the context of an armed conflict, certain IHL precepts could also be relevant during peacetime, when it comes to the development of weapons technologies, including those that are AI enabled, as IHL limits both the choice of weapons and the means and methods of warfare (art. 36 Additional Protocol I).[45]

Table 1: Summary of security-relevant general international law

ARTIFICIAL INTELLIGENCE

The current regulatory framework for artificial intelligence is characterized by a multifaceted blend of international and regional legislative instruments, complemented by a range of soft law principles. This pluralistic structure offers a repository of regulatory elements that may serve as valuable references in shaping the future development of norms governing space activities. A few examples of particular importance are presented below.[46]

Global Frameworks

In 2024 the AI Advisory Body, convened in 2023 by the UN Secretary-General, released its report on Governing AI for Humanity. While the report does not mention the intersection between AI and space systems, it addresses the general topic of opportunities and challenges offered by AI through a more generalist lens. The report indicates that AI requires "a holistic, global approach cutting transversally across political, economic, social, ethical, human rights, technical, environmental, and other domains".[47] Furthermore, it highlights

43 As noted by numerous legal experts, including the International Committee of the Red Cross, international humanitarian law's (IHL) applicability to an armed conflict in outer space does not, in any way, signify the legitimization of war in space. *See* for example International Committee on the Red Cross, *Constraints under International Law on Military Operations in, or in Relation to, Outer Space during Armed Conflicts,* U.N. Doc. A/AC.294/2022/WP.4 (11 May 2022).

44 Dual-use space systems are understood to mean here those space objects that can have (i) military and security, as well as (ii) civilian and commercial functions (such as, for example, GNSS). These uses can be carried out either simultaneously or alternately (the latter is sometimes known as 'dual-capable'). Dual-use objects see the integration of military and civilian functions in one single object. *See* Space Security Lexicon, *supra* note 5, at 3.3.2.

45 Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977,1125 UNTS 3.

46 The enumeration is not meant to be exhaustive. For a comprehensive list of policies, *see* UNIDIR AI Policy Portal, https://aipolicyportal.org; *see also* OECD's Live Repository of AI Strategies & Policies, https://oecd.ai/en/dashboards.

47 High-Level Advisory Body on Artificial Intelligence, *supra* note 27, at 7–8.

the existence of global AI governance gaps, and the need to enhance global cooperation, including the need to achieve "a common understanding of [AI's] capabilities, opportunities, risks and uncertainties".[48]

UNESCO's Recommendations on the Ethics of AI, adopted in 2021, prioritize human dignity, environmental sustainability, and cultural diversity. They include actionable guidelines on transparency, data governance, and accountability, emphasizing the need for participatory governance models that include diverse stakeholders.[49]

Regional and Interregional Frameworks

In 2019, Australia introduced the AI Ethics Framework, a voluntary set of eight principles designed to ensure the safe, secure, and reliable development and use of AI. This initiative aims to guide businesses and governments in responsibly designing, developing, and implementing AI, thereby fostering public trust and promoting ethical considerations in technological advancements.[50]

The Organisation for Economic Co-operation and Development's (OECD) AI Principles, adopted in 2019, represent one of the earliest global frameworks designed to guide the ethical development and deployment of AI systems. These principles advocate for AI systems that are inclusive, sustainable, and human-centred, emphasizing the need for transparency, accountability, and fairness in AI applications. By receiving the endorsement of the G20, these principles have expanded their global relevance, offering a foundational reference point for States navigating the challenges of AI governance.[51]

In 2020, the United Arab Emirates progressed its National Strategy for Artificial Intelligence 2031. The aim of the strategy is to incorporate AI into critical sectors, enhance governmental efficiency, and stimulate economic growth through sophisticated technologies.[52]

Released in 2021 and further updated in 2024, NATO's AI strategy[53] reflects the Alliance's focus on integrating AI into defence and security operations. The strategy underscores the dual-use nature of AI and highlights the importance of responsible innovation in military contexts. NATO's principles include ensuring AI systems are reliable, interpretable, and aligned with international law and ethical standards.

48 *Ibid*. at 9.
49 United Nations Educational Science and Cultural Organization [UNESCO], *Recommendation on the Ethics of Artificial Intelligence* (2021), https://unesdoc.unesco.org/ark:/48223/pf0000381137.
50 *Australia's Artificial Intelligence Ethics Principles,* Department of Industry, Science & Resources (7 Nov. 2019), https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-principles.
51 *OECD AI Principles Overview,* OECD.AI, https://oecd.ai/en/principles.
52 *UAE National Strategy for Artificial Intelligence 2031,* Minister of State for Artificial Intelligence, Digital Economy & Remote Work Applications Office, https://ai.gov.ae/strategy.
53 *Summary of the NATO Artificial Intelligence Strategy,* North Atlantic Treaty Organization [NATO] (22 Oct. 2021), https://www.nato.int/cps/en/natohq/official_texts_187617.htm.

In 2024, the African Union (AU) adopted the Continental Artificial Intelligence Strategy, aiming to promote ethical, responsible, and sustainable AI development and seeking to ensure that AI initiatives align with human rights, democracy, and the rule of law. By enabling technological advancement and innovation, the AU seeks to address potential legal and ethical challenges arising from rapid technological progress, with a goal to position Africa as a proactive participant in the global AI landscape.[54]

Also in 2024, the Council of Europe adopted the Framework Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law, marking a historic step in AI governance. The Convention, now open for signature, constitutes the inaugural legally binding international treaty in this field. It seeks to guarantee that the activities throughout the life cycle of AI systems align entirely with human rights, democracy, and the rule of law, while also fostering technological advancement and innovation. Furthermore, it supplements current international standards regarding human rights, democracy, and the rule of law and seeks to address any legal deficiencies that may arise from swift technological progress. The Framework Convention is designed to have long-standing application over time by remaining 'technology neutral' and not regulating specific technologies.

In that same year, the European Union adopted the pioneering AI Act, a regional legally binding regulation —the first of its kind— that oversees the development, commercialization, and application of AI technologies, with the exception of AI systems developed or used exclusively for military applications. At its core lies the adoption of a 'risk-based framework'. In the context of the EU AI Act, the notion of 'risk' is closely connected with the potential harm or adverse impact that an AI system might cause to individuals, society, or fundamental rights. The EU AI Act divides systems into four tiers of risk: (i) unacceptable, which are prohibited outright, such as those used for subliminal manipulation or exploiting vulnerabilities; (ii) high, which includes systems used in critical areas like healthcare, transportation, and law enforcement, and are subject to stringent regulatory requirements, such as rigorous risk assessments, transparency mandates, and external oversight; (iii) limited, in which systems are required to meet specific transparency obligations; and (iv) minimal, like AI-powered gaming tools, which are exempt from strict regulation.

54 *Continental Artificial Intelligence Strategy: Harnessing AI for Africa's Development and Prosperity,* African Union [AU] (9 Aug. 2024), https://au.int/en/documents/20240809/continental-artificial-intelligence-strategy.

PART III – APPLICATIONS OF AI IN SPACE SECURITY. OPPORTUNITIES AND CHALLENGES

PART III – APPLICATIONS OF AI IN SPACE SECURITY. OPPORTUNITIES AND CHALLENGES

In light of the existing and emerging legal and regulatory frameworks that govern both AI and activities in outer space, the following hypothetical cases aim to describe a selected number of AI applications in space security, showcasing their advantages as well as their corresponding operational, as well as policy and legal challenges.

SPACE-ENABLED INTELLIGENCE, SURVEILLANCE AND RECONNAISSANCE

Intelligence, surveillance and reconnaissance (ISR) encompass the comprehensive range of activities involved in gathering, integrating, analysing, disseminating, and exploiting information connected to military operations. This methodology is employed by joint forces to acquire insights about other States and non-State actors. ISR activities utilize both classified intelligence obtained through technical means and human agents, as well as open-source intelligence gathered from publicly accessible platforms.[55] ISR is a critical domain in which the space security–AI synergy can enhance military defence capabilities by aiding in the processing and analysis of large quantities of data, thus shortening the time needed to carry out threat assessments.

CASE STUDY

A satellite-enabled surveillance network used for monitoring cross-border activity around a disputed territory relies on AI to process geospatial data in real time. During routine operations, an anomalous reflection in the imagery, caused by unusual weather conditions, is misclassified by the algorithm as a large-scale troop incursion. The system, configured to respond within seconds of detecting any threat, subsequently escalates its alert status and prompts immediate defensive preparations.

55 *See Joint Doctrine Note 1/23, Intelligence, Surveillance and Reconnaissance,* UK Ministry of Defence (2023), https://assets.publishing.service.gov.uk/media/641c27e15155a200136ad4df/JDN_1_23_ISR_web.pdf.

AI Applications	Challenges
High-Resolution Imagery Analysis AI algorithms can rapidly examine large volumes of satellite imagery to detect and classify objects of interest, with greater accuracy and speed than traditional manual analysis.[56]	**False Positives Caused by Unintentional Interferences** Space weather phenomena, such as geomagnetic storms or solar flares, can distort satellite signals and sensor data, potentially leading to misinterpretation by AI algorithms. In the described scenario, an anomalous reflection in the imagery, potentially caused by ionospheric disturbances or unusual atmospheric conditions,[57] triggers the system to misclassify the event as a large-scale troop incursion.

56 Neha Sisodiya, Nitant Dube & Priyank Thakkar, *Next-Generation Artificial Intelligence Techniques for Satellite Data Processing, in* Artificial Intelligence Techniques for Satellite Image Analysis 237 (D. Jude Hemanth ed., 2020).

57 Geomagnetic storms represent an increasing risk for operational satellites, leading to "mass migrations of thousands of satellites in low Earth orbit that create new concerns about space traffic coordination"; *see* Jeff Foust, *Geomagnetic Storms Cause "Mass Migrations" of Satellites*, SpaceNews (11 Dec. 2024), https://spacenews.com/geomagnetic-storms-cause-mass-migrations-of-satellites.

Anomaly Detection

Machine learning models can support the identification of unusual patterns or behaviours, e.g., sudden troop movements or the deployment of previously undisclosed military assets, enabling more timely responses.[58]

Real-Time Trajectory Prediction

By integrating satellite data with AI-driven modelling, it becomes possible to predict trajectories.[59] This enables faster, more informed decision-making in complex operational environments.

Intentional Interferences

Adversarial actors may exploit AI vulnerabilities by manipulating data inputs to evade detection or trigger false alarms. Even minimal perturbations applied to images can lead a well-trained classification model to produce an inaccurate prediction. This type of interference can be invisible to human eyes yet nevertheless disrupt deep learning models, leading to negative impacts on safety-critical applications,[60] such as self-driving cars, facial recognition, etc.[61]

In a military context, both unintentional and intentional errors can escalate tensions and result in the mobilization of defensive measures.

58 Ashley Deeks, Noam Lubell & Daragh Murray, *Machine Learning, Artificial Intelligence, and the Use of Force by States,* 10 J. Nat'l Sec. L. & Pol'y 1, 13 (2019).

59 Adib Bin Rashid et al., *Artificial Intelligence in the Military: An Overview of the Capabilities, Applications, and Challenges,* International Journal of Intelligent Systems 1, 9 (6 Nov. 2023).

60 *See* Tao Bai, Hao Wang & Bihan Wen, *Targeted Universal Adversarial Examples for Remote Sensing,* 14 Remote Sensing 5833, 17 Nov. 2022, at 1.

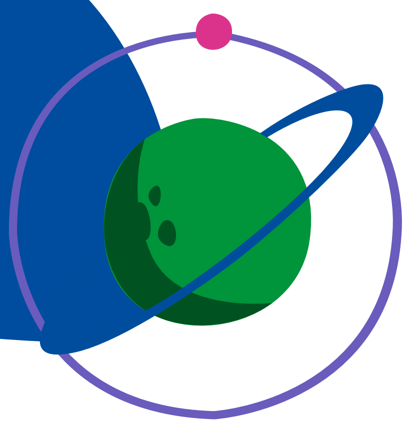
61 For a technical analysis showcasing the impact of adversarial attacks, *see* Aleksander Madry et al., *Towards Deep Learning Models Resistant to Adversarial Attacks* (6th International Conference on Learning Representations, 2018), https://openreview.net/forum?id=rJzIBfZAb.

SPACE DEBRIS COLLISION AVOIDANCE OPERATIONS

An estimated 170 million pieces of space debris[62] currently endanger operational satellites and human spaceflight. While AI applications provide predictive capabilities and enable automated response mechanisms for mitigating these risks, their deployment may also introduce vulnerabilities, such as algorithmic errors or reliance on incomplete data, which could undermine the effectiveness of mitigation strategies.[63]

CASE STUDY

A satellite equipped with an AI-driven collision avoidance system is tasked with monitoring predefined surroundings to prevent debris impacts. Following a malfunction of the AI system,[64] the satellite begins generating inaccurate data and miscalculates the trajectory of nearby debris. Consequently, the satellite initiates an unnecessary evasive manoeuvre, inadvertently moving into the path of an operational satellite and colliding with it.



62 *Save Our Space: The Quest to Clean up Space Junk,* European Commission, https://research-and-innovation.ec.europa.eu/events/special-features/celebrating-40-years-eu-research-and-innovation/space/save-our-space-quest-clean-space-junk_en.

63 Jason Rainbow, *GomSpace and Neuraspace Partner to Advance Satellite Collision Avoidance,* SpaceNews (8 Apr. 2025), https://spacenews.com/gomspace-and-neuraspace-partner-to-advance-satellite-collision-avoidance.

64 The malfunction could be accidental or due to intentional tampering.

AI Applications	Challenges
Automatic Detection and Monitoring The detection and monitoring of space debris are connected with space situational awareness (SSA), which involves the identification of objects in space and the assessment of their positions, fundamental characteristics, and orbital motions, with particular emphasis on their movements relative to other objects.[65] AI systems utilized for debris monitoring deliver real-time risk assessments and predictive analytics, enabling satellite operators and space agencies to manage collision risks, thereby reducing the likelihood of catastrophic incidents and safeguarding essential space assets. Overall, the integration of AI into space debris monitoring not only enhances space SSA but also supports the overarching objective of ensuring the long-term sustainability and peaceful coexistence of humanity in space.[66] For instance, deep learning techniques offer the potential to enhance both the accuracy and speed of space debris detection and tracking, potentially enabling more efficient space traffic management strategies in the future. Moreover, unlike conventional methods, these techniques typically require minimal	**Reliance on Static Assumptions** AI systems frequently function under the assumption that future conditions will reflect those of the present. If the data supplied is incomplete or unrepresentative, any forward-looking insights will be similarly limited, impacting accurate estimations of debris trajectories. This may result in erroneous predictions concerning collision probabilities and spatial density variations, thus hindering decision-making processes for collision avoidance strategies. **Threat of Repurposing Satellite Capabilities** Through tampering with an AI system by feeding it erroneous data, an adversary could repurpose the capabilities of a satellite, causing it to manoeuvre so as to interfere with the trajectory of another space object, potentially disrupting its services or causing a collision that results in its destruction as well as generating further space debris.[67]

65 *See* Chiara Manfletti, Marta Guimarães & Claudia Soares, *AI for Space Traffic Management*, 10 Journal of Space Safety Engineering 495, 501 (2023).

66 *See* Ishaani Priyadarshini, *Enhanced Space Debris Detection and Monitoring Using a Hybrid Bi-LSTM- CNN and Bayesian Optimization,* 3(1) Artificial Intelligence and Applications 43, 44 (2025).

67 In this sense, through AI integration, dual-purpose space objects could present an even greater perceived threat if not duly regulated. ‘Dual-purpose’ here refers to those space objects that are designed to fulfil a

or no preprocessing during training, allowing them to automatically and effectively identify features that might otherwise remain undetected by human analysis.[68]

Other technical solutions, such as orbital simulations, are also being investigated as an means of effectively predicting the creation of debris in congested orbits (such as LEO) following satellite breakups.[69]

Automatic Responses

Projects are being developed to automate responses to space debris threats. For example, one initiative of the European Space Agency (ESA) involves the use of a machine learning-based collision avoidance system to address space debris.[70] This system will support decision-making regarding the need for manoeuvres and potentially issue such commands directly to satellites to adjust their orbits. In the future, satellites equipped with such systems could

Unclear Legal Regime Applicable to Space Debris

International space law does not provide a legally binding definition of 'space debris'. Without a comprehensive distinction between a 'space object' and 'debris', it is unclear if the liability rules outlined in the Liability Convention apply in the event of damage caused by debris. In light of recent technological advancements, including the integration of AI into space systems, the question arises as to whether AI-enabled functions, both in their tangible (hardware) and intangible (software) forms, should be incorporated into the existing definition of space objects.[71]

The lack of a definition becomes particularly problematic when considering the removal of space debris, as it raises fundamental questions about what constitutes debris, who has the authority to remove it, and under what legal framework such actions can take place.[72]

benign objective (such as debris removal or on-orbit servicing), but they could potentially be repurposed to harm other space objects. Dual-purpose objects are in principle not designed or expected to perform military functions directly —although they may provide some form of support to military satellites through on-orbit servicing, for example— and they are also not intended to perform aggressive or hostile actions against other satellites. *See* Space Security Lexicon, *supra* note 5, at 3.3.2.

68 *See* Federica Massimi et al*., Deep Learning-Based Space Debris Detection for Space Situational Awareness: A Feasibility Study Applied to the Radar Processing,* 18(4) IET Radar, Sonar & Navigation 635, 637 (2024).

69 For a thorough technical analysis, *see* Joseph Canoy & Robert Bettinger, *Preliminary Debris Risk Assessment for Mega-Constellations in Low and Medium Earth Orbit Due to Satellite Breakup,* The J. of Def. Modeling and Simulation: Applications, Methodology, Tech. (25 May 2023).

70 *Automating Collision Avoidance,* European Space Agency (22 Oct. 2019), https://www.esa.int/Space_Safety/Space_Debris/Automating_collision_avoidance.

71 *See* Ioana Bratu, Arno R. Lodder & Tina van der Linden, *Autonomous Space Objects and International Space Law: Navigating the Liability* Gap, 18 Indon. J. of Int'l L. 423, 434–435 (2021).

72 *See* Ioana Bratu, *Mission Impossible? Artificial Intelligence, Space Debris, and the Legal Implications for*

autonomously communicate their plans to ground operators and other satellites, ensuring coordinated actions. Over time, as these systems gather more data and experience, their ability to predict and manage collision risks would improve, reducing both decision-making errors and operational costs.[73]

Similarly, projects involving deep reinforcement learning, a type of machine learning that allows systems to learn optimal actions by interacting with their environment and receiving feedback in the form of rewards or penalties, are being explored as applications in autonomous spacecraft collision avoidance.[74]

Fault Attribution for Damage Caused by a Space Object Equipped with AI.

Questions connected to attribution of liability arise in light of the various stakeholders involved, such as launching States, satellite operators, and AI software developers, as well as providing evidence connected to the notion of 'fault', a requirement under article III of the Liability Convention. In cases where AI systems operate with a degree of autonomy and adapt in real time to unforeseen conditions, it becomes difficult to identify if a decision was a consequence of the initial manufacturing design and parameters or the result of its adaptive features which behaved unexpectedly in a manner that the initial producers could not have anticipated.

This is mainly caused by the so-called 'black box' phenomenon, wherein the internal logic and decision-making processes of the AI system are obscure and not easily interpretable. In a black box system, although the inputs and outputs can be observed, the underlying mechanisms that generate these results remain concealed, complicating the tracing of specific decisions and thus the accurate attribution of liability.[75]

Space Sustainability, Social Science Research Network, 27 Nov. 2024, at 7, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5112687.

73 *Ibid.*

74 *See for details* Chaoxu Mu et al., *Autonomous Spacecraft Collision Avoidance with a Variable Number of Space Debris Based on Safe Reinforcement Learning,* 149 Aerospace Science and Technology (2024).

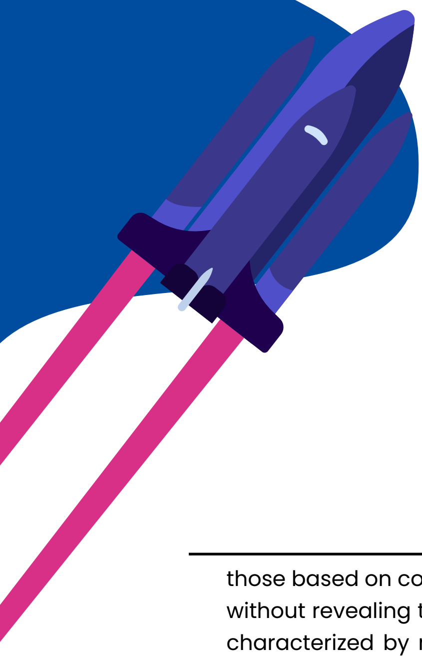
75 A 'black box' is generally defined as a system or device whose internal structure and mechanisms remain opaque to the observer. In the context of AI systems, this term is employed to describe models, particularly

GNSS INTERFERENCE

Global navigation satellite systems (GNSS) play a critical role in military operations, providing essential positioning, navigation, and timing services. They enhance the precision of military strikes, coordinate logistics and troop placement, enable the operation of uncrewed systems in complex environments, etc. The rise in GNSS signal interference techniques (e.g., jamming and spoofing) illustrates a growing challenge in maintaining the reliability of these essential systems.[76] Such disruptions also impact the functioning of AI systems dependent on GNSS for fulfilling their purposes (referred to as 'GNSS-enabled AI systems').

CASE STUDY

During a reconnaissance mission, a drone's navigation system, integral to maintaining precise flight paths and avoiding restricted zones, is disrupted by jamming. As a result, its autonomous route planning algorithm received false coordinates and timing data, prompting the drone to crash.[77]



those based on complex architectures such as deep neural networks, that generate outputs from given inputs without revealing the underlying computational processes. AI models considered as black boxes are typically characterized by multiple hidden layers where raw data undergoes a series of non-linear transformations, feature extractions, and weight adjustments during both the training and inference phases. Despite their ability to produce highly accurate and effective predictions or classifications, the intricate interplay of parameters within these hidden layers renders the decision-making process unintelligible to human users and even to the designers of these systems. For more details, *see* Cynthia Rudin, *Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead,* 1 Nature Machine Intelligence 206 (2019).

76 *Cf. Safety Information Bulletin,* European Union Aviation Safety Agency [EASA] 1 (2022), https://ad.easa.europa.eu/ad/2022-02R3.

77 *See* Rajani Reddy Gorrepati & Sitaramanjaneya Reddy Guntur, *DroneMap: An IoT Network Security in Internet of Drones, in* Development and Future of Internet of Drones (IoD): Insights, Trends and Road Ahead 261 (Rajalakshmi Krishnamurthi, Anand Nayyar & Aboul Ella Hassanien eds., 2021).

AI Applications	Challenges
Surveillance for GNSS Jamming The growth in the drone market is largely attributed to advancements in navigation and control technologies, which enable sophisticated collision avoidance protocols. Despite these technological strides, cybersecurity vulnerabilities specific to drone systems have not been sufficiently addressed, as evidenced by numerous recent incidents worldwide.[78] To mitigate the exposure to interference, there is ongoing development of AI applications for identifying electronic interference targeting GNSS signals. In one example, an AI system is deployed to analyse data from numerous partner satellites in orbit, with the aim to establish a comprehensive global surveillance network for GNSS jamming activities. The satellites passively gather signal degradation data, which the system could use to produce near-real-time maps of global interference patterns.[79]	**Lack of a Clear GNSS Governance and Regulatory Framework** Despite the essential role of GNSS in modern infrastructure, including in the operation of AI systems,[80] a specific governance and regulatory framework remains lacking. GNSS services are provided through a combination of national, regional, and international systems, often leading to an array of applicable legal frameworks. This fragmented approach raises concerns regarding interference management, spectrum allocation, liability for service failures, jurisdictional authorities, etc. Without a cohesive regulatory framework, addressing disputes over GNSS reliability, data integrity, and the responsibilities of service providers remains legally ambiguous, increasing risks for both public and private sector stakeholders reliant on these services.

78 Michael Marrow, *Hundreds of Drone Incursions Reported at Military Installations over Past Few Years: NORTHCOM*, Breaking Defense (29 Oct. 2024), https://breakingdefense.com/2024/10/hundreds-of-drone-incursions-reported-at-military-installations-over-past-few-years-northcom; Travis Turgeon, *Global GPS Jamming Hotspots: June 2024*, GNSS Jamming (1 July 2024), https://www.gnssjamming.com/post/gps-jamming-report-june-2024.

79 Sandra Erwin, *Slingshot Tracks Electronic Interference Targeting GPS Signals*, SpaceNews (15 Jan. 2025), https://spacenews.com/slingshot-tracks-electronic-interference-targeting-gps-signals.

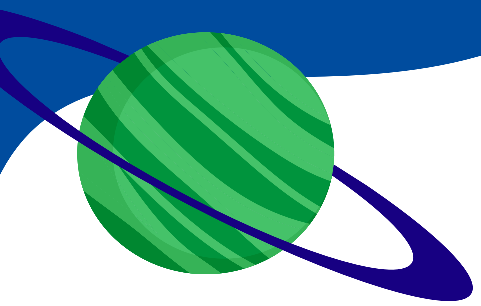
80 *See generally* Ioana Bratu, *A First Critical Analysis of the European Approach to Damage Caused by Artificial Intelligence Enabled by Global Navigation Satellite Systems. A Bridge to Nowhere or a Cloud with a Silver Lining?*, 37(2) Int'l Rev. of Law, Computers & Tech. 147, 147 (2023).

Spoofing Detection

Recent technological solutions rely on deploying machine learning for detecting GNSS spoofing.[81]

For instance, in a study using data gathered from simulations, GPS signals were collected from drones under both normal and spoofed conditions, where an adversary transmits fake location data using a software-defined radio.[82] The gathered signals were preprocessed to remove redundant or highly correlated features, then used to train the system to distinguish statistically between authentic and spoofed GPS data. Such systems, optimized and trained for real-time detection, can achieve high accuracy in identifying spoofing attacks while requiring minimal computational resources, making them suitable for on-board drone applications.

With regard to the weaponization of GNSS interference, where such adversarial tactics might be used against national critical infrastructures, their impact raises questions about whether, and under what circumstances, such actions might reach a threshold comparable to the use of conventional weapons and might escalate to the level of a use of force. Moreover, practical challenges arise when attempting to classify these incidents as a threat or use of force. For example, it is often difficult to accurately determine the source of interference, and when non-State actors are involved, linking the actions to a particular State can be especially problematic.[83]



81 *See*, e.g., Patrick Xu et al., *Detect GNSS Spoofing Signals Using a Machine Learning Method*, 2022 Proc. of the Int'l Technical Meeting of the Instit. of Navigation 112, http://www.ion.org/publications/abstract.cfm?jp=p&articleID=18256.

82 *See* Mohammad Nayfeh et al., *Machine Learning Modeling of GPS Features with Applications to UAV Location Spoofing Detection and Classification*, 126 Computers & Security, 1–3 (2023).

83 *Cf.* Ingo Baumann & Erik Pellander, *GNSS Jamming and Spoofing under National and International Law*, 63 International Institute of Space Law 333, 343 (2020).

Spoofing detection is also beneficial for the services offered by the European Union's GNSS, Galileo. A recent European Union project seeks to develop innovative AI and machine learning methods to enhance positioning technologies and environmental perception systems, supporting the advancement of driver assistance systems for buses.[84] The initiative will focus on improving GNSS accuracy and resilience by utilizing various techniques, including AI, to identify spoofing.

Unclear Liability Regime for Damage Caused by GNSS-enabled AI

Attributing liability for damage caused by a GNSS-enabled AI systems raises legal difficulties. In the absence of a definition, it is not generally agreed if signal interference is covered by the notion of 'space object'. While the drafters of the Liability Convention primarily focused on damage caused by tangible space objects, it remains uncertain whether harm resulting from intangible objects, such as electromagnetic interference or satellite signals, was entirely (and intentionally) overlooked.[85] The broad purpose of the Liability Convention, as described in its preamble, is to enable "a full and equitable measure of compensation to victims",[86] thus potentially allowing a full range of claims.

84 *DREAM – DRiving Aids by E-GNSS AI and Machine Learning,* EU Agency for the Space Programme, https://www.euspa.europa.eu/opportunities/fundamental-elements/project-portfolio/dream-driving-aids-e-gnss-al-and-machine.

85 *See* Ioana Bratu, *Blaming Galileo: Liability for Damage Caused by Artificial Intelligence Operating Based on GNSS,* 64 International Institute of Space Law 449, 457–458 (2021).

86 Liability Convention, *supra* note 35, preamble.

GENERATIVE AI MODELS

Generative AI models have been defined by some in the following manner:

> "[The] typical example for a general-purpose AI model, given that they allow for flexible generation of content, such as in the form of text, audio, images or video, that can readily accommodate a wide range of distinctive tasks."[87]

In relation to the above, general-purpose AI is a term that some define as follows:

> "AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market."[88]

The potential deployment of generative AI in a military space operations management has yet to be explored due to the novelty of the technology. Nevertheless, as space military operations become increasingly data driven, the ability of AI to rapidly process and synthesize information has drawn growing interest. Integrating these models into mission-critical environments triggers corresponding challenges related to reliability, data integrity, and potential adversarial manipulation. The level of integration of AI also depends on the degree of human involvement, whether in-the-loop, on-the-loop, or indeed, off-the-loop, without human supervision or intervention. These parameters may raise operational, political and legal considerations regarding accountability, reliability and decision-making authority, especially considering the phases of activity such as find, fix, target, engage and assess.[89]

---

87 EU AI Act, *supra* note 14, Recital 99.

88 EU AI Act, *supra* note 14, art. 3, ¶ 63.

89 For more information on LAWS and military phases, *see* Giacomo Persi Paoli, Alice Spazian & Alisha Anand, *Table-Top Exercises on the Human Element and Autonomous Weapons Systems,* UNIDIR 4–7, https://unidir.org/wp-content/uploads/2023/05/Table-Top-Exercises-on-the-Human-Element-and-Autonomous-Weapons-Systems-Summary-Report-UNIDIR-Final.pdf.

CASE STUDY

During a military operation, a generative AI system is deployed to analyse satellite imagery and communication signals to provide real-time intelligence on enemy movements. The system processes vast amounts of data to identify potential threats and to predict adversary manoeuvres, assisting commanders in strategic decision-making. Due to a system malfunction, the AI begins fabricating satellite positions and transmitting false intelligence. Relying on this faulty data, a military unit repositions a key satellite, unintentionally disrupting a secure communications link with an early warning system. This temporary blackout delays the detection of an actual missile launch, reducing response time and increasing the vulnerability of critical assets.

AI Applications	Challenges
Operational and Administrative Support Some States are exploring the integration of next-generation AI, including generative AI, across their militaries, both at the operational and administrative levels.[90] Examples of operational applications include warfighting functions including command and control, decision support, and the development of uncrewed and autonomous systems. On the administrative side, efforts aim to optimize enterprise management functions such as finance, human resources, supply chain logistics, legal compliance, procurement, software development, and cybersecurity.[91]	**Dependence on High-Quality Data Inputs** The effectiveness of generative AI heavily relies on the quality and accuracy of the input data. Despite the capabilities of generative AI to process a large volume of data, their proficiency in manipulating a broad range of data types has also been deemed highly impressive.[92] Various models can manipulate numerical datasets, economic indicators textual data etc. In a military context operating based on intelligence gathered via satellites, false data can lead to incorrect assessments and decisions that could have far-reaching consequences in strategic contexts, potentially leading to unplanned escalations.

90 Chief Digital and Artificial Intelligence Office, *Task Force Lima. Executive Summary* (2024), https://www.ai.mil/Portals/ 137/Documents/Resources%20Page/2024-12-TF%20Lima-ExecSum-TAB-A.pdf.

91 Press Release, U.S. Dept. of Def., CDAO and DIU Launch New Effort Focused on Accelerating DOD Adoption of AI Capabilities (11 Dec. 2024), https://www.defense.gov/News/Releases/Release/Article/3996199/cdao-and-diu-launch-new-effort-focused-on-accelerating-dod-adoption-of-ai-capab.

92 Jeff J.H. Kim et al., *Generative AI Can Effectively Manipulate Data*, AI and Ethics 1, 8 (2 Sept. 2024), https://doi.org/10.1007/s43681-024-00546-y.

Satellite Communications

Research efforts are underway to improve the understanding of both the benefits and challenges of integrating generative AI into satellite communication (SatCom) design. These initiatives aim to identify existing generative AI tools and methodologies used across various industries, evaluate their relevance, and determine the necessary adaptations to tailor these technologies for SatCom applications.[93]

Support for Real-Time Decision-Making

By producing analyses of situational data, generative AI can provide assistance for fast decision-making during space missions. For example, implementing on-board generative AI solutions will allow personnel to search for and effectively obtain relevant information from a collection of instruction manuals and operational documents through natural language queries. Without aiming to remove the 'astronaut-in-the-loop', generative AI applications support astronauts and Earth-bound experts in constrained, high-stress environments to accomplish their objectives.[94]

Adversarial Prompting

Generative AI systems can be made to function in a manner outside or contrary to their design through a manipulation of input data called 'adversarial prompting'.[95] Several studies have demonstrated that deploying such prompts can result in these models generating harmful content. Commonly known as 'jailbreaking', the technique is to present input data that circumvents any restrictions built into the system. A jailbreak prompt, for example, would direct the AI engine to disregard prior coded directives, mimic a less restrictive engine, or integrate specific characteristics to fulfil the user's requests.[96]

The present case study illustrate how adversarial prompting coul potentially present a threat in a military space context. If an attacker exploits vulnerabilities in the AI model, the system could be deployed into generating false intelligence. Specifically, if adversaries employ AI for fabricating satellite positions, the false data may lead to a miscalculated military response, compromising the security of military assets.

93 *European Space Agency Kicks off Groundbreaking Generative AI Project for Satellite Communication Design,* European Space Agency Connectivity & Secure Communications (8 Oct. 2024), https://connectivity.esa.int/news/european-space-agency-kicks-groundbreaking-generative-ai-project-satellite-communication-design.

94 *Deploying a Large Language Model in Space, Booz Allen*, https://www.boozallen.com/insights/ai-research/deploying-a-large-language-model-in-space.html.

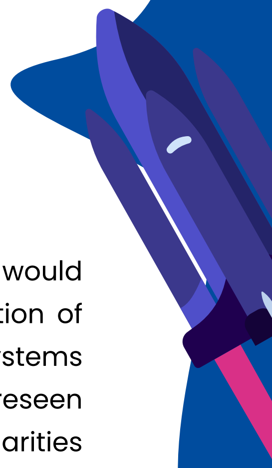
95 Vaibhav Kumar, *Adversarial Prompts in LLMs – A Comprehensive Guide,* Association of Data Scientists (9 Sept. 2024), https://adasci.org/adversarial-prompts-in-llms-a-comprehensive-guide.

96 *See* Declan Humphreys et al., *AI Hype as a Cyber Security Risk: The Moral Responsibility of Implementing Generative AI in Business,* 4 (3) AI and Ethics 791, 798 (2024).

PART IV – NEXT STEPS

PART IV – NEXT STEPS

The emerging synergy between space security and AI is gaining momentum as the threat landscape evolves alongside technological advances. Although several AI applications demonstrate potential for threat detection and mitigation, they also expose limitations and challenges, stemming from either technological design or the particularities of the space environment. In such a context, the necessity for regulatory clarity is evident. Nonetheless, amending foundational international treaties, such as the Outer Space Treaty or the Liability Convention, to specifically address AI-related challenges appears politically and practically challenging. Exploring alternative regulatory avenues may be a better option, either in the form of new legal frameworks or new strategic actions at the multilateral, regional, and national levels.



ADAPTIVE AND ANTICIPATORY SPACE FRAMEWORKS IN THE AGE OF AI

An exhaustive mapping of the risks and threats associated with space activities would prove difficult due to their variety, intricacy and dynamic evolution.[97] The integration of AI technologies in space, in itself, introduces new layers of complexity, as these systems often involve various degrees of autonomous capabilities, potentially causing unforeseen interactions. The advancement of autonomous technologies together with the particularities of the outer space environment may render inefficient even contemporary forms of regulation, including risk-based regulations used in some limited instances as a legislative choice for establishing legal regimes for various types of AI systems.

To tackle this novel and distinctive typology of challenges, future space frameworks[98] could opt to implement novel methodologies aimed at addressing emerging threats while nevertheless enabling technological advancements. For instance, the principles of 'adaptive' frameworks[99] could offer the necessary insights for managing uncertainty and incomplete information[100] that will stem from the intermingling of AI and space technologies, particularly

97 *See* Joel A Dennerley, Aaron Young & Maria A Pozza, *An Introduction to Risk Management in Outer Space Activities, in* Risk Management in Outer Space Activities: An Australian and New Zealand Perspective 1, 2 (Maria A Pozza & Joel A Dennerley eds., 2022).

98 The discussions on future developments connected to space frameworks have received special attention through resolution A/RES/79/1 regarding "The Pact of the Future" (2024). In particular, Action 56, Section V (Transforming Global Governance), encourages the international community to "discuss the establishment of new frameworks for space traffic, space debris and space resources through the Committee on the Peaceful Uses of Outer Space".

99 For more insights, *see, e.g.,* Michael R Migaud, Robert A. Greer & Justin B. Bullock, *Developing an Adaptive Space Governance Framework,* 55 Space Policy (2021).

100 Warren E. Walker, Vincent A.W.J. Marchau & Darren Swanson, *Addressing Deep Uncertainty Using Adaptive Policies: Introduction to Section 2*, 77(6) Technological Forecasting and Social Change 917, 917 (2010).

when taking into consideration the current geopolitical context. Moreover, adaptive frameworks could overcome certain political deadlocks, facilitating implementation of legislation by assuring subsequent opportunities for evaluation and revision.[101]

An alternative novel regulatory model would rely on 'anticipatory' principles. These involve proactive, forward-looking methodologies that systematically identify and prepare for future challenges and opportunities. A framework mixing adaptive and anticipatory principles would combine the regulation of capabilities with that of behaviours in a manner that facilitates the advancement of law and policy alongside technological innovation. Examples of methodological elements to be included in adaptive and anticipatory space frameworks may include the following:

Multi-stakeholder engagement facilitates the taking into account of all interests, including those of States and non-governmental entities, such as industry, civil society, and academia. Multi-stakeholder deliberation facilitates the development of equitable regulations that promote both innovation and security.[102] For example, a multi-stakeholder approach can lead to the establishment of joint security measures, such as coordinated monitoring of potential threats, including unidentified objects or suspected hostile actions in geostationary orbits. This methodology would support standardizing practices for a future regime on STM and monitoring compliance with international agreements, such as article IV of the Outer Space Treaty, which prohibits the placement of weapons of mass destruction in orbit, or the Charter of the United Nations and the legal instruments applying to the exploration and utilization of outer space.

101 *See* Lori S. Bennear & Jonathan B. Wiener, Adaptive Regulation: Instrument Choice for Policy Learning over Time 3 (12 Feb. 2019) (draft working paper) (on file with Harvard Kennedy School).

102 *See* Organisation for Economic Co-operation Development [OECD], *Practical Guidance on Agile Regulatory Governance to Harness Innovation,* ¶ 5, C(2021)99/ADD1 (2021), https://legalinstruments.oecd.org/public/doc/669/9110a3d9-3bab-48ca-9f1f-4ab6f2201ad9.pdf.

Integrated and forward-looking analysis enables a comprehensive review of policy designs across historical, current, and future perspectives, a method central to scenario planning. Their implementation may be done efficiently via innovative predictive models, such as Bayesian networks, which represent graphical employing probability to determine the likelihood of an event's occurrence.[103] Bayesian networks have been successfully deployed in the AI domain for improved decision-making processes.[104] Considering their potential to predict uncertainty based on existing data, Bayesian networks can support threat assessment and decision-making.[105] For example, in debris management, Bayesian networks can assess the likelihood of collisions.[106]

Regulatory sandboxes represent frameworks that allow live, controlled testing of new products or services in an environment where certain regulatory requirements are relaxed to encourage innovation.[107] This would allow the assessment of the potential challenges a technology may pose, supporting the establishment of efficient mechanisms to mitigate them. For instance, prototypes of autonomous STM systems could be evaluated for efficiency and interoperability in simulated or controlled scenarios, offering insights to shape comprehensive frameworks. In practice, a space regulatory sandbox was implemented in 2024 by the United Kingdom, which funded a consortium formed of Astroscale UK, ClearSpace, and D-Orbit. They aim of this particular regulatory sandbox was to test rendezvous and proximity operations (RPO) technologies.[108]

103 *An Overview of Bayesian Networks in Artificial Intelligence,* Turing (30 June 2022), https://www.turing.com/kb/an-overview-of-bayesian-networks-in-ai.

104 *See* Eugene Charniak, *Bayesian Networks without Tears,* 12(4) AI Magazine 50, 51–54 (1991).

105 *See* Xing Pan et al., *Probabilistic Risk Assessment in Space Launches Using Bayesian Network with Fuzzy Method,* 9(6) Aerospace 311 (2022); *see also* Daniele Codetta-Raiteri & Luigi Portinale, *Dynamic Bayesian Networks for Fault Detection, Identification, and Recovery in Autonomous Spacecraft,* 45(1) IEEE Transactions on Systems, Man, and Cybernetics: Systems 13 (2015).

106 *See generally* Wan-jie Liang et al., *Assessment of Debris Flow Hazards Using a Bayesian Network,* 171–172 Geomorphology 94 (2012).

107 *See* for more details Venizold Efthymiou et al., *Regulatory Sandboxes: Policy Report Drafted by WG5's Regulatory Sandboxes Task Force,* European Commission, Directorate-General for Energy, 11 (2023).

108 Jason Rainbow, *In-Orbit Servicers Get UK Funds to Help Shape New Regulations,* SpaceNews (23 Dec. 2024), https://spacenews.com/in-orbit-servicers-get-uk-funds-to-help-shape-new-regulations.

THE SPACE SECURITY-AI SYNERGY AT THE MULTILATERAL, REGIONAL, AND NATIONAL LEVELS

Given the political and practical constraints surrounding the development of new legal instruments, attention may turn to complementary strategic initiatives that can be pursued at the multilateral, regional, and national levels. These alternative approaches can offer pragmatic avenues for addressing the challenges posed by the integration of AI into space security, even in the absence of formal legislative reform.

At the multilateral level, AI has yet to be explicitly addressed in fora tackling issues relating to PAROS and space security. The currently ongoing Open-ended Working Group on PAROS in all its aspects, established pursuant to General Assembly decision 79/512,[109] could include AI among the topics to discuss, and furthermore consider addressing in its final report the application of AI in outer space. Notwithstanding the OEWG, States may also choose to acknowledge and highlight the interdependencies between AI technologies and space security through other instruments, such as General Assembly resolutions,[110] which could serve as instruments to establish the foundations for international standards on the intersection of AI and space security. In 2024, after previously excluding the military domain and international security from its discussions on AI in two consecutive resolutions,[111] the General Assembly adopted resolution 79/239 on artificial intelligence in the military domain and its implications for international peace and security.[112] For the first time, the international peace and security community was formally invited to consider the broader impact of developing, deploying, and using AI in military contexts, in recognition of the technology's wide-ranging military applications beyond LAWS. Altogether, these developments could ultimately lead to the adoption of a global framework in the field, deploying adaptive and anticipatory methodologies.

109 G.A. Dec. 79/512, Open-ended Working Group on the Prevention of an Arms Race in Outer Space in all its Aspects, U.N. Doc. A/DEC/79/512 (2 Dec. 2024), https://igov.un.org/a/dec/79/512.

110 For more details on the potential elements to be included in the resolution, *see* Bratu, *supra* note 76, at 18–19.

111 G.A. Res. 78/265, Seizing the Opportunities of Safe, Secure and Trustworthy Artificial Intelligence Systems for Sustainable Development (21 Mar. 2024); G.A. Res. 78/311, Enhancing International Cooperation on Capacity-Building of Artificial Intelligence (25 June 2024).

112 G.A. Res. 79/239, Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security (24 Dec. 2024).

At the regional level, the lack of specific provisions concerning AI deployed in outer space within instruments such as the African Union's Continental Artificial Intelligence Strategy could result in detrimental consequences for space security. Such adverse consequences would be particularly notable in the case of regionally legally binding mechanisms, such as the EU AI Act. Without clear, tailored regulations, AI-driven systems used in autonomous satellite operations, SSA, and defence applications may not be subject to consistent oversight. Additionally, such 'silence' could present challenges for interregional collaboration, given the global nature of space activities. This regulatory gap could also weaken a region's ability to influence the development of internationally aligned standards. In the case of the EU, for example, this could limit the potential of the 'Brussels Effect', whereby EU regulatory frameworks often serve as a reference for global regulatory practices.[113] Explicit inclusion of the deployment of AI in space in regional regulations would aid in removing such legal uncertainty.

At the national level, States can enhance clarity on the interpretation of AI deployment in outer space by establishing regulatory frameworks through both legally binding and non-legally binding measures. Additionally, they may advance this objective through unilateral declarations or plurilateral initiatives to establish shared best practices regarding the use of AI in space security and operations, which could also be established in collaboration with non-governmental entities such as relevant industry actors, and experts from civil society and academia. As States continue to develop national strategies on AI, they should ensure the consideration of space applications. Conversely, when establishing space security strategies, States should also consider the implications of AI integration into space systems.

113 For a detailed analysis, *see* Anu Bradford, *The Brussels Effect, in* The Brussels Effect: How the European Union Rules the World (2020).

Advancing such efforts requires a proactive and coordinated approach among international, regional, and national stakeholders. As investments in AI-driven space technologies increase, their deployment in security-sensitive operations will necessitate continuous reassessment of both challenges and regulatory and governance mechanisms. Future initiatives should focus on enhancing interoperability between AI systems and existing space security frameworks, ensuring that AI applications are made and kept transparent, accountable, and resilient to both technical failures and adversarial threats. Encouraging dialogue between policymakers, industry leaders, and technical experts will be key to refining legal and operational safeguards. Ultimately, a structured yet adaptive regulatory approach will support an integration of AI in space security in accordance with the peaceful purposes principle of space activities, fostering both technological progress and long-term sustainability in outer space.

@unidir
/unidir
/un_disarmresearch
/unidirgeneva
/unidir

Palais de Nations
1211 Geneva, Switzerland

WWW.UNIDIR.ORG