BALANCING INNOVATION AND RESPONSIBILITY:

INTERNATIONAL RECOMMENDATIONS FOR AI REGULATION IN SPACE

REPORT BY THE

IIISL WORKING GROUP ON THE LEGAL ASPECTS OF AI IN SPACE

CHAIR: TUANA YAZICI

MEMBERS: ANNE-SOPHIE MARTIN, STEVEN WOOD, ROSER ALMENAR, LISA KUCHER, LAURA YVONNE ZIELINSKI, STEFAN-MICHAEL WEDENIG, AND GIOVANNI TRICCO

PUBLISHED BY THE
INTERNATIONAL INSTITUTE OF SPACE LAW | DECEMBER 2024

BALANCING INNOVATION AND RESPONSIBILITY:
INTERNATIONAL RECOMMENDATIONS FOR AI REGULATION IN SPACE

Report of the Working Group on Legal Aspects of AI in Space

International Institute of Space Law

2024

Table of Contents

PART I: LEGAL FRAMEWORKS

PART II: REGULATORY CONSIDERATIONS

Section 3: Overview of Procurement Frameworks in Regulating the Use of AI for Space Applications……………….……………………………………..………………………125

PART III: ETHICAL IMPLICATIONS

Section 2: Recommendations for International Guidelines on Ethical AI Governance Across Emerging Sectors……………………………………………………………...………185

PART IV: ENVIRONMENTAL RESPONSIBILITY

PART V: VOIDS AND RECOMMENDATIONS

Foreword

The International Institute of Space Law (IISL) is proud to present this study on legal aspects of the use of Artificial Intelligence (AI) in space activities. With it, IISL intends to analyse a field of high topicality and relevance and propose ways forward. Based on its more than sixty years history and its global membership from the governmental, academic, and industrial fields as well as private practice, the Institute is in a perfect position to thoroughly, neutrally, and creatively contribute to the understanding of legal issues as well as to the development of space law and regulation. In this context, the present study is not a formal position paper by the Institute but should be regarded as a specific contribution by our study group to shaping the nascent international discussion in this field. In view of its traditional role, IISL will also provide these findings to the relevant international forums.

These general characteristics are extremely well fulfilled in the present study, "Balancing Innovation and Responsibility: International Recommendations for AI Regulation in Space," which has been prepared by an international team initiated and led by Tuana Yazici. This team not only frames and covers the issue in a detailed and comprehensive way, but it also completed this pioneering study following an impressively dense schedule, guaranteeing that the information and perspectives contained are meeting the speed with which this field is developing. We are therefore confident that the study provides understanding of and timely answers to the pressing questions and issues in the field of AI applied in space activities and will be a useful resource for all actors in this field.

Kai-Uwe Schrogl
President, IISL

Acknowledgement

This study, conducted by the Working Group on Legal Aspects of AI in Space, was initiated in April 2024 and achieved an important milestone with the presentation of our preliminary findings during the International Astronautical Congress (IAC) in October 2024 in Milan, Italy. Completing a study of this scope and significance within just six months reflects the shared commitment of all involved to addressing the rapid advancements in artificial intelligence in space applications and the significant risks posed by the slower pace of regulation. Following comments and revisions, the study was finalized in December 2024, ensuring a comprehensive and polished outcome. Our accelerated timeline was guided by a sense of responsibility to ensure this study could contribute effectively and efficiently to the discourse.

I would like to express my deepest gratitude to Prof. Dr. Kai-Uwe Schrogl for encouraging the initiation of this study and for providing continuous support throughout the process. Additionally, we are grateful to the International Institute of Space Law (IISL) for providing a platform to conduct this important work and for fostering an environment of collaboration. We also thank the experts and practitioners in the field, as well as representatives from international organizations, whose valuable insights greatly enriched our research.

Furthermore, we are appreciative of the thoughtful feedback shared during the IISL lunchtime event in Milan and the constructive comments and suggestions provided by the IISL Board following our presentation at the IISL Board meeting. These interactions contributed meaningfully to refining the final outcomes of this manuscript.

We hope this study underscores the urgency of the highlighted issues and serves as a valuable resource for advancing discussions and collaborative solutions in this evolving field.

Tuana Yazici
Chair, IISL Working Group on Legal Aspects of AI in Space

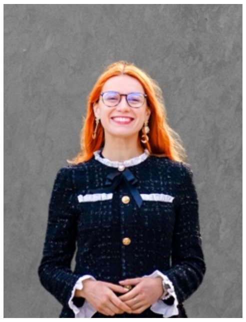
Authors' Biographies

Tuana Yazici (Study Chair): Tuana Yazici is the Founder, Chair, and CEO of the holding company Tuana Group, its two subsidiaries, AeroAI Voyages and AeroAI DesignLab, and the nonprofit AeroAI Global Solutions. These entities aim to leverage AI and space technologies to improve global living conditions. As a third-year law student and member of the International Institute of Space Law, Yazici holds a Master's in International Administration and a B.A. in PPE (Philosophy, Politics, and Economics). With publications in space law, she is a featured speaker at global conferences, sharing insights into her work. Her upcoming research focuses on regulating space technology and AI for responsible use.

Anne-Sophie Martin: Anne-Sophie Martin is a research fellow at the Institute for International Legal Studies of the National Research Council (ISGI-CNR) in Rome, Italy. She holds an LL.M. in Space Law from the University of Paris-Saclay and a PhD from Sapienza University of Rome. She is a member of the Legal Council of For All Moonkind and Senior Fellow at its Institute on Space Law and Ethics. She is affiliated with the International Astronomical Union's Centre for the Protection of Dark and Quiet Skies and the Open Lunar Foundation. Since 2021, she participates in the Global Expert Group on Sustainable Lunar Activities.

Steven Wood: Steven Wood is a US patent attorney and space lawyer. He serves as a member of the Advisory Board for the Association of Commercial Space Professionals, Senior of Counsel with Vela Wood, and Senior Director of Neuromorphic Technology Development with the SUNY Albany College of Nanotechnology, Science, and Engineering. He focuses his VW legal practice on IP due diligence, building in-house patent filing capabilities, and IP strategies and transactions. In his SUNY role, Steven coordinates research efforts with the Air Force Research Labs funded NeuroPipes neuromorphic computing research group. He has also held prior roles as an Adjunct Professor at Albany Law School and US Patent Examiner.

Roser Almenar: Roser Almenar is a PhD Candidate in AI & Space Law at the University of Valencia (Spain) and serves as a member of the ITU Secretary-General's Youth Advisory Board, representing Europe. She is currently Co-Lead of the "AI and Space Law" Research Group, hosted by the Space Law and Policy Project Group of the Space Generation Advisory Council (SGAC). Moreover, she is involved in several policy divisions under the Space Generation Advocacy and Policy Platform (SGAPP), dealing with European space exploration and responsible space behavior. Her research focuses on the discipline of Space Law and Policy, especially dealing with remote sensing, AI, data protection, and telecommunications.



Lisa Kucher: Lisa Kucher is an interdisciplinary space legal professional with over five years of experience, specializing in export control compliance. Based in Toulouse, she focuses on legal and regulatory policy matters. Previously she was also advising space companies on cross-border transactions involving civil, dual-use, and military products, as well as license management and technology transfer. Lisa has drafted over 100 due diligence reports and created educational videos and boutique workshops covering the fundamentals of export control, due diligence, and risk assessment. Passionate about the intersection of legal technology, AI, and sustainable development, she is eager to see their responsible evolution in the space industry.

Laura Yvonne Zielinski: Laura Yvonne Zielinski is a PhD candidate at the Air and Space Law Institute of the University of Cologne. Her thesis covers the protection of space investments under international law. Laura has over ten years of private practice experience in international law firms and is admitted to practice in Paris, New York and Mexico. She is a member of the International Institute of Space Law, Vice-Chair of the Space Law Committee of the International Bar Association, and founder and president of the Space Arbitration Association.

Stefan-Michael Wedenig: Stefan-Michael Wedenig is Executive Director of the McGill Institute of Air and Space Law, where he also pursues a Doctorate. He earned a Master of Laws in Air & Space Law from McGill Institute of Air and Space Law and a Masters in Law at Johannes Kepler University. Before joining McGill, he worked in the diplomatic field at the Austrian Embassy in Canada. He was involved in the McGill Manual on International Law Applicable to Military Uses of Outer Space and was Assistant Editor Annals of Air and Space Law. His current research focuses on new approaches on AI in the air and space domains.

Giovanni Tricco: Giovanni Tricco is a PhD candidate at the University of Bologna and Vrije Universiteit Brussel. His research explores the legal and policy implications of integrating AI into space activities. He co-leads the AI and Space Law Research Group within the Space Law and Policy Project Group at the Space Generation Advisory Council (SGAC). Previously, Giovanni worked in tech policy at the Center for European Policy Analysis (CEPA), a Washington, DC-based think tank, focusing on research at the intersection of AI, data, and cloud governance.

List of Abbreviations

ADS	Automated Decision Systems
AECA	Arms Export Control Act
AI	Artificial Intelligence
ALD	Atomic Layer Deposition
ASMS	Automated Spectrum Management Systems
BIS	Bureau of Industry and Security
CARA	Conjunction Assessment Risk Analysis
CCL	Commerce Control List
CCSDS	Consultative Committee for Space Data Systems
CCW	Convention on Certain Conventional Weapons
CJEU	Court of Justice of the European Union
CMOS	Complementary Metal Oxide Semiconductor
COCOM	Coordinating Committee for Multilateral Export Controls
CPU	Central Processing Unit
CSO	Commercial Solutions Opening
CWC	Chemical Weapons Convention
DCNN	Deep Convolutional Neural Networks
DDTC	Directorate of Defense Trade Controls
DFARS	Defense FAR Supplement
DL	Deep Learning
DNI	Director of National Intelligence
DOC	Department of Commerce
DOD	Department of Defense
DOE	Department of Energy
DOJ	Department of Justice
DOS	Department of State
DPIA	Data Protection Impact Assessments
DTSA	Defense Technology Security Administration
DTSF	Disruptive Technology Strike Force
EAD	Ethically Aligned Design
EAR	Export Administration Regulations
EC	European Commission
ECCN	Export Control Classification Number
EHS	Environmental Health and Safety
EU	European Union
FAR	Federal Acquisition Regulation

FCC	Federal Communications Commission
FDA	Food & Drug Administration
FLOPS	Floating-point Operations Per Second
FRE	Fundamental Research Exclusion
GDPR	General Data Protection Regulation
GE	Genetic Engineering
GNSS	Global Navigation Satellite System
GPA	Government Procurement Agreement
GPS	Global Positioning System
GPU	Graphics Processing Unit
HNS	Hazardous and Noxious Substances
HPET	Human Performance Enhancement Technology
IEEE	Institute of Electrical and Electronics Engineers
IC	Integrated Circuit
ICAO	International Civil Aviation Organization
ICT	Information & Communication Technology
IDL	Interface Definition Language
IEC	International Electrotechnical Commission
IHL	International Humanitarian Law
IMO	International Maritime Organization
IRIS	Infrastructure for Resilience, Interconnectivity and Security by Satellite
ISO	International Organization for Standardization
ISS	International Space Station
ITAR	International Traffic in Arms Regulations
ITU	International Telecommunications Union
JARUS	Joint Authorities for Rulemaking on Unmanned Systems
LAWS	Lethal Autonomous Weapon Systems
LEO	Low Earth Orbit
LLM	Large Language Model
MHC	Meaningful Human Control
MIFR	Master International Frequency Register
ML	Machine Learning
MTCR	Missile Technology Control Regime
NAIAC	National AI Advisory Committee
NAIIO	National AI Initiative Office
NASA	National Aeronautics and Space Administration
NETTEM	New and Evolving Technologies Technical Experts Meeting
NN	Neural Network
NPT	Nuclear Non-Proliferation Treaty
NSG	Nuclear Suppliers Group

OECD	Council of the Organization for Economic and Development
OFAC	Office of Foreign Assets Controls
OMB	Office of Management and Budget
OTA	Other Transaction Authority
OTE	Office of Technology Evaluation
PDL	Product Liability Directive
RFI	Request for Information
RFQ	Request for Quote
RPO	Rendezvous and Proximity Operations
SaaS	Software as a Service
SDG	Sustainable Development Goals
SDR	Software Defined Radio
SME	Small and Medium-Sized Enterprises
SMS	Safety Management Systems
SOLAS	Safety of Life at Sea
SSN	Space Surveillance Network
SSO	Standards Setting Organization
STM	Space Traffic Management
STMA	Space Traffic Management Authority
TCMB	Transparency and Confidence Building Measures
TPU	Tensor Processing Unit
TRL	Technology Readiness Level
TSA	Technology Safeguard Agreement
UNCOPUOS	United Nations Committee on the Peaceful Uses of Outer Space
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNOOSA	United Nations Office for Outer Space Affairs
USML	United States Munitions List
VHR	Very High-Resolution
WA	Wassenaar Arrangement
WMD	Weapons of Mass Destruction
WTO	World Trade Organization
XAI	Explainable AI

Executive Summary

Introduction and Scope

As Artificial Intelligence (AI) becomes integral to space activities, its transformative potential underscores the urgent need for a comprehensive regulatory framework to ensure ethical, sustainable, and collaborative applications. Given the rapid evolution of AI technologies, periodic re-evaluations every 12–24 months, with interim reviews every 6–12 months, could help assess progress. Agencies could evaluate hardware performance against pre-established thresholds, triggering immediate regulatory reviews if those standards are met, with timelines negotiated by stakeholders based on technical advancements. This study examines the legal, regulatory, and ethical challenges posed by AI in space, offering actionable recommendations grounded in international treaties and established frameworks. Structured slightly differently from the main manuscript, this executive summary highlights practical solutions aimed at initiating discussions and guiding policymakers, industry stakeholders, and international organizations in navigating the complexities of AI in the evolving space domain to balance ethics and innovation.

1. *Liability Frameworks in Space Law: AI Challenges in International Treaties*

The Liability Convention and the Outer Space Treaty define "launching state" through four criteria: the state that launches, procures the launch, launches from its territory, or launches from its facility. This framework ensures that victims can identify at least one launching state, such as the registry state, to seek compensation.

This approach remains effective even with the incorporation of AI into space objects, as the clear definitions and joint and several liability provisions enable victims to seek redress despite the complexities of modern space operations. While complex launch arrangements, like those involving maritime platforms, may complicate jurisdictional issues, these treaties provide flexibility and clarity in attributing liability.

However, the increasing use of AI in space activities introduces new dimensions to this framework. AI systems that autonomously conduct space object docking and other rendezvous and proximity operations (RPO) or independently determine launch timing, configurations, or other launch parameters challenge typical assumptions and could redefine the analyses required to determine states liable for damages caused by space objects and states responsible for the procurement of a launch. Moreover, AI's reliance on globally distributed components raises intricate questions about liability attribution.

2. *Standardizing Contractual Procurement Frameworks for AI in Space Operations*

Below the state level, liability frameworks for private operators and AI developers must address the distinct risks posed by AI in space activities. Procurement contracts serve as a critical tool for managing risks in AI-driven space activities, particularly as these systems become more complex and autonomous. Effective contractual provisions can help allocate and limit liability, e.g., via indemnity obligations, establish robust risk management protocols, and ensure ongoing accountability for AI performance.

Some of the efforts in progress include the European Commission Proposal for Standard Contractual Clauses for the Procurement of Artificial Intelligence by Public Organisations, the US Office of Management and Budget (OMB) Memorandum M-24-18 on Advancing the Responsible Acquisition of Artificial Intelligence in Government, and the Institute of Electrical and Electronics Engineers (IEEE) P3119 Draft Standard for the Procurement of Artificial Intelligence and Automated Decision Systems (in development). These initiatives emphasize transparency, risk management, and clear allocation of responsibilities.

A prime example is the European Commission's Proposal for Standard Contractual Clauses for AI Procurement, which outlines specific measures to address foreseeable risks associated with AI systems. For instance, suppliers must implement a comprehensive risk management system prior to delivering the AI system, encompassing the identification, evaluation, and mitigation of risks to health, safety, and fundamental rights. These measures must include detailed documentation, which suppliers are obligated to provide at the time of delivery and keep updated throughout the lifecycle of the agreement. The risk management process should remain continuous and iterative, with ongoing testing, reviews, and updates to address evolving risks. The iterative risk management process ensures systems remain responsive to evolving threats.

For instance, a clause might specify that: "The supplier ensures that the AI system undergoes rigorous testing to verify compliance with risk management measures and provides updated documentation to the contracting entity at the time of delivery."

Similarly, indemnification clauses are essential in delineating responsibilities between parties. For example: "The supplier indemnifies the operator against third-party claims arising from intellectual property or privacy rights breaches related to the AI system."

In addition to European Commission guidelines, industry practices provide additional insights into limitation of liability provisions. For example, CEGSOFT, a provider of cloud applications and AI Software as a Service (SaaS), includes clauses emphasizing the limitations of AI systems. Their contracts acknowledge AI's probabilistic nature, encouraging cautious reliance on AI-generated outputs.

A limitation of liability clause for a license to use an AI system or platform might read: "Licensor will neither be liable nor responsible in any way for any direct, indirect, incidental, special, or consequential damages resulting from the Licensee's use or inability to use the AI Solutions. Licensee agrees to use the AI Solutions responsibly and not to rely solely on its responses for making critical or legally binding decisions."

Similarly, Checksum AI, a platform specializing in AI-powered software testing services, includes provisions addressing the probabilistic nature of AI outputs: "While Checksum is always working to improve its AI Functions, due to the probabilistic nature of the AI Functions, the services may provide inaccurate output or otherwise not always produce the intended results.

Checksum will have no liability with respect to the output (or any Party's use thereof) of Checksum's AI Functions."

These examples illustrate how procurement agreements can incorporate specific contractual language to address the inherent risks and uncertainties of AI systems. Key elements include risk identification, mitigation, testing, indemnification, and liability limitation to ensure comprehensive safeguards.

In the context of AI in space, such provisions can be further standardized and tailored to meet the unique challenges of this domain. For instance, agreements should account for scenarios where AI malfunctions lead to damages in space, ensuring that liability is clearly allocated between operators and developers. This approach ensures that operators remain responsible for in-mission risks, while developers are held accountable for defects in design or updates.

To achieve international standardization, procurement frameworks must be adapted to the space domain. Collaborative efforts should prioritize reviewing existing regulations to identify gaps specific to AI technologies in space. Developing standardized guidelines that emphasize transparency, competition, ethical considerations, and risk management will promote equitable frameworks. Moreover, fostering international cooperation for information sharing and best practices will ensure consistent and reliable AI integration in space missions.

3. *Applying International Human Rights Law to AI Applications in Space*

To ensure that AI-driven technologies in space comply with International Humanitarian Law (IHL), specific measures must be prioritized within the frameworks of the Geneva Conventions and the Convention on Certain Conventional Weapons (CCW). These frameworks provide critical guidelines for managing the dual-use nature of AI technologies, mitigating risks, and promoting accountability.

The CCW's emphasis on Meaningful Human Control (MHC) underscores the importance of enforceable standards for human oversight of autonomous systems in space. Guidelines should require that human operators maintain supervisory control, particularly in high-risk missions. However, MHC faces challenges, including differing interpretations, responsibility gaps, and the impracticality of real-time human intervention due to space communication delays. Pre-deployment reviews must evaluate compliance with IHL principles such as distinction and proportionality to mitigate these limitations and prevent errors, misuse, or unintended harm.

Addressing the dual-use nature of satellites and AI systems requires a clear framework to differentiate civilian from military applications. To protect critical civilian infrastructure, international regulations should establish clear definitions of dual-use technologies and prohibit the targeting of civilian assets, even when they serve auxiliary military purposes. Enhanced cybersecurity measures must also be implemented to safeguard against electronic attacks that could lead to indiscriminate harm. When utilizing the term "attack," we refer to "acts of violence against the adversary, whether in offence or in defence," as stipulated by Article 49(1) of the 1977 Protocol Additional to the Geneva Conventions of 12 August 1949 and relating to the Protection of Victims of International Armed Conflicts (Protocol I).

The Geneva Conventions' principles of distinction, proportionality, and necessity must guide the design and deployment of AI systems in space. Strict adherence to these principles can be supported by comprehensive risk assessments and safeguards to prevent excessive or unintended collateral damage during operations. Harmonizing international standards requires collaboration. This includes expanding dialogue under CCW auspices to address the unique challenges posed by space-based AI technologies, encouraging transparency in military AI applications, and creating mechanisms to hold states accountable for violations of IHL in space operations.

By implementing these measures, the international community can balance innovation with responsibility, ensuring AI-driven technologies in space uphold IHL principles while contributing to the security and sustainability of outer space.

4. *Advancing Standardization in Space: Regulatory Considerations for GDPR, Export Control, and Telecommunications*

GDPR

The General Data Protection Regulation (GDPR) is the EU's comprehensive framework for safeguarding personal data. Applying its principles to AI-driven space systems introduces unique challenges, including ensuring lawful data processing, adhering to data minimization and transparency, and maintaining accountability in autonomous operations. While AI's reliance on vast datasets often conflicts with GDPR norms, risks such as biases and inaccuracies can be mitigated through safeguards like pseudonymization, encryption, and Data Protection Impact Assessments (DPIAs). Clear regulatory guidance is essential to align GDPR with the evolving demands of AI in space.

To address these challenges, targeted recommendations are critical. Data protection measures must be embedded throughout the lifecycle of AI systems, from design to deployment, ensuring compliance with GDPR principles such as data minimization, purpose limitation, and transparency. Robust pseudonymization and encryption protocols are vital to safeguard personal data, particularly in high-risk contexts involving autonomous decision-making in space.

Clear guidelines should define lawful bases for data processing in emergencies, such as humanitarian crises or mission contingencies, balancing the need for rapid action with GDPR compliance. Advanced tools like DPIAs, algorithm impact assessments, and model cards should be mandated to identify, mitigate, and demonstrate accountability for risks tied to AI-driven data processing.

Data processing must remain strictly necessary, even as AI systems adapt to unexpected scenarios. Flexible yet precise definitions of "necessary data" and periodic reviews should prevent excessive or unwarranted processing. Transparency mechanisms must log and explain AI decision-making in accessible terms, enabling human operators to review, understand, and challenge automated decisions when needed.

Lastly, collaboration among regulatory bodies and industry stakeholders is essential to harmonize GDPR interpretations across jurisdictions, fostering innovation while ensuring compliance. By implementing these measures, operators can effectively address regulatory challenges, protect individual rights, and harness AI's transformative potential in space.

Export Control

The integration of AI into space exploration brings unique complexities to export control frameworks, demanding a refined approach to balance technological progress, national security, and international cooperation. This discussion examined the interplay between international agreements like the Wassenaar Arrangement and the Missile Technology Control Regime (MTCR) and national regulations such as the US International Traffic in Arms Regulations (ITAR) and Export Administration Regulations (EAR). Key issues include the dual-use nature of AI technologies, challenges in harmonizing export control lists, and the limitations of current frameworks in addressing AI's rapid evolution.

AI technologies present regulatory challenges due to their adaptability and dual-use potential. For instance, systems that improve satellite communications or enable autonomous spacecraft navigation can be repurposed for military objectives, underscoring the critical need for robust regulatory oversight. While existing frameworks provide a foundation, they often fall short in addressing the dynamic and rapidly advancing capabilities of AI. Furthermore, inconsistencies in export control laws across jurisdictions impede global alignment, complicating efforts to regulate sensitive technologies.

To address these challenges, a balanced approach is necessary. Strengthening national compliance measures, including advanced licensing processes for dual-use AI systems, is essential. Simultaneously, international organizations like the Wassenaar Arrangement must adapt their control lists to incorporate emerging AI capabilities and promote consistency among member states. Greater transparency and collaboration between nations will also be crucial in addressing regulatory gaps and preventing the unauthorized export of sensitive AI technologies.

Key recommendations emerge to guide future efforts. Developing AI-specific control lists with clear criteria for dual-use classification and risk assessment will help align regulations with technological advancements. Encouraging a bottom-up approach to harmonization, where states adopt agile national measures while aligning with international standards, can create a more cohesive framework. Enhanced transparency and information-sharing among nations are vital to close existing gaps and prevent undercutting of regulations. Lastly, forming expert teams and employing detailed risk matrices will strengthen oversight, ensuring export control systems evolve in step with advancements in AI and space technologies.

Telecommunications

The integration of AI into telecommunications for space applications requires targeted action to address communication protocols, spectrum allocation, and security. Standardized communication protocols should prioritize interoperability, resilience, and efficiency, particularly in the context of deep-space missions. Dynamic spectrum management frameworks must adapt to the unique demands of AI-driven systems, including real-time spectrum allocation, enhanced monitoring, and interference management, to optimize the use of limited radio frequencies and satellite orbits. Future updates to International Telecommunications Union (ITU) Radio Regulations and agenda items for upcoming World Radiocommunication Conferences should incorporate specific provisions addressing AI technologies.

The ITU, as a key international body, could lead by forming a study group on AI in space telecommunications and conducting a comprehensive review of existing spectrum standards to align them with the evolving needs of AI-driven systems. Collaborative efforts with other international entities, such as the United Nations Committee on the Peaceful Uses of Outer Space (UNCOPUOS), should focus on promoting equitable and secure spectrum allocation, facilitating the development of AI-enabled satellite frameworks, and addressing governance gaps. Strengthened security measures, including encryption and advanced data-sharing protocols, are vital to mitigate risks of cyberattacks and ensure confidentiality in AI-driven space systems.

By addressing these areas through enhanced cooperation, regulatory adaptation, and technical innovation, the ITU and international stakeholders can support the sustainable and equitable integration of AI into space telecommunications, ensuring its alignment with broader goals of innovation, security, and sustainability.

5. *International Guidelines for Ethical AI Governance in Space*

The integration of AI into genetics, robotics, border control, and space exploration showcase immense potential but also introduces significant ethical and governance challenges. In genetics, AI enhances genome sequencing, optimizes CRISPR-Cas9, and advances targeted therapies. Tools like SPROUT predict off-target effects, improving precision medicine.

However, these innovations raise risks, including unintended mutations, concerns about human enhancement, and inequitable access. Ethical frameworks must prioritize informed consent, data security, and equitable access, requiring transparency in gene-editing processes, public reporting on objectives, and safeguards to mitigate risks. In space, proposals for gene-editing astronauts for radiation resistance or microgravity adaptation must undergo rigorous ethical reviews to ensure safety and fairness.

In robotics, the rise of lethal autonomous weapon systems (LAWS) exemplifies AI's dual-edged capabilities. These systems, capable of independently targeting and engaging, pose challenges regarding accountability, compliance with international humanitarian law, and the necessity of human oversight. While not a complete solution, it remains a critical safeguard, ensuring accountability and compliance while research into AI's reliability and decision-making progresses.

Autonomous systems in space face similar risks, including unauthorized satellite maneuvers and algorithmic bias. Recommendations include robust accountability measures, such as decision logs, clear command chains, and rigorous testing to meet international standards. Frameworks like those from the European Commission's High-Level Expert Group on AI should be expanded under UNCOPUOS to address space-specific concerns, including collision avoidance and unauthorized surveillance.

AI applications in border control, such as drones, facial recognition, and surveillance robots, bring efficiency but risk privacy violations, bias, and human rights abuses. These risks mirror space surveillance challenges, where unauthorized data collection and regional bias are significant issues.

Recommendations include diverse training datasets to mitigate bias, strict data governance for transparency and accountability, and fail-safes to prevent unapproved actions. AI systems should prioritize fairness, with algorithms regularly audited for compliance with ethical standards. Collaborative international agreements under UNCOPUOS should codify these practices into enforceable regulations to ensure responsible AI deployment.

By addressing shared ethical challenges across genetics, robotics, border control, and space systems, stakeholders can ensure AI fosters innovation responsibly while protecting human rights, equity, and global security.

OECD, UNESCO, EU AI Act

The OECD Principles on AI, UNESCO's Recommendations on the Ethics of Artificial Intelligence, and the EU Artificial Intelligence Act form a robust framework for ethical AI governance in space exploration, emphasizing transparency, accountability, and collaboration to address challenges while fostering innovation.

Key recommendations include internationally standardized AI regulations to ensure compliance with ethical and safety standards. Interoperable global standards are essential for seamless collaboration, while mandating explainable AI (XAI) enables accountability through transparent reporting and human oversight.

Data governance policies must prioritize privacy, fairness, and compliance with frameworks like the GDPR. Sustainability efforts should leverage AI for debris management and energy-efficient satellite operations, minimizing environmental impact.

Inclusivity is critical, with global consortia promoting equitable access to AI benefits and bridging technological gaps between nations. Educational programs and international workshops should raise awareness and embed ethical AI practices, ensuring space technologies responsibly serve humanity.

Integrating these recommendations into national and international policies will balance innovation with transparency, equity, and sustainability in AI-driven space exploration, advancing technologies responsibly while preserving the space environment.

6. *Standardizing AI Regulation for Environmental Protection in Space*

The rapid growth of space exploration and satellite deployment demands innovative strategies to ensure safety, sustainability, and functionality. AI is pivotal in addressing these challenges, offering capabilities for autonomous monitoring, satellite control, collision avoidance, and debris mitigation. However, the absence of standardized frameworks for AI in space introduces significant risks, including ethical concerns and operational inefficiencies, necessitating comprehensive regulation.

AI-controlled systems have demonstrated their value in automating spacecraft operations, as seen with SpaceX's Dragon spacecraft and ESA's Sentinel-1 satellite, which use AI for autonomous docking, collision avoidance, and precise monitoring. AI algorithms process massive datasets in real time, improving trajectory estimation, orbit determination, and risk assessment, vital for navigating increasingly congested Low Earth Orbit (LEO), especially with mega-constellations like SpaceX's Starlink. ESA's automated systems and NASA's Conjunction Assessment Risk Analysis (CARA) enhance collision avoidance, reducing false alerts and improving decision-making, while reinforcement learning algorithms further optimize maneuvers, balancing collision avoidance with fuel efficiency. These systems also improve debris tracking and real-time navigation, reducing reliance on manual interventions.

Sustainability remains crucial, with AI strategies like satellite fuel optimization, lifespan extension, and debris removal aligning with UNCOPUOS Space Debris Mitigation Guidelines. Initiatives such as ESA's Zero Debris Charter and updated FCC regulations emphasize integrating sustainability into AI-driven technologies to protect the orbital environment.

To address the regulatory void, the establishment of a Working Group on AI Governance in Space within the Legal Subcommittee of UNCOPUOS is essential. This body could develop phased guidelines and standards, building on successful models like the Space Resources Working Group, while ensuring stakeholder participation. Subcommittees could address targeted issues, such as satellite communications, debris management, and planetary exploration.

Complementing this, interoperable protocols and robust data governance frameworks are essential to prevent bias, protect privacy, and enhance accountability. Regulatory sandboxes can refine AI applications, and transparent reporting mechanisms with human oversight are critical for systems managing tasks like collision prediction and autonomous maneuvers.

International collaboration remains pivotal. Non-binding guidelines, developed through UNCOPUOS or facilitated by the United Nations Office for Outer Space Affairs (UNOOSA), can harmonize AI standards globally, with potential to evolve into binding agreements over time. Educational initiatives and public-private partnerships can further build capacity and foster ethical practices, ensuring equitable access to AI advancements for all space-faring nations.

Through these combined measures—clear governance structures, technical innovation, and international collaboration—the integration of AI in space can be guided toward ethical, sustainable, and secure applications. These steps are essential not only for advancing space exploration but also for ensuring the long-term safety and shared benefits of humanity's activities in outer space.

7. *Soft Law and the Role of International Standards*

In the absence of binding international agreements, soft law and technical standards are indispensable tools for shaping the governance of AI in space. Organizations like the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) have already developed AI risk management and ethical standards that can be applied to space activities. These standards, though non-binding, provide a framework for promoting the responsible development and use of AI technologies in space.

These soft law instruments may gradually evolve into customary international law as more nations adopt them. By adhering to these standards, spacefaring nations can ensure the development and deployment of AI technologies that uphold human rights, enhance safety, and foster international cooperation. This adaptable regulatory approach is essential for addressing the rapidly evolving technological landscape of space activities.

8. *Applying Aviation and Maritime Frameworks to AI Regulation in Space*

The International Civil Aviation Organization (ICAO) and the International Maritime Organization (IMO) offer proven frameworks that can guide the regulation of AI systems in space, emphasizing safety, security, and operational efficiency through internationally accepted standards.

From ICAO, Safety Management Systems (SMS) and collision avoidance protocols provide key models. SMS uses a structured approach to risk identification, safety monitoring, and corrective measures, which could enhance orbital safety and system reliability if adapted to space operations. ICAO's standardized collision avoidance protocols, emphasizing clear communication and procedures, could improve coordination among AI systems in space.

IMO's standardized navigation and communication systems under the Safety of Life at Sea (SOLAS) Convention highlight the value of consistent guidelines across jurisdictions. Applying these principles to space operations would support the development of interoperable communication protocols and AI system standards. IMO's emergency response protocols also offer a model for managing contingencies like satellite malfunctions or debris collisions.

Insurance frameworks from aviation and maritime sectors, such as the Montreal Convention and the Carriage of Hazardous and Noxious Substances (HNS) Convention, provide structures for liability and compensation that could address financial risks in space activities. A liability regime similar to the Montreal Convention could impose strict liability on operators, requiring claimants to prove only an incident occurred and its connection to damage. Exoneration clauses, like Article 20 of the Montreal Convention, could limit liability when claimants' negligence contributes to the damage.

Mandatory insurance requirements, modeled after the HNS Convention, would ensure operators of AI-powered spacecraft maintain sufficient coverage for personal injury, property damage, environmental harm, and AI malfunctions. An international space liability fund, akin to the IOPC Funds in maritime law, could provide financial security for damages exceeding operator coverage, encouraging safety adherence and ensuring compensation.

A robust liability framework should define tiers of responsibility, placing primary accountability on operators while supplementing coverage through an international fund. Liability would include personal injury, property damage, environmental harm, and AI-driven impacts, with seamless transfer provisions during ownership changes. Explicitly incorporating AI into the definition of space objects, by clarifying that AI software should be regarded as a component of a "space object," within liability frameworks would enhance regulatory clarity and accountability. This may need revisiting once AI evolves to perform autonomous functions beyond the control of parent space objects, to address emerging challenges and ensure relevance.

Both ICAO and IMO emphasize the importance of regularly updating standards to keep pace with technology. Similarly, the space sector can adopt periodic revisions to AI regulations, along with robust insurance mechanisms, to enhance safety, promote international collaboration, and address challenges such as orbital congestion, system interoperability, and financial risks.

9. *Strategic Phases for Space Traffic Management: Short and Long-Term Approaches*

AI plays a critical role for Space Traffic Management (STM), enhancing situational awareness, improving collision avoidance, and automating complex processes in increasingly congested orbital environments. Effective STM requires a phased approach, addressing immediate challenges while building a comprehensive governance framework.

Short-term priorities include creating standardized data-sharing protocols for satellite operators, fostering collaboration, and addressing immediate risks like collisions and orbital debris. Models like Global Navigation Satellite Systems (GNSS) and Joint Authorities for Rulemaking on Unmanned Systems (JARUS) offer frameworks to enhance transparency, interoperability, and situational awareness, building the foundation for structured STM practices.

Over the long term, STM governance should progress through defined phases. Initially, nations can adopt domestic regulations tailored to their operational contexts, supporting bilateral agreements that facilitate data exchange and cooperative cross-border STM efforts. These bilateral initiatives can evolve into international technical standards, developed under organizations like the ISO, ensuring consistency in processes, reducing conflicts, and supporting global interoperability.

Ultimately, these incremental steps would pave the way for establishing a Space Traffic Management Authority (STMA), an independent international body modeled after ICAO and IMO. The STMA would oversee global STM, regulate AI-driven technologies, coordinate international efforts, and enforce compliance through technical and legal mechanisms. Its duties could include collaborating with ISO to refine technical standards for collision avoidance, debris mitigation, and satellite maneuvering; addressing transparency protocols; detecting dual-use risks; preventing misuse of autonomous systems; and tackling military and security challenges in STM.

Two governance models are proposed for long-term STM. A centralized model, embodied by the STMA, would provide a cohesive, enforceable framework, ensuring global consistency. Alternatively, a decentralized approach would allow nations or regions to manage STM standards independently through bilateral or sector-specific agreements, which would then need to be aligned to minimize friction and ensure interoperability. While decentralization offers flexibility and quicker implementation, it risks fragmentation and inconsistencies, complicating coordination. The centralized model requires more diplomatic effort but ensures unified traffic management.

By integrating short-term actions with a long-term vision, these recommendations encourage international cooperation, balance innovation with regulatory consistency, and ensure the safe, sustainable integration of AI technologies in increasingly crowded orbital environments.

Overview

AI is increasingly integrated into various space activities, revolutionizing the industry by enabling advanced autonomous systems and improving efficiency. AI has become essential across all facets of space operations, spanning exploration and applications, and affecting both civilian and military domains. It is already playing crucial roles in STM, satellite operations, communications, cybersecurity, robotics, novel materials and component designs, genetics, synthetic biology, healthcare, and data analytics.

For instance, AI algorithms optimize mission planning, enhance navigation accuracy, monitor satellite health, predict failures, manage space debris tracking, and improve communication systems. Drawing regulatory lessons from fields like robotics, genetics, autonomous vehicles, algorithmic decision-making, and privacy and data protection provides a pathway to inform the development of safeguards and guidelines, ensuring ethical deployment of AI systems in space. These efforts must conform to current legislative frameworks that protect human rights and privacy.

This research study explores the legal, regulatory, and ethical dimensions surrounding the use of AI in space, identifying key legally binding international instruments such as the 1967 Outer Space Treaty and the 1972 Liability Convention. These form the foundation for governing AI-driven activities in space.

In addition, the study delves into CCW and the Geneva Conventions, focusing on military uses of AI in space. In particular, the concept of MHC is critical to ensuring that AI systems adhere to IHL principles, including distinction, proportionality, and accountability. This enables AI systems used in military operations to be governed responsibly and effectively, minimizing the risks of unintended consequences.

GDPR, adopted by the European Union, provides a model for data privacy and security, especially concerning the collection and processing of personal data in space applications. GDPR Recital 71 and Article 25 emphasize the importance of privacy, and these principles are vital in the context of AI systems collecting data from space, particularly in the use of Very High-Resolution (VHR) cameras, which capture personal data without the data subject's explicit consent.

By addressing privacy, confidentiality, and automated decision-making as defined in GDPR, this study advocates for clear legal protections to safeguard individuals against potential data breaches and misuse by AI systems in space.

The regulatory section addresses export controls, procurement issues, and telecommunications, discussing the role of international organizations such as the Wassenaar Arrangement, the ITU, and the World Trade Organization (WTO). These agreements and frameworks can support the responsible development, export, and deployment of AI technologies in space. For instance, the WTO's Agreement on Government Procurement (GPA) fosters open competition in AI-enabled goods and services, while the ITU sets technical standards and facilitates global communication infrastructures. IEEE's AI procurement standards offer additional guidelines, particularly for government AI procurement, ensuring transparency, fairness, and compliance.

Additionally, this study explores case studies in genetics and robotics, drawing parallels between AI's application in these fields and its use in space. AI has revolutionized genetics through improved genome sequencing, enhanced genetic therapies, and augmented CRISPR technologies. However, this progress raises concerns over data privacy, discrimination, and the risks of destabilizing the human genome through excessive manipulation.

Similarly, AI-powered LAWS and robotic systems in military operations highlight the ethical dilemmas of delegating life-and-death decisions to machines. These case studies emphasize the importance of thoughtful governance in regulating AI in sensitive areas like genetics and autonomous robotics.

This study also delves into the environmental impact of AI-driven space systems, focusing on the growing issue of space debris. AI can significantly mitigate this threat by tracking and removing debris, predicting collisions, and enhancing space traffic management systems. The study analyzes the potential of implementing mandatory international regulations for AI-driven space debris removal operations, ensuring transparency and accountability. It also explores the development of real-time AI-driven monitoring systems with standardized data sharing protocols across nations to optimize debris mitigation efforts. Robust regulatory frameworks are required to govern these activities and ensure the long-term sustainability of space operations.

Finally, the study examines voids and recommendations, discussing the role of soft law and international standards in governing AI technologies in space, drawing lessons from aviation and maritime law, and exploring approaches to STM. These insights emphasize the importance of non-binding frameworks like those from ISO and Consultative Committee for Space Data Systems (CCSDS), proven models from ICAO and IMO, and phased STM strategies, including short-term data-sharing measures and long-term governance options such as a centralized Space Traffic Management Authority (STMA) or decentralized national agreements aligned for interoperability, to ensure safe and sustainable space operations.

In conclusion, this research outlines the legal, regulatory, and ethical frameworks required to govern the growing role of AI in space, advocating for international cooperation to address voids in regulation and oversight. It provides actionable recommendations aimed at promoting the responsible, transparent, and ethical use of AI technologies in space. This study calls for continued exploration and regulation in these areas to ensure that AI contributes to the advancement of space activities while protecting human rights, the environment, and global security.

PART I: LEGAL FRAMEWORKS

Section 1: Legal Binding International Laws Relating to Use of AI in Space

Introduction

This section delves into the intersection of international legal frameworks and the evolving role of AI technologies in space. As humanity continues to explore outer space, the use of AI introduces new challenges, particularly in how established legal principles—originally designed for human decision-making—apply to autonomous systems.

Central to this discussion are key international treaties such as the 1967 Treaty on Principles Governing the Activities of State in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies (Outer Space Treaty) and the 1972 Convention on International Liability for Damage Caused by Space Objects (Liability Convention), which serve as the backbone of space law. These agreements outline principles related to the peaceful use of space, non-appropriation, state responsibility, and liability for damages, all of which are critically relevant as AI becomes more prevalent in space operations.

In addressing the growing presence of AI in space, this section explores how the foundational principles set forth in the Outer Space Treaty are applied to autonomous systems. The Treaty, adopted in 1967, establishes the fundamental rules for space exploration and usage, emphasizing peaceful purposes, equality of access, and the prohibition against national appropriation of celestial bodies.

These norms are particularly significant when applied to AI, as the autonomy of such systems can potentially blur the lines of state responsibility and sovereign claims in space. A deeper examination of Article II and Article VI of the Outer Space Treaty will highlight how the Treaty addresses issues such as non-appropriation and state responsibility, as well as the potential complications that arise when AI technologies are employed in space missions.

Beyond the Outer Space Treaty, the Liability Convention offers a more detailed framework concerning the responsibilities and liabilities of states for damages caused by space objects. As AI takes on greater roles in tasks such as launch operations and satellite management, questions of liability become increasingly complex.

The Convention’s provisions on absolute and fault-based liability provide a structure for determining responsibility, but the introduction of AI adds a layer of ambiguity regarding fault, particularly in scenarios where AI operates autonomously and across multiple jurisdictions. This section will explore these challenges, especially how AI may impact the interpretation of terms like “launching state” and the criteria for assigning liability in international disputes.

1. *Outer Space Treaty*

The Outer Space Treaty is the bedrock of international space law and stands at the beginning of what is often considered a "golden era" of treaty creation relating to space.[1]

The period of the 1960s and 70s saw the conclusion of the five UN space treaties, the Outer Space Treaty, the Agreement on the Rescue of Astronauts, the Return of Astronauts, and the Return of Objects Launched into Outer Space (Rescue Agreement), the Liability Convention, the Convention on Registration of Objects Launched into Outer Space (Registration Convention), and the Agreement Governing the Activities of States on the Moon and Other Celestial Bodies (Moon Agreement).

As the name of the Outer Space Treaty suggests, it lays down principles governing the use and exploration of space by States. Subsequent legal agreements—including the other four UN space treaties—accordingly build and expand upon the principles laid down in the Outer Space Treaty.

Given the overarching nature of the Outer Space Treaty, it will also apply to the ever-increasing use of AI in space applications. Therefore, a discussion of the relevant principles of the Outer Space Treaty is necessary.

Before analyzing certain provisions of the Outer Space Treaty in more detail, it is interesting to mention that the above-cited international space treaties do not cover all aspects of activities in space. This raises the question whether other legal rules can be applied to those situations in space, which are not regulated by specific rules of international space law.

The Outer Space Treaty provides in Article I that "[o]uter space, including the Moon and other celestial bodies, shall be free for exploration and use by all States without discrimination of any kind, on a basis of equality and in accordance with international law, [...]" and adds in Article III that space activities shall be carried out "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." This is understood to imply the application of general international law to those aspects of space activities not otherwise addressed by the space treaties.

Space law may thus be designated as *lex specialis* in relation to general international law. International space law is not a self-contained regime as it does not exclusively decide on all possible measures of implementation and procedure and on all respective primary rules on all problems.[2] For situations not addressed by space law, general international law serves as a fallback option.

---

[1] See Stefan-Michael Wedenig & Jack Wright Nelson, "Moon Agreement: Hanging by a. Thread?" (26 January 2023), online: *McGill Institute of Air and Space Law, https://www.mcgill.ca/iasl/article/moon-agreement-hanging-thread*

[2] Stephan Hobe, *Space Law* (Beck,2023) at. 55.

1.1 *Article II*

Article II of the Outer Space Treaty encompasses the "Non-Appropriation" principle and states that "[o]uter space, including the Moon and other celestial bodies, is not subject to national appropriation by claim of sovereignty, by means of use or occupation, or by any other means." The drafters opted to formulate this principle very broadly, declaring that no national appropriation by any means is permissible under the Outer Space Treaty.

This equally means that States cannot employ AI to make claims of sovereignty, occupy a certain area of a celestial body or appropriate territory by using AI over or within a specific area. For example, AI could not be used to autonomously set up a perimeter, control access to that area and prevent others from utilizing a certain area on a celestial body beyond the regulated concept of "safety zones," which could not entail any claim of sovereignty.

The question of whether resource extraction is prohibited under Article II is subject to debate. On the one hand, Article II is quite clear in its wording prohibiting appropriation "by any other means" but, on the other hand, States seem to apply a different interpretation.

For instance, the 43 States subscribing to the Artemis Accords affirm "that the extraction of space resources does not inherently constitute national appropriation under Article II of the Outer Space Treaty, and that contracts and other legal instruments relating to space resources should be consistent with that Treaty."[3]

Likewise, national laws enacted by the United States, Luxembourg, the United Arab Emirates, Japan, and Brazil authorize private resource extraction in outer space, while affirming that they consider this interpretation of Article II as compliant with the Outer Space Treaty. For space actors deploying AI applications or otherwise using AI in space, those activities remain subject to and must comply with the non-appropriation principle of Article II.

1.2 *Article VI*

Article VI of the Outer Space Treaty is an innovative provision, insofar as it does something remarkable in international law: it makes States directly responsible for the activities of non-governmental entities (i.e., private actors); an aspect that is treated differently under the norms of general international law.

Under public international law, a State is generally only internationally responsible if there has been a breach of an international obligation and that violation is attributable to the State.[4] Under the existing attribution rules, actions of organs, persons and entities equipped with governmental authority or actions of persons or entities under the control of a State, are attributable to the State.[5] The actions of private actors generally do not impute the State and give rise to

---

[3] Artemis Accords, Section 3

[4] Responsibility of States for internationally wrongful acts, GA Res 56/83, 56th Sess, UN Doc A/Res/56/83 (2002), art 2 (ILC Articles).

[5] *Ibid.*, arts 4, 5, 8

responsibility, unless a State was under the obligation to prevent a certain conduct of private actors.[6]

Article VI removes the attribution problem and imputes the State directly, even when the activities are conducted by private actors, as long as the activity can be considered a "national activity in outer space." The responsibility of a State for its national activities in outer space is only one part of the responsibility aspect prescribed by Article VI.

A State party is further responsible for "assuring that national activities are carried out in conformity with the provisions set forth" in the Outer Space Treaty and that all activities of non-governmental activities are authorized and fall under the continuing supervision by the appropriate State party.

Although Article VI of the Outer Space Treaty seems to be straightforward at first glance, the wording is far from clear and multiple aspects of this provision have been debated in the literature. First, the article does not define what should be considered a "national activity in outer space."

It is not settled whether "national activities" for the purpose of Article VI of the Outer Space Treaty refer to activities of nationals, to activities conducted from national territory, any combination thereof, or yet a third criterion for determining its scope.[7] In addition, the treaty does not tell us whether an activity conducted on the surface of the earth can also be considered a "national activity in outer space."

In this regard, Wassenbergh, in the minority, favors a definition contained within domestic law, that is, domestic law is to determine what should be considered "national" in the context of Article VI of the Outer Space Treaty.[8] The majority of authors, however, seem to lean towards the understanding that a definition of the term "national" should revolve around the Outer Space Treaty directly, without reference to domestic law.[9]

Arguably, the use of AI in support of an activity that would qualify as a national activity under the current legal framework would also qualify as a national activity according to these same criteria. For example, collision avoidance systems operated as part of a satellite system would qualify as national activity similarly to the operation of the satellite system itself.

However, whether the use and development of AI can be considered a national activity in outer space is not clear and will need to be determined in due course.[10] This determination will likely depend on the context of the space activity in which AI is being used. In other words, it

---

[6] See e.g. Timo Koivurova,"What is the Principle of Due Diligence" in Jarna Petman & Jan Klabbers, eds, *Nordic cosmopolitanism: essays in international law for Martti Koskenniemi,* (Leiden: Martinus Nijhoff Publishers, 2003).

[7] Frank von der Dunk, International Satellite Law (2019) at 8.

[8] See Henri Wassenbergh, "An International Institutional Framework for Private Space Activities" (1997) 22 Ann Air & Space L 529.

[9] See e.g. Bin Cheng, "The Commercial Development of Space: The Need for New Treaties" (1990) 19:1 J Space L 17 at 36ff; He Qizhi, "Certain Legal Aspects of Commercialization of Space Activities" (1990) 15 Ann Air & Space L 333; Hanneke Louise van Traa-Engelman, "Problems of State Responsibility in International Space Law" (1983) 26 Proc on L Outer Space 139;

[10] See e.g. Stefan-Michael Wedenig & Jack Wright Nelson, "Artificial Intelligence in Outer Space: The Responsibility of the Software Developer under Article VI Outer Space Treaty" (paper delivered at the 75th International Astronautical Congress 2024), online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5084614

would be appropriate to consider the use of AI as "national activity in outer space" if the underlying activity can be considered a national activity in outer space. For instance, the use of AI in operating a ground station to communicate, track, and position a satellite would likely qualify as a "national activity in outer space" because the operation of the ground station itself is considered such an activity.

Conversely, it could also be posited that the development of AI is not, in and of itself, a "space activity"; as in other domains, such as medicine or automotive technology, AI is a supporting element within a larger system, but it is not an independent activity. For example, in the case of an AI-equipped satellite for autonomous navigation, the development of the autonomous navigation technology itself does not fall within the remit of space activity, but rather within that of terrestrial activity, which, if applicable, could be applied to outer space. The integration of AI within satellite, probe, or rover systems is analogous to the integration of other components such as propulsion systems, solar panels, or cameras. Consequently, it could likewise be contended that the development and incorporation of AI in space objects should not be regarded as a "national space activity" in the terms of Article VI of the Outer Space Treaty.

Second, the Outer Space Treaty does not determine which State is the "appropriate State party." This, rather unfortunate, issue is crucial as Article VI links the obligation to authorize and to provide continuing supervision to the appropriate State party. To date, this term has not been defined but the majority of the literature leans towards the understanding that the appropriate State is at the very least the State that has jurisdiction over the space object.[11]

Third, whether States accept full responsibility for all national activities conducted in outer space by non-governmental entities or whether responsibility only extends to ensuring compliance with the provisions of the Outer Space Treaty, has also been debated in literature.[12] Nevertheless, the majority of the literature seems to lean towards the understanding that a State would accept

---

[11] See e.g. Imre Anthony Csabafi, The concept of state jurisdiction in international space law: a study in the progressive development of space law in the United Nations, (The Hague: Nijhoff, 1971) at 122–23; Karl-Heinz Böckstiegel, "Term Appropriate State in International Space Law" (1994) 37 Proc on L Outer Space 77; Sergio Marchisio, "National Jurisdiction for Regulating Space Activities of Governmental and Non-Governmental Entities" (Paper delivered at the United Nations Workshop on Space Law held in Bangkok, Thailand on 16 – 19 November, 2010) at 4; Tanja L Masson-Zwaan & Mahulena Hofmann, Introduction to space law, 4th ed (Alphen aan den Rijn: Kluwer Law International, 2019) at 25; Frans Von der Dunk, "Private enterprise and public interest in the European 'spacescape'" (PhD Leiden University, 1998) [unpublished] at 19, 297; Michael Gerhard, "Article VI" in Stephan Hobe et al, eds, Cologne commentary on space law, vol 1 (Köln: Heymanns, 2009) 103; Bin Cheng, "Article VI of the 1967 Space Treaty Revisited: "International Responsibility," "National Activities," and "the Appropriate State" (1998) 26:1 J Space L 7 at 9 at 30.

[12] See e.g. Irmgard Marboe, "National Space Law" in Frans G von der Dunk & Fabio Tronchetti, eds, Handbook of space law, (Cheltenham: Edward Elgar Publishing, 2015) 127 at 131–32; Gina Petrovici & Antonio Carlo, in P J Blount et al, eds, Proceedings of the International Institute of Space Law 2018, (The Hague: Eleven International Publishing, 2019) 71, s 5.1.1; Armel Kerrest, "Remarks on the Responsibility and Liability for Damages Caused by Private Activity in Outer Space" in Proceedings of the Fortieth Colloquium on the Law of Outer Space, (Reston: American Institute of Aeronautics and Astronautics, 1998) 134; Ricky J Lee, "Liability Arising from Article VI of the Outer Space Treaty: States, Domestic Law and Private Operators" in Proceedings of the Forty-Eighth Colloquium on the Law of Outer Space 2005, (Reston: American Institute of Aeronautics and Astronautics, 2006) 216 at 217; Cheng, "Article VI revisited," supra note 3, s 3.1.4.

full responsibility. This would also include any responsibility that might arise from the use of AI within the context of a national activity in outer space.

The introduction of AI-enabled components will add another, more ambiguous, layer to this problematique. Imagine, for example, an AI-enabled communication satellite that utilizes a cloud computing element. The satellite is operated by State A, the AI was developed and is maintained by a company incorporated in State B, the cloud server is situated in State C and the data center in State D.

Which State will then be the "appropriate State party" to authorize and provide continuing supervision? AI has the ability to go beyond our traditional conceptions and could drastically change the number of States potentially involved in a space activity.

Opinions amongst the IISL working group are divided. One member argues that as long as the development of AI software for outer space applications falls within the definition of national activity in outer space, particularly in light of the likely scenario that only the State in which the software developer is situated has the ability to provide continuing supervision and authorization (e.g., for AI software maintenance, updates, and support), then the State of the software developer could be held responsible under Article VI.

Other members are of the view that the State which deploys and actually controls and operates the AI-enabled space object should be held responsible rather than a State which merely provides the servers in which the AI is hosted. At the date of writing of this report, the question remains open and further research is needed.

1.3 *Article VII*

Whereas Article VI is concerned with the responsibility of States, Article VII addresses the liability of the launching State for damage caused by a space object. The liability in Article VII of the Outer Space Treaty is anchored in the "launching State," that is, the State that launches, procures the launching of a space object and from whose territory or facility a space object is launched (art. I(a) LC). Article VII comes with a myriad of interpretational problems.[13]

There is a substantial body of literature that has developed concerning the correct interpretation of the wording "to procure" and to date there is no definitely accepted definition.

---

[13] Ricky J Lee, Law and Regulation of Commercial Mining of Minerals in Outer Space, (Dordrecht: Springer Netherlands, 2012) at 140; Ian Awford, "Legal Liability Arising from Commercial Space Activities in Outer Space" in Said Mosteshar, ed, Research and invention in outer space: liability and intellectual property rights, (Dordrecht: M. Nijhoff Publishers, 1995) 95 at 102; Armel Kerrest & Caroline Thro, "Liability for damage caused by space activities" in Ram S Jakhu & Paul Stephen Dempsey, eds, Routledge handbook of space law, (New York: Routledge, 2017) 59 at 61; Kai-Uwe Schrogl & Charles Davies, "A New Look at the 'Launching State': The Results of the UNCOPUOS Legal Subcommittee Working Group "Review of the Concept of the 'Launching State' 2000-2002" (2002) 45 Proc on L Outer Space 286; Armel Kerrest, "Launching Spacecraft from the Sea and the Outer Space Treaty: The Sea Launch Project" (1997) 40 Proc on L Outer Space 264; Bruce A Hurwitz, State liability for outer space activities: in accordance with the 1972 Convention on International Liability for Damage Caused by Space Objects, (Dordrecht: M. Nijhoff Publishers, 1992) at 21–23; Nicolas Matte Mateesco, Space Activities and Emerging International Law, (Montreal: Centre for Research of Air and Space Law, 1984) at 305; Kayser, supra note 42 at 305; Frans Von der Dunk, "International Space Law" in Frans G von der Dunk & Fabio Tronchetti, eds, Handbook of space law, (Cheltenham: Edward Elgar Publishing, 2015) 29 ay 83ff.

Suggestions in this regard range from the State that finances, requests, or benefits from the launch, to the States whose nationals order or provide financing for it.[14] It is not hard to imagine that AI will eventually take over some launch procedures.

Investigating what the drafters of the 1972 Liability Convention discussed about procurement, there are some remarks that are worth mentioning. First, the US proposal to consider a State which *"actively and substantially participates in the launching of an object into outer space"*[15] as a "launching State" fulfilling the procurement criterion. Second, building upon this perspective, the Japanese delegation's nuance provides a more detailed explanation:

> *"'Procure' consists of two requirements mentioned in the United States draft convention (A/AC.105/0.2/L.19), namely "actively and substantially participate." The Japanese delegation interprets that "actively participate" means participation in the decision of launching through agreement or consultation with the launching state, and "substantially participate" means participating in the substantial part of the project. The Japanese delegation also interprets that the manufacture of space objects or technical assistance for the manufacture of them or for the drawing up of a plan of space object is not, by itself, included in "actively and substantially participate."*[16]

However, it remains an open question as to whether AI systems will ultimately be given appropriate authority needed to take over all steps of launch procurement, potentially including determining that a new space asset is needed (e.g., a communications or debris remediation satellite), requisitioning the production or manufacture of the space asset, and undertaking the decision of whether and when to conduct a launch for the new space asset, and also especially executing a purchase order for the required launch services.

Nevertheless, the concept of "launching state," as enshrined by the Liability Convention, clearly lies in the four aforementioned criteria, according to which multiple states can be classified as "launching states" under the Outer Space Treaty and the Liability Convention. In consequence, as the Convention permits victims to pursue compensation from either any or all launching states involved (art. IV(2)), they are generally able to identify at least one entity to which they can submit a claim.

In cases involving complex launch arrangements, such as those from maritime platforms —which can get complicated between nationality of ops team, flag of vessel, nationality of vessel owner, etc.—challenges in discerning the appropriate territory of a launching state might arise. However, other launching states, such as the registry state, can typically be traced back, thereby enabling the victim to seek redress for the damage incurred. These treaties allow the launching

---

[14] Rebecca J Martin, "Legal Ramifications of the Uncontrolled Return of Space Objects to Earth" (1980) 45:2 J Air L & Com 457; Karl-Heinz Böckstiegel, "The Term 'Launching State' in International Space Law" (1994) 37 Proc on L Outer Space; Peter D Nesgos, "International and Domestic Law Applicable to Commercial Launch Vehicle Transportation" (1984) 27 Proc on L Outer Space 98; William B Wirin, "Practical Implications of Launching State - Appropriate State Definitions" (1994) 37 Proc on L Outer Space 109; Hurwitz, supra note 67; Mateesco, supra note 67; Stephan Hobe, Space law, 1st ed (Baden-Baden: Nomos Verlagsgesellschaft, 2019) at 80.

[15] United States of America: proposal (A/AC.105/0.2/L.19 and L.58).

[16] Japan: Working Paper (A/AC.105/C.2/L.61 and Corr. 1).

states to decide among themselves how liability is allocated, with the possibility that all may be jointly and severally liable (arts. IV and V).

A minor complication can occur when the launching state is not a treaty party, or when tracing authorization and supervision becomes challenging. While this does not prevent the process, it can make it more difficult to execute. Against this backdrop, it is possible to trace at least one launching state easily and this is the registering state. The concept's flexibility generally benefits victims by providing multiple avenues for reimbursement.

Translating this narrative into an AI-powered space context, identifying the launching state (or states) seems not to be ambiguous given the straightforward nature of the Convention's provisions to locate at least one launching state (e.g., registering state) even when AI tools in outer space are involved. In addition, it is important to outline that inter-party agreements must address AI-related provisions, which will be further analyzed in the procurement section, specifically regarding product liability and private operators, along with the provision of targeted recommendations on contract clauses for holding operators and entities accountable for AI-driven activities in space.

Indeed, AI seems to be uniquely suited for this as it is capable of real-time data collection taking into account various parameters such as weather data, space debris, technical issues, load factors, etc. This will inevitably raise the question whether the use of AI could be seen as to "procure the launch" of a space object by the State that uses the AI application. It would be advisable for States to introduce appropriate national legislation to account for this development, especially due to the fact that AI could entirely replace human involvement in a launch.

With the increased use of AI in launch procedures, the question arises how the development and operation of the corresponding AI system fits into the current definition of a "launching State." If, for example, the AI system that is being used was developed in a State which would otherwise not fall under the definition of a launching State, could that State escape all liability for damage by the space object operated by that AI system on the basis of not being a launching State? And, if so, are other areas of private international law and national laws able to fill potential resulting gaps, for example through contract and products liability laws?

Or, alternatively, does the interpretation of a "launching State" need to be adapted to include the development and operation of corresponding AI systems? How would the legal framework be applied to situations in which an AI system is constituted of different parts developed in different countries?

Under the Liability Convention, the pertinence of this inquiry may be restricted to AI systems that control operations of a satellite or other space asset that cause or otherwise result in physical damage, in space, in the air, or on the Earth. But what about situations where an AI error or malfunction causes or otherwise results in damages deriving from the monitoring of and/or providing communications in a hazardous disaster area, whether natural or man-made?

As use of AI in launch procedures becomes more prevalent, the interpretational difficulties will only increase, and it may be advisable to introduce concrete regulation in this regard. Further, once a space object is launched, the launching State will not change, even when the ownership of a space object is transferred while it is in orbit. Therefore, once a State is a launching State, it remains a launching State and continues to be liable.[17]

To fall within the scope of Article VII, the damage has to be caused by a "space object," albeit no definition of this term has been included in the Outer Space Treaty.[18] A vivid literary discussion has evolved around this terminology insofar as it is not clear from the treaties whether a space object needs to be tangible or intangible, a debate that was in the past mainly confined to Global Navigation Satellite System (GNSS) enabled space systems but has started to gain traction again in the discussion surrounding emerging technologies.[19]

2. *Liability Convention*

The question of liability has been on the agenda of UNCOPUOS almost as long as it has existed. Liability for damage and responsibility of States for their space activities was a pressing issue on the minds of the drafters of the Outer Space Treaty and the Liability Convention. Although Article VII Outer Space Treaty established the rules relating to the liability of the launching State, the need for a more detailed and dedicated liability regime became evident.

The Liability Convention was largely negotiated in parallel to the Outer Space Treaty, albeit it took longer to be concluded. This dedicated regime builds upon Article VII of the Outer Space Treaty, adopts largely the same wording and elaborates the more general provisions of the Outer Space Treaty.

Mutatis mutandis to the Outer Space Treaty, the Liability Convention also anchors liability in the launching State (i.e., the State that launches or procures the launching of a space object, or from whose territory or facility a space object is launched). However, there are some noteworthy differences between Article VII Outer Space Treaty and the Liability Convention.

First, the Liability Convention covers damage caused by "a space object on the surface of the earth or to an aircraft in flight" while the Outer Space Treaty does not use such a precise formulation but simply speaks of damage to another State Party on Earth or in airspace (in addition to damage in outer space).

This would mean that the Outer Space Treaty would also apply to damage to objects that are not considered "aircraft." Albeit at present it would be very hard to imagine an object traversing

---

[17] Upasana Dasgupta, "On-Orbit Transfer of Satellites between States: Legal Issues – with Special Emphasis on Liability and Registration," Proceedings of the International Institute of Space Law 59 (2016): 641, 648; Frans Von der Dunk, "Transfer of Ownership in Orbit: from Fiction to Problem," in Ownership of Satellites: 4th Luxembourg Workshop on Space and Satellite Communication Law, ed. Mahulena Hofmann and Andreas Loukakis, 1st ed., Luxemburger Juristische Studien - Luxembourg Legal Studies, vol. 9 (Baden-Baden: Hart Publishing, 2017), 29, 36.

[18] The subsequent Liability Convention only clarifies that the "term "space object" includes component parts of a space object as well as its launch vehicle and parts thereof". See *Liability Convention*, Art I

[19] B. D. Kofi Henaku, "The International Liability of the GNSS Space Segment Provider," Annals of Air and Space Law 21 (1996): 143, 165.

Lesley Jane Smith, "Legal Aspects of Satellite Navigation," in Handbook of Space Law, ed. Frans G. von der Dunk and Fabio Tronchetti, Research Handbooks in International Law (Cheltenham, England: Edward Elgar Publishing, 2015), 554.

Stephen Gorove, "Toward a Clarification of the Term Space Object: An International Legal and Policy Imperative," Journal of Space Law 21, no. 1 (1993): 11, 25.

the airspace of a State which would not fall under the definition of aircraft as stipulated in Annex 7 to the Chicago Convention.[20]

First, the Outer Space Treaty applies to damage caused by a space object "to another State Party to the Treaty or to its natural or juridical persons" on Earth, in airspace or in outer space, including the Moon and other celestial bodies. The Liability Convention, on the other hand, speaks of damage to "a space object of one launching State or to persons or property on board such a space object." This formulation makes the scope of the Outer Space Treaty broader, and it would also cover damage to things or persons outside of space objects, such as an astronaut on a spacewalk.

Second, and perhaps most strikingly, the Liability Convention introduces a two-tiered liability regime for damage caused by a space object. For damage on the surface of the Earth or to an aircraft in flight, the Liability Convention prescribes absolute liability of the launching State whereas for damage anywhere else (i.e., outer space), the Liability Convention requires fault of the launching State. This division into fault and non-fault-based liability is unique in public international law and has rendered it one of the most discussed provisions within the space law community.

Absent a reverse onus clause, the burden of proof lies with the claimant State, and proving fault for damage caused by a space object in outer space is difficult. It is often not entirely possible to determine how the damage occurred and what led to the damage.[21] The introduction of AI will add another layer of complications to this conundrum. This is so because the nature of AI challenges the principles that underlie a fault analysis.

While in the past, the challenge of proving fault of the launching State was mostly related to the difficulty of gaining access to adequate evidence, AI could challenge negligence and intent directly, as it introduces the AI developer as a potential additional party. For example, it remains to be determined how the concept of "fault" would be applied to an autonomous decision taken by an AI that caused the damage (let's say, in the context of a failed collision avoidance maneuver).

In such a situation, it would need to be determined whether the fault lies with the operator of the spacecraft or with the AI developer. While this has already caused an academic debate, it will only increase the more AI develops and States will have to find new solutions (e.g., a new legal instrument, or a specifically curated insurance product) to deal with this emerging technology. In this regard, it is interesting to recall what the drafters of the Liability Convention assessed back in the day: *"excludes from the field of application of the convention damage sustained by equipment and personnel connected with the space object, but only where damage is the result of endogenous causes, i.e. defects of construction or of handling of the space object in question. This is logical, because the space object belongs to the Launching State and the*

---

[20] Annex 7 defines aircraft as "any machine that can derive support from the reactions of the air other than the reaction of the air against Earth's surface."

[21] For example, it is difficult and cost-intensive to track space objects. This is particularly pressing in the realm of space debris.

personnel involved are, consequently, protected by the labour legislation of the State in question and, in any event, by agreements previously concluded between them and the State."[22]

A hypothetical scenario could involve an AI-controlled space probe malfunctioning and damaging another country's satellite. Unlike traditional software, AI systems operate autonomously, making decisions without human input based on real-time data and learning algorithms. This adaptive and unpredictable nature of AI, some argue, goes beyond a simple software glitch, raising questions about the adequacy of existing product liability frameworks.

Another nuanced and related question is whether a State that hosts an AI (e.g., the AI resides on data servers located within that State) that was created by a developer domiciled in a third State may be susceptible to any claims of liability due to a malfunction of the AI (e.g., resulting from a failure or fault in the data server hardware or software infrastructures) or security breach (e.g., an attack or hack against an AI where it becomes controlled by an adversary).

On the other hand, it can be viewed that AI malfunctions, even in complex systems, should be treated similarly to software glitches. Under this perspective, product liability laws would cover AI failures, whether the malfunction was caused by faulty programming, poor training data, or hardware issues like radiation exposure. This view suggests that AI systems should not introduce new legal complexities but rather fall under existing regimes for product liability.

Nevertheless, the autonomous and evolving capabilities of AI present challenges for assigning liability, especially in situations where fault must be proven by the claimant state. This highlights the need to explore whether existing frameworks are sufficient or if new legal instruments are necessary.

Recommendations include developing accountability frameworks that account for AI's autonomy, establishing international standards for AI in space missions, creating robust testing and validation protocols, and implementing clear liability and compensation mechanisms tailored to AI systems.

Moreover, scenarios where AI-controlled space objects change ownership or control mid-mission introduce further complications. For instance, if an AI-controlled satellite is sold to another operator while in orbit, the liability for any subsequent damages caused by the AI system must be clearly assigned.

The question of who owns the AI is crucial: is it the operator using the AI system on their satellite, or the developers who own the copyright and remotely update, service, and maintain the AI? To ensure accountability, liability should transfer with ownership of the satellite, meaning that the new operator assumes responsibility for the AI's actions.

This approach is necessary because the operator has direct control over the satellite and its mission parameters and is in the best position to manage and mitigate risks during its operation. Developers, while responsible for the maintenance and updates of the AI, do not have operational control and may not be fully aware of the satellite's specific mission context. Therefore, agreements should clearly delineate the roles and responsibilities of both operators and developers,

---

[22] Working Paper submitted by the Italian delegation (A/AC.105/0.2/L.40/Rev.1, Annex).

ensuring that the party in control of the satellite during the mission is held accountable for any damages.

As for damage on the surface of the Earth or to an aircraft in flight, the liability is absolute which circumvents the need to prove fault. The claimant State will only need to prove that the damage was caused by a space object, its State of registry, and that the damage occurred to an aircraft in flight or on the surface of the Earth, which will ordinarily be quite easy for larger space objects that do not burn up entirely and disintegrate during re-entry.

A potential open question would be how liability may be attributed in the instance where an AI system controlling, say, a satellite module or component responsible for the failure and resulting damages were sourced from a foreign AI developer. And could some level of control over the space object, and its respective component or module, potentially be construed as residing with the AI developer?

Conclusions

The Outer Space Treaty provides a critical foundation for regulating the conduct of States in space. As Article III of the Outer Space Treaty mandates compliance with international law, including the United Nations Charter, the Treaty incorporates general principles of international humanitarian law into space activities.

The relevance of the Outer Space Treaty to AI systems in space is significant, particularly concerning the principles of non-appropriation (art. II) and state responsibility (art. VI). States cannot use AI to circumvent the fundamental rules against national appropriation, such as by using autonomous systems to claim territory or control access to celestial bodies.

Moreover, states remain responsible for all national space activities, including those conducted by private actors utilizing AI technologies. The introduction of AI into space heightens the complexity of ensuring compliance with these principles, especially given the potential for AI to operate autonomously across multiple jurisdictions, making the determination of the "appropriate state" for authorization and supervision an ongoing challenge.

Similarly, the Liability Convention expands upon Article VII of the Outer Space Treaty by establishing a detailed liability regime for damage caused by space objects. This regime is particularly relevant as AI takes on a greater role in space activities, including launch procedures and satellite operations.

The Convention's two-tier liability system—imposing absolute liability for damage on the surface of the Earth or to aircraft in flight (art. II) and fault-based liability for damage in outer space (art. III)—creates specific challenges when AI systems are involved. For example, the difficulty of assigning fault when an AI system autonomously controls a space object adds a new layer of complexity to liability disputes. The level of autonomy in AI systems is decisive, as it influences decision-making processes that could lead to incidents. AI systems might act beyond the direct control or intent of the State or entity responsible, complicating fault attribution. For example, an autonomous collision-avoidance system in a satellite may misinterpret sensor data, leading to an unintended maneuver that damages another satellite. In such a case, the fault might lie in the AI's programming, testing, or deployment process; not in the "behavior" of the State or of the persons for whom it is responsible (unless legal personality is attributed to the AI system, which is a different discussion).

Regarding the concept of "launching state," provided for by Article I of the Liability Convention, the unambiguous identification of at least one of the launching states allows for the victim to seek compensation from any launching states involved (art. IV(2)), so that with the incorporation of AI in the space industry, it will always be possible to locate a minimum of one of those launching states (most easily, the state of registry of the AI-powered space object).

It is also essential to highlight that inter-party agreements in these cases must encompass provisions pertaining to AI, such as risk management, indemnification, and limitation of liability, which will be subjected to further study in the procurement section. This examination will focus on matters pertaining to product liability and private operators, in addition to the formulation of specific recommendations concerning contract clauses that hold operators and entities accountable for AI-driven activities in space.

Additionally, the question of how AI developers and operators should be held accountable for damages caused by AI-controlled space objects requires further regulatory clarification. As AI systems continue to advance, particularly in launch procedures and in-orbit operations, states may need to amend national legislation to address these emerging liabilities, ensuring that AI's growing autonomy does not undermine the existing legal framework.

Section 2: Applying International Humanitarian Law to AI-Driven Applications in Space: Perspectives from the CCW and Geneva Conventions

Introduction

The dual-use nature of AI, where technologies can serve both civilian and military purposes, adds complexity to the legal landscape surrounding space operations. This section explores how International Humanitarian Law (IHL) and the Convention on Certain Conventional Weapons (CCW) address the challenges posed by autonomous systems, particularly in armed conflict.

As the CCW governs lethal autonomous weapons systems (LAWS), these frameworks must evolve to address AI-driven technologies in space. Key principles such as meaningful human control, accountability, and weapons reviews—central to the CCW—are essential to ensuring the responsible deployment of AI-powered satellites and other dual-use technologies with potential military applications.

The Geneva Conventions add another layer of oversight, ensuring that core IHL principles like distinction, proportionality, and military necessity are upheld, even in the realm of space operations. This section will examine how these principles are affected by the rise of autonomous systems, particularly in military decision-making where human intervention may be minimal or absent.

By analyzing these legal frameworks and their relevance to AI-driven technologies in space, this section highlights the evolving challenges AI presents. As autonomous systems continue to grow in complexity, new interpretations of existing laws—and possibly new regulations—will be necessary to fill gaps in the current legal regime. This analysis aims to propose solutions to ensure that AI technologies in space operate within a responsible and effective legal framework.

1. *Certain Conventional Weapons (CCW)*

The inherent nature of any technology is dual use, meaning that alongside civil usage, it can be implemented for military purposes.[23] When the technology in question is AI, its autonomy and unpredictability features raise important questions to be analyzed.

According to the International Committee of the Red Cross an autonomous weapon system is defined as: "*Any weapon system with autonomy in its critical functions—that is, a weapon system that can select (search for, detect, identify, track or select) and attack (use force against, neutralize, damage or destroy) targets without human intervention.*"[24]

---

[23] Almudena Azcárate Ortega, Not a Rose by Any Other Name: Dual-Use and Dual-Purpose Space Systems, Lawfare, available at: https://www.lawfaremedia.org/article/not-a-rose-by-any-other-name-dual-use-and-dual-purpose-space-systems.

[24] Neil Davison, A legal perspective: Autonomous weapon systems under international humanitarian law, 2018: 5-18.

Issues regarding the Meaningful Human Control (MHC) over the entire technological cycle and functioning of such autonomous systems is of particular interest for the scope of this report.[25] Discussion regarding this topic was carried in the scope of the 1980 United Nations CCW convention and protocols.

Within the framework of the UN, the CCW plays an important role in the governance of autonomous lethal weapons and how these regulations should be applied in outer space, alongside some relevant UN resolutions such as on the "Prevention of an arms race in outer space" and on the "No first placement of weapons in outer space."[26] It is of extreme importance to understand whether the CCW presents any limits on placing autonomous arms in outer space.

1.1 *The Importance of Meaningful Human Control (MHC*)

The concept of MHC is essential in maintaining control over AI systems, particularly in space operations where communication delays can complicate real-time decision-making. MHC ensures that human judgment and oversight are integrated into the lifecycle and operational decision-making of AI systems.

Indeed, several ethical concerns have been raised regarding the risk of creating "responsibility gaps" when AI systems are implemented in military operations. MHC has been introduced into the debate to ensure that humans, rather than computers, remain ultimately in control and can be held morally responsible for decisions involving lethal autonomous weapons in military operations.[27]

However, reaching a consensus on what constitutes meaningful human control is challenging. One state might interpret it as requiring a "human in the loop"[28] throughout the technology lifecycle, ensuring human supervision of a weapon system's actions. Another state might consider that responsible programming alone provides sufficient meaningful human control.[29] Notably, varying levels of human involvement in decision-making can be implemented in AI systems, ranging from requiring human approval for every decision the AI makes to merely overseeing the AI's actions and retaining the authority to intervene in case of malfunction.[30]

The first approach would be suitable for various applications, as it ensures real-time human intervention and decision-making. However, this approach can be particularly challenging for space missions due to significant communication delays between Earth and space-based assets, especially in deep space exploration. Nonetheless, some form of human oversight may be a

---

[25] Ugo Pagallo, Eleonora Bassi, and Massimo Durante, The normative challenges of AI in outer space: law, ethics, and the realignment of terrestrial standards, *Philosophy & Technology* 36.2 (2023): 23.

[26] United Nations, Prevention of an arms race in outer space, United Nations resolution; United Nations, No first placement of weapons in outer space, Resolution 69/32.

[27] Filippo Santoni de Sio, Jeroen Van Den Hoven, Meaningful human control over autonomous systems: A philosophical account, *Frontiers in Robotics and AI*, 2018, 5: 15.

[28] Human in the loop refers to learning models that include human interaction in the entirety of the technology life-cycle, for an extensive overview of the term see here: Human in the Loop, Science Direct, https://www.sciencedirect.com/topics/computer-science/human-in-the-loop.

[29] Rebecca Crootof, "A meaningful floor for meaningful human control," *Temp. Int'l & Comp. LJ*, 2016, 30: 53.

[30] Robert Mazzolini, "Artificial Intelligence and Keeping Humans 'in the Loop,'" in *Modern Conflict and Artificial Intelligence*, Center for International Governance Innovation, 2020, 48.

necessary safeguard, depending on the type of AI-driven mission, particularly when a malfunction in AI systems could jeopardize the operation of critical space-based services.

The second approach could support MHC in space activities. During the development phase of AI for specific missions, particular attention could be given to various elements to ensure the responsible deployment of systems.[31] In a study, Riebe asserts that time, predictability, and reliability—interconnected to varying degrees—alongside accountability and explainability, are instrumental in supporting MHC.[32]

Defining and implementing MHC during the development phase allows for better regulation of AI systems' behavior in space, addressing risks tied to their autonomy and unpredictability. This is particularly critical in the context of space operations, where the physical distance and environmental conditions complicate real-time human intervention. Currently, a balanced strategy that evaluates the specific risks of each AI-driven operation appears to be the most prudent course of action. For high-risk scenarios, stricter adherence to MHC can ensure compliance with ethical principles, whereas scenarios with greater risk tolerance may permit a more flexible yet responsible developmental approach. Establishing an internationally recognized baseline for MHC would foster consensus on its application to AI in space activities. Ultimately, determining the appropriate level of MHC for each unique context is imperative.

1.2 *CCW Relevance for AI in Space*

The applicability of IHL to outer space is well-established and will be explored in detail in a dedicated section of this report. However, this section examines how IHL provisions and principles apply to space weapon systems, particularly in relation to AI systems in outer space, with a specific focus on LAWS.

The CCW, grounded in the principles of IHL, is articulated through five protocols. It includes a set of guiding principles regarding the autonomy of intelligent systems, which can be effectively extended to address the military challenges posed by AI systems in space.

This includes clear prohibitions. Certain autonomous technologies should unequivocally be banned, particularly any "weapons system... of a nature to cause superfluous injury or unnecessary suffering, or if it is inherently indiscriminate, or is otherwise incapable of being used in accordance with the requirements and principles of IHL."[33]

The challenge in achieving a balanced approach to banning dual-use technologies arises from the blurred line between dangerous AI weapons and AI applications designed for intelligence, surveillance, and reconnaissance (ISR) missions.[34]

---

[31] Thomas Graham, Kathiravan Thangavel, "Artificial Intelligence in Space: An Analysis of Responsible AI Principles for the Space Domain," In: *International Astronautical Congress, Baku, Azerbaijan*. 2023.

[32] Thea Riebe, "Meaningful Human Control of LAWS: The CCW-Debate and its Implications for Value-Sensitive Design," In: Technology Assessment of Dual-Use ICTs, Springer Vieweg, Wiesbaden, 2023, https://doi.org/10.1007/978-3-658-41667-6_10

[33] Ugo Pagallo, The New Laws of Outer Space: Ethics, Legislation, and Governance in the Age of Artificial Intelligence, Hart Publishing, 2024.

[34] *Ibid*.

Indeed, AI in space can be employed for legitimate military purposes. However, clear guidance is lacking on how, when, and for what reasons AI in space should be banned. A critical first step is the development of a comprehensive framework to clarify AI's military applications, providing insight into the current state of the technology and its potential future capabilities. This framework would serve as a foundation for determining which uses should be prohibited under the CCW.

One real-life example of this complexity is autonomous navigation and collision avoidance systems in satellites. While these systems are primarily designed for civilian purposes, such as preventing space debris collisions and ensuring safe satellite operations, their dual-use potential adds to the challenges of regulation and oversight.

However, their potential military applications cannot be ignored. An AI-powered satellite designed for autonomous collision avoidance could also serve as a strategic military asset. For example, it could be employed to evade anti-satellite weapons or to maneuver strategically for enhanced surveillance capabilities.

An AI-controlled satellite with defensive capabilities raises significant concerns regarding MHC. Such a satellite could be programmed to autonomously identify and neutralize potential threats. However, in an unforeseen incident, the AI might misinterpret a civilian satellite's maneuvers as hostile behavior and destroy it, leading to serious consequences.

Such scenarios underscore the significant risks associated with the proliferation of autonomous military systems in space and the severe consequences of errors in threat identification. They highlight the critical need for stringent review and control mechanisms to prevent such incidents, ensuring that AI applications in space remain safe and reliable.

These examples illustrate the delicate balance required in regulating AI technologies in space. While the potential benefits for both civilian and military applications are considerable, the risks of misuse or error demand thorough consideration and the establishment of robust regulatory frameworks. In this context, the use of autonomous weapons in space could fall under the ban category, particularly in scenarios involving:

- Kinetic and non-kinetic attacks halting the functioning of space infrastructure/satellites considered critical infrastructure or enabling national security or essential services.
- The destruction of spacecraft in orbit, following with the creation of space debris leading to the so-called Kessler effect, in which circumstances the LEO space environment will become relatively unusable.

Some of the principles of the CCW that can help tackle such challenges include:

- Applicability of IHL to LAWS: Ensuring that AI systems comply with IHL principles (CCW Protocol I, Article 1).
- Non-Delegation of Human Responsibility: Maintaining human oversight over AI systems (CCW Preamble, reflecting customary IHL principles).
- Accountability for Use of Force: Ensuring accountability under international law (CCW Protocol I, Article 36).

- Weapons Reviews Before Deployment: Conducting thorough reviews before deployment (CCW Article 36).
- Incorporation of Safeguards: Including physical, non-proliferation, and cybersecurity safeguards (CCW Protocol V, Article 3).
- Risk Assessment and Mitigation: Assessing and mitigating risks during development (CCW Protocol V, Article 9).
- Non-Harm to Civilian Research: Ensuring that military use of AI does not harm civilian research and development (CCW Protocol III, Article 2).

These principles provide an essential foundation for regulating AI systems in space, particularly regarding dual-use technologies and the risks associated with autonomous military systems. However, a more comprehensive legal analysis under the Geneva Conventions is needed to fully address the implications for IHL, particularly in ensuring adherence to fundamental norms such as distinction and proportionality.

2. *Geneva Conventions (1949)*

At the present time, concepts that have traditionally been confined to the realms of science fiction, such as humanoid robots, laser weapons, and counter space weapons, are becoming progressively tangible, substantively influencing the trajectory and character of warfare. The integration of sophisticated cutting-edge technologies, such as AI systems, offers a promising avenue for the development of novel defense applications, which could potentially enhance a country's military advantage over possible rivals.[35]

Specifically, AI technology, due to its embedded autonomy and decision-making capabilities, presents unique challenges in outer space, particularly with respect to observing the fundamental norms of IHL. This is why, building upon the previous examination of the CCW, this section serves as a complementary legal analysis of the utilization of AI-enabled space systems in the context of an armed conflict from an IHL perspective under the 1949 Geneva Conventions and their corresponding Additional Protocols. By way of example, autonomous satellites or AI-enabled drones used in military operations must, in any case, adhere to the latter.

The topic of LAWS has been under debate within the framework of the aforementioned CCW since 2014, being discussed within an informal "meeting of experts." In 2017, a formal "Group of Governmental Experts" (GGE) was constituted, aimed at investigating the technological, military, ethical and regulatory dimensions of LAWS.[36] In its 2019 session, some of the utmost IHL principles were reiterated in the conclusions, making it expressly clear that the latter should be considered when examining the potential deployment of AI-enabled LAWS:

---

[35] Raska, M., and Davis, M. 2024. "The 'AI Wave' in Space Operations Implications for Future Warfare." in Pekkanen, S. M., and Blount, P. J. (eds.), *The Oxford Handbook of Space Security*, Oxford, Oxford University Press, p. 596.

[36] Martin, A.-S., and Freeland, S. 2021. "The Advent of Artificial Intelligence in Space Activities: New Legal Challenges." *Space Policy* 55, p. 3. https://doi.org/10.1016/j.spacepol.2020.101408

> "On the agenda item 5 (a) "An exploration of the potential challenges posed by emerging technologies in the area of lethal autonomous weapons systems to International Humanitarian Law" the Group concluded as follows:
>
> (a) The potential use of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems must be conducted in accordance with applicable international law, in particular IHL and its requirements and principles, including inter alia distinction, proportionality and precautions in attack;
>
> (b) IHL imposes obligations on States, parties to armed conflict and individuals, not machines;
>
> (c) States, parties to armed conflict and individuals remain at all times responsible for adhering to their obligations under applicable international law, including IHL. States must also ensure individual responsibility for the employment of means or methods of warfare involving the potential use of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems in accordance with their obligations under IHL;
>
> (d) The IHL requirements and principles including inter alia distinction, proportionality and precautions in attack must be applied through a chain of responsible command and control by the human operators and commanders who use weapons systems based on emerging technologies in the area of lethal autonomous weapons systems;
>
> (e) Human judgment is essential in order to ensure that the potential use of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems is in compliance with international law, and in particular IHL;
>
> (f) Compliance with the IHL requirements and principles, including inter alia distinction, proportionality and precautions in attack, in the potential use of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems requires inter alia that human beings make certain judgements in good faith based on their assessment of the information available to them at the time (…)."[37]

Additionally, the GCE developed some Guiding Principles[38] with respect to "Emerging Technologies in the Area of Lethal Autonomous Weapons Systems." Among these, it is interesting to highlight that:

- IHL remains fully operational with regards to all weapons systems, including those with the potential for LAWS.
- It is imperative that the human element retains responsibility for decisions pertaining to the utilization of weapons systems, as accountability cannot be

---

[37] CCW/GGE.1/2019/3. 2019. *Report of the 2019 Session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems*, 25 September 2019, p. 4. https://undocs.org/en/CCW/GGE.1/2019/3

[38] *Ibid.*, p. 13, Annex IV.

delegated to machines. This must be borne in mind throughout every phase of the weapons system's life cycle.

- The interaction between humans and machines in the context of weaponry may manifest in various forms and occur at different stages of the weapon's life cycle. It is of paramount importance for this interaction to ensure that the potential use of weapon systems based on emerging technologies in the domain of LAWS complies with applicable international law, in particular IHL.
- It is recommended that consideration be given to the potential use of emerging technologies in LAWS with a view to ensuring adherence to IHL and other applicable international legal obligations.

The significance of space in modern warfare is increasing, concretely in light of the growing prevalence and sophistication of autonomous space systems, which are set to become a dominant force on future battlefields. As such, space has emerged as a crucial element to armed forces and the future of force structures worldwide, thus demanding the adherence to the core principles of IHL in the operation of such systems.[39]

As the preceding sections have shown, the introduction of AI into space operations is likely to challenge the existing legal framework and will require the adaptation of existing binding laws governing outer space activities to take into account the emergence of AI. In particular, it will be necessary to analyze and determine whether the existence of AI challenges traditional understandings of responsibility and fault or whether AI can be understood and interpreted within the traditional notions of international space law.

2.1 *International Humanitarian Law (IHL) or Jus in Bello*

2.1.1 Conditions Triggering Applicability

To begin with, it is imperative to provide a brief definition of the discipline of IHL, also referred to as the law of armed conflict or *jus in bello*, which is the body of rules designed to restrict the effects of armed conflict on humanitarian grounds. It aims to regulate the methods and means of warfare employed by combatants, as well as to safeguard the wellbeing and humane treatment of individuals who are not directly engaged in combat or have ceased to be involved in hostilities. In essence, IHL encompasses those norms of international law that provide a baseline standard of humanity that must be upheld in the framework of any armed conflict.[40]

The United Nations Security Council has repeatedly reaffirmed the obligations arising under the fundamental principles of the *jus in bello*, particularly those outlined in the 1949 Geneva Conventions and their Additional Protocols,[41] by demanding and urging all parties concerned to comply strictly with their commitments under these instruments. AI-driven space systems are not

---

[39] See Raska, M., and Davis, M. "The "AI Wave"... *op. cit.*, p. 601.

[40] Melzer, N., and Kuster, E. (coord.) 2019. *International Humanitarian Law: A Comprehensive Introduction*, International Committee of the Red Cross (ICRC), November 2019, p. 17. https://library.icrc.org/library/docs/DOC/icrc-4231-002-2019.pdf

[41] In this regard, see United Nations Security Council (UNSC) Resolutions 1265 (1999), 1296 (2000), and 1674 (2006), among others, on the protection of civilians in armed conflict.

the exception to the rule; therefore, they must comply with these obligations in scenarios of armed conflict, both in outer space and on Earth.

In the context of international armed conflicts, Article 2 common to all Geneva Conventions foresees that, "In addition to the provisions which shall be implemented in peacetime, the present Convention *shall apply to all cases of declared war or of any other armed conflict which may arise between two or more of the High Contracting Parties*, even if the state of war is not recognized by one of them" (emphasis added).

Therefore, in order for IHL to be applicable in outer space, there are certain conditions that should be met.

1.º The main criterion for the applicability of IHL is the existence of an armed conflict. In this sense, it is important to resort to the definition of "attack" provided by Article 49(1) of the 1977 Protocol Additional to the Geneva Conventions of 12 August 1949 and relating to the Protection of Victims of International Armed Conflicts (Protocol I), according to which: "'Attacks' means acts of violence against the adversary, whether in offence or in defence."

By making use of this term, it is evidenced that the application of IHL is dependent on a factual basis, that is, on the grounds of the prevailing facts, as opposed to the interpretation given to it by the parties.[42] In this sense, the recognition of States and Governments involved in a conflict as such is irrelevant to the question of the validity of international humanitarian law.[43]

To provide an example, an AI-controlled satellite attacking another space object by "violent" means,[44] with a clear offensive nature, would fall under this definition, thus being labeled as "attack." However, it should be likewise mentioned that the word "violent" in the context of AI applications is losing its significance, as this may not be primarily physical.

So stated the International Criminal Tribunal for the former Yugoslavia (ICTY) in *Milutinović et al.*: "the existence of an armed conflict does not depend upon the views of the parties to the conflict."[45] Otherwise, "If the application of international humanitarian law depended solely on the discretionary judgment of the parties to the conflict, in most cases there would be a tendency for the conflict to be minimized by the parties thereto."[46]

Moreover, in the *Boškoski & Tarčulovski* case,[47] the ICTY expressed that "the question of whether there was an armed conflict at the relevant time is a factual determination to be made by the Trial Chamber upon hearing and reviewing the evidence admitted at trial."

Two possibilities arise out of this concept when dealing with an armed conflict where space systems may be involved. Firstly, the occurrence of an attack from one space object to another in

---

[42] David, É. 2012. *Principes de droit des conflits armés*. 5th edition, Bruylant, pp. 102 ff.

[43] Fleck, D. (ed.) 2008. *The Handbook of International Humanitarian Law*. Oxford, Oxford University Press, p. 46.

[44] It is interesting to mention in this regard, however, that the word "violent" in the context of AI is losing its significance, as it is not primarily physical; in contrast with the traditional conceptualization.

[45] *The Prosecutor v. Milutinović et al.* [2009]. ICTY, Case No. IT-05-87-T, Judgment (Trial Chamber), 26 February 2009, para. 125. https://www.refworld.org/jurisprudence/caselaw/icty/2009/en/78020

[46] *The Prosecutor v. Jean-Paul Akayesu* [1998]. ICTR, Case ICTR-96-4-T, Judgment (Trial Chamber I), 2 September 1998, para. 603. https://www.refworld.org/jurisprudence/caselaw/ictr/1998/en/19275

[47] *The Prosecutor v. Boškoski & Tarčulovski* [2008]. ICTY, Case No. IT-04-82-T, Judgment (Trial Chamber II), 10 July 2008, para. 174. https://www.refworld.org/jurisprudence/caselaw/icty/2008/en/61641

the outer space environment, and, secondly, the attack to a space object in the middle of an armed attack taking place on Earth.

This clause continues, in its third paragraph, by stipulating that: "The provisions of this Section apply to any land, air or sea warfare which may affect the civilian population, individual civilians or civilian objects on land. They further apply to all attacks from the sea or from the air against objectives on land but do not otherwise affect the rules of international law applicable in armed conflict at sea or in the air."

2.º The (international) armed conflict shall take place between two or more opposed High Contracting Parties to the Geneva Conventions, which are understood to be States, regardless of the underlying causes or the intensity of the conflict. However, it is possible that important IHL rules are applicable even in scenarios where there are no open hostilities. In addition, it is not necessary to declare war or to recognise the situation formally.[48]

3.º In relation to the previous points, it should be added the wording of Article 1, common to the four Geneva Conventions, which states that: "The High Contracting Parties undertake to respect and to ensure respect for the present Convention in all circumstances." The term "in all circumstances" signifies that, upon the occurrence of any of the bases for application as outlined in Article 2, no Contracting Party may offer any justifiable rationale, whether legal or otherwise, for not adhering to the Convention in toto.

The aforementioned proposition also indicates that the Convention is applicable irrespective of the nature of the conflict. The legitimacy of a war has no bearing on the treatment that protected persons should enjoy, no matter whether it is deemed just or unjust, whether it is characterized as an act of aggression or resistance to aggression, and notwithstanding the specific intention behind the conflict.[49]

2.1.2 IHL Fundamental Principles

Over time, state practice has led to the establishment of three IHL overriding principles: the principle of distinction, the principle of proportionality, and the principle of (military) necessity. In particular, the autonomous nature of AI-enabled space systems presents significant challenges in the application of these principles.

Regarding the principle of distinction, it is forbidden to willfully attack civilians or civilian objects in isolation or to strike at both military and civilian targets indiscriminately.[50] This cornerstone principle of IHL was first enshrined in the 1868 St. Petersburg Declaration, where it was provided that "the only legitimate object which States should endeavor to accomplish during war is to weaken the military forces of the enemy."[51]

---

[48] ICRC. 2008. *How is the Term "Armed Conflict" Defined in International Humanitarian Law?*. Opinion Paper, March 2008, p. 1. https://www.icrc.org/en/doc/assets/files/other/opinion-paper-armed-conflict.pdf

[49] Pictet, J. (ed.) 1958. *Commentary to the Fourth Geneva Convention Relative to the Protection of Civilian Persons in Time of War*. International Committee of the Red Cross (ICRC), pp. 16-17.

[50] Cassese, A. 2005. *International Law*. 2nd edition, Oxford, Oxford University Press, p. 415.

[51] *St. Petersburg Declaration 1868*, preamble.

On another note, the Hague Conventions of 1899 and 1907 do not explicitly require the differentiation between civilians and combatants, albeit Article 25 is founded upon this fundamental principle, thereby prohibiting "the attack or bombardment, by whatever means, of towns, villages, dwellings, or buildings which are undefended."

The principle of distinction is currently codified in the following articles of the 1977 Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I):

- Article 48: "In order to ensure respect for and protection of the civilian population and civilian objects, the Parties to the conflict shall at all times distinguish between the civilian population and combatants and between civilian objects and military objectives and accordingly shall direct their operations only against military objectives."
- Article 51(2): "The civilian population as such, as well as individual civilians, shall not be the object of attack. Acts or threats of violence the primary purpose of which is to spread terror among the civilian population are prohibited."
- Article 52(2): "Attacks shall be limited strictly to military objectives. In so far as objects are concerned, military objectives are limited to those objects which by their nature, location, purpose or use make an effective contribution to military action and whose total or partial destruction, capture or neutralization, in the circumstances ruling at the time, offers a definite military advantage."

For AI systems, ensuring the distinction between combatants and civilians remains a key problem, which transfers to the outer space domain where AI-driven spacecraft might struggle to accurately identify military versus civilian targets, leading to potential violations of this IHL core principle.

In terms of the principle of proportionality, in the event of launching an attack against a military objective, it is prohibited to cause the unintended loss of civilian life or the destruction of civilian objects disproportionate to the direct and tangible military advantage anticipated.[52]

The principle of proportionality in attack is codified in Article 51(5)(b) of Additional Protocol I, according to which: "Among others, the following types of attacks are to be considered as indiscriminate: b) an attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated."

Moreover, this passage is reiterated in Article 57(2)(b) of the same Protocol, which states that "With respect to attacks, the following precautions shall be taken: b) an attack shall be canceled or suspended if it becomes apparent that the objective is not a military one or is subject to special protection or that the attack may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated."

---

[52] See Cassese. *International Law… op. cit.,* pp. 415-416.

In this scenario, an illustrative example might be an autonomous space weapon system targeting a military installation yet inadvertently causing excessive damage to nearby civilian infrastructure.

Furthermore, with regard to the proportionality principle, it is pertinent to inquire whether the deployment of AI in response to non-AI aggression could be deemed proportional. In light of the exponential growth in AI capabilities, it may be advisable to refrain from deploying certain classes or categories of AI in military contexts.

By way of example, some AI systems possess the capacity for self-reprogramming, which could potentially lead to undesirable outcomes. It is crucial to ensure that AI systems are not capable of assuming full control of their own operations, as this could result in unintended consequences, including the targeting of humans.

It may be beneficial to utilize AI systems with exceptional coding capabilities in high-energy particle physics, as this could enhance their ability to transcend human-imposed limitations on the understanding of physical laws. This could potentially facilitate the development of a more comprehensive and accurate model encompassing subatomic particle physics, galactic superclusters, and even the enigmatic dark matter and dark energy.

Finally, the principle of (military) necessity allows the implementation of measures that are genuinely indispensable for the attainment of a legitimate military objective, provided that they are not otherwise proscribed by IHL.[53] In the context of an armed conflict, the sole legitimate military objective is the diminution of the opposing parties' military capabilities.

In consequence, the parties to an armed conflict may only employ the specified means and methods for as long as they are required to attain such a legitimate objective.[54] The principle of necessity frequently conflicts with humanitarian considerations, therefore, the objective of IHL is to achieve a balance between these two.[55]

The balance between humanity and necessity serves as the foundation and pervades the entirety of the normative framework of IHL. The context established by the concept of military necessity serves to inform the interpretation of the rules and principles set forth in IHL, including those of distinction and proportionality.

The principle of necessity does not abrogate or supersede the specific tenets of IHL. Instead, they act as a framework for interpreting the rights and obligations of parties to armed conflicts within the confines of these established rules.[56]

In regard to AI, ensuring that autonomous systems adhere to this principle can be challenging, particularly in determining what constitutes a legitimate military necessity versus excessive or unnecessary use of force.

---

[53] "Military necessity or expediency do not justify a violation of positive rules." See *United States v. Wilhelm List* [1948]. United States Military Tribunal at Nuremberg, *Hostages case*, pp. 66-67. https://casebook.icrc.org/case-study/united-states-military-tribunal-nuremberg-united-states-v-wilhelm-list

[54] *St. Petersburg Declaration 1868*, preamble.

[55] International Committee of the Red Cross. Military Necessity. https://casebook.icrc.org/a_to_z/glossary/military-necessity.

[56] ICRC. 2009. *Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law*. ICRC DPH Guidance, pp. 78-79.

2.2 *IHL and Space Law: A Case for "Humane" Autonomous Space Weapons*

2.2.1 Applicability of IHL to Armed Conflicts in Outer Space

Although not directly related to the conventional regulation of outer space, IHL is applicable to the field of outer space as any other branch of public international law as per Article III of the 1967 Outer Space Treaty, which establishes 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 cooperation and understanding."

It can be posited that IHL principles, as a fundamental element of public international law, are theoretically applicable to the utilization of outer space for military purposes. The laws and customs of war do not have a concrete "territorial" limitation. Conversely, they apply not only to the geographical area where the hostilities are materializing, but also to other areas which have been affected by these conflicts.

However, it is still under debate what constitutes an armed conflict in outer space.[57] The unique environment of space and the dual-use nature of many space technologies add layers of complexity to this issue.

For instance, the distinction between civilian and military satellites is often blurred, especially when private satellites that handle civilian data and/or communications can or are contracted to provide military services-grade reconnaissance materials (e.g., high-resolution imagery made available to military or intelligence services).

Actions that may be considered hostile in terrestrial conflicts, such as jamming communications or disabling satellites, may have different implications in space. This ongoing debate highlights the need for clear policies and definitions to effectively apply IHL to armed conflicts taking place in outer space.

In the event that military action occurs in outer space and the consequences thereof impact on civilians on Earth, such military activity shall be considered to fall under the rubric of *jus in bello*. This is not merely in regard to the direct action in question in space but also as to its effects and repercussions on Earth.[58]

2.2.2 The Observance of IHL by AI-enabled Spacecraft

When dealing with AI-based space systems, the observance of IHL norms should be investigated on a two-fold basis. On the one hand, with respect to the autonomous space object being the attacker (that is, targeting objectives and taking decisions such as to which attacks to make and how such attacks should be directed), and, on the other, concerning the AI-enabled

---

[57] In this respect, see Yan, W. 2023. "Definition of 'Armed Conflict' in Outer Space." *Beijing Law Review* 14, pp. 287-299. https://doi.org/10.4236/blr.2023.141016

[58] Freeland, S., and Gruttner, E. 2020. "The Laws of War in Outer Space." in Schrogl, K.-U. (ed.) *Handbook of Space Security: Policies, Applications and Programs*. 2nd edition, Cham, Springer, p. 85.

spacecraft as the attacked, especially taking into consideration the prevalent dual-use nature of space systems. For instance, an AI-driven satellite designed for military surveillance must ensure it does not engage in indiscriminate attacks.

In the first place, the acquisition of LAWS, especially by Western liberal democracies, will mandate the observation of IHL principles in the formulation of engagement norms for the deployment of lethal force by means of autonomous systems. This will necessitate the presence of a human "on the loop"[59] to supervise and direct the AI-based system and, in certain instances, such presence will be demanded "in the loop"[60] to enable direct control (e.g., when exercising the use of force).

It is plausible, in the future, that fully autonomous systems integrating advanced AI technologies permit humans to become progressively and potentially fully disengaged from the decision-making process, with machine intelligence assuming responsibility for determining tactics and weaponry deployment, which would be at odds with the current conditions enshrined by IHL. The stance of authoritarian states, nevertheless, remains somewhat opaque.

The development of AI-enabled systems for tactical command and control purposes in the absence of human oversight could potentially confer a tactical military leverage upon a determined adversary,[61] therefore being a cause of significant concern. This could lead to scenarios where autonomous space systems make decisions that violate IHL principles, thus necessitating strict regulatory frameworks.

Accordingly, human operators should be always kept in the loop to ensure the respect of IHL principles. The use of AI-driven satellites to target systems should be overseen by a human operator in all instances to avoid any indiscriminate attack to civilian population in the framework of an armed conflict, hence, complying with the fundamental principle of distinction.

Autonomous space objects, even though endowed with different decision-making capabilities in such circumstances, can hardly acquire the 'humanity' IHL requires in these situations, and the military objective of defeating the enemy may take precedence over the principle of humanity and the distinction between civilians and military.

It may also be susceptible to cyber or electronic attacks (counterspace weapons) that affect its essential components and impact not only the satellite infrastructure itself, but also human lives by interfering with these decision-making processes and target identification, potentially wreaking havoc and causing mayhem in the process. For example, a hacked AI-driven satellite could cause

---

[59] "Human-on-the-loop" means that *"the device has more autonomy but must be monitored by a human who can redirect, authorize or undo a specific action, and thus it is a supervised control regime."* See Martin and Freeland. "The Advent of… *op. cit.*, p. 3. Building up on this definition, we consider "on the loop" as meaning that the AI system will provide alerts when a complication arises for a prior defined objective. If the alert fails for some reason, the human may very well in all likelihood miss the "signal" in the ocean of data based on which the AI was supposed to trigger the alert. Therefore, it is our view that "on the loop" could also encompass autonomous weapons systems that can make independent kill determinations, unless it is required to pull the human into the loop on a particular decision.

[60] The "human-in-the-loop" entails that "the human retains control over the device function, which means that there is direct human control over the system." See *ibid.*

[61] See Raska and Davis. "The "AI Wave"... *op. cit.*, p. 601.

indiscriminate damage, which highlights the need for robust cybersecurity measures to be put in place.

In the second place, as a consequence of the dual nature of satellites, for both military and civilian purposes, the non-peaceful use of spacecraft has prompted questions regarding the circumstances under which these may be regarded as legitimate targets during an armed conflict, thereby heightening the risk of critical civilian infrastructure being attacked.[62]

In this vein, some States have already made statements to this effect (e.g., Russia), stating that, in their view, dual-use satellites would be considered legitimate military targets.[63] This ambiguity necessitates clear guidelines on targeting dual-use technologies to protect civilian infrastructure.

Conclusions

With respect to the observance of both the CCW and the Geneva Conventions, their applicability is feasible by means of Article III OST, due to the lack of concrete provisions in International Space Law on these subjects. In the event an armed attack occurred in outer space, abiding by the definition of "attack" provided by Article 49(1) of the I Additional Protocol to the Geneva Conventions, the norms of these two instruments should be respected.

On the one hand, the CCW provides an important framework for addressing the unique challenges posed by the dual-use nature of AI technologies in space. As the boundaries between civilian and military uses of AI blur, the CCW's guiding principles offer valuable mechanisms for ensuring the responsible deployment of autonomous systems. In particular, the focus on Meaningful Human Control (MHC), accountability, and thorough weapons reviews highlights the importance of maintaining human oversight in the responsible use of AI-driven space assets.

However, the complexity of balancing technological innovation with regulatory constraints underscores the need for continued dialogue, especially as AI in space evolves. Ultimately, the extension of CCW principles to space operations is essential to safeguard against the risks posed by autonomous military systems while promoting the safe and responsible use of AI technologies in this new frontier.

On the other hand, the Geneva Conventions, which enshrine the most fundamental IHL principles, are undoubtedly relevant to the weaponization of the outer space environment due to the absence of specific territorial legal constraints. In the context of AI-powered satellites being used for aggressive purposes, it is imperative that human operators are kept informed and involved at all stages to ensure compliance with IHL principles.

---

[62] delle Fave, D. et al. 2023. "The non-peaceful use of commercial satellites: existing issues and new challenges from a legal and policy perspective." IAC-23,E7,1,14,x77789, 74th International Astronautical Congress (IAC), Baku, Azerbaijan, 2–6 October 2023, p. 3.

[63] See the statement by Mr. Konstantin Vorontsov, Deputy Head of the Delegation of the Russian Federation and Deputy Director of the Department for Non-Proliferation and Arms Control of the Ministry of Foreign Affairs of the Russian Federation, at the Thematic Discussion on Outer Space (Disarmament Aspects) in the First Committee of the 77th Session of the UN General Assembly (New York, 26 October 2022). https://estatements.unmeetings.org/estatements/11.0010/20221026/5yPwCsESxyBr/N5pGP22K6MRm_en.pdf

The use of AI-driven space objects for targeting systems must be continuously supervised by a human responsible to prevent arbitrary attacks on civilian infrastructure, thereby aligning with the distinction principle. In particular, in instances of cyber or electronic attacks that compromise the integrity of their core components, underscoring the necessity for the implementation of robust cybersecurity measures.

PART II: REGULATORY CONSIDERATIONS

Section 1: Applying General Data Protection Regulation (GDPR) Principles to AI-Driven Space Systems

Introduction

This section explores the regulatory lessons that can be drawn from the General Data Protection Regulation (GDPR), a comprehensive legal framework for data privacy in the European Union, and how its key principles can be adapted to address the challenges posed by AI systems in space.

Given the unique context of space activities—characterized by long-distance data transmission, real-time decision-making by autonomous systems, and operations in multiple jurisdictions—the need to ensure the lawful processing and protection of personal data is paramount.

AI-driven satellites, probes, and space-based systems are increasingly tasked with processing large amounts of personal data, often without immediate human intervention. This raises questions about how to apply GDPR principles like lawfulness, fairness, transparency, data minimization, and security in such unregulated environments.

This section delves into the most pertinent GDPR provisions, including Article 5 (principles relating to the processing of personal data), Article 6 (lawfulness of processing), Article 22 (automated individual decision-making), Article 25 (data protection by design and by default), and Article 32 (security of processing).

By examining these specific articles, the analysis highlights how their provisions can be used to create safeguards and guidelines for the ethical use of AI in space, ensuring privacy protections are maintained even when data is processed autonomously by AI systems.

In doing so, this section presents a detailed framework for addressing the regulatory gaps and privacy risks that emerge when AI operates independently in space. It also emphasizes the importance of embedding GDPR principles throughout the lifecycle of AI systems, from their initial development to their deployment and operational phases in space.

By drawing these connections, the analysis provides a roadmap for ensuring that AI-driven space operations respect privacy laws and align with the best practices of data protection.

1. *Article 5: Principles Relating to Personal Data Processing*

Within Chapter II of the GDPR, Article 5 provides a series of core principles which shall be considered when processing personal data under this regime, thereby influencing, both directly and indirectly, the set of rules contained in this instrument. In order for the processing to be deemed legitimate, these principles are to be applied cumulatively.

These principles were informed by the initial content of Article 5 of the Council of Europe's Convention 108,[64] as it encompasses the same principles pertaining to the lawfulness and fairness of processing, purpose limitation, data minimization, accuracy of data, and storage limitation. Nevertheless, the accountability principle was not explicitly referenced in the Convention until the revised version, which incorporated the accountability principle into the newly introduced Article 10(1).[65]

In the case where an AI-enabled space object processes information that is considered "personal data" (as per Article 2 GDPR), following the territoriality criterion laid down in Article 3,[66] such processing will necessarily have to comply with these principles to be deemed lawful in conformity with the present instrument.

1.1 *Principle of Lawfulness, Fairness, and Transparency*

In accordance with Article 5(1), letter a), "Personal data shall be: (a) processed lawfully, fairly and in a transparent manner in relation to the data subject ('lawfulness, fairness and transparency');" a mandate which is enunciated previously in the preamble to the Regulation, concretely, in Recital (39): "Any processing of personal data should be lawful and fair."

Consequently, there are three constraints on this first principle: (1) lawfulness; (2) fairness; and (3) transparency. A detailed analysis of these notions independently is shown as necessary in order to understand the scope of the principle.

1.1.1 Lawfulness

The "lawfulness" of the processing is defined by its compliance with the GDPR and other regulatory regimes. Accordingly, this precept shall be read in connection with Recital (40), establishing that "In order for processing to be lawful, personal data should be processed on the basis of the consent of the data subject concerned or some other legitimate basis, laid down by law, either in this Regulation or in other Union or Member State law as referred to in this Regulation."

---

[64] *Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (ETS No. 108) 1985* (5 ratifications so far). This Convention represents the first ever legally binding international instrument designed to safeguard individuals against potential infringements associated with the collection and processing of personal data. Additionally, it endeavors to regulate the transfrontier flow of personal data. https://www.coe.int/en/web/conventions/full-list?module=treaty-detail&treatynum=108

[65] De Terwangne, C. 2020. "Article 5: Principles relating to processing of personal data." in Kuner, C., Bygrave, L. A., and Docksey, C. (eds.), *The EU General Data Protection Regulation (GDPR): A Commentary*. Oxford, Oxford University Press, p. 312.

[66] Some examples include instances where a data controller or processor based in the EU processes personal data by means of satellite technology in outer space, or where satellite internet, GPS, and media providers offer their services to customers in the EU, or even where an EU-based data center handles data concerning individuals in the EU. The applicability of the GDPR in such situations appears to be well-founded. See Zoltick, M. M., and Colgate, J. L. 2019. "The Application of Data Protection Laws in (Outer) Space." *International Comparative Legal Guide to: Data Protection 2019*, Global Legal Group Ltd, p. 9. https://www.rothwellfigg.com/assets/htmldocuments/ICLG__Data_Protection_2019_RothwellFigg_Outer_Space.pdf

For the processing of "personal data" under the GDPR to be lawful, the substantiation of the processing on one of the legal grounds provided for by Article 6 GDPR is thus required; otherwise, such processing shall be deemed unlawful. These different legitimating bases regulated under Article 6 GDPR will be carefully considered in the following section.

When interpreting this provision, especially in the context of the development and deployment of AI systems, the question of what processes or procedures are covered by the "processing of personal data" under the GDPR becomes crucial. For instance, in addition to the typical processing of personal data for the benefit of the data subject, it could also be interpreted that training of AI systems using large datasets of anonymized or de-identified data collected from millions of data subjects comprises processing of their personal data.

1.1.2 Fairness

In some legal systems, the "fairness" requirement might be compared to the general requirement of good faith.[67] In the context of data protection, fairness can be defined as a standard of conduct that requires the responsible handling of personal data in a manner consistent with the reasonable expectations of individuals and the avoidance of any actions resulting in unjustified harm.[68]

As per Recital (60) GDPR, the principle of fair (and transparent) processing demands an informational duty of the controller towards the data subject, in a way that the latter shall be duly informed as to the existence of the processing operation and its intended purposes, providing the data subject with all the necessary information to guarantee the criterion of fairness in the processing of the personal data (e.g., the consequences of not providing the personal data), and contextualizing it with respect to the specific conditions thereof.

1.1.3 Transparency

The concept of "transparency" is not defined *per se* in the GDPR.[69] Nevertheless, Recital (39) GDPR offers valuable insights on the meaning and impact of the data transparency principle within the context of data processing. Accordingly, this principle encompasses several core elements: firstly, data subjects must be informed of the collection, use, consultation, or processing

---

[67] Hoofnagle, C. J., Van Der Sloot, B., and Zuiderveen Borgesius, F. 2019. "The European Union general data protection regulation: what it is and what it means." *Information & Communications Technology Law* 28(1), p. 77. https://doi.org/10.1080/13600834.2019.1573501 Good faith can be described as: "The expression "good faith and fair dealing" refers to a standard of conduct characterised by honesty, openness and consideration for the interests of the other party to the transaction or relationship in question" (Draft Common Frame of Reference (DCFR), I. – 1:103(1), Book I: General provisions). See von Bar, C. et al. 2009. *Principles, Definitions and Model Rules of European Private Law – Draft Common Frame of Reference (DCFR)*, prepared by the Study Group on a European Civil Code and the Research Group on EC Private Law (Acquis Group), Munich, sellier. European law publishers, p. 178.

[68] Information Commissioner's Office (ICO). *What do we need to know about accuracy and statistical accuracy?* https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/what-do-we-need-to-know-about-accuracy-and-statistical-accuracy/

[69] Article 29 Data Protection Working Party. 2018. *Guidelines on transparency under Regulation 2016/679*, 17/EN WP260, p. 6. https://www.edpb.europa.eu/our-work-tools/our-documents/article-29-working-party-guidelines-transparency-under-regulation_en

of their personal data; secondly, they must be made aware of the extent to which such data is or will be processed; and thirdly, they must be provided with the identity of the controller.

In addition, all this information shall be delivered in a readily discernible, intelligible,[70] and legibly presented format. It is thus imperative that clear and unambiguous language be employed,[71] complemented where appropriate by the use of visual aids. This approach ensures that the information pertaining to the processing of personal data is readily comprehensible.[72] Such information could be provided in an electronic medium, for instance, when directed to the general public, via a website.[73]

Yet, there is no obligation to provide information when the data subject already has access to it, when the recording or disclosure of personal data is explicitly permitted by law, or in situations where it would be impossible or entail an undue burden to share the information with the data subject (e.g., where data is processed for archiving purposes in the public interest, scientific or historical research purposes, or statistical purposes).[74]

The transparency principle is further elaborated on by Articles 12-14 GDPR, dealing with transparent information, communication and modalities for the exercise of the rights of the data subject, depending on whether personal data have been collected from the data subject or not.

Accordingly, when personal data is processed by an AI-driven space object, the compliance of such processing with the principle of lawfulness, fairness and transparency demands an explanation of the way the AI system makes decisions in a transparent and open manner, as well as clearly outlining how personal data may be used as input to train the AI system.

The explainability or interpretability of AI decisions can be conducted through process-based explanations, which illustrate the adherence to GDPR and best practices throughout the design and operationalization of the system, and through outcome-based explanations, which serve to elucidate the results of a particular decision taken by the AI software. The latter may be focused on clarifying the degree of meaningful human involvement in the decision-making process, or on whether the actual outcome –the decision made by AI– observes the criteria predetermined in the design process.[75]

---

[70] The requirement that information is "intelligible" demands it to be comprehensible to an average member of the target public. The controller must initially identify the latter and determine the average member's level of comprehension. It is important to note that the designated audience may differ from the actual one, being therefore the controller's responsibility to assess on a regular basis the suitability of the information or communication for the actual audience, particularly in cases where the children are involved. If necessary, adjustments should be made to ensure compliance with the transparency principle. One effective method is through user panels, which can test the intelligibility of the information and the effectiveness of user interfaces, notices, policies, and other relevant materials. See *ibid.*, pp. 7-8.

[71] The requirement for clear and plain language implies the provision of information in a straightforward manner, eschewing complex sentence structures and language. The content should be specific and definitive, avoiding abstract or ambivalent phrasing and interpretations. Concretely, the purposes and legal basis for processing personal data must be explicitly stated. See *ibid.*, p. 9.

[72] Recitals (40) and (58), *General Data Protection Regulation (GDPR) 2016.*

[73] Recital (58), *GDPR 2016.*

[74] Recital (62), *GDPR 2016.*

[75] ICO. *How do we ensure fairness in AI?*

Moreover, the "transparency" requirement of the principle, as aforementioned, implies that, when breaking down the decision-making process of AI systems in outer space, the explanations shall be given in a plain, straightforward, and readily comprehensible language.

1.2 *Principle of Purpose Limitation*

In addition to being lawful, fair and transparent, personal data needs to be "collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes," as Article 5(1)(b) enshrines. Consequently, the principle of purpose limitation is constituted of two components: (a) the data must collect data solely for the purposes specified, explicitly and legitimately, and (b) once data have been collected, these must not be further processed in a manner that is incompatible with those purposes.

The principle of purpose limitation is intended to delineate the scope of permissible processing and reuse of personal data collected for a specific purpose. This underscores the importance of "specification of purpose" as a prerequisite for applying other data quality criteria, such as the adequacy, relevance, proportionality, and accuracy of data collected, and the requirements concerning data retention periods.[76] The "specification of purpose" shall take place at the time of the personal data collection, and these specific purposes, for which personal data are processed, must be explicit and legitimate.[77]

In the event that the controller aims to provide the data subject with information regarding the additional purpose for which the data is to be processed, as well as any other relevant information, prior to such further processing.[78] In order to ascertain the compatibility of a novel purpose with the original, the controller is advised to consider a number of factors. These include the relationship between the initial and new purposes, the contextual factors, the reasonable expectations of data subjects, the nature and sensitivity of the data in question, and the potential repercussions for data subjects of the envisaged further processing.[79]

The rule also introduces that "further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes." The GDPR does not view further processing as inherently incompatible with the original purposes for which data was collected, provided that it is conducted for archiving purposes in the public interest, scientific or historical research purposes, or statistical purposes. Nevertheless, the GDPR mandates additional

---

https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/how-do-we-ensure-fairness-in-ai/

[76] Article 29 Data Protection Working Party. 2013. *Opinion 03/2013 on purpose limitation*, 00569/13/EN WP 203, adopted on 2 April 2013, p. 4. https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf

[77] Recital (40), *GDPR 2016*.

[78] Recital (61), *GDPR 2016*.

[79] See Article 6(4) and Recital (50), *GDPR 2016*.

safeguards be put in place for such further processing (e.g., pseudonymization), which correspond to those underlined by Article 89(1) GDPR.[80]

Concerning scientific research as one of the exceptions to the general norm of purpose limitation, it is interesting to resort to Recital (159) GDPR which provides that "the processing of personal data for scientific research purposes should be interpreted in a broad manner including for example technological development and demonstration, fundamental research, applied research and privately funded research."

These elements suggest that, in certain instances, the development of AI may be regarded as scientific research,[81] and therefore the reutilization of personal data for these purposes should be considered to still be compatible with the original purposes for data collection.

The principle of purpose limitation, which is likewise set forth in Article 8(2) of the Charter of Fundamental Rights of the European Union ("Such data must be processed fairly for specified purposes…"), has a statutory development in Articles 18 and 19, relating to the "Right to restriction of processing" and the "Notification obligation regarding rectification or erasure of personal data or restriction of processing," respectively.

Regarding the application of this principle to autonomous space systems, the purpose limitation principle can be interpreted in an approach that is consistent with AI and Big Data. In this respect, the concept of "compatibility" shall be applied in a flexible manner, which would enable the reuse of personal data when it does not contravene the initial purposes for which the data were collected. In addition, the reuse of data for statistical purposes is presumed to be compatible and, as a result, would typically be permissible, unless it poses undue risks for the data subject.[82]

It is undeniable that AI systems may process personal data throughout their lifecycle for a multitude of purposes, varying according to their specific stages and functions. This is why, in order to observe the principle of purpose limitation, it is crucial to determine these potential purposes in advance to then decide the appropriate legal basis for each of them, thereby ensuring the fairness of the processing of these data and the resulting outcome for individuals.

To this effect, it might be interesting to differentiate between specific phases of the AI lifecycle (e.g., between system development and deployment) and delineate the objectives of data processing at each stage.[83]

---

[80] "Processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, shall be subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject. Those safeguards shall ensure that technical and organizational measures are in place in particular in order to ensure respect for the principle of data minimization. Those measures may include pseudonymization provided that those purposes can be fulfilled in that manner. Where those purposes can be fulfilled by further processing which does not permit or no longer permits the identification of data subjects, those purposes shall be fulfilled in that manner."

[81] Datatilsynet. 2018. *Artificial intelligence and privacy*, Report, January 2018, p. 17. https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf

[82] European Parliament. *The impact of… op. cit.*, p. II.

[83] ICO. *How do we… op. cit.*

1.3 *Principle of Data Minimization*

Adding up to the previous requirements, Article 5(1)(c) imposes the obligation for personal data to be "adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation')."[84] Data minimization is the natural corollary to purpose limitation.[85]

In this regard, it is crucial that personal data be sufficient, pertinent, and confined to what is essential for fulfilling the intended purposes of processing. In particular, this entails ensuring that the retention period for personal data is strictly limited to a minimal duration. Personal data should only be processed insofar as the purpose of the processing could not be reasonably achieved by any alternative means.[86]

Accordingly, this provision stipulates that only the minimum amount of personal data required to achieve the intended purpose should be processed, which constitutes, in turn, one of the technical and organizational measures outlined in Article 25(2) GDPR, which reiterates that controllers may only process personal data that is "necessary for each specific purpose of the processing."

Data minimization is a risk-management measure.[87] The processing of excess data is unnecessary and, as a result, creates an array of potential, unnecessary risks. These encompass a spectrum of issues, from the vulnerability to hacking to the formation of unreliable inferences, which may culminate in erroneous, injurious, and potentially perilous decisions.[88]

It is only permissible to obtain data that are necessary for the specific purpose in question. The data minimization principle thus precludes the practice of collecting as much personal data as possible, under the basis that such data could prove advantageous in future scenarios. This approach effectively challenges the tenets of many Big Data business models,[89] among which the space sector could be included.

Primarily, due to the lack of a clear and definitive objective for the processing of data while developing AI. The nature of the algorithmic processes involved entails that it is inherently difficult to ascertain precisely what the machine will learn. As the learning and development process continues, the purpose may also evolve. This introduces a challenge to the data minimization principle, as it becomes increasingly intricate to determine which data is critical and which can be regarded as superfluous.[90]

---

[84] The principle of data minimization is also reflected in Recital (78) GDPR, which includes minimizing the processing of personal data as an organizational measure for data protection by design and by default.

[85] Biega, A. J., and Finck, M. 2021. "Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems." *Technology and Regulation*, p. 55. https://doi.org/10.26116/techreg.2021.004

[86] Recital (39), *GDPR 2016*.

[87] See Biega and Finck. "Reviving Purpose Limitation… *op. cit.*, p. 55.

[88] Hildebrandt, M. 2018. "Primitives of Legal Protection in the Era of Data-Driven Platforms." *Georgetown Law Technology Review* 2(2), p. 267. https://georgetownlawtechreview.org/primitives-of-legal-protection-in-the-era-of-data-driven-platforms/GLTR-07-2018/

[89] See Hoofnagle, Van Der Sloot, and Zuiderveen Borgesius. "The European Union… *op. cit.*, p. 78.

[90] See Datatilsynet. *Artificial intelligence and... op. cit.*, p. 18.

In light of these considerations, it may appear that complying with the principle of data minimization is an elusive goal in the context of AI. Nevertheless, this principle does not imply that an AI system incorporated into a space object is prevented from processing any personal data in absolute terms. Conversely, the objective is to ascertain what personal data is adequate, relevant, and limited, based on the specific use case of the AI system in question.[91]

There is this common misconception that the principle of data minimization has no bearing on this particular domain. In fact, however, this obligation extends throughout the entirety of the AI system's lifecycle, including the testing, acceptance, and production release phases. This ensures that personal data is not indiscriminately collected, processed,[92] or used for training AI systems.

In applying this principle also to the development and training of AI systems, it is pertinent to note that there is a natural tension between the principle of data minimization and the development of highly accurate AI systems, which typically relies on maximizing the amount of data (i.e., big data) to train the deep neural nets used in LLM transformer architectures.

Therefore, the conclusion is that data minimization does not preclude the incorporation of supplementary personal data into a processing operation, provided that the inclusion of such data confers an advantage in relation to the objectives of the processing, which offsets the additional risks for the data subjects.[93]

The continued relevance of the data for future processing may also justify the retention of the same, on the condition that appropriate security measures are implemented. Concretely, the combined use of pseudonymization with other security measures may serve to mitigate risks and, consequently, boost the compatibility of retention with minimization.[94]

Within the space context, in future space missions, AI might need to collect more data to adapt to unexpected situations or optimize its performance. This could challenge the principle of data minimization, since the definition of "necessary" data may expand. As AI systems evolve, defining what data is truly necessary for their operation becomes complex, making it crucial to establish clear guidelines on what constitutes necessary data.

Future scenarios could involve AI deciding to collect additional data autonomously to enhance its operations, which might challenge the principle of data minimization. To illustrate, in a scenario where an AI system on a satellite autonomously decides to collect more detailed personal data to improve its predictive models, it would need to ensure that this data collection adheres to the principle of data minimization and is justifiable under the GDPR framework.

While an AI system on a satellite might propose collecting detailed personal data from Earth-based sensors to enhance its models, it must limit this collection to only what is necessary for the specific purpose of maintaining satellite operations and ensuring its efficiency and safety.

---

[91] ICO. *How do we… op. cit.*

[92] European Data Protection Supervisor (EDPS). 2024. *Generative AI and the EUDPR. First EDPS Orientations for ensuring data protection compliance when using Generative AI systems*, 3 June 2024, p. 14.

[93] See, in this regard, the legitimate interest as a legal basis for the processing of personal data as per Article 6(1)(f) GDPR.

[94] See European Parliament. *The impact of... op. cit.*, p. 47.

This way, an AI system on a satellite should only collect data essential for its mission objectives, avoiding the accumulation of unnecessary personal information.

1.4 *Principle of Accuracy*

Article 5(1)(d) GDPR provides that personal data shall be "accurate and, where necessary, kept up to date," thereby embodying the principle of accuracy. Pursuant to the same, data controllers must take "every reasonable step (…) to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay."[95]

The accuracy principle does not mandate full accuracy in and of itself; rather, it is contingent on accuracy being demonstrated in a manner that is commensurate with the intended purposes for which personal data are processed. It is incumbent upon data controllers to proactively ensure the accuracy of the data in their possession and to afford data subjects the opportunity to correct any inaccuracies.[96]

The accuracy principle is applicable to all forms of personal data, including information about an individual which is employed as "input" into an AI system or generated as an output of the system. Nonetheless, it is not a requirement for an AI system to be completely "statistically accurate"[97] so as to observe such principle. It is thus essential that the AI system in question is "sufficiently statistically accurate" to guarantee that any personal data yielded by it is processed in accordance with the principle of lawfulness, fairness and transparency.[98]

This is the reason why the personal data utilized for the training of AI models, or their outputs must be kept current and maintained in a consistent state. Inaccurate input data has the potential to result in erroneous inferences about individuals which may be at odds with their reasonable expectations or culminate in unfavorable outcomes,[99] both in terms of development/training and operational processes.

It is the responsibility of data controllers to verify the accuracy of data at each stage of the development and deployment of an AI system. The implementation of the requisite measures to incorporate data protection by design is of paramount importance for the enhancement of data accuracy across all stages.[100]

Such an approach entails the examination of both the structure and content of the datasets utilized for training models, encompassing those procured from third parties. In regard to the output data, it is recommended that developers utilize validation sets as part of the training process

---

[95] See also Recitals (40) and (71) GDPR, the latter concerning the principle of accuracy in the context of profiling.

[96] See Hoofnagle, Van Der Sloot, and Zuiderveen Borgesius. "The European Union… *op. cit.*, p. 78. In this regard, see also Article 16 GDPR on the right to rectification: "The data subject shall have the right to obtain from the controller without undue delay the rectification of inaccurate personal data concerning him or her. Taking into account the purposes of the processing, the data subject shall have the right to have incomplete personal data completed, including by means of providing a supplementary statement."

[97] "Statistical accuracy refers to the proportion of answers that an AI system gets correct or incorrect." See ICO. *What do we… op. cit.*

[98] See *ibid.*

[99] ICO. *How do we… op. cit.*

[100] See EDPS. *Generative AI and… op. cit.*, p. 15.

and employ separate testing sets for the final evaluation in order to ascertain an estimation of the system's anticipated performance.[101]

In the context of machine learning (ML) systems, it is essential to differentiate between two distinct scenarios: the utilization of personal data solely for the purpose of training a model to identify general statistical correlations, and the integration of personal data as input to a profiling algorithm. It should be acknowledged, however, that these are not the only functions that AI neural network (NN) models are trained to perform.

The majority of NNs in current AI systems are relatively opaque, and, as a result, while training may be conducted to meet specific statistical correlation performance metrics, this does not necessarily reflect their underlying capabilities. For example, inferential and deductive reasoning applications extend beyond the detection of statistical correlations.

After the data have been made accessible for the training set, the inclination to utilize the same data for individualized inferences will be considerable. The implementation of anonymization or pseudonymization procedures, coupled with robust security measures, can facilitate the mitigation of associated risks.[102]

1.5 *Principle of Storage Limitation*

As per Article 5(1)(e) GDPR, personal data is to be "kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed." By formulating this principle of storage limitation, the GDPR imposes strict restrictions on the duration for which personal data may be stored: a "no longer than necessary" standard is mandated.[103] Recital (39) GDPR further specifies that, in order to meet this objective, data controllers should establish in advance either periodic reviews or specific timelines for the prospective deletion of personal data.

However, the clause contains an exception to this general rule, by which "personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) subject to implementation of the appropriate technical and organizational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject."

It cannot be denied that there exists a significant conflict between the utilization of AI-based processing techniques on extensive collections of personal data and the principle of storage limitation. Nonetheless, this inherent discord can be mitigated to a certain extent when, first, the data are employed for statistical purposes and, second, the appropriate measures are implemented at the national level.[104]

Therefore, in the event that an AI system retains data for a period longer than is necessary to achieve the desired outcome, this processing is not only unwarranted but also constitutes an

---

[101] *Ibid.*

[102] See European Parliament. *The impact of... op. cit.*, p. 48.

[103] See Hoofnagle, Van Der Sloot, and Zuiderveen Borgesius. "The European Union… *op. cit.*, p. 78.

[104] See European Parliament. *The impact of... op. cit.*, p. 49.

unfair practice. It is imperative that a proportionate approach be taken with regard to the retention periods in question, whereby the needs of data controllers are carefully weighed against the potential implications that such data retention may have on the privacy rights of the data holders.[105]

1.6 *Principle of Integrity and Confidentiality*

With respect to the principle of integrity and confidentiality, it is Article 5(1)(f) which prescribes that personal data shall be "processed in a manner that ensures appropriate security [and confidentiality][106] of the personal data, including protection against unauthorized or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organizational measures." This principle is further concretized in Articles 32 and following of the Regulation, regarding the security of processing.

The principle of integrity and confidentiality places an onus on those responsible for data security. Security measures must be adequate and designed to protect against loss, destruction, damage and unlawful processing. Consequently, the use of data within an organization, for instance, may even constitute a security violation.[107]

1.7 *Principle of Accountability*

Finally, Article 5 GDPR, now in its second paragraph, provides that "The controller shall be responsible for, and be able to demonstrate compliance with, paragraph 1." This corresponds to the principle of accountability, which, in turn, is primarily developed in Articles 24 and 25 GDPR.

It is crucial that accountability be firmly established for any party engaged in the processing of personal data, that being the controller or someone on behalf of the latter.[108] This responsibility extends to the implementation of robust and effective safeguards, and to the ability to substantiate compliance with the provisions outlined in the GDPR, particularly with regard to the effectiveness of said measures. Furthermore, these measures must align with the specific characteristics of the data in question, including its intended use, the purposes and context in which it is processed, and the potential risks to the rights and freedoms of individuals.[109]

In this way, the GDPR reiterates the accountability of the data controller as the entity entrusted with the responsibility of data management. It also introduces heightened regulatory constraints, obligations, and potential liability for processors.[110] The objective is to encourage entities that process personal data to prioritize prevention.

---

[105] ICO. *How do we… op. cit.*

[106] Recital (39), *GDPR 2016*.

[107] See Hoofnagle, Van Der Sloot, and Zuiderveen Borgesius. "The European Union…*op. cit.*, p. 78.

[108] "The protection of the rights and freedoms of data subjects as well as the responsibility and liability of controllers and processors, also in relation to the monitoring by and measures of supervisory authorities, requires a clear allocation of the responsibilities under this Regulation, including where a controller determines the purposes and means of the processing jointly with other controllers or where a processing operation is carried out on behalf of a controller." Recital (79), *GDPR 2016*.

[109] Recital (74), *GDPR 2016*.

[110] See Hoofnagle, Van Der Sloot, and Zuiderveen Borgesius. "The European Union…*op. cit.*, p. 85.

This entails taking a proactive approach, anticipating and addressing potential risks before a breach occurs. The term "proactive responsibility" emphasizes the need for data controllers to implement measures that reasonably ensure their compliance with the principles, rights, and guarantees set forth in the Regulation.[111]

Pursuant to Articles 24 and 25 GDPR, in light of the prevailing state of the art, the associated costs of implementation, the nature, scope, context, and purposes of processing, in addition to the potential risks to the rights and freedoms of individuals posed by the processing, the controller is obliged to consider the following at two distinct stages: firstly, when determining the means of processing, and secondly, when processing is underway.

It is incumbent upon the controller to put in place the requisite technical and organizational measures, namely pseudonymization, which are intended to ensure the effective implementation of data protection principles. Moreover, these measures must be integrated into the processing in a manner that meets the requirements of the GDPR and safeguards the rights of data subjects. Such measures shall be subject to periodic review and updating as necessary.

In this statement, the GDPR does not establish a list of security measures. Instead, it allows data controllers a wide margin of flexibility in determining the specific security measures that they will implement. These measures must be consistent with the structure of the data processing operations, the type of data being processed, and the specific regulations that the data controllers are subject to.[112]

In order to prove compliance, it is advisable to make use of a Data Protection Impact Assessment (DPIA), which offers an optimal avenue for such a demonstration. In accordance with the GDPR, a DPIA is mandatory whenever a new project is initiated, in the event that is presumed to entail a considerable risk to the personal data of other individuals.[113]

This information can be found in Article 35 GDPR, which regulates DPIAs: "Where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data. A single assessment may address a set of similar processing operations that present similar high risks."

Specifically, in accordance with the established criteria,[114] DPIAs are required in instances where a comprehensive and far-reaching assessment of personal aspects pertaining to natural persons is conducted through automated processing, including profiling, and upon which decisions are based that have legal implications for the individual or exert a considerable influence on them. Furthermore, it is mandated for the large-scale processing of special categories of data (Art. 9(1)

---

[111] See Beltrán Aguirre, J. L. 2018. "Reglamento General de Protección de Datos: novedades. Adaptación de la normativa española: el proyecto de LOPD." Vol. 28 Extraordinario XXVII Congreso, p. 77.

[112] *Ibid.*, pp. 77-78.

[113] Wolford, B. *Data Protection Impact Assessment (DPIA)*, https://gdpr.eu/data-protection-impact-assessment-template/

[114] Article 35(3), *GDPR 2016*.

GDPR) and of personal data concerning criminal convictions and offenses (Art. 10 GDPR), and in cases of systematic surveillance of a publicly accessible area on a vast scale.

In this context, it is important to note that conducting DPIAs for AI should not be regarded as mere perfunctory compliance exercises. Conversely, these can serve as comprehensive roadmaps for effectively identifying and mitigating the potential risks to rights and freedoms stemming from the use of AI. They also offer a valuable platform for assessing and demonstrating accountability for the decisions made in the design and procurement of AI systems.[115]

In addition to conducting a DPIA, other types of impact assessments may also be demanded in conjunction to fulfill the necessary requirements. By way of illustration, some organizations elect to conduct "algorithm impact assessments" (AIAs).[116] Likewise, the machine learning community has proposed the use of "model cards"[117] and "datasheets."

These documents describe how machine learning models may perform under different conditions and provide information about the datasets on which they are trained. This information may assist in conducting an impact assessment.[118]

2. *Article 6: Lawfulness of Processing*

Efficient and effective deployment of AI at the edge will unlock new process capabilities for the space industry, extracting insight from vast amounts of data collected by satellites. Until recently, satellite data had to be analyzed by operators on Earth.

However, such capabilities raise questions connected to the risk of having personal data processed by AI on board satellites. Art. 6 concerns the lawfulness of processing of personal data.[119] Indeed, it requires that any processing of personal data has a legal basis. This is particularly crucial for AI systems in space, which may need to process data under urgent or emergency conditions, such as during a space mission crisis or to address a natural disaster.

It is important to highlight the significance of establishing legal bases for data processing in space, especially in emergency scenarios, and to discuss potential ambiguities. For example, during an emergency, obtaining consent may not be feasible, and reliance on "vital interests" or "public interest" could lead to ambiguities. Establishing clear protocols for such scenarios is essential to ensuring compliance.

---

[115] ICO. *What are the accountability and governance implications of AI?* https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/what-are-the-accountability-and-governance-implications-of-ai/

[116] For instance, in the case of Canada, see https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html

[117] Model cards are concise documents that accompany trained machine learning models. They provide information regarding the intended context of model usage, the specifics of the performance evaluation procedures, and other pertinent details. This framework may be utilized for the documentation of any trained machine learning model. To see more information, visit https://oecd.ai/en/catalogue/tools/model-cards

[118] *Ibid.*

[119] GDPR, Art.6, https://gdpr-info.eu/art-42-gdpr/.

In addition, the requirement is constitutionalized in Article 8 of the European Charter of Fundamental Rights, according to which personal data "must be processed […] on the basis of the consent of the person concerned or some other legitimate basis laid down by law."[120]

The processing of personal data in the context of AI application brings complications related to the existence of a lawful legal basis. To understand when a legal basis may support AI-based processing, it is needed to consider the different legal bases set forth in Article 6 GDPR, which states that the processing of personal data is lawful under the following conditions: (a) consent of the data subject, or necessity (b) for performing or entering into a contract, (c) for complying with a legal obligation, (d) for protecting vital interests(e) for performing a task in the public interest or in the exercise of public authority, or (f) for a legitimate interest.[121]

Of particular interest for the use of AI in outer space is recital 46 connected to art.6(1)(d) GDPR that states: "The processing of personal data should also be regarded to be lawful where it is necessary to protect an interest which is essential for the life of the data subject or that of another natural person. […] Some types of processing may serve both important grounds of public interest and the vital interests of the data subject as for instance when processing is necessary for humanitarian purposes, including for monitoring epidemics and their spread or in situations of humanitarian emergencies, in particular in situations of natural and man-made disasters."[122]

It follows, that in the GDPR we find a lawful basis for processing data, allowing AI to extract insight to intervene in specific humanitarian emergencies, such as climate disaster mitigations plan where space data is of extreme importance. GDPR broadly supports processing of personal data by AI for any legitimate public interest including scientific research.

Indeed, Articles 89[123] and 9(2)(j) GDPR[124] provide some specific clauses that support the process of personal data for scientific research purposes. But, in any particular situation, according to art. 6(1)(f), a balancing test must be carried out to ensure that the rights and freedoms of individuals are not overridden by the public interest claim.

In a hypothetical situation where an AI system on a satellite detects a natural disaster on Earth and autonomously begins processing data to assist in disaster response, the legal basis for this data processing must be clearly established, potentially under the "vital interests" or "public interest" grounds.

3. *Article 22: Automated Individual Decision-Making, Including Profiling*

Lessons for AI regulation from GDPR Art. 22 involve data subjects and data controllers and include data subject rights that are independent of computing system architectures, i.e., whether a data aggregation and processing platform is a conventional computing system, an

---

[120] Charter of Fundamental Rights of the European Union 2012/C 326/02 of 26 October 2012 [2012] OJ C326/391.

[121] European Parliament, 'The impact of the General Data Protection Regulation (GDPR) on artificial intelligence' (2020).

[122] GDPR, Recital 46, https://gdpr-info.eu/art-42-gdpr/.

[123] GDPR, Art. 89, https://gdpr-info.eu/art-42-gdpr/.

[124] GDPR, Art. 9, https://gdpr-info.eu/art-42-gdpr/.

artificial intelligence computing system, or other advanced computing platform, e.g., quantum, neuromorphic, biocomputing, etc.

And, to the extent that AI systems perform the activities described in GDPR Art. 22, they would be subject to regulation under GDPR. For example, an AI system on a spacecraft making autonomous navigational decisions must still comply with these regulations, ensuring human oversight and the ability to challenge automated decisions.

As AI systems become more autonomous, ensuring meaningful human oversight will become more challenging. Future scenarios might involve AI making split-second decisions that humans cannot review in real time, raising questions about accountability and compliance with GDPR's requirement for human intervention.

Take for example a scenario where an AI system on a satellite must autonomously decide at the last minute with minimal warning to adjust its orbit to avoid a potential collision with space debris. This decision, made without immediate human intervention, must still comply with GDPR requirements for transparency and the ability for human review and intervention.

Meaning, at the very least the AI must log its decision-making process, providing a clear rationale and allowing operators on Earth to review and, to the extent possible, if necessary, override the AI's actions to ensure accountability and compliance with GDPR.

Of direct importance to AI systems, the first paragraph of GDPR Art. 22 begins by stipulating that "[t]he data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her," such as automatic refusals of an online applications (e.g., credit) or e-recruiting practices lacking human intervention.[125]

GDPR Art. 22 immediately thereafter caveats that this does not apply to decisions: 1) necessary for entering into or performance of a contract between the data subject and a data controller, 2) authorized by the EU or appropriate Member State under law that provides sufficient safeguards to data subject rights, freedoms and legitimate interests, or 3) based on explicit consent of the data subject.[126]

Otherwise, under paragraph 3, data controllers must "implement suitable measures to safeguard data subject rights, freedoms and legitimate interests," including the right to obtain intervention by a human representative of the data controller to enable data subjects to express their perspectives, obtain specific information and explanations regarding decisions made, and contest decisions.[127]

Further, the exempted decisions provided under paragraph 2 must "not be based on special categories of personal data as provided for under Art. 9(1),"[128] unless excluded under Art.

---

[125] EU Regulation 2016/679, "General Data Protection Regulation," Official Journal L 119 (2016), 1-88 [hereinafter GDPR], Art. 22(1), https://gdpr-info.eu/art-22-gdpr/.

[126] GDPR, Art. 22(2), https://gdpr-info.eu/art-22-gdpr/.

[127] GDPR, Art. 22(3), https://gdpr-info.eu/art-22-gdpr/.

[128] GDPR, Art. 9(1), https://gdpr-info.eu/art-9-gdpr/, *See* "Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation shall be prohibited."

9(2)(a)[129] or 9(2)(g)[130] and "suitable measures to safeguard data subject rights and freedoms and legitimate interests" have been implemented.[131]

Recital 71 of Art. 22 elaborates that automated data processing involves "profiling" when it includes "any form of automated processing of personal data evaluating the personal aspects relating to a natural person, in particular to analyse or predict aspects concerning the data subject's performance at work, economic situation, health, personal preferences or interests, reliability or behaviour, location or movements" where it produces legal or similar effects upon the data subject.[132]

Decisions based on automated data processing including profiling are allowed in certain circumstances, including expressly consented to by the data subject or explicitly authorized by EU or member state laws, such as "for fraud and tax-evasion monitoring and prevention purposes conducted in accordance with the regulations, standards and recommendations of Union institutions or national oversight bodies […] to ensure the security and reliability of a service provided by the [data] controller."[133]

In many cases, apps and online platforms, e.g., social media, and web portals of other kinds, e.g., product warranty registration, etc., rely on the explicit consent of the user to access, process, and commercialize their data.

However, when it comes to deploying AI systems in space or capturing data to be processed by AI systems from space (e.g., via Very High-Resolution (VHR) cameras), data subjects may not be presented with an obvious opportunity to become informed that their data has been captured and processed, i.e., explicit consent is not necessarily part of the equation when images or video of a person is captured from space.

With the ongoing exponential proliferation of satellites in LEO, including VHR camera equipped satellites, the rapid technological advancement and broad expansion of these intrusive capabilities present clear privacy risks to individuals and organizations. The integration of data analytics software directly on board these satellites further amplify these risks, as it could potentially lead to prohibited types of automated decision making under Article 22.

Recital 71 further elaborates that data controllers should use appropriate mathematical and statistical tools in profiling, ensure error minimization and correction, and appropriately secure personal data in accordance with risks to interests and rights of data subjects.[134] Identified as of special importance, data controllers must prevent discriminatory effects on natural persons (i.e., data subjects) "on the basis of racial or ethnic origin, political opinion, religion or beliefs, trade

---

[129] GDPR, Art. 9(2)(a), https://gdpr-info.eu/art-9-gdpr/, *See* "the data subject has given explicit consent to the processing of those personal data for one or more specified purposes, except where Union or Member State law provide that the prohibition referred to in paragraph 1 may not be lifted by the data subject."

[130] GDPR, Art. 9(2)(g), https://gdpr-info.eu/art-9-gdpr/, *See* "processing is necessary for reasons of substantial public interest, on the basis of Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject."

[131] GDPR, Art. 22(4), https://gdpr-info.eu/art-22-gdpr/.

[132] GDPR, Recital 71 – Profiling, https://gdpr-info.eu/recitals/no-71/.

[133] *Ibid*.

[134] *Ibid*.

union membership, genetic or health status or sexual orientation" and automated decision making based on such special categories of personal data is allowed only under particular conditions.[135]

Recital 91 of Art. 22 requires data protection impact assessments, especially for large-scale data processing operations affecting large populations of data subjects (e.g., regional, national, supranational) that are likely to generate high degrees of risk, e.g., on account of the special character or sensitivity of the personal data.[136]

Recital 91 also explicitly calls out exemplary new technologies (opto-electronic devices, e.g., cameras) that would generate such high potential risks, especially in cases of systematic and extensive evaluation of personal data, e.g., for profiling or processing of special categories of personal data, biometric data, criminality and security data.[137] Recital 91 exempts data processing that "concerns personal data from patients or clients by an individual physician, other health care professional or lawyer."[138]

4. *Article 25: Data Protection by Design and by Default*

Lessons for AI regulation from GDPR Art. 25 involve requirements for data controllers to implement appropriate technical and organizational measures, e.g., data protection policies such as pseudonymization and data minimization, to effectively provide necessary data processing safeguards to protect data subject rights.[139]

Data controllers must implement these safeguards both in making determinations regarding specific data processing means (e.g., data processing system size, architecture, operations, etc.) and during/throughout the data processing activity. In the context of AI systems in space, this means incorporating data protection measures from the design phase to ensure that personal data is safeguarded throughout the mission lifecycle.

In implementing these safeguards, the data controllers must take into account the current state of the art, cost, nature, scope, context and purpose(s) of the data processing, and risk posed by the data processing to rights and freedoms of natural persons (i.e., data subjects).[140]

The safeguards must ensure that data processing systems default to processing only personal data necessary to a specific identified purpose for the data processing operation, including minimizing amount of personal data collected, extent of data processing, duration of storage and accessibility of the data.[141] The safeguards must also require data subjects to intervene or take affirmative action to allow their personal data to be made more widely accessible, with the system default to lock down data.[142]

---

[135] *Ibid.*

[136] GDPR, Recital 91 – Necessity of a Data Protection Impact Assessment, https://gdpr-info.eu/recitals/no-91/.

[137] *Ibid.*

[138] *Ibid.*

[139] GDPR, Art. 25(1), https://gdpr-info.eu/art-25-gdpr/.

[140] *Ibid.*

[141] GDPR, Art. 25(2), https://gdpr-info.eu/art-25-gdpr/.

[142] GDPR, Art. 25(1), https://gdpr-info.eu/art-25-gdpr/.

Recital 78 of Art. 25 states that additional safeguard measures may further include "pseudonymising personal data as soon as possible, transparency with regard to the functions and processing of personal data, enabling the data subject to monitor the data processing, enabling the controller to create and improve security features."[143]

Future AI systems in space may need to adapt to new threats and operational changes dynamically. Ensuring data protection by design will require AI systems to be flexible yet secure, capable of integrating new protection measures without compromising their functionality or security.

Consider the development of a new AI-controlled satellite where data protection measures such as encryption and pseudonymization are built into the design. As the satellite's mission evolves, these measures must be updated to address new security threats without compromising system integrity.

Art. 25 further states that "[a]n approved certification mechanism pursuant to Article 42 may be used as an element to demonstrate compliance with the [Art. 25] requirements."[144] Art. 42 highlights the special needs of micro, small and medium-sized enterprises,[145] requires Member States, the supervisory authorities, the Board, and the European Commission to encourage, especially at Union level, establishment of data protection certification mechanisms and data protection seals and marks for demonstrating compliance.[146]

Further, in addition to other stipulations regarding certification mechanisms, Art. 42 requires the Board to collate all certification mechanisms and data protection seals and marks in a register to be made publicly available by any appropriate means.[147]

5. *Article 32: Security of Processing*

Regarding the security of the processing of personal data, Article 32 GDPR reads as follows:

> "1. Taking into account the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons, the controller and the processor shall implement appropriate technical and organizational measures to ensure a level of security appropriate to the risk, including inter alia as appropriate:
>
> (a) the pseudonymization and encryption of personal data;
> (b) the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services;
> (c) the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident;

---

[143] GDPR, Recital 78 – Appropriate Technical and Organisational Measures, https://gdpr-info.eu/recitals/no-78/.
[144] GDPR, Art. 25(3), https://gdpr-info.eu/art-25-gdpr/.
[145] GDPR, Art. 42(1), https://gdpr-info.eu/art-42-gdpr/
[146] GDPR, Art. 42(2), https://gdpr-info.eu/art-42-gdpr/
[147] GDPR, Art. 42(8), https://gdpr-info.eu/art-42-gdpr/

(d) a process for regularly testing, assessing and evaluating the effectiveness of technical and organizational measures for ensuring the security of the processing."

When speaking of "data security," it is inevitable to address personal data breaches as the opposite case. As per Article 4(12) GDPR, a 'personal data breach' is understood as a "breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to, personal data transmitted, stored or otherwise processed."

The consequences of data breaches can include the inadvertent loss of partial and comprehensive databases, posing significant risks for the unauthorized disclosure of personal data. To forestall such breaches, Article 33 GDPR[148] establishes an overarching obligation to implement security measures on both data controllers and data processors.[149]

This demonstrates the interconnection between Articles 32, 33 and 34 GDPR, the latter stipulating that, in instances where a personal data breach is likely to pose a significant risk to the rights and freedoms of individuals, the controller is obliged to promptly communicate the breach to the data subject. This communication must be conveyed in a clear and straightforward manner, providing a comprehensive description of the nature of the personal data breach.

Data security is a fundamental aspect of the obligations imposed on data controllers and data processors within the regulatory framework set forth by the GDPR. For instance, both the integrity and confidentiality principle (Art. 5(1)(f) GDPR) and the accountability principle (Art. 5(2) GDPR) cannot be upheld without the implementation of robust security measures to safeguard the data being processed.[150]

In this vein, Recital (83) GDPR alludes to the responsibility of the data controller to assess the inherent risks associated with data processing and to adopt measures to mitigate those risks (e.g., encryption) with the aim of safeguarding the security of data. Such measures must guarantee an adequate level of security, including confidentiality, while considering the current technological standards and the costs of implementation relative to the risks and the nature of the personal data to be protected.

For space missions, these measures are critical to protect data transmitted across vast distances, where the risks of interception and unauthorized access are heightened. Implementing advanced encryption, for instance, by using quantum- or AI-based cryptographic networks and systems and employing regular security audits can mitigate these risks.

The Court of Justice of the European Union (CJEU), in turn, has ruled accordingly in its judgment *Worten – Equipamentos para o Lar SA v Autoridade para as Condições de Trabalho*

---

[148] "In the case of a personal data breach, the controller shall without undue delay and, where feasible, not later than 72 hours after having become aware of it, notify the personal data breach to the supervisory authority competent in accordance with Article 55, unless the personal data breach is unlikely to result in a risk to the rights and freedoms of natural persons. Where the notification to the supervisory authority is not made within 72 hours, it shall be accompanied by reasons for the delay," para. 1.

[149] Burton, C. 2020. "Article 33: Notification of a personal data breach to the supervisory authority." in Kuner, C., Bygrave, L. A., and Docksey, C. (eds.). *The EU General Data Protection Regulation (GDPR): A Commentary*. Oxford, Oxford University Press, p. 631.

[150] *Ibid.*, pp. 631-632.

(ACT): "It must be recalled that, in accordance with Article 17(1) of Directive 95/46 concerning security of processing, Member States are to provide that the controller must implement appropriate technical and organizational measures which, having regard to the state of the art and the cost of their implementation, are to ensure a level of security appropriate to the risks represented by the processing and the nature of the data to be protected."[151]

Moreover, in evaluating data security risks and the requisite level of security, it is essential to consider the potential hazards presented by personal data processing. These include risks of accidental or unlawful destruction, loss, alteration, or unauthorized disclosure or access to personal data in transit, storage, or otherwise processed.

Such risks may result in physical, material, or non-material damage,[152] particularly in instances where the processing may compromise the confidentiality of personal data afforded professional secrecy, result in the unwarranted de-anonymization of data, or otherwise lead to adverse economic or social consequences.[153]

In light of the aforementioned dangers, a critical assessment of the probability and gravity of these risks to the rights and freedoms of the data subject is essential. Due to the rapidly evolving nature of AI systems, this appraisal should be conducted not only with due consideration to the nature, extent, context, and objectives of the data processing, but also with an awareness of the network security offensive capabilities that AI systems can currently and will soon be able to perform. This evaluation and determination of the presence and degree of risk associated with data processing operations should be evidence-based.[154]

In case such a peril is identified, the safeguarding of the data subjects' fundamental rights and freedoms with respect to the processing of their personal data necessitates the implementation of suitable technical and organizational measures. The controller is duty-bound to adopt internal policies and put in place measures that align with the tenets of data protection by design and data protection by default. This encompasses, but is not limited to, the minimization of personal data processing, the pseudonymization of personal data, and the facilitation of the controller's ability to develop and enhance security features.[155]

In the context of AI deployment in spacecraft, satellites, or other space systems, for instance, it would be feasible to implement de-identification methodologies (e.g., removing specific data features or applying privacy-enhancing technologies (PETs)) on training data prior

---

[151] *Worten – Equipamentos para o Lar SA v Autoridade para as Condições de Trabalho (ACT)* [2013]. Court of Justice of the European Union (CJEU), Case C-342/12, 30 May 2013 (ECLI:EU:C:2013:355), para. 24. https://curia.europa.eu/juris/document/document.jsf;jsessionid=0782E57CCDDCC543272C244528E4C625?text=&docid=137824&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=4371952 See also *College van burgemeester en wethouders van Rotterdam v M.E.E. Rijkeboer* [2009]. CJEU, Case C-553/07, 7 May 2009 (ECLI:EU:C:2009:293), para. 62. https://curia.europa.eu/juris/document/document.jsf?text=&docid=74028&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=4372402

[152] Recital (83) and Article 32(2), *GDPR 2016*.

[153] Recital (75), *GDPR 2016*.

[154] Recital (76), *GDPR 2016*.

[155] Recital (78), *GDPR 2016*.

to its retrieval from the source and subsequent dissemination within or beyond the enterprise, contingent upon the probability and gravity of the potential impact on individuals.

Furthermore, establishing a secure pipeline from the development stage to implementation can serve to further reduce security risks. It is feasible to train an ML model utilizing a programming language and framework that are conducive to exploratory development, but then transform the model into a different, more secure format for deployment.[156]

Besides the conventional security measures, controllers should adopt specific protocols targeted at the particular vulnerabilities of AI systems (i.e., model inversion attacks, side-channel attacks,[157] prompt injection, and jailbreaks). It is recommended that controllers resort to datasets obtained from trusted sources and conduct periodic verification and validation procedures, encompassing in-house datasets as well.

Given the rapid evolution of AI and associated risks landscape, there is a compelling need for regular monitoring of and updates to the risk assessment. Similarly, the shifting modalities of attacks demand a commitment to acquiring and retaining advanced knowledge and expertise. In addressing uncertain risks, a potential strategy is to employ "red teaming" techniques to identify and expose vulnerabilities.[158]

Space missions involve unique security challenges such as exposure to cosmic radiation, potential physical attacks on satellites, and the difficulty of updating security protocols remotely. Future scenarios could involve AI systems needing to autonomously update and adapt their security measures to counteract new threats, necessitating robust and flexible security frameworks. For example, a satellite equipped with an AI system must ensure that all data transmissions are encrypted to prevent interception. Regular security assessments and updates to the AI's security protocols would be necessary to address emerging threats and ensure the continued protection of personal data.

Conclusions

The utilization of AI systems presents a potential avenue for the processing of personal data, occurring both during the development and training of an AI model and its further deployment. As has been discussed throughout this section, there are a number of prospective risks associated with the processing of personal data, depending on the context and scope of the processing.

Such risks include the use of large amounts of personal data during any development phases which are unnecessary, or which lack the corresponding legal basis for such use. This further begs the question of how to determine the necessity of training for the functioning of the

---

[156] ICO. *How should we assess security and data minimisation in AI?* https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/how-should-we-assess-security-and-data-minimisation-in-ai/

[157] B. Vigliarolo. "ChatGPT side-channel attack has easy fix: Token obfuscation," 18 March 2024, https://www.theregister.com/2024/03/18/chatgpt_sidechannel_attack_has_easy/

[158] See EDPS. *Generative AI and… op. cit.*, p. 23.

model, for instance, where additional data and training no longer improve the accuracy of the AI model.

The legitimacy of data processing may be also called into question when referring to the accuracy or retention of personal data, which is particularly relevant in situations where the outputs of AI systems are employed as part of a decision-making process. Furthermore, the potential for bias caused by inaccurate or incomplete training data must also be considered and appropriately managed.[159]

Given the natural tension between the need of AI systems to train on ultra-large datasets and the principle of data minimization under GDPR, the extensive use of AI and big data can be seen as incompatible with GDPR. Nevertheless, it can be argued that their reconciliation is feasible through the implementation of adequate security protocols. While AI is not explicitly referenced in the GDPR, many of its provisions are pertinent to the field of AI.

Indeed, some of these norms are being called into question as a result of the advent of new methods for processing personal data that are enabled by AI. There is a discrepancy between the conventional tenets of data protection –i.e., purpose limitation, data minimization, and automated decisions restrictions– and the full implementation of big data-based AI. Yet, there are methods of interpreting, applying, and developing such data protection principles that are consistent with the beneficial application of AI.[160]

The GDPR permits the development of AI-based space technologies, effectively reconciling the data protection rights of individuals with the economic interests of operators. However, it offers only limited guidance on the means of accomplishing this objective. The GDPR is replete with vague clauses and open standards, the application of which often necessitates the balancing of competing interests.[161]

Consequently, the successful implementation of the GDPR in the context of AI applications in the space industry hinges on the guidance provided by data protection authorities and other relevant bodies to controllers and data subjects, with the aim of mitigating legal uncertainty and enhancing the current regulatory landscape.

---

[159] Data Protection Commission. "AI, Large Language Models and Data Protection." *DPC Guidance Blog*. July 18, 2024. https://www.dataprotection.ie/en/dpc-guidance/blogs/AI-LLMs-and-Data-Protection.

[160] European Parliament. *The impact of… op. cit.*, p. II.

[161] *Ibid.*, p. III.

Section 2: Overview of Export Controls in Regulating the Use of AI for Space Applications

Introduction

This section provides a comprehensive examination of the complexities of export control regulations and their implications for the integration and regulation of AI technologies in space exploration.

As AI continues to revolutionize satellite operations, data analysis, and autonomous decision-making in space missions, it becomes critical to identify and create new international laws and standards for export controls that ensure the responsible development and use of AI in the domain of outer space.

This section explores how dual-use AI technologies are treated by existing international agreements, such as the Wassenaar Arrangement, and national laws, such as the US International Traffic in Arms Regulations (ITAR) and Export Administration Regulations (EAR), while identifying gaps in these frameworks that need to be bridged to accommodate the complexities of AI in outer space.

The analysis begins with export control measures, discussing the international agreements and regimes that regulate dual-use AI technologies, such as the Wassenaar Arrangement, the Missile Technology Control Regime (MTCR), the Australia Group, and the Nuclear Suppliers Group. It highlights the challenges of harmonizing control lists across jurisdictions and explores the top-down and bottom-up approaches to regulating AI used in space.

The section further explores how these international export control regimes regulate AI technologies that may be used in both civilian and military applications. For instance, AI systems that enhance satellite communication or provide autonomous spacecraft navigation pose challenges in terms of categorizing them under current export control frameworks.

With a focus on the evolving nature of AI, the section examines the difficulty in categorizing these technologies under existing export control frameworks, particularly when dealing with AI systems that have both civilian and military applications, such as AI-enabled satellites. The discussion references key regulatory frameworks such as Regulation EU 2021/821 and updates like Regulation 2023/2616, as well as the US ITAR and EAR regimes.

As noted in relevant articles, the dual-use nature of many AI systems requires countries to adopt robust licensing processes, which can be seen in the updates to the EU's Regulation 2023/2616 on dual-use technologies. In this context, the analysis examines the implications of export control regimes for both the development and deployment of AI in space systems.

Articles and documents such as the Office of Management and Budget (OMB) memorandum Advancing Governance, Innovation, and Risk Management for Agency Use of AI are discussed in relation to how procurement strategies can align with emerging AI practices. The section also elaborates on the need for comprehensive technology assessments, both in the European Union and the United States, to ensure that AI components used in space missions meet dual-use regulatory requirements, referencing relevant EU and US export control guidelines and exceptions.

This section draws on a range of articles and documents, including key export control regulations, government memoranda, etc., to provide a detailed understanding of how export control regulations intersect with AI technologies in space exploration.

Through a meticulous review of relevant international agreements and regulations, this section underscores the importance of creating a harmonized legal framework that not only facilitates the advancement of AI in space but also ensures its responsible and secure deployment.

1. *Frameworks for Harmonizing Export Controls*

1.1 *Harmonization of Export Controls Lists*

The fast development and implementation of new technologies, such as AI systems, and their potential use for military purposes urge regulators to timely assess their capabilities and restrict their proliferation for security reasons. Controls are implemented from the moment a relevant product or technology described in the control list is created, including the designs of such product or technology, as well as prototypes, and commercial embodiments.

International export control law is composed of key multilateral export control regimes, each targeting a different type of technology and focuses on their controls. These are the Australia Group,[162] the Wassenaar Arrangement,[163] the Nuclear Suppliers Group,[164] and the Missile Technology Control Regime.[165]

The principal objective of Australia Group participants is to control the export of chemicals, biological agents, and related items so that they do not contribute to the spread of

---

[162] The 43 participating states in the Australia Group are Argentina, Australia, Austria, Belgium, Bulgaria, Canada, Chile, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Italy, Japan, Latvia, Lithuania, Luxembourg, Malaysia, Malta, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Romania, Russia, Slovakia, Slovenia, South Africa, South Korea, Spain, Sweden, Switzerland, Turkey, Ukraine, and the United States.
Australia Group. "Member States." Accessed August 31, 2024. https://www.australiagroup.net/en/members.html.

[163] The 42 participating states in the Wassenaar Arrangement are Argentina, Australia, Austria, Belgium, Bulgaria, Canada, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, India, Ireland, Italy, Japan, Latvia, Lithuania, Luxembourg, Malta, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Republic of Korea, Romania, Russian Federation, Slovakia, Slovenia, South Africa, Spain, Sweden, Switzerland, Türkiye, Ukraine, United Kingdom and United States.
Wassenaar Arrangement. "About Us." Accessed August 31, 2024. https://www.wassenaar.org/about-us/.

[164] The 48 participating states in the Nuclear Suppliers Group are Argentina, Australia, Austria, Belgium, Brazil, Bulgaria, Canada, Chile, China, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, India, Ireland, Israel, Italy, Japan, Kazakhstan, Latvia, Lithuania, Luxembourg, Malaysia, Malta, Mexico, Netherlands, New Zealand, Norway, Pakistan, Poland, Portugal, Romania, Russia, Slovakia, Slovenia, South Africa, South Korea, Spain, Sweden, Switzerland, Turkey, Ukraine, the United Arab Emirates, and the United States.
Nuclear Suppliers Group. "Participating Governments." Accessed August 31, 2024.
https://www.nuclearsuppliersgroup.org/en/participating-governments.

[165] The 37 participating states in the Missile Technology Control Regime are Argentina, Australia, Austria, Belgium, Brazil, Canada, Chile, China, Cyprus, Czech Republic, Denmark, Egypt, Estonia, Finland, France, Germany, Greece, Hungary, India, Israel, Italy, Japan, Kazakhstan, Latvia, Lithuania, Luxembourg, Malaysia, Malta, Mexico, Netherlands, New Zealand, Norway, Pakistan, Poland, Portugal, Romania, Russia, Saudi Arabia, Slovakia, Slovenia, South Africa, South Korea, Spain, Sweden, Switzerland, Turkey, Ukraine, the United Arab Emirates, and the United States.
Missile Technology Control Regime. "Members." Accessed August 31, 2024. https://www.mctp.org/about/members.

chemical and biological weapons. To achieve this, the Group seeks to harmonize participating countries' national export licensing measures pertaining to these products and technologies.

The Wassenaar Arrangement controls the transfer of conventional arms and dual-use goods and technologies. It aims to "complement and reinforce, without duplication, the existing control regimes for weapons of mass destruction and their delivery systems."[166]

The Nuclear Suppliers Group contributes to the non-proliferation of nuclear weapons through the implementation of two sets of guidelines for nuclear exports. The Missile Technology Control Regime coordinates national export licensing efforts aimed at preventing the proliferation of unmanned delivery systems capable of delivering weapons of mass destruction.

These regimes, especially the Wassenaar Arrangement and the Missile Technology Control Regime, play a crucial role in regulating AI technologies used in space, as these present unique challenges due to their military potential, such as integration into unmanned delivery systems capable of delivering weapons or into military reconnaissance satellites.

1.1.1 Wassenaar Arrangement and Related International Agreements

Succeeding the Cold War-era Coordinating Committee for Multilateral Export Controls (COCOM), the Wassenaar Arrangement (WA) was created in 1996 to encourage non-proliferation of advanced weaponry and dual-use tech.[167] It provides a voluntary multilateral export control regime for member states to mutually share information regarding exports of sensitive conventional weapons and dual-use goods and technologies.[168]

WA member states "must implement national export control laws that prohibit the export and sale of arms or sensitive dual-use goods to [countries] of concern."[169] Further, members must act "in accordance with international non-proliferation norms and standards, like the Nuclear Non-Proliferation Treaty (NPT), the Missile Technology Control Regime (MTCR), the Chemical Weapons Convention (CWC), and the UN Register of Conventional Arms. WA members are also expected to maintain export controls based on the control lists of the WA: the Munitions List and the Dual-Use List [,]" which member state experts regularly review and update.[170]

To foster transparency, responsibility and accountability, the current 42 WA member states share periodic reports on licenses, transfers, and denials to non-members of technologies covered by the Wassenaar Dual-Use List (and also weapons covered by the UN Register of Conventional

---

[166] Wassenaar Arrangement. Export Controls for Conventional Arms and Dual-Use Goods and Technologies: Public Documents Volume I, Founding Documents. Compiled by the Wassenaar Arrangement Secretariat. WA-DOC (19) PUB 007. December 2019.
https://www.wassenaar.org/app/uploads/2021/12/Public-Docs-Vol-I-Founding-Documents.pdf.

[167] Wassenar Arrangement on Export Controls for Conventional Arms and Dual-Use Goods and Technologies, Founding Documents, Public Documents, Vol. 1, Wassenaar Arrangement Secretariat, December 2019 [hereinafter WA], https://www.wassenaar.org/app/uploads/2021/12/Public-Docs-Vol-I-Founding-Documents.pdf; Farah Sonde, "Fact Sheet: The Wassenaar Arrangement," Center for Arms Control and Non-Proliferation, March 13, 2023, https://armscontrolcenter.org/fact-sheet-the-wassenaar-arrangement.

[168] WA; Sonde, "Fact Sheet: The Wassenaar Arrangement."

[169] *Ibid.*

[170] *Ibid.*

Arms).[171] Transparency among the members enables them to better avoid providing destabilizing influxes of arms and dual-use goods and technologies to other states, especially states of particular concern like Iran, Iraq, Syria, Libya, and North Korea.[172]

The Wassenaar Dual-Use List includes a Basic List of controlled technologies, further including a subset Sensitive List, with its own Very Sensitive List subset.[173] Semiannual reports are requested of members for license denials under the Basic List and on licenses issued and transfers conducted for Sensitive List technologies.[174]

Reports for license denials of Sensitive List technologies are requested within 30-60 days of such denial. Members are requested to report any export licenses denied for proposed transfers to non-Wassenaar members within 60 days.[175]

Unlike COCOM, the WA has no undercut rule. Correspondingly, member states may independently deny or approve licenses and transfers entirely at their own discretion, and if one member denies a transfer another member may approve it.[176] The WA neither includes any kind of veto, nor does it provide for any prior reviews of proposed exports.[177]

However, although WA member states are not obligated to deny transfers previously denied by other members, they are requested to report any approvals of export licenses for transfers "essentially identical" to a transfer denied by another WA member within the prior three years.[178] Wassenaar Very Sensitive List technologies, such as stealth technology materials and advanced radar, are expected to be licensed for transfer only with the utmost discretion and vigilance.[179]

Russian membership in the WA has proven controversial along these lines, for instance.[180] Although it provides other members a level of insight into Russian export activities, their membership has failed to produce the hoped-for reductions in Russian approvals for exports to unstable regions, they have obstructed updates to control lists, and they have used the information exchanges as intelligence gathering opportunities.[181] Israel, China and Belarus, three major arms exporters, are notable non-members to the WA.[182]

Similar to export control laws enacted by the US and EU to fulfill their respective WA obligations, WA export controls on dual-use items relevant to AI and space applications of AI primarily focus on listing items comprising the advanced semiconductor processors and hardware, as well as the equipment for designing and manufacturing these advanced chips and computers.

---

[171] *Ibid.*

[172] *Ibid.*

[173] *Ibid.*

[174] *Ibid.*

[175] *Ibid.*

[176] *Ibid.*

[177] *Ibid.*

[178] *Ibid.*

[179] *Ibid.*

[180] Sonde, "Fact Sheet: The Wassenaar Arrangement."

[181] *Ibid.*

[182] *Ibid.*

To some extent, the software side is also regulated, exemplified in existing controls on encryption software and more recent regulations (e.g., in the U.S.) regarding geospatial AI.

Additionally, maintaining up-to-date control lists is especially challenging given such rapid advancements in both AI and space tech, demonstrated by the recent EU decision to accelerate their next export control list review to Q1 of 2025 from the originally scheduled timeframe of 2026-2028. WA also specifically includes autonomous submersibles and UAVs.

National dual-use export control regimes, such as US EAR, and supra-national regimes, such as the EU Dual-Use Export Control Regulations implement the various respective international agreements. "Both include control lists that are organized by categories and individual control entries that generally correspond when present on both lists. For example, the top US Export Control Classification Number (ECCN) by shipment count to the EU is 5A002 for information security commodities. The EU list contains an equivalent entry under the same 5A002 designation."[183]

ECCNs are five-character alphanumeric codes designating various dual-use item categories.[184] The first character of an ECCN identifies one of the ten CCL categories:

Category 0: Nuclear Materials, Facilities, and Equipment (and Misc. Items)
Category 1: Special Materials and Related Equipment, Chemicals, "Microorganisms," and "Toxins"
Category 2: Materials Processing
Category 3: Electronics
Category 4: Computers
Category 5: Telecommunications and "Information Security"
Category 6: Sensors and Lasers
Category 7: Navigation and Avionics
Category 8: Marine
Category 9: Aerospace and Propulsion[185]

The second ECCN character identifies the Product Group within the primary Category:

A. End items, equipment, accessories, attachments, parts, components, and systems
B. Test, inspection, and production equipment
C. Materials

---

[183] Kimberly Munch, "A Comparison Between U.S. Export Controls and EU Export Controls," Export Compliance Training Institute, January 9, 2024, https://www.learnexportcompliance.com/a-comparison-between-u-s-export-controls-and-european-export-controls/.

[184] "Introduction to Commerce Department Export Controls"; "Export Control Classification Number (ECCN)"; "Technology Classification."

[185] EU Regulation 2021/821; EU Recommendation 2021/1700; "Dual-Use Export Controls, Summaries of EU Legislation"; "Exporting Dual-use Items"; Jasper, "Explained: EU Export Control Lists;" "Introduction to Commerce Department Export Controls," Office of Exporter Services, Bureau of Industry and Security, US Department of Commerce, Revised November 2018, https://www.bis.doc.gov/index.php/documents/regulations-docs/142-eccn-pdf/file; "Export Control Classification Number (ECCN)," Bureau of Industry and Security, US Department of Commerce, 2024, https://www.bis.doc.gov/index.php/licensing/commerce-control-list-classification/export-control-classification-number-eccn; "Technology Classification."

D. Software
E. Technology[186]

The final three characters of an ECCN identify the "Reason for Control":

001 - 099: Wassenaar Arrangement (dual use)
101 - 199: Missile Technology Control Regime (dual use)
201 - 299: Nuclear Suppliers Group (dual use)
301 - 399: Australian Group
401 - 499: Chemical Weapons Convention
501 - 899: (reserved)
901 - 999: National controls[187]

Additional designation identifiers include: (1) Anti-Terrorism (AT), (2) Chemical & Biological Weapons (CB), (3) Chemical Weapons Convention (CW), (4) Crime Control (CC), (5) Encryption Items (EI), (6) Firearms Convention (FC), (7) Missile Technology (MT), (8) National Security (NS), (9) Nuclear Nonproliferation (NP), (10) Regional Stability (RS), (11) Short Supply (SS), (12) Significant Items (SI), (13) Surreptitious Listening (SL) and (14) United Nations sanctions (UN).[188]

"Other dual-use items, including any associated brokering services or technical assistance, need export authorization if they are intended, entirely or in part, for:

- chemical, biological or nuclear weapons;
- military use in countries subject to an arms embargo;
- components of military items already exported from an EU Member State without the necessary authorization.

Authorization is [further] required for:

- the export of cybersurveillance items likely to be used for internal repression or serious violations of human rights and international humanitarian law;
- the transfer of dual-use items listed in Annex IV, such as stealth technology and strategic control, from one Member State to another.[189]nuclear materials, facilities and equipment."[190]

As AI technologies continue to rapidly evolve, they pose increasingly unique challenges to existing export control frameworks. The dynamic nature of AI necessitates a more agile and coordinated approach to international export controls to address these challenges effectively.

One of the biggest challenges in the international realm is agreeing on which items to include on the control list. Currently, there are two approaches to tackling this issue: top-down, where national regulators integrate technologies selected internationally through targeted export

---

[186] *Ibid.*

[187] *Ibid.*

[188] *Introduction to U.S. Export Controls for the Commercial Space Industry*; "Technology Classification."

[189] EU Regulation 2021/821; EU Recommendation 2021/1700; "Dual-Use Export Controls, Summaries of EU Legislation"; "Exporting Dual-use Items"; Jasper, "Explained: EU Export Control Lists".

[190] EU Regulation 2021/821, Art. 4; EU Recommendation 2021/1700; *See quote in* "Dual-Use Export Controls, Summaries of EU Legislation"; "Exporting Dual-use Items"; Jasper, "Explained: EU Export Control Lists".

control groups; or bottom-up, where international groups include items controlled at the national level. Each of these approaches has its conveniences and inconveniences, analyzed below:

1.1.2 Top-Down Approach to Harmonization

In the top-down approach, export control lists are first agreed upon at the international level, then subsequently adopted, fully or partially, by individual states. This process appears to be the most efficient in terms of harmonization; however, it also presents significant challenges.

The first of such challenges pertains to the identification and selection of products and technologies to be controlled. As per the Wassenaar Arrangement, for an item to be included in the Dual-Use control list, the following criteria are considered: 1) foreign availability outside Participating States; 2) the ability to control the export of the goods effectively; 3) the ability to make a clear and objective specification of the item; and 4) whether it is controlled by another regime.[191]

Additionally, the Wassenaar Arrangement has created special Sensitive and Very Sensitive lists of dual-use items. Sensitive items include those from the Dual-Use List that are "key elements directly related to the indigenous development, production, use, or enhancement of advanced conventional military capabilities, the proliferation of which would significantly undermine the objectives of the Wassenaar Arrangement."[192] The list is not intended to include general commercially applied materials or components.

Moreover, the Wassenaar Arrangement specifies that the relevant threshold parameters regarding this list should be developed on a case-by-case basis. Very Sensitive items are further defined as "key elements essential for the indigenous development, production, use, or enhancement of advanced conventional military capabilities, the proliferation of which would significantly undermine the objectives of the Wassenaar Arrangement."[193] They are also assessed on a case-by-case basis.

Given the highly technical nature of AI systems and the need for specialized international experts, combined with the rapidly evolving landscape of AI technologies, conducting case-by-case assessments and performance benchmarking becomes exceptionally challenging.[194]

One can imagine a hypothetical scenario where a new AI algorithm designed for satellite collision avoidance is developed. While this technology is intended for peaceful space operations, it could also be adapted for military satellite maneuvers, e.g., to intercept a satellite rather than avoid it. The complexity of its dual-use nature and the divergent technical views within the Technical Expert Group (WA-EG) can make it difficult to reach an international consensus on the inclusion of AI systems in control lists.

---

[191] Wassenaar Arrangement. Criteria for the Selection of Dual-Use Goods, including Sensitive and Very Sensitive Items. Accessed August 31, 2024. https://www.wassenaar.org/control-lists/.

[192] *Ibid.*

[193] *Ibid.*

[194] Ritwik Gupta, and Andrew W. Reddie, "Accelerating the Evolution of AI Export Controls," September 21, 2023, https://www.techpolicy.press/accelerating-the-evolution-of-ai-export-controls/.

The second major challenge is the time required to make amendments to the lists. Coming back to the Wassenaar Arrangement example, its main decision-making body is the Wassenaar Arrangement Plenary, composed of the participating member states. The Plenary usually meets once, at the end of the year, and takes decisions by consensus.

Given this strict voting process, the decisions are often getting blocked. The overall process of reviewing the lists is protracted, as multiple consultations must take place. Consequently, these complications have led to a growing shift from a top-down to a bottom-up approach.

1.1.3 Bottom-Up Approach to Harmonization

Individual states are more flexible and reactive when it comes to updating their own national lists, especially regarding emerging technologies. For example, in March 2024, the UK expanded its export control regime to include quantum computing, semiconductor technologies, and additive manufacturing.[195] Similar developments have recently occurred in France,[196] Italy, the Netherlands,[197] and Spain.[198]

While this flexibility allows for rapid adaptation to emerging technologies, it complicates international harmonization, particularly for AI systems used in space where consistent regulations are critical for global cooperation and security. It is important to consider the rapid advancements in AI-driven satellite imaging technologies.

While individual countries may quickly update their export controls to include these technologies, the lack of international harmonization can lead to discrepancies and potential conflicts in their use and transfer across borders. This can already be vividly seen at the EU level.

As indicated in the EU White Paper on Export Controls,[199] several shortcomings come into play: lack of transparency and insufficient consultation; uncertainty regarding the implementation of controls; and constraints dictated by existing national laws, to name a few. These issues are particularly challenging for AI technologies used in space, where clear and consistent regulations are essential for safe and effective operations.

To mitigate this challenge, the EU proposed an immediate solution: expanding Annex I of the EU's Dual-Use Regulation to include items not adopted by the multilateral export control regimes. The Dual-Use Regulation is scheduled for evaluation between 2026 and 2028. However,

---

[195] UK Government. "UK Expands Export Controls to Include Quantum Computing and Other Advanced Technologies." GOV.UK. March 2024.
https://www.gov.uk/government/news/uk-expands-export-controls-to-include-quantum-computing.

[196] France. Arrêté du 2 février 2024 relatif aux exportations vers les pays tiers de biens et technologies associés à l'ordinateur quantique et à ses technologies habilitantes et d'équipements de conception, développement, production, test et inspection de composants électroniques avancés. NOR: ECOI2401482A. JORF no. 0034, February 10, 2024, text no. 8. https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000049120866.

[197] Netherlands. Minister for Foreign Trade and Development Cooperation. Letter to the Dutch Parliament, 8 March 2023. https://open.overheid.nl/documenten/oep-7b25ba07017fbcc4a9d285cc013849f6516f03bd/pdf.

[198] Spain. Order ICT/534/2023 of 26 May 2023. Boletín Oficial del Estado, no. 12785 (May 31, 2023).
https://www.boe.es/boe/dias/2023/05/31/pdfs/BOE-A-2023-12785.pdf.

[199] *White Paper on Export Controls*. COM(2024) 25 final. Brussels, January 24, 2024. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=COM:2024:25:FIN.

due to the rapid pace of developments in the last three years, the Commission has concluded that the evaluation should be advanced to the first quarter of 2025.

The second significant challenge is to increase the number of countries participating in the multilateral regimes. Although the Wassenaar Arrangement is effective, it only involves 42 countries, leaving a considerable number of actors unaddressed.

For a country to join the Arrangement, it must meet several criteria, including the capacity to handle sensitive goods and adhere to specific principles. These include being a producer or exporter of arms or dual-use goods, maintaining national non-proliferation policies, upholding commitments to international non-proliferation treaties, and having a fully functional export control system in place.

Furthermore, adding new members is complicated because it requires unanimous agreement from all current participants.[200] Furthermore, adding new members is complicated because it requires unanimous agreement from all current participants.

For AI technologies used in space, broader participation is crucial to ensure global security and cooperation. Strategies to achieve this could include creating incentives for new members to join, simplifying the admission process, and enhancing collaboration with non-member states on AI technology standards and security protocols.

One can also envision the establishment of an international consortium focused on AI in space, promoting broader participation and facilitating the exchange of best practices and technologies among member and non-member states alike.

To conclude, harmonization is more likely to succeed through a bottom-up approach, starting with individual countries rather than the other way around. To facilitate this process, a separate body is needed to integrate individual developments.

A great example of this is the Australia Group. Although specifically dedicated to biological and chemical weapons, the group has the New and Evolving Technologies Technical Experts Meeting (NETTEM), tasked with identifying emerging technological trends that may have proliferation impacts.

Adopting a similar model for AI export controls, particularly for space applications, could provide the necessary structure and coordination to address the unique challenges posed by emerging AI technologies. This body could focus on developing standardized protocols for AI in space, fostering international cooperation, and ensuring that export controls keep pace with technological advancements.

---

[200] Wassenaar Arrangement, Guidelines for Applicant Countries (Agreed at the 2014 Plenary), accessed September 26, 2024, https://www.wassenaar.org/app/uploads/2019/consolidated/11Guidelines-for-Applicant-Countries.pdf.

2. *Export Control Regulations and Technology Assessments*

2.1 *European Union Export Control Regulations and Technology Assessments*

Under EU dual-use export control regulations, companies that export services, products, and any components or materials used therein must comply with licensing requirements for any dual-use products or services that the business intends to export.[201]

These regulations are particularly relevant for AI technologies used in space, as both space and AI systems often involve components and software with potential dual-use capabilities. On the European Union level, AI-related technologies, as well as space-related items, are controlled by the EU Dual-Use Regulation.[202]

EU Regulation 2021/821 governs the EU regime for regulating exports, brokering, technical assistance, and transfers of dual items under several dual-use item lists;[203] Regulation 2023/2616 (September 2023) amended the list of dual-use items provided under Regulation 2021/821.[204]

Annex I to EU 2021/821 provides lists of the dual-use items that require export license authorizations.[205] Further, Annex IV of the EU Dual-Use Regulation presents a subset listing of all the dual-use items from Annex I that are so highly sensitive that they require licensing for transits and transfers within the EU.[206]

---

[201] EU Regulation 2021/821, “Dual-Use Export Controls Regulation,” Official Journal L 206 (2021), 1-461 [hereinafter EU Regulation 2021/821], https://eur-lex.europa.eu/eli/reg/2021/821/; EU Recommendation 2021/1700, “Internal Compliance Programmes for Controls of Research Involving Dual-Use Items Under Regulation (EU) 2021/821,” Official Journal L 338 (2021), 1-52 [hereinafter, EU Recommendation 2021/1700], https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2021.338.01.0001.01.ENG&toc=OJ%3AL%3A2021%3A338%3AFULL; “Dual-Use Export Controls, Summaries of EU Legislation,” Eur-Lex, November 17, 2023, https://eur-lex.europa.eu/EN/legal-content/summary/dual-use-export-controls.html; “Exporting Dual-use Items,” European Commission Trade Department, February 23, 2024, https://policy.trade.ec.europa.eu/help-exporters-and-importers/exporting-dual-use-items_en; Ulrike Jasper, “Explained: EU Export Control Lists and Dual-use Goods Classification,” AEB SE, January 2, 2023, https://www.aeb.com/en/magazine/articles/export-control-lists-classification-dual-use.php#classification.

[202] Regulation (EU) 2021/821 of the European Parliament and of the Council of 20 May 2021 setting up a Union regime for the control of exports, brokering, technical assistance, transit and transfer of dual-use items (OJ L 206, 11.6.2021, p. 1–461)

[203] EU Regulation 2021/821; EU Recommendation 2021/1700; “Dual-Use Export Controls, Summaries of EU Legislation”; “Exporting Dual-use Items”; Jasper, “Explained: EU Export Control Lists.”

[204] EU Regulation 2023/2616, “Amending Regulation (EU) 2021/821 as regards the list of dual-use items,” Official Journal L series (2023), 1-244, https://eur-lex.europa.eu/eli/reg_del/2023/2616.

[205] EU Regulation 2021/821; EU Recommendation 2021/1700; “Dual-Use Export Controls, Summaries of EU Legislation”; “Exporting Dual-use Items”; Jasper, “Explained: EU Export Control Lists.”

[206] *Ibid.* Annex IV specifies that the following items of the MTCR technology are not controlled: (1) items transferred on the basis of orders pursuant to a contractual relationship placed by the European Space Agency (ESA) or transferred by ESA to accomplish its official tasks; (2) items transferred on the basis of orders pursuant to a contractual relationship placed by a Member State's national space organization or transferred by it to accomplish its official tasks; (3) items transferred on the basis of orders pursuant to a contractual relationship placed in connection with a Community space launch development and production program signed by two or more European governments; and (4) items transferred to a State-controlled space launching site in the territory of a Member State, unless that Member State controls such transfers within the terms of this Regulation.

Aside from the EU export control lists, national dual-use lists may also apply to respective sensitive services, products, components, and materials, depending on the country of origin.[207] This makes export control compliance even more complicated in cases of intra-EU international technology manufacturing.

In addition, all EU member states have their own national export control laws and lists for military services, products, components, and materials.[208] Consequently, all EU companies must comply with at least two sets of export control laws and dual-use classification lists for the services, products, components, and materials they export.[209]

For example, a company seeking to export AI-driven satellite navigation systems must conduct a thorough technical review to ensure that the system and each component complies with dual-use regulations. If different components are sourced from various countries within the EU, the national export controls of the respective member state(s) will also apply.

To conduct a compliant dual-use goods classification the company exporting the technology must obtain the classification (e.g., from the manufacturer) or perform its own detailed technical review of the specific item.[210] Performing this technical review involves mapping elements and components of the item in question against the relevant export control classification lists to determine whether it may require an export license.[211]

Conducting a technology assessment under the EU dual-use regulations starts classifying the "product master," or "material master" under the Annex I dual-use technology descriptions of the EU Dual-Use Regulations.[212] Classifying the "product master" or "material master" requires identifying and classifying all services, products, components, and materials the company makes or procures and sells, stores, services, maintains, or which the company otherwise intends to or may share with a national of a non-EU member State.[213]

For AI technologies used in space, this process is particularly challenging due to the complex and rapidly evolving nature of these technologies, which often blur the lines between civilian and military applications. Satellite remote sensing and imaging systems designed for environmental monitoring are often also potentially useful in military reconnaissance contexts, and this dual-use potential complicates classification under export control regulations.

Establishing and maintaining reliable export control procedures under a compliant export control program requires conducting a thorough and detailed technical review of all dual-use item classification listings in Annex I and Annex IV and executing efficient and effective checks against all company services, products, components, and materials.[214]

---

[207] *Ibid.*

[208] *Ibid.*

[209] *Ibid.*

[210] Jasper, "Explained: EU Export Control Lists."

[211] *Ibid.*

[212] *Ibid.*

[213] *Ibid.*

[214] EU Regulation 2021/821; EU Recommendation 2021/1700; "Dual-Use Export Controls, Summaries of EU Legislation"; "Exporting Dual-use Items"; Jasper, "Explained: EU Export Control Lists."

Services, technologies, hardware, products, components, or materials comprise dual-use items if they include all the relevant technical elements, components, or features listed in an Annex I dual-use item classification description.[215] Further, products must be evaluated for their dual-use potential independent of any known or intended end-use and recipient individual or entity.[216]

Furthermore, simply because an item becomes incorporated into military technology or used for military purposes does not in and of itself result in a dual-use classification; only items described in the Annex I dual-use classification list are controlled.[217] Especially in view of the potential penalties, appropriate classification of items requires a deep knowledge of and familiarity with the technical specification of the particular product in question—typically at the level of a manufacturer, inventor, or creator of the tech.[218]

Conducting dual-use item classifications and performing the required comparison of product technical details with pertinent export control lists can implicate significant burdens for many companies in the way of administrative overhead on time and effort of company staff, expenses, and other challenges, especially for those developing AI technologies for space applications.

These challenges include keeping up with rapidly evolving technologies and regulations, managing complex supply chains, and finding and working with appropriate experts to help ensure that all products and components meet stringent export control requirements.[219]

Companies with large or frequently changing product lines, those exporting replacement parts, and startups that typically lack large amounts of administrative bandwidth to support export control reviews may be particularly affected.[220] Both domains of AI and space tend to be very capital intensive and space ratings for products (e.g., satellites and components) typically demand long development timeframes. Consequently, startups developing AI-driven space products can be profoundly impacted.

For instance, dual-use items integrated within a larger piece of equipment or machinery are typically subsumed under the component provision.[221] Consequently, so long as the component does not exceed a threshold amount of contribution to the monetary value or overall product proportion, many types of machines and equipment may be exported without a license.[222]

However, of utmost importance to exporters of replacement part components, the component provision does not apply when exporting replacement parts, e.g., software updates and modules, filters, pumps, valves, detectors, sensors, printed circuit boards, or frequency changers.[223] Therefore, companies exporting replacement parts must conduct dual-use classification for all such

---

[215] *Ibid.*

[216] *Ibid.*

[217] *Ibid.*

[218] Jasper, "Explained: EU Export Control Lists."

[219] *Ibid.*

[220] *Ibid.*

[221] EU Regulation 2021/821; EU Recommendation 2021/1700; "Dual-Use Export Controls, Summaries of EU Legislation"; "Exporting Dual-use Items"; Jasper, "Explained: EU Export Control Lists."

[222] *Ibid.*

[223] *Ibid.*

replacement parts independent of any underlying product classification, and all dual-use replacement part components must obtain export licenses under EU Dual-Use Regulation Art. 3.[224]

No universal, "one-size fits all" approach to dual-use classification exists, and each company must build their own export control procedures optimized to their specific business considerations, e.g., corporate structure, product development, component suppliers, service providers, and export sales operations.[225] Dual-use items are nearly always high-tech and Annex I generally does not list any mass-produced products readily obtainable in major global markets.[226]

Classification of products, components, and materials must often be requested from the original manufacturer or supplier, as the complexity of technical details included in the Annex I dual-use item listings requires a level of familiarity commensurate with having designed and developed the product.[227] For companies that integrate materials or components sourced from other manufacturers into downstream products, the ability to rely on the attestation of an unrelated manufacturer or supplier is also advantageous from a legal risk perspective.

Exporters may rely on manufacturer information provided regarding export control classification, but only if the information is plausible and cannot be disproved through simple means.[228] If any information provided by the manufacturer is easily recognized or identifiable as implausible using simple means, the exporter may either work with the manufacturer to correct any errors resulting in the implausibility or conduct an independent reclassification.[229]

Potentially of more fundamental importance than any specific policies, procedures, and/or guidelines developed or selected by an EU company engaged in export operations, are the consistent and clear documentation of all products technical assessments, diligently keeping and securing all records, and the meticulous maintenance and periodic updating of all master data.[230]

The list of dual-use items contained in Annex I implements internationally agreed dual-use controls including the Australia Group (4), the Missile Technology Control Regime (MTCR) (5), the Nuclear Suppliers Group (NSG) (6), the Wassenaar Arrangement (7) and the Chemical Weapons Convention (CWC) (8).[231]

Irrespective that certain items may not meet the technical specifications of any items described in Annex I, they may nevertheless be sensitive due to their technical capabilities or a suspected end-use of concern.[232]

Under Article 4 of the Dual-Use Regulation national authorities can impose an authorization requirement for dual-use items not listed in Annex I if there is a (suspected) use in a

---

[224] EU Regulation 2021/821; EU Recommendation 2021/1700; "Dual-Use Export Controls, Summaries of EU Legislation"; "Exporting Dual-use Items"; Jasper, "Explained: EU Export Control Lists."

[225] Jasper, "Explained: EU Export Control Lists."

[226] *Ibid.*

[227] *Ibid.*

[228] *Ibid.*

[229] *Ibid.*

[230] *Ibid.*

[231] EU Regulation 2021/821; EU Recommendation 2021/1700; "Dual-Use Export Controls, Summaries of EU Legislation"; "Exporting Dual-use Items"; Jasper, "Explained: EU Export Control Lists".

[232] *Ibid.*

weapons of mass destruction (WMD) program, (suspected) military end-use in a country subject to an arms embargo, or (suspected) use of the item as a component in military equipment that has been exported without or in violation of an authorization.[233] This provision is known as the "catch-all control" and in the case where the company suspects an end use of company products under any of the above-mentioned circumstances, the company should contact their national authority for further information.[234]

In addition to the list of dual-use items present in Annex I, under Article 9 of the Regulation EU Member States may also prohibit or impose authorization requirements for the export of dual-use items not listed in Annex I for reasons of public security, including the prevention of acts of terrorism, or human rights considerations.[235] A list of such national measures is compiled by the European Commission (EC), published in their Official Journal, and also made available on the EC website.[236]

This right was recently exercised by several countries, such as the Netherlands, Spain, and France, to control various semiconductors and quantum equipment and technologies not included in the Annex. These updates are particularly relevant given the strategic importance of such technologies in the fields of AI and space, where quantum computing and semiconductors underpin many mission-critical systems.

The ramifications of these regulatory shifts are far-reaching, especially for companies engaged in AI-driven space technologies, as they must navigate increasingly complex compliance landscapes.

On 7 June 2023, the Spanish Ministry of Industry, Commerce and Tourism amended Annex III.5[237] to Royal Decree 679/2014.[238] Article 2(3)(6º) of Royal Decree 679/2014 establishes the obligation of requiring authorization in the definitive and temporary exports of dual-use items and technologies included in Annex III.5, which encompasses the dual-use items subject to export control, by virtue of Article 9 the EU Dual-Use Regulation, yet not included in the Annex I to the Regulation.

These are dual-use items and technologies identified by the European Commission and the Council of the EU as being of particular relevance from the point of view of non-proliferation criteria, which are not included in the multilateral non-proliferation regimes. They are adopted by virtue of the aforementioned article through national measures taken by the Member States.

It is therefore that the proposed amendment seeks to include such dual-use items and technologies. Accordingly, the following items, among others, were added:

---

[233] *Ibid.*

[234] *Ibid.*

[235] *Ibid.*

[236] EU Regulation 2021/821; EU Recommendation 2021/1700; "Dual-Use Export Controls, Summaries of EU Legislation"; "Exporting Dual-use Items"; Jasper, "Explained: EU Export Control Lists".

[237] By means of Order ICT/534/2023, of May 26th, amending Annexes I.1, III.2 and III.5 of the Regulation on the control of foreign trade in defense material, other material and dual-use products and technologies, approved by Royal Decree 679/2014, of August 1st.

[238] Royal Decree 679/2014, of August 1st, approving the Regulation on the control of foreign trade in defense material, other material and dual-use products and technologies.

- Surveillance systems, equipment, and components for public information and communication networks, not specified in the EU Dual Use Regulation.
- Computer programs not specified in the EU Dual Use Regulation.
- Quantum computers and related electronic assemblies and their components.

A few weeks later, on 23 June 2023, the Netherlands issued a similar legal text – a Regulation on Advanced Production Equipment for Semiconductors.[239] The Regulation mentions the list of goods and technologies requiring a license. Among the listed products are lithographic equipment, equipment for atomic layer deposition (ALD) of "work function" metals, and equipment designed for epitaxial growth of silicon (Si), carbon-doped silicon, silicon germanium (SiGe), or carbon-doped SiGe.

According to the Dutch Government, these measures were adopted in recognition of the crucial contribution of these items to certain advanced military applications as well as their potential use in the development of high-quality military (weapon) systems and weapons of mass destruction.[240]

France has also followed up with such developments, and on 2 February 2024, has adopted an Order on the export of goods and technologies related to quantum computers and quantum technologies, as well as advanced technology equipment.[241] The list covers, similarly to Spain and the Netherlands, complementary Metal Oxide Semiconductor (CMOS) integrated circuits, equipment designed for dry etching, quantum computers, "electronic assemblies," and components designed for them.

In conclusion, the recent actions by Spain, the Netherlands, and France to tighten controls on semiconductors and quantum technologies reflect a broader trend of regulatory adaptation to the dual-use nature of emerging technologies. As AI continues to play an increasingly central role in space exploration, it is critical for stakeholders to remain vigilant and well-informed regarding export control developments.

By fostering a culture of compliance and engagement with regulatory authorities, companies can not only navigate these challenges but also position themselves at the forefront of innovation in the AI and space sectors.

2.2 *United States Export Control Regulations and Technology Assessments*

Comparable to EU export control technology assessments, proper identification of relevant US export control regulations and appropriate classification of dual-use items requires deep familiarity with the technical details of services, products, components, and materials design and

---

[239] Minister of Foreign Trade and Development Cooperation. Ministerial Decree on Advanced Semiconductor Manufacturing Equipment. Staatscourant 2023, No. 18212. June 23, 2023. https://zoek.officielebekendmakingen.nl/stcrt-2023-18212.html.

[240] *Ibid.*

[241] France. Arrêté du 2 février 2024 relatif aux exportations vers les pays tiers de biens et technologies associés à l'ordinateur quantique et à ses technologies habilitantes et d'équipements de conception, développement, production, test et inspection de composants électroniques avancés. NOR: ECOI2401482A. JORF no. 0034, February 10, 2024, text no. 8. https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000049120866.

manufacture. Given this information can often be proprietary, exporters may correspondingly need to obtain export classifications from manufacturers, vendors, or inventors in such circumstances.

Where manufacturers, vendors, or inventors are unable or unwilling to provide an export control classification or the information required to appropriately evaluate and classify the technology, the exporter may either seek an official commodity jurisdiction determination from the Directorate of Defense Trade Controls (DDTC) or conduct an independent self-classification.[242] If self-classification attempts prove unsuccessful, the exporter must pursue a commodity jurisdiction request with DDTC to determine potential export restrictions or requirements.[243]

For instance, classifying an AI-driven satellite imaging system under US regulations requires detailed technical knowledge. Challenges include the difficulty of identifying all relevant technical specifications for compliance, understanding the nuanced differences between similar technologies, and ensuring that updates to existing AI systems do not inadvertently violate export control laws.

Potential solutions include establishing a dedicated compliance team within the company or to work with an external expert to regularly liaise with export control authorities and industry experts, stay abreast of regulatory changes, and ensure that innovation continues without legal interruptions.

In the US, three primary regimes of export control laws exist regulating export of covered products, components, materials, and related services to foreign persons: 1) the International Traffic in Arms Regulations (ITAR) regulating export of military products, 2) the Export Administration Regulations (EAR) regulating export of dual-use products and specific recipients, including individuals and entities, and 3) the US Sanctions Laws.[244]

Although EAR dual-use item descriptions may technically cover certain technologies, because ITAR takes precedence and prevails over the EAR the product export control, the classification process should always start with evaluating the services, products, components and materials comprising its "product master" vis a vis ITAR.[245] ITAR regulates exports of military products, components, and materials to foreign persons, i.e., anyone not a US citizen or green card holder, under the U.S. Munitions List (USML).[246]

---

[242] "Commerce Control List Classification," Bureau of Industry and Security, US Department of Commerce, 2024, https://www.bis.doc.gov/index.php/licensing/commerce-control-list-classification; "Technology Classification," Florida Atlantic University, 2024, https://www.fau.edu/research-admin/export-control/technology-classification/; "Commodity Jurisdiction," Bureau of Industry and Security, US Department of Commerce, 2024, https://www.bis.doc.gov/index.php/licensing/commerce-control-list-classification/commodity-jurisdiction.

[243] "Commerce Control List Classification"; "Technology Classification"; "Commodity Jurisdiction."

[244] Thomas B. McVey, "Tutorial 1: Introduction to ITAR and the U.S. Munitions List," SBIR.gov, n.d., https://legacy.www.sbir.gov/tutorials/itar/tutorial-1; "Export Controls: Overview," Stanford University, n.d., https://doresearch.stanford.edu/resources/topics/export-controls-overview.

[245] "Technology Classification."

[246] *Ibid.*

2.2.1 International Traffic in Arms Regulations and the US Munitions List

Section 38 of the Arms Export Control Act[247] (AECA) authorizes the President to control the export and import of defense articles and services, as described and implemented by ITAR.[248] Executive Order 13637[249] delegated this authority to the Department of State (DOS), where DDTC administers ITAR export control licenses.[250] The Defense Technology Security Administration (DTSA) and other program offices within the Department of Defense are also closely engaged, and when the Department of State reviews an ITAR case they generally share it freely with DoD.[251]

It is essential to recognize that certain AI technologies, particularly those used for military applications, could fall under ITAR's strict controls. AI systems involved in satellite-based defense operations, autonomous decision-making for weapon systems, or encrypted communications via military satellites may be considered "defense articles" or "defense services."

These systems would require compliance with ITAR's export control measures, ensuring that sensitive technologies are not transferred to foreign entities without proper authorization. The Cornell Legal Information Institute describes ITAR as follows:

> "The International Traffic in Arms Regulations (ITAR) is a set of US Government regulations that control the import and export of defense products. The purpose of ITAR is to safeguard national security, and to further American foreign policy interests.
>
> ITAR (22 CFR parts 120-130) governs the manufacture, export, and temporary import of defense articles, the furnishing of defense services, and brokering activities involving items described on the USML (ITAR section 121.1).
>
> ITAR is authorized by Section 38 of the Arms Export Control Act (22 U.S.C. 2778), which authorizes the President to control the export and import of defense articles and services. This authority was delegated to the State Department by Executive Order 13637.
>
> ITAR is currently administered by the Directorate of Defense Trade Controls (DDTC) which operates within the State Department. ITAR applies to any defense products found on the United States Munitions List (USML).

---

[247] 22 U.S. Code § 2778 - Control of arms exports and imports, 2013, https://www.govinfo.gov/content/pkg/USCODE-2013-title22/pdf/USCODE-2013-title22-chap39.pdf.

[248] U.S. Code of Federal Regulations, Title 22 Subchapter M - International Traffic in Arms Regulations, Parts 120-130, October 28, 2024 [Hereinafter ITAR], https://www.ecfr.gov/current/title-22/chapter-I/subchapter-M; McVey, "Tutorial 1: Introduction to ITAR and the U.S. Munitions List"; "The International Traffic in Arms Regulations (ITAR)," Directorate of Defense Trade Controls, US Department of State, n.d., https://www.pmddtc.state.gov/ddtc_public/ddtc_public?id=ddtc_kb_article_page&sys_id=24d528fddbfc930044f9ff621f961987; "International Traffic in Arms Regulations (ITAR)," Legal Information Institute, Cornell, June 2023, https://www.law.cornell.edu/wex/international_traffic_in_arms_regulations_(itar).

[249] Executive Order 13637 - Administration of Reformed Export Controls, March 8, 2013, https://www.govinfo.gov/content/pkg/DCPD-201300143/pdf/DCPD-201300143.pdf.

[250] McVey, "Tutorial 1: Introduction to ITAR and the U.S. Munitions List"; "The International Traffic in Arms Regulations (ITAR)," DDTC; "International Traffic in Arms Regulations (ITAR)," LII Cornell.

[251] McVey, "Tutorial 1: Introduction to ITAR and the U.S. Munitions List."

> The USML is divided into three subcomponents, defense articles, defense services, and related technical data. Defense Articles under the USML include:
>
> - Guns, ammunition/ordnance
> - Launch vehicles, rockets (including satellites), torpedoes, bombs, and mines
> - Ground vehicles
> - Aircraft and related articles
> - Personal protective equipment
> - Toxicological agents, including chemical agents, biological agents, and associated equipment
>
> Defense Services under the USML include:
>
> - Providing assistance, including training, to foreign persons on anything related to defense articles, including design, development, manufacturing, maintenance, etc.
> - Providing foreign persons with controlled related technical data.
> - Military training of foreign units and forces
>
> Related Technical Data under the USML include:
>
> - Blueprints, drawings, documentation
> - Classified information about the defense articles and defense services
> - Software directly related to defense articles
>
> All manufacturers, exporters, and brokers of defense products under the USML must be compliant with the ITAR, or they face fines and/or incarceration. To comply with ITAR, a company must register with the DDTC and apply for an export license or temporary import license."[252]

Specific applications of AI in space that may be classified as defense articles under ITAR include AI technologies used in military satellite communications or autonomous defense systems on satellites. Companies and individuals developing such AI technologies must comply with stringent requirements, including the need for registration and licensing through the DDTC, to avoid severe penalties and ensure responsible development and deployment of AI in space missions.

Beyond physical products, components, and materials, ITAR also controls software and technical data.[253] For instance, if a GPS navigational device is included in the USML, the software used to run the device and the technical data, e.g., drawings, electronic files, algorithms, specifications, technical manuals, training materials, and any other information on the design, manufacture, or use, related to the device are also on the list and subject to the same ITAR controls as the underlying physical product.[254]

---

[252] "International Traffic in Arms Regulations (ITAR)," Legal Information Institute, Cornell University, n.d., https://www.law.cornell.edu/wex/international_traffic_in_arms_regulations_(itar).

[253] McVey, "Tutorial 1: Introduction to ITAR and the U.S. Munitions List"; "International Traffic in Arms Regulations (ITAR)," LII Cornell.

[254] *Ibid*.

Further, performing services related to any USML listed items, e.g., design, engineering, installation, retrofitting, consulting, troubleshooting, and repair services (warranty or otherwise), etc., must comply with the same controls as any products covered by such item descriptions.[255]

For proper USML classification, exporters must first review the twenty-one USML categories that cover the service, product, component, or material in question.[256] Next, exporters need to consider the function of the product to identify the appropriate sub-category.[257] Companies seeking to conduct a U.S. export control tech assessment are encouraged to use the Department of State's ITAR-USML Decision Tool.[258]

If the technology or its technical data are not described in the ITAR USML, the exporter must perform a similar assessment under the EAR Commerce Control List (CCL).[259]

2.2.2 United States Export Administration Regulations and the Commerce Control List

"The Export Administration Regulations (EAR) is authorized by the Export Control Reform Act of 2018 (ECRA), which replaced the Export Administration Act of 1979 (EAA) as the primary authority for U.S. export control regulations. The EAR also derives its authority from several other statutes and Executive Orders."[260]

EAR governs the export and re-export of dual-use items, including products, software, and technology.[261] And according to the International Trade Administration:

> - "The U.S. Department of Commerce's Bureau of Industry and Security (BIS) administers U.S. laws, regulations and policies governing the export and reexport of commodities, software, and technology (collectively "items") falling under the jurisdiction of the [EAR]. The primary goal of BIS is to advance national security, foreign policy, and economic objectives by ensuring an effective export control and treaty compliance system and promoting continued U.S. strategic technology leadership. BIS also enforces anti-boycott laws and coordinates with U.S. agencies and other countries on export control, nonproliferation, and strategic trade issues.
> - BIS is responsible for implementing and enforcing the EAR, which regulate the export, reexport, and transfer (in-country) of items with commercial uses that can also be used in conventional arms, weapons of

---

[255] McVey, "Tutorial 1: Introduction to ITAR and the U.S. Munitions List"; "International Traffic in Arms Regulations (ITAR)," LII Cornell.

[256] "The International Traffic in Arms Regulations (ITAR)," DDTC; ITAR, Part 121, https://www.ecfr.gov/current/title-22/chapter-I/subchapter-M/part-121.

[257] "Technology Classification."

[258] "The International Traffic in Arms Regulations (ITAR)," DDTC.

[259] "Technology Classification."

[260] "About Export Administration Regulations (EAR)," Bureau of Industry and Security, US Department of Commerce, n.d., https://www.bis.gov/regulations.

[261] "U.S. Export Controls," International Trade Administration, Trade.gov, n.d., https://www.trade.gov/us-export-controls.

mass destruction, terrorist activities, or human rights abuses, and less sensitive military items.

- BIS's Export Administration (EA) reviews license applications for exports, reexports, transfers and deemed exports (technology transfers to foreign nationals in the United States) subject to the EAR. Through its Office of Exporter Services, EA provides information on BIS programs, conducts seminars on complying with the EAR, and provides guidance on licensing requirements and procedures.
- EA's Office of Technology Evaluation (OTE) analyzes U.S. export data on items subject to the EAR, BIS license application data, and global trade information to assess data trends. OTE's data portal provides excerpts from statistical reports, along with data sets to enable the public to perform analyses of exports and licensing on its own."[262]

EAR prohibits exports, reexports, and transfers by US and foreign persons of any US-origin commodity, software, or technology if they know it will support use, development, or production of a missile, chemical, biological, or nuclear weapon.[263] EAR also prohibits US persons from providing support for use, development, or production of such weapons even if all the underlying commodities, technologies, and software are not US-origin or otherwise subject to any export controls.[264]

The EAR CCL lists categories of services, commercial products, components (e.g., replacement parts), and materials that are regulated due to a potential military or intelligence use. The EAR and CCL designate dual-use item categories using five-character alphanumeric codes known as Export Control Classification Numbers (ECCNs).[265]

Companies seeking to conduct a U.S. export control tech assessment under EAR are encouraged to use the List of Publicly Available ECCN Classifications,[266] and the Department of Commerce's EAR-CCL Decision Tool.[267]

EAR99 is a "catch-all" ECCN that is used to designate the lowest level of oversight and restrictions imposed by export control regulations for technologies not specifically described in

---

[262] "U.S. Export Controls."

[263] *Introduction to U.S. Export Controls for the Commercial Space Industry*, 2nd ed., Office of Space Commerce, US Department of Commerce, Office of Commercial Space Transportation, Federal Aviation Administration, November 2017, https://www.space.commerce.gov/wp-content/uploads/2017-export-controls-guidebook.pdf

[264] *Introduction to U.S. Export Controls for the Commercial Space Industry.*

[265] "Introduction to Commerce Department Export Controls"; "Export Control Classification Number (ECCN)"; "Technology Classification."

[266] "Publicly Available Classification Information," Bureau of Industry and Security, US Department of Commerce, 2024, https://www.bis.doc.gov/index.php/licensing/commerce-control-list-classification/publicly-available-classification-information.

[267] "Export Control Classification Interactive Tool," Bureau of Industry and Security, US Department of Commerce, 2024, https://www.bis.doc.gov/index.php/export-control-classification-interactive-tool; "Introduction to Commerce Department Export Controls"; *Introduction to U.S. Export Controls for the Commercial Space Industry*; "Technology Classification."

the ITAR USML, EAR CCL, or other agency list.[268] Typically, EAR99 items do not require an export license and may utilize the “No License Required” or “NLR” designation on the shipping label.[269]

However, an export license will almost always be required for an EAR99 item destined for a sanctioned or embargoed country, to a prohibited party, or in support of a prohibited end-use.[270] As discussed in further detail in Section 1.3, End-User and End-Use Monitoring, EAR Part 744 stipulates regulations related to end-user and end-use controls.[271] The consolidated U.S. government screening list includes many prohibited end-users, drawn from proscribed party lists provided by the Department of State and the Department of Commerce.[272]

Companies seeking to export AI applications for space and space applications in general should review the FAA guidebook, "Introduction to U.S. Export Controls for the Commercial Space Industry"[273] as well as the Bureau of Industry and Security export control classification tool.[274]

2.2.3 United States Export Control Guidelines

On December 5, 2022, the Office of Defense Trade Controls Compliance under the Department of State Directorate of Defense Trade Controls issued a first version of the “International Traffic in Arms Regulations (ITAR) Compliance Program Guidelines,”[275] specifying DDTC expectations and best practices for an effective ITAR Compliance Program.[276]

Similar compliance program guidelines have been issued respectively by the US Department of Commerce Bureau of Industry and Security, with their “Export Compliance Guidelines,”[277] and the US Department of the Treasury Office of Foreign Assets Controls, with

---

[268] “Commerce Control List (CCL),” Bureau of Industry and Security, US Department of Commerce, 2024, https://www.bis.doc.gov/index.php/licensing/commerce-control-list-classification/commerce-control-list-ccl/17-regulations/139-commerce-control-list-ccl; “Introduction to Commerce Department Export Controls”; *Introduction to U.S. Export Controls for the Commercial Space Industry*; “Export Controls: Overview”; “Technology Classification.”

[269] “Commerce Control List (CCL)”; *Introduction to U.S. Export Controls for the Commercial Space Industry*; “Export Controls: Overview”; “Technology Classification.”

[270] “Export Controls: Overview”; “Technology Classification.”

[271] *Introduction to U.S. Export Controls for the Commercial Space Industry*.

[272] “Consolidated Screening List,” International Trade Administration, n.d., https://www.trade.gov/consolidated-screening-list.

[273] *Introduction to U.S. Export Controls for the Commercial Space Industry.*

[274] “Export Control Classification Interactive Tool.”

[275] “International Traffic in Arms Regulations (ITAR) Compliance Program Guidelines,” Directorate of Defense Trade Controls, Department of State, September 19, 2023, https://deccspmddtc.servicenowservices.com/sys_attachment.do?sys_id=cc037c571bdd7150c6c3866ae54bcbc6.

[276] Alison J. Stafford Powell and Eunkyung Kim Shin, “DDTC Issues ITAR Compliance Program Guidelines,” Baker McKenzie, December 14, 2022, https://sanctionsnews.bakermckenzie.com/ddtc-issues-itar-compliance-program-guidelines/.

[277] Export Compliance Guidelines, The Elements of an Effective Export Compliance Program, Bureau of Industry and Security, US Department of Commerce, 2017, https://www.bis.doc.gov/index.php/documents/pdfs/1641-ecp/file.

"A Framework for OFAC Compliance Commitments,"[278] relating to their administration of US sanctions laws.[279]

These guidelines are crucial for companies developing AI technologies for space, as they outline the necessary steps to ensure compliance with ITAR, EAR, and other export controls. For example, a company exporting an AI-enabled communication system for satellites must follow these guidelines to avoid penalties and ensure compliant exports and other regulated transfers.

2.2.4 United States Space-Specific Export Controls

The United States has recently introduced significant changes to its export controls for space-related items through a series of regulatory updates, reflecting the government's commitment to promoting international collaboration and commercial space activities. These updates include the Final Rule (Space 1)[280] and the Interim Final Rule (Space 2),[281] both published and effective on October 23, 2024. Additionally, the Proposed Rule (Space 3)[282] and the State Categories IV and XV Proposed Rule[283] were published on the same date, with public comments for these proposals open until November 22, 2024.[284]

The Final Rule eliminates license requirements for certain spacecraft and related items classified under ECCN 9A515. Specifically, the Rule applies to remote sensing spacecraft designed to meet specific imaging requirements, spacecraft for in-orbit servicing and logistics, spacecraft intended for in-space assembly, and spacecraft with specialized equipment. Additionally, optical sensors for specific imaging applications are included. These changes facilitate exports to trusted partners, namely Australia, Canada, and the United Kingdom, by removing regulatory barriers for items that meet defined criteria.[285]

---

[278] A Framework for OFAC Compliance Commitments, US Department of the Treasury, Office of Foreign Assets Control, May 2, 2019, https://ofac.treasury.gov/media/16331/download.

[279] Powell and Shin, "DDTC Issues ITAR Compliance Program Guidelines."

[280] Bureau of Industry and Security. Final Rule: Revisions to Controls for Spacecraft and Related Items. Federal Register 89, no. 84766 (October 23, 2024). https://public-inspection.federalregister.gov/2024-23932.pdf.

[281] Bureau of Industry and Security. Interim Final Rule: Updates to ECCN Classifications for Space-Related Items (Space 2). Federal Register 89, no. 84770 (October 23, 2024). https://public-inspection.federalregister.gov/2024-23958.pdf.

[282] Bureau of Industry and Security. Export Administration Regulations: Revisions to Space-Related Export Controls, Including Addition of License Exception Commercial Space Activities (CSA) (Space 3). Federal Register 89, no. 84784 (October 23, 2024). https://public-inspection.federalregister.gov/2024-23975.pdf.

[283] Department of State. International Traffic in Arms Regulations (ITAR): U.S. Munitions List Categories IV and XV. Federal Register 89, no. 84482 (October 23, 2024). https://www.federalregister.gov/documents/2024/10/23/2024-24091/international-traffic-in-arms-regulations-itar-us-munitions-list-categories-iv-and-xv.

[284] Office of Space Commerce. Export Administration Regulations (EAR): Final Rule, Interim Final Rule, and Proposed Rule: Public Briefing on Revisions to Space-Related Export Controls under the Export Administration Regulations and International Traffic in Arms Regulations. November 6, 2024. https://www.space.commerce.gov/wp-content/uploads/11-6-2024-US-Space-Export-Control-Stakeholder-Briefing.pdf.

[285] Bureau of Industry and Security. Commerce Announces Series of Rules to Modernize Space-Related Export Controls. U.S. Department of Commerce, 2024. https://www.bis.gov/press-release/commerce-announces-series-rules-modernize-space-related-export-controls.

The Interim Final Rule implements significant revisions to export controls for space-related items, focusing on reducing license requirements for less-sensitive components. It reclassifies items under ECCNs 9A515.x and 9A004.v/.x to National Security Level 2 (NS2) and Regional Stability Level 2 (RS2), reflecting a lower level of control, while reserving 9A515.w for higher-sensitivity items requiring stricter NS1 and RS1 controls. Additionally, the Rule removes license requirements for certain low-sensitivity items, such as solar arrays under 9A515.x, allowing their export to 40 Wassenaar Arrangement countries, including EU member states, Japan, South Korea, and South Africa. New classifications further facilitate license-free exports of low-risk components, such as connectors and parts for landing leg assemblies.[286]

The Interim Final Rule also expands exemptions for standards-related activities under ECCNs 9D515 (Software) and 9E515 (Technology), supporting the development of global space safety standards. Moreover, it refines controls on space-related items within the Commerce Control List (CCL), including clarifications for spacecraft operating on celestial bodies and the introduction of new classifications, such as 9A004.r, which covers in-space habitats.[287]

The Proposed Rule introduces a Commercial Space Activities (CSA) License Exception to facilitate exports supporting official space agency programs, space tourism, and research. Eligible programs include NASA's Lunar Gateway, the Mars Sample Return, the Nancy Grace Roman Telescope, the Orion spacecraft, the Habitable Worlds Observatory, and the Commercial Low Earth Orbit Development Program. It also authorizes exports for space tourism and research activities. [288]

To accommodate emerging technologies, ECCN 9A515 has been broadened to include logistics and rendezvous/proximity operations spacecraft, advanced space-qualified optics, electric thrusters, control moment gyroscopes, and separation mechanisms. Additionally, ECCNs 9C515, 9D515, and 9E515 address materials for reducing in-orbit signatures, space situational awareness software, and technology controls for these advanced systems.[289]

Applied to the subject of this study, several observations and recommendations arise from the recent regulatory updates. Specifically, the CSA License Exception, which facilitates exports for space agency programs, tourism, and research, could potentially extend to AI-powered space systems if these systems are integrated into controlled items. Such inclusion would automatically classify AI space systems under lower levels of export control, thereby simplifying their export and enabling broader international collaboration.

To ensure clarity and robust oversight of AI system exports in the future, a tiered licensing framework could be introduced. This framework would assess factors such as the system's autonomy level, dual-use potential, and the sensitivity of its training data. Furthermore, ECCNs like 9D515 and 9E515 could be explicitly expanded to cover AI algorithms, training datasets, and neural network hardware dependencies. These measures would enhance regulatory oversight while

---

[286] *Ibid.*
[287] *Ibid.*
[288] *Ibid.*
[289] *Ibid.*

fostering the responsible and innovative development of AI technologies in the space sector, ensuring a balance between technological advancement and national security.

2.2.5 United States AI-Specific Export Controls

2.2.5.1 Executive Order 14110

With respect to US AI-specific export controls, on October 30, 2023 President Biden released EO 14110, for "Safe, Secure, and Trustworthy Development and Use of AI."[290] It mandated about 150 concrete requirements across some 50 agencies, engaging a whole-of-government approach, set ambitious deadlines, most of which fall within 90 days to a year, and aims to manage AI security and risks.[291]

EO 14110 relies on Defense Production Act (50 U.S.C. §§4501-4568) to require responses,[292] and at the end of January 2024, the White House announced completion of all 90-day required actions.[293] Of particular relevance to companies and private sector entities, under §4.2(a), the US Department of Commerce must "require:

1) companies developing, or intending to develop, dual-use AI models to report to the government on model training, testing, and data ownership; and
2) entities that acquire, develop, or possess large computing infrastructure to report on the location and amount of computing power."[294]

§4.2(b) of EO 14110 requires USDOC (in collaboration with DOS, DOD, DOE, DNI) to define technical criteria for which specific models and resources must be reported under 4.2(a).[295]

Researchers have estimated the EO minimum computational power threshold triggering reporting requirements for companies, set at 10^{26} floating-point operations per second (FLOPS), as greater than any fully trained models yet developed—though OpenAI's GPT-4 model was just

---

[290] "Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence," The White House, October 30, 2023, https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/; "FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence," The White House, October 30, 2023, https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/; "Highlights of the 2023 Executive Order on Artificial Intelligence for Congress," CRS Report, Congressional Research Service, April 3, 2024, https://crsreports.congress.gov/product/pdf/R/R47843.

[291] "Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence"; "Highlights of the 2023 Executive Order on Artificial Intelligence for Congress."

[292] *Ibid.*

[293] "Fact Sheet: Biden-Harris Administration Announces Key AI Actions Following President Biden's Landmark Executive Order," The White House, January 29, 2024, https://www.whitehouse.gov/briefing-room/statements-releases/2024/01/29/fact-sheet-biden-harris-administration-announces-key-ai-actions-following-president-bidens-landmark-executive-order/.

[294] "Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence"; "Highlights of the 2023 Executive Order on Artificial Intelligence for Congress."

[295] *Ibid.*

under.[296] Aurora-M is the first open-source multilingual language model red-teamed under EO 14110.[297]

Given that GPT-4 and competing AIs are already nearly at the computational power threshold, in all likelihood many contemporary AIs in the near future will exceed the computational power threshold, unless specifically downgraded to enable export. In addition, although AI system components, including hardware, software modules, or technical data regarding AI system design and development may be covered by the USML, ITAR does not presently provide any specific definitions for AI items, products, technology or software.[298]

Further, as AI systems become more autonomous and capable of critical decision-making, new export control measures will become necessary. Such measures could include mandatory real-time reporting to a central oversight authority for any autonomous actions taken by AI systems, implementing fail-safes that require human authorization for critical decisions, and developing international agreements to monitor and regulate the deployment of AI technologies in space to ensure they are not repurposed for military applications.

Additionally, establishing a certification process for AI systems to verify their compliance with military and dual-use export control regulations prior to export could further mitigate the risk of misuse.

2.2.5.2 Bureau of Industry and Security AI-Specific Export Control Updates

On May 23, 2019, the Department of Commerce Bureau of Industry and Security (BIS) added 5 emerging techs essential to US national security to the EAR CCL, implementing amendments to Wassenaar Arrangement List of Dual Use Goods and Technologies made at the December 2018 plenary meeting.[299] Among other changes, BIS amended ECCN 5A002 to add certain types of post-quantum cryptographic algorithms, specifically including certain post-quantum asymmetric cryptographic algorithms.[300]

Examples of AI space applications that may fall under EAR regulations include AI systems used for autonomous satellite operations or real-time space-mission data analytics. Such technologies must comply with EAR and BIS must evaluate and grant appropriate licenses for such AI technologies.

---

[296] Markus Anderljung et al., "Frontier AI Regulation: Managing Emerging Risks to Public Safety," arXiv, July 6, 2023, https://arxiv.org/abs/2307.03718; Rishi Bommasani et al., "Decoding the White House AI Executive Order's Achievements," Stanford University Institute for Human-Centered AI, November 2, 2023, https://hai.stanford.edu/news/decoding-white-house-ai-executive-orders-achievements.

[297] Tiashi Nakamura et al., "Aurora-M: The First Open Source Multilingual Language Model Red-teamed under EO 14110," arXiv, April 23, 2024, https://arxiv.org/abs/2404.00399.

[298] David Plotinsky and Giovanna Cinelli, "Existing and Proposed Federal AI Regulation in the United States," Morgan Lewis, April 09, 2024, https://www.morganlewis.com/pubs/2024/04/existing-and-proposed-federal-ai-regulation-in-the-united-states.

[299] "Implementation of Certain New Controls on Emerging Technologies Agreed at Wassenaar Arrangement 2018 Plenary," Federal Register, May 23, 2019, https://www.federalregister.gov/documents/2019/05/23/2019-10778/implementation-of-certain-new-controls-on-emerging-technologies-agreed-at-wassenaar-arrangement-2018.

[300] "Implementation of Certain New Controls on Emerging Technologies."

The potential penalties and sanctions emphasize the importance of compliance to avoid such penalties and help ensure the responsible use of AI in space exploration and operations. To help support these efforts, BIS provides many resources for guidance and support to companies navigating these regulations, including practical steps necessary for compliance.

On January 6, 2020, BIS added new regional stability export controls on AIs designed to automate analysis of geospatial imagery under ECCN 0Y521.[301] Under regional stability, export licenses are required for export to any country except Canada, and the only applicable license exception is for exports, reexports, and transfers (in-country) made by US Government (License Exception GOV, section 740.11(b)(2)(ii)).[302]

Under ECCN 0Y521, the AI must be designed to train Deep Convolutional Neural Networks (DCCNs) to automate analysis of geospatial imagery and point clouds ("digital surface models" or collections of data points within a coordinate system).[303]

Further, the AI must meet all of the following technical requirements for controls to apply: 1) it provides a graphical user interface (GUI) enabling users to identify objects (e.g., vehicles, houses, etc.) within geospatial imagery and point clouds and extract positive and negative samples of an object of interest; 2) it reduces pixel variation by performing scale, color, and rotational normalization; 3) it trains DCNNs to detect objects of interest from positive and negative samples; and 4) it identifies objects in geospatial imagery using the trained DCNN by matching the rotational patterns from positive samples and objects in the geospatial imagery.[304]

These controls cover and impact exports of AI systems that perform the requisite autonomous geospatial analysis on satellite imagery and other remote sensing "imaging" data to any country other than Canada. To address this regulation, companies seeking to export such AI systems should engage in early consultations with the appropriate regulatory personnel to ensure compliance and obtain support in applying for specific export licenses.

Internal procedures should include implementing robust compliance checks during the development phase and considering the use of export control classification tools to identify and manage dual-use items effectively.

---

[301] "Addition of Software Specially Designed to Automate the Analysis of Geospatial Imagery to the Export Control Classification Number 0Y521 Series," Federal Register, January 6, 2020, https://www.federalregister.gov/documents/2020/01/06/2019-27649/addition-of-software-specially-designed-to-automate-the-analysis-of-geospatial-imagery-to-the-export; "Export Control Classification Number 0Y521 Series Supplement-Extension of Controls on an Emerging Technology (Software Specially Designed to Automate the Analysis of Geospatial Imagery Classification)," Federal Register, January 6, 2022, https://www.federalregister.gov/documents/2022/01/06/2021-28444/export-control-classification-number-0y521-series-supplement-extension-of-controls-on-an-emerging; "Supplement No. 5 to Part 774—Items Classified Under ECCNS 0A521, 0B521, 0C521, 0D521 and 0E521," Bureau of Industry and Security, US Department of Commerce, n.d., https://www.bis.gov/ear/title-15/subtitle-b/chapter-vii/subchapter-c/part-774/supplement-no-5-part-774-items-classified.

[302] "Addition of Software Specially Designed to Automate the Analysis of Geospatial Imagery"; "Export Control Classification Number 0Y521 Series Supplement-Extension of Controls"; "Supplement No. 5 to Part 774."

[303] *Ibid.*

[304] *Ibid.*

In October 2022, BIS issued new export controls on advanced semiconductors to further limit China's access to advanced processors and semiconductors chips needed to build supercomputers and train advanced AI systems.[305] These new export controls provide new restrictions and licensing requirements on advanced computing ICs (ECCN 3A090) and computers and machines containing those ICs (ECCN 4A090), as well as software (4D090) and semiconductor manufacturing equipment (3B090) related to design and development of items listed in 4A090 and 3A090.[306]

Additionally, these export controls regulate export of software and technologies used to manufacture items under ECCNs 3A090 and 4A090 and put end-use controls on any of these items if the exporter knows they will be used to build a supercomputer.[307]

To cover the broad industry sector and not just companies such as NVIDIA or AMD, the interim final rule provided performance thresholds and any integrated circuit (IC) computer chip that surpasses these thresholds is subject to export controls under ECCN 3A090 if: 1) aggregate bidirectional data transfer speeds exceed 600 GB/s or more; 2) comprises a digital processor that can meets or exceeds processing speeds of 4,800 TOPS (trillion operations per second); 3) comprises a digital primitive computational unit that can meets or exceeds 4,800 TOPS; or 4) comprises an analog/multi-level computational unit that can meets or exceeds 38,400 TOPS.[308]

These performance thresholds matched almost exactly the performance of the NVIDIA A100 chip and were downgraded by BIS in October 2023 to cover NVIDIA's slower A800 and H800 chips.[309] Any company building comparable ICs would be bound to the same controls. As

---

[305] "Implementation of Additional Export Controls: Certain Advanced Computing and Semiconductor Manufacturing Items; Supercomputer and Semiconductor End Use; Entity List Modification," Federal Register, October 13, 2022, https://www.federalregister.gov/documents/2022/10/13/2022-21658/implementation-of-additional-export-controls-certain-advanced-computing-and-semiconductor. As updated, *see* "Implementation of Additional Export Controls: Certain Advanced Computing Items; Supercomputer and Semiconductor End Use; Updates and Corrections; and Export Controls on Semiconductor Manufacturing Items; Corrections and Clarifications," Federal Register, April 4, 2024, https://www.federalregister.gov/documents/2024/04/04/2024-07004/implementation-of-additional-export-controls-certain-advanced-computing-items-supercomputer-and; "Commerce Implements New Export Controls on Advanced Computing and Semiconductor Manufacturing Items to the People's Republic of China (PRC)," Bureau of Industry and Security, US Department of Commerce, October 7, 2022, https://www.bis.doc.gov/index.php/documents/about-bis/newsroom/press-releases/3158-2022-10-07-bis-press-release-advanced-computing-and-semiconductor-manufacturing-controls-final/file; William Reinsch et al., "Insight into the U.S. Semiconductor Export Controls Update," Center for Strategic and International Studies, October 20, 2023, https://www.csis.org/analysis/insight-us-semiconductor-export-controls-update; Gupta and Reddie, "Accelerating the Evolution of AI Export Controls."

[306] "Implementation of Additional Export Controls: Certain Advanced Computing"; "Commerce Implements New Export Controls on Advanced Computing"; Reinsch, "Insight into the U.S. Semiconductor Export Controls Update"; Gupta and Reddie, "Accelerating the Evolution of AI Export Controls."

[307] "Implementation of Additional Export Controls: Certain Advanced Computing"; Reinsch, "Insight into the U.S. Semiconductor Export Controls Update"; Gupta and Reddie, "Accelerating the Evolution of AI Export Controls."

[308] "Implementation of Additional Export Controls: Certain Advanced Computing"; Reinsch, "Insight into the U.S. Semiconductor Export Controls Update"; Gupta and Reddie,"Accelerating the Evolution of AI Export Controls"; "U.S. AI Chip Export Restrictions: Impact on Nvidia, AMD," Cimphony, 2023, https://www.cimphony.ai/insights/us-ai-chip-export-restrictions-impact-on-nvidia-amd.

[309] Reinsch, "Insight into the U.S. Semiconductor Export Controls Update"; Gupta and Reddie, "Accelerating the Evolution of AI Export Controls."

detailed elsewhere herein, CCL and ITAR decision support tools exist to help companies determine whether an export license is required and to support licensing applications.

Despite these restrictions prohibiting export of these advanced processors and supercomputers to China, a network of buyers, sellers, and couriers operates to circumvent the U.S. export controls.[310] More than 70 Chinese vendors openly market and sell export restricted chips, computers, and servers online, in many cases with delivery in just weeks.[311]

ECCN 4A090 controls computers if they contain ICs exceeding the performance thresholds specified under ECCN 3A090 above. Additionally, ECCN 4A004 controls specific kinds of advanced computers, including "optical computers" and "neural computers"—AI systems that learn like NNs and store complex data structures like computers.[312]

In view of these controls, companies developing software and hardware AI systems for space applications would either have to downgrade their engineering specifications and performance metrics to meet the export control thresholds required for export without a license or obtain any required export licenses. However, in many cases launch of a space object does not constitute an export of the item, as under U.S. law.

From an R&D perspective, investing time and effort into downgrading engineering specs to enable export is not an interesting and worthwhile endeavor to advance technological developments. However, where markets will support the commercial activity, such endeavors can certainly make sense from a business perspective.

For AI technologies used in space, such as computing processors onboard satellites, companies must implement a rigorous compliance strategy that includes regular audits and collaboration with export control experts to navigate these restrictions effectively. For instance, at some inflection point within the growth trajectory of a company it begins to make better sense to develop dedicated in-house export compliance teams than to rely solely on external export compliance professionals.

Further, specialized products and trading systems capable of monitoring product specifications against evolving regulations can help alleviate some of the administrative burden.

2.2.5.3 The Disruptive Technology Strike Force

On February 16, 2023, the Department of Justice (DOJ) and Commerce Department jointly created the Disruptive Technology Strike Force (DTSF) to prevent foreign competitors from obtaining advanced technologies.[313]

---

[310] Anton Shilov, "Underground Network Smuggles Nvidia's AI GPUs Into China Despite US Sanctions &Mdash; Some Smugglers Even Sell Entire Servers," *Tom's Hardware*, July 3, 2024, https://www.tomshardware.com/tech-industry/artificial-intelligence/underground-network-smuggles-nvidia-gpus-into-china-despite-us-sanctions-some-smugglers-even-sell-entire-servers.

[311] Shilov, "Underground Network Smuggles Nvidia's AI GPUs Into China."

[312] "Implementation of Additional Export Controls: Certain Advanced Computing"; Emily Weinstein and Kevin Wolf, "For Export Controls on AI, Don't Forget the "Catch-All" Basics," Center for Security and Emerging Technology, Georgetown University, July 5, 2-23, https://cset.georgetown.edu/article/dont-forget-the-catch-all-basics-ai-export-controls/.

[313] "Justice and Commerce Departments Announce Creation of Disruptive Technology Strike Force," US Department of Justice, US Department of Commerce, February 16, 2023,

The strike force represents a collaborative effort of the DOJ National Security Division (NSD), BIS, the Federal Bureau of Investigation, Homeland Security Investigations, and 14 US Attorneys' Offices in 12 major US cities, and is co-led by the DOJ Assistant Attorney General for NSD and Assistant Secretary for Export Enforcement at BIS.[314]

The strike force's work will cover:

> "[1)] investigating and prosecuting criminal violations of US export laws;
> [2)] enhancing administrative enforcement of US export controls;
> [3)] fostering partnerships with the private sector;
> [4)] leveraging international partnerships to coordination law enforcement actions and disruption strategies;
> [5)] utilizing advanced data analytics and all-source intelligence to develop and build investigations;
> [6)] conducting regular trainings for field offices; and
> [7)] strengthening connectivity between the strike force and Intelligence Community."[315]

Since the DTSF was formed, it has established more than a dozen local investigative cells, opened scores of investigations, and brought criminal charges, including sanctions violations, export controls and other offenses involving unlawful transfer of sensitive information and technology, against more than a dozen individuals and companies, including corporate executives, engineers, distributors, and other high-profile targets associated with nation-state adversaries, such as China and Iran.[316]

Applied Materials, Inc., the largest U.S. semiconductor equipment maker, is one of the high-profile companies currently under criminal investigation, and it disclosed receipt of another subpoena from the U.S. Department of Commerce on May 23, 2024, as regulators seek additional details on shipments to China.[317]

As reported, Applied Materials has allegedly engaged in shipments of hundreds of millions of dollars' worth of chip-making equipment comprising unauthorized re-exports via South Korea

---

https://www.bis.doc.gov/index.php/documents/about-bis/newsroom/press-releases/3222-distruptive-tech-strike-force/file; Brian J. Fleming et al., "Justice and Commerce Departments Announce Creation of Disruptive Technology Strike Force," International Compliance Blog, Steptoe LLP, February 7, 2023, https://www.steptoe.com/en/news-publications/international-compliance-blog/justice-and-commerce-departments-announce-creation-of-disruptive-technology-strike-force.html.

[314] "Justice and Commerce Departments Announce Creation of Disruptive Technology Strike Force"; Fleming, "Justice and Commerce Departments Announce Creation of Disruptive Technology Strike Force."

[315] *Ibid.*

[316] "Justice Department Announces Charges and Arrest in Two Separate Illicit Technology Transfer Schemes to Benefit Governments of China and Iran" Bureau of Industry and Security, US Department of Commerce, February 7, 2024, https://www.bis.gov/press-release/justice-department-announces-charges-and-arrest-two-separate-illicit-technology.

[317] Jaspreet Singh, "Applied Materials Gets Another Subpoena on China Customer Shipments," Reuters, May 23, 2024, https://www.usnews.com/news/technology/articles/2024-05-23/applied-materials-says-received-subpoena-from-us-commerce-department.

to China's top chipmaker, Semiconductor Manufacturing International Corporation (SMIC), in violation of export control restrictions imposed against that entity.[318]

Applied Materials has also confirmed receipt of an SEC subpoena and two from the U.S. Attorney's Office for the District of Massachusetts. In November 2023 the Commerce Department sent an initial subpoena "requesting information relating to certain China customer shipments."[319]

2.2.5.4 Distinguishing Software vs. Hardware in US AI-Specific Export Controls

The US understands that it faces a fiercely competitive geopolitical landscape and that it must strategically invest, regulate, and restrict access to maintain its AI leadership while also counteracting opponent advantages.[320] According to the Stanford 2023 AI Index report,[321] in the development of state of the art AI systems, the US is slightly ahead on AI models and further ahead on leading-edge AI hardware (e.g., ICs and other devices) needed to train and deploy advanced AI models.[322]

Fortunately for the US, its semiconductor industry maintains a majority share of the world market in advanced integrated circuits (ICs) such as CPUs, GPUs, TPUs, etc., which provide the foundation for nearly the entire global AI market.[323] Advanced ICs and AI hardware have become the dominant focus of export controls, in part because AI models can be more freely shared online under permissive licenses (e.g., open source).[324] Further, the manufacturing skills and expertise required to produce advanced ICs and AI hardware tend to be very limited, e.g., geographically.[325]

And consequently, the US is able to leverage this discrepancy to advantage, due to the current critical fact that creating and fielding the most advanced AI models without access to appropriately advanced ICs and AI hardware remains relatively implausible, if not impossible.[326] However, this present asymmetry could potentially be disrupted at any time by any number of new developments currently on the verge, such as neuromorphic hardware, liquid neural networks, and biocomputing.

Furthermore, deploying AI systems in space needs to take account of the special consideration and fact that hardware operating such AI systems must be radiation hardened, and high-end rad hard electronics are almost always export controlled.

---

[318] *Ibid*.

[319] *Ibid*.

[320] Gupta and Reddie, "Accelerating the Evolution of AI Export Controls."

[321] "AI Index Report 2023," Artificial Intelligence Index, Stanford University, 2023, https://aiindex.stanford.edu/ai-index-report-2023/.

[322] Gupta and Reddie, "Accelerating the Evolution of AI Export Controls."

[323] *Ibid.*

[324] *Ibid.*

[325] *Ibid.*

[326] *Ibid.*

3. *End-User and End-Use Monitoring*

Large language models (LLMs) and other AI-enabled systems and capabilities are poised to have military applications in the near future, if they do not already.[327] Specific applications, such as autonomous drones or AI-driven surveillance systems, underscore the urgency of rigorous end-user and end-use checks. These checks are crucial steps in the export control compliance process, along with product classification, risk assessment, and license determination.

End-user checks are typically tied to sanctioned countries, persons, and entities. For example, AI software used to run the autonomous weapons systems must be scrutinized to prevent proliferation. The goal of end-user checks is to ensure that these items do not reach unauthorized parties and do not contribute to the proliferation of weapons of mass destruction.

While many entities are listed on the United Nations Security Council Consolidated List,[328] each state maintains its own list of sanctioned entities. In the EU, these lists are associated with specific sanctioned countries or activities, such as terrorism, human rights violations, or cybersecurity threats.[329]

As per the end-use checks, in the case of emerging technologies, such as AI it is often a challenge to properly classify it and hence determine the level of controls. Multiple AI-related technology items are simply not part of the lists, as they evolve faster than can be regulated.

For instance, AI algorithms used in predictive maintenance for military equipment may not be explicitly listed. To address this, in order to avoid accidental exports to malicious end-users, multiple jurisdictions implement “catch-all” controls.

The end-use “catch-all” controls were created in the early 1990s to regulate the export, reexport, and transfer of widely available items that were still nonetheless useful in the development, production, or use of missiles, chemical and biological weapons, or nuclear weapons.

In today's context, these controls are crucial for AI technologies. They enable authorities to “catch all” exports of unlisted items that might be diverted to such end-uses, thereby closing regulatory gaps.

In the US export control regime, for instance the EAR sections 744.2, 744.3, and 744.4,[330] prohibit any U.S. or foreign person from exporting, without a license, any type of commodity, software, or technology that is US-origin to a foreign country, if there is “knowledge” that the item will be used directly or indirectly in the production, development, or use of:

- Nuclear energy or explosives;

---

[327] CSET. "Don't Forget the Catch-All Basics: AI Export Controls." Georgetown University Center for Security and Emerging Technologies. Accessed September 27, 2024. https://cset.georgetown.edu/article/dont-forget-the-catch-all-basics-ai-export-controls/; U.S. Department of Defense. 2023. DoD Data Analytics and AI Adoption Strategy. November. https://media.defense.gov/2023/Nov/02/2003333300/-1/-1/1/DOD_DATA_ANALYTICS_AI_ADOPTION_STRATEGY.PDF.

[328] United Nations Security Council. "UN Security Council Consolidated List." Accessed September 27, 2024. https://main.un.org/securitycouncil/en/content/un-sc-consolidated-list.

[329] EU Sanctions Map, Accessed September 27, 2024. https://www.sanctionsmap.eu/#/main.

[330] U.S. Department of Commerce. *Export Administration Regulations (EAR)*. 15 C.F.R. §§ 744.2, 744.3, 744.4. Accessed September 27, 2024. https://www.ecfr.gov/current/title-15/subtitle-B/chapter-VII/part-744.

- Rocket systems, missiles, certain unmanned aerial vehicles; or
- Chemical or biological weapons.

BIS also mandates end-user controls via the EAR Entity List included in Supplement No. 4 to Part 744 of the EAR.[331] The Entity List contains a list of names of foreign persons "including businesses, research institutions, government and private organizations, individuals, and other types of legal persons,"[332] and additional specific requirements imposed beyond EAR standard requirements for the license of export, reexport and/or transfer (in-country) of specified items.[333]

According to BIS:

> "BIS first published the Entity List in February 1997 as part of its efforts to inform the public of entities that have engaged in activities that could result in an increased risk of the diversion of exported, reexported or transferred (in-country) items to weapons of mass destruction (WMD) programs. Since its initial publication, grounds for inclusion on the Entity List have expanded to activities sanctioned by the State Department and activities contrary to U.S. national security and/or foreign policy interests."[334]

For instance, in 2022 BIS added Chinese chip maker Cambricon, a "major AI chip research and development company whose manufacturing and sales entities are, or have close ties to, government organizations that support the Chinese military and defense industry,"[335] to the EAR Entity List after it sought "to acquire US-origin items in support of China's military modernization."[336]

A similar concept is present at the European level. Article 4 of the EU Dual Use Regulation imposes an authorization requirement for the export of dual-use items not listed in Annex I if the exporter has been informed by the competent authority that the items in question are or may be intended, in their entirety or in part: 1) for use in connection with the development, production, handling, operation, maintenance, storage, detection, identification, or dissemination of chemical, biological, or nuclear weapons, or other nuclear explosive devices, or the development, production, maintenance, or storage of missiles capable of delivering such weapons; 2) for a military end-use if the purchasing country or country of destination is subject to an arms embargo; 3) for use as parts or components of military items listed in the national military list that have been exported

---

[331] "Implementation of Additional Export Controls: Certain Advanced Computing"; Weinstein, "For Export Controls on AI, Don't Forget the "Catch-All" Basics"; "Entity List," Bureau of Industry and Security, US Department of Commerce, 2024, https://www.bis.doc.gov/index.php/policy-guidance/lists-of-parties-of-concern/entity-list.

[332] "Entity List," Bureau of Industry and Security.

[333] *Ibid*.

[334] *Ibid*.

[335] "Additions and Revisions to the Entity List and Conforming Removal From the Unverified List," Federal Register, December 19, 2022, https://www.federalregister.gov/documents/2022/12/19/2022-27151/additions-and-revisions-to-the-entity-list-and-conforming-removal-from-the-unverified-list.

[336] "Additions and Revisions to the Entity List," Federal Register; Weinstein, "For Export Controls on AI, Don't Forget the "Catch-All" Basics"; EAR Entity List, Supplement No. 4 to 15 CFR Part 744, https://www.ecfr.gov/current/title-15/subtitle-B/chapter-VII/subchapter-C/part-744.

from the territory of a Member State without authorization or in violation of an authorization prescribed by the national legislation of that Member State.[337]

For example, an AI-powered drone navigation system intended for a country under an arms embargo would require thorough scrutiny and authorization before exporting.

The EU Dual Use Regulation also further explains the meaning 'military end-use', comprising: (i) incorporation into military items listed in the military list of Member States; (ii) use of production, test or analytical equipment and components therefor, for the development, production or maintenance of military items listed in the military list of Member States; or (iii) use of any unfinished products in a plant for the production of military items listed in the military list of Member States.

For AI technologies, this means an AI system for predictive maintenance of military equipment would fall under these regulations. Companies should establish compliance protocols, including rigorous vetting of end-users, mandatory documentation and reporting of AI deployment, and regular audits to ensure AI systems are not repurposed for unauthorized military applications.

To help companies in this task, a series of "red flag" indicators has been developed by different jurisdictions. While the exact wording varies from country to country, the core content of the red flags remains the same. Here are the most common ones:

- Reluctance to provide information about the end-use of items is a significant concern.
- Requests for unusual modifications or customizations can indicate potential risks.
- A lack of familiarity with the product's performance characteristics by the customer raises alarms.
- Customers willing to pay cash for high-value items, contrary to expected financing terms, may signal potential issues.
- Limited business backgrounds of customers can be a red flag.
- Proposing unusual shipping arrangements or vague delivery dates raises concerns.
- The involvement of freight forwarding firms or distributors as the final destination creates uncertainty.
- Orders for items that appear incompatible with the technical capabilities of the destination country should be scrutinized.
- A lack of interest in spare parts or requests for an excessive number of them warrant caution.
- Unclear intentions regarding whether a purchased product is for domestic use or export, particularly when its stated use is civilian but linked to military or defense entities, raises significant concerns.

---

[337] Regulation (EU) 2021/821 of the European Parliament and of the Council of 20 May 2021 setting up a Union regime for the control of exports, brokering, technical assistance, transit and transfer of dual-use items (OJ L 206, 11.6.2021, p. 1, ELI: http://data.europa.eu/eli/reg/2021/821/oj).

In conclusion, efficient end-user and end-use monitoring is crucial to preventing the misappropriation or diversion of AI technology, especially those with military uses, to unapproved parties. Robust compliance procedures, such as stringent end-user vetting and the deployment of catch-all controls, are essential as the landscape of developing technologies rapidly changes.

Due to the dual-use nature of many AI systems, export controls must be implemented proactively, stressing the importance of thorough processes and being aware of any warning signs that might point to possible misuse. By implementing these strategies, businesses can more adeptly negotiate the intricacy. By taking these steps, businesses may protect both international stability and national security by more adeptly navigating the complicated regulatory landscape.

4. *Exemptions to Export Controls*

4.1 *European Union Export Control Exemptions*

The EU export control law de-control notes describe circumstances wherein specific listed items may be excluded from controls, with no license required for exports or intra-EU transfers.[338] The EU dual-use regulations provide further exemptions for "basic scientific research" and technology already "in the public domain."[339]

Whether one of the de-control notes should apply must be evaluated on a case-by-case basis via company internal export screening procedures and, where appropriate, in consultation with competent authorities. In many cases, export controls do not apply to teaching activities, such as university lectures on publicly available scientific knowledge.[340]

However, when teaching involves sensitive technologies related to AI and space to foreign nationals within the EU, it may fall under the notion of "technical assistance" as provided by restrictive measures (sanctions) regulations.[341]

4.1.1 Basic Scientific Research

The EU dual-use regulation defines "basic scientific research" as "experimental or theoretical work undertaken principally to acquire new knowledge of the fundamental principles of phenomena or observable facts, not primarily directed towards a specific practical aim or

---

[338] EU Regulation 2021/821, Annex I, Nuclear Technology Note and General Technology Note; EU Recommendation 2021/1700, Sections 2.3.4 and 2.3.5; "EU Compliance Guidance for Research Involving Dual-use Items," EU Dual Use Research Guidance-Draft Version for Targeted Consultation, European Commission Trade Department, August 5, 2019, https://trade.ec.europa.eu/consultations/documents/consul_183.pdf.

[339] EU Regulation 2021/821, Annex I, Nuclear Technology Note and General Technology Note; EU Recommendation 2021/1700, Sections 2.3.4 and 2.3.5; "EU Compliance Guidance for Research Involving Dual-use Items."

[340] EU Recommendation 2021/1700, note 11; "EU Compliance Guidance for Research Involving Dual-use Items," note 7.

[341] EU Recommendation 2021/1700, note 11; "EU Compliance Guidance for Research Involving Dual-use Items," note 6.

objective."[342] The reference to fundamental research implicitly excludes non-fundamental forms of research, such as translational or applied research.[343]

For the use of AI in space, this could include research on fundamental AI algorithms for data analysis in space environments, e.g., for debris detection, as opposed to applied research directly aimed at developing commercial-ready space technologies, e.g., for debris mitigation or remediation.

"Basic scientific research" operates as a de-control only for listed dual-use technology research outputs and does not cover intent to conduct "basic scientific research," i.e., conversations at the research funding application stage.[344] This distinction is key in determining when it is appropriate to apply an exemption under this de-control note, e.g., when an academic may safely share their research and results in publications or at conferences.[345]

The fact that AI research, and especially research of AI in space applications, advances so rapidly only make the challenge of assessing the state of the art in current basic scientific research more difficult.

Further, basic scientific research also cannot be used to enable export of physical equipment, components, and materials.[346] For instance, theoretical research on AI algorithms to analyze cosmic radiation data may fall under this exemption and be eligible for export without a license.

Whereas this is not the case, and a license would be needed for the physical export of an AI-powered satellite navigation system that, for example, may employ a chart and analysis of the cosmic background radiation for navigation purposes, e.g., along with a star chart, etc., especially for location and navigation in non-GPS enabled and GPS-denied space environments.

Technology Readiness Level (TRL) and amount of industry funding are two factors important in assessing applicability of the basic scientific research de-control note.[347] TRL is a scale (from 1 to 9) originally developed by National Aeronautics and Space Administration (NASA), and later adopted by the U.S. government, European Space Agency and others, to assess and indicate developmental status for space technologies.[348]

Further, TRL is commonly used by the investment and research communities, including venture capital, U.S. federal funding agencies, and EU funding organizations such as the European Research Agency, to assess and indicate the commercial readiness of new and emerging technologies.[349]

---

[342] EU Recommendation 2021/1700, Glossary; "EU Compliance Guidance for Research Involving Dual-use Items," Section 4.2.2.

[343] "EU Compliance Guidance for Research Involving Dual-use Items," Section 4.2.2.

[344] EU Recommendation 2021/1700, Section 2.3.5.

[345] *Ibid.*

[346] *Ibid.*

[347] *Ibid.*

[348] *Ibid.*

[349] *Ibid.*

Research outputs and technology developments at TRL 1 and 2 generally comprise basic scientific research.[350] The eligibility of research output producing TRL 3 and 4 technology, involving initial prototype developments, requires assessment on a case-by-case basis. And research outputs for technology developments above TRL 4, such as operational AI-powered spacecraft systems, fall outside the definition of basic scientific research.[351]

Furthermore, academic research funded by external industry partners often aims to create commercial developments, which would exclude it from the basic scientific research exemption.[352] For AI in space, this might involve partnerships with large aerospace companies. To mitigate conflicts, researchers should document the fundamental nature of their work and ensure transparent agreements regarding the dissemination of research results.

Often, industry sponsors may be given the opportunity to review, comment on, and in some cases require removal of confidential information from publications, presentations, and other disseminations of the research results prior to such public disclosures, e.g., via the funding contract.[353] They are also often accorded the right to delay any such prospective public disclosures (typically by up to 60 days or so) to allow the industry sponsor to determine whether to pursue a patent application and to file such application if so.[354]

Such restrictions in these types of collaborations can be indicative that the resulting research output will not be basic scientific research and thus may require a license before publication or patenting.[355]

4.1.2 In the Public Domain

The EU dual-use regulation defines "in the public domain" as: "technology" or "software" made available without any restrictions on further dissemination.[356] Logically, the de-control note "in the public domain" only applies to controlled dual-use technology already publicly available.[357]

This "means that a to-be research output (open-source software, publication, conference material, ...) can only benefit from this de-control if the listed dual-use software or technology that it contains is already in the public domain."[358]

Therefore, careful evaluation prior to disseminating research outputs is of critical importance to avoiding the potential risks and legal implications of sharing controlled AI technologies used in space research. And, such efforts can be especially challenging, particularly given the rapid advancement of AI research for space applications.

---

[350] *Ibid.*

[351] *Ibid.*

[352] *Ibid.*

[353] *Ibid.*

[354] *Ibid.*

[355] *Ibid.*

[356] EU Recommendation 2021/1700, Glossary and Section 2.3.5.

[357] EU Recommendation 2021/1700, Section 2.3.5.

[358] *Ibid.*

For instance, AI algorithms published in academic journals or presented at international conferences qualify, whereas proprietary space mission AI systems data does not. To be eligible to export without a license under the "public domain" exemption, the exporter must ensure the information is genuinely accessible without restriction or significant barriers to further dissemination:[359]

> "If a researcher refers to or integrates proliferation sensitive information from other sources already in the public domain, the research output does not automatically comprise controlled dual-use technology. The fact that such listed dual-use technology became available in the public domain without a license is a violation of export control regulations, but this violation cannot be attributed to the researcher under such circumstances."[360]

To elaborate on the definition of "in the public domain," "without restrictions" means access is not limited to some restricted group of persons.[361] Thus, if the owner or exporter of the information must make a decision regarding access by any particular individual or entity to the information, with some being excluded on a non-commercial basis, then the information is not freely available without restrictions and the information therefore cannot be considered to be in the public domain.[362]

The "in the public domain" de-control note applies to software and technology, irrespective of whether in tangible or intangible form, e.g., a computer chip vs. an electronic file for a semiconductor lithography mask work.[363]

Subcategory D "software" is defined under the EU dual-use export controls as "a collection of one or more "programs" or "microprograms" fixed in any tangible medium of expression" (which includes computer hard drives and other digital memory storage).[364]

Subcategory E "technology" means "specific information necessary for the 'development,' 'production' or 'use' of goods."[365] This information is typically considered to be either "technical data" or "technical assistance."[366]

Only listed dual use "software" or "technology" can benefit from the respective decontrols included in the Software and Technology Notes. It is therefore of critical importance to determine whether a research output, and if so which specific parts, meet the technology control entry in subcategory E and satisfy the definition of technology.[367]

---

[359] *Ibid.*

[360] *Ibid.*

[361] *Ibid.*

[362] *Ibid.*

[363] *Ibid.*

[364] *Ibid.*

[365] *Ibid.*

[366] *Ibid.*

[367] *Ibid.*

This can require extensive documentation, comprehensive internal reviews and regular consultations with export control authorities to ensure compliance,[368] especially when considering space applications for cutting-edge AI technologies, which are advancing exponentially and not yet fully understood.

De-control notes for "software" or "technology" are included in the Nuclear Technology Note (NTN), the Nuclear Software Note (NSN), the General Technology Note (GTN), and the General Software Note (GSN) in Annex I to the EU dual-use Regulation.[369] These de-control notes only apply to listed software and technology, not to listed goods such as equipment, materials, samples and components.[370]

The GTN contains an exemption for the minimum necessary information for patent applications.[371] This minimum information needed to submit a patent application is thus exempt from export controls.[372] This de-control makes no distinction between national, EU or international patent applications, and once the patent information is published in the public domain, it is no longer subject to export controls.[373]

To summarize, for listed dual-use software, except for category 0, there are two main de-controls possible: software that is "generally available to the public" and software that is already "in the public domain."

For listed dual-use technology there are three de-controls possible: "technology" that is the result from "basic scientific research," "technology" that is already "in the public domain" and the minimum necessary information for patent applications (inapplicable for category 0 nuclear technology).[374] This highlights the disparate requirements between patent protection for intellectual property and export control licensing requirements for technology products.

For AI applications in space, de-controls described above may apply to AI software deployed on satellites or other spacecraft, which poses additional challenges in distinguishing between controlled and decontrolled technology within a rapidly evolving technological landscape. For example, open-source AI software for satellite image processing may be made publicly available within certain constraints, and research on AI navigation systems may fall under basic scientific research.

Further, legal rights such as copyrights do not affect eligibility for de-control under export control laws, i.e., copyright restrictions do not remove "technology" or "software" from being "in

---

[368] *Ibid.*

[369] EU Regulation 2021/821, Annex I, Nuclear Technology Note, Nuclear Software Note, and General Technology Note; EU Recommendation 2021/1700, Sections 2.3.4 and 2.3.5; "EU Compliance Guidance for Research Involving Dual-use Items."

[370] EU Regulation 2021/821; EU Recommendation 2021/1700, Section 2.3.5; "EU Compliance Guidance for Research Involving Dual-use Items."

[371] EU Regulation 2021/821, Annex I, General Technology Note; EU Recommendation 2021/1700, Sections 2.3.4 and 2.3.5; "EU Compliance Guidance for Research Involving Dual-use Items."

[372] EU Regulation 2021/821; EU Recommendation 2021/1700, Sections 2.3.4 and 2.3.5; "EU Compliance Guidance for Research Involving Dual-use Items."

[373] *Ibid.*

[374] EU Recommendation 2021/1700, Section 2.3.5.

the public domain."[375] For instance, requesting a fee for access or requiring registration prior to access, so long as every person is allowed to pay the fee or register, is not incompatible with de-control eligibility.[376]

For AI applications in space, software could include AI programs that control satellite operations, while technology could encompass computer processors and algorithms for data analysis, or the hardware specifications for the machine operating the AI control system.

Researchers developing AI applications for space should clearly differentiate between what is covered by export controls and what is governed by intellectual property laws. They should also establish clear protocols for sharing research internationally, such as securing appropriate licenses and conducting thorough compliance checks to avoid inadvertent violations.

As indicated above, "technology" means specific information necessary for the "development," "production" or "use" of goods, though exactly what is covered by "specific" information necessary for these purposes is not defined.[377]

The following non-exhaustive examples comprise types of information that is typically considered to lack sufficient detail and specifications to fall under the definition of technology: 1) sales brochures, catalogs and excerpts thereof, which, in their respective form, are intended or may be intended for an indefinite number of interested parties and which are made available to them without individual changes to the contents; 2) schematic diagrams, block diagrams, process diagrams lacking detailed data; and 3) technical performance data and key performance indicators.[378]

For instance, a block diagram of an AI system architecture, function, and/or operation may lack sufficient detail so as to be exempt from export controls, whereas the source code for the AI algorithms employed within the system may be controlled.

4.2 *United States Export Control Exemptions*

"Both the munitions and dual-use export control systems of the United States allow for license exemptions (or exceptions) when the government has determined that the particular item, value, end-use and end-user do not represent a risk of sufficient threat to require an export license."[379]

Export control regulations provide exceptions, exclusions, and exemptions regarding export control regulations, including the: 1) Fundamental Research Exclusion (FRE); 2) Public Domain/Publicly Available Exclusion; 3) Temporary Imports, Exports, and Re-exports; 4) Educational Instruction Exclusion; and 5) Baggage (BAG) License Exemptions.[380]

---

[375] EU Recommendation 2021/1700, Section 2.3.5; "EU Compliance Guidance for Research Involving Dual-use Items."

[376] *Ibid.*

[377] EU Recommendation 2021/1700, Section 2.3.4; "EU Compliance Guidance for Research Involving Dual-use Items."

[378] *Ibid.*

[379] "Overview of U.S. Export Control System," Export Control and Related Border Security Program, US Department of State, 2011, https://2009-2017.state.gov/strategictrade/overview/index.htm.

[380] "Export Control - Exclusions, Exemptions, and Exceptions."

The U.S. Educational Instruction Exclusion operates similarly to the analogous EU exclusion of course curricula and subject matter content, and the first three exclusions in the list are further elaborated, below.[381]

For example, universities and government labs can benefit from the FRE in circumstances where the research is eligible and immediate commercial applications are not enabled by the research. Similarly, research that is publicly available or which comprises materials typically included in educational instruction are likewise exempt from licensing.

Often, within the contexts of the practice of academic research and technology transfer, even research that involves proof of principle or bench-scale prototypes can benefit from fundamental research and other exclusions.

Whereas, when a commercial entity takes a license to commercialize that same technology, its contributions to R&D and advancement of the technology are subject to export controls. This can make things especially challenging for would-be academic entrepreneurs when they are also foreign nationals, especially those from sanctioned countries.

With respect to AI systems, including AI systems for space applications, the demarcation between research and commercial-ready product can be challenging to pinpoint in that as soon as an AI is trained to adequately begin performing the functions required for a commercial implementation, much of the details remaining for successful commercialization are on the business side rather than technical development.

4.2.1 Fundamental Research Exclusion

Research comprising "Fundamental Research" as defined by Export Control Regulations is eligible for the FRE, which permits results of this type of research to be shared without the need for an export license, even if they relate to items or technologies that are otherwise controlled.[382] However, although the results of this type of research may be exempt under export controls, the methods and equipment used to produce the research results do not benefit from this exclusion.[383]

For instance, with respect to AI systems development, U.S. export controls focus on the hardware needed to train large neural networks, whereas the academic research that employs such export-controlled hardware to develop an AI NN model configured to detect, monitor, learn, predict, and allocate wireless spectrum usage may be eligible for exclusion from export control licensing.

Further, such researchers need to be wary of the changing export control landscape as evidenced by recent BIS updates to the CCL (e.g., covering geospatial AI). For research to be eligible under the FRE, it must meet all of the following criteria:

- satisfy the definition of "Fundamental Research" as described by relevant US Government Agencies, as different agencies provide varying definitions of Fundamental Research;

---

[381] *Ibid.*
[382] *Ibid.*
[383] *Ibid.*

- comprise basic or applied research in science and engineering NSDD-189, National Policy on the Transfer of Scientific, Technical and Engineering Information;
- conducted at a US based "accredited institution of higher education" (EAR) or place of "higher learning" (ITAR);
- conducted with the intent to publish the results and share broadly within the scientific community; and
- is not publication-restricted (either by written agreement or by informal understanding) for proprietary reasons or specific national security controls, or subject to specific US Government access and dissemination controls.[384]

The ITAR Exemption for fundamental research is not as broad as the EAR FRE.[385] ITAR and the EAR both govern physical export and sharing (e.g., online) of software and encryption, whether independently developed or obtained from another party.[386]

To adhere to required compliance under ITAR Fundamental Research on encryption software, researchers developing ITAR-related encryption software for fundamental research must make their code publicly available online to show publication.[387]

Further, the code must be freely downloadable by all interested members of the scientific community at no charge and without knowledge by whom or from where the data is being downloaded.[388] No logins, passwords, or other authentication are allowed, as the government could potentially view any of these items as an "access control" that destroys the ability to characterize as fundamental research.[389]

And of course, these same restrictions for distribution of open source apply to open-source AI models that may implement covered encryption protocols, e.g., for space-based communications, data collection, data storage, etc.

Consequently, ensuring compliance with ITAR can be challenging, especially when dealing with advanced AI encryption technologies and protocols used in space, e.g., for data security and/or compression in storage, communications, and processing. This emphasizes the need for clear guidelines and processes to verify that all publicly shared AI encryption software meets ITAR requirements, especially given the potential for these technologies to be considered dual use.

Some of the most frequent FRE invalidators causing the loss of FRE status for research results deriving from Fundamental Research include that it:

- "Contains publication/dissemination or access restrictions (i.e. your project is not subject to publication approval by sponsors or the government) even if funded by the US government, private, or nonprofit sponsor.

---

[384] ITAR, Part 123.16; "Export Control - Exclusions, Exemptions, and Exceptions."

[385] "Export Control - Exclusions, Exemptions, and Exceptions."

[386] "Controlled Software and Encryption," Export Control, Florida International University, December 21, 2020, https://exportcontrol.fiu.edu/export/topics/controlled-software-and-encryption/.

[387] *Ibid.*

[388] *Ibid.*

[389] *Ibid.*

- Involves proprietary research and from industrial development, design, production, and product utilization, the results of which ordinarily are restricted for proprietary or national security reasons. Which is a reason for a restriction on publication.
- Includes restriction on publication to ensure the patent rights of the sponsor.
- Sponsor includes requirement to approve publication.
- Forbids participation of foreign national in research activity.
- Contains citizenship-based restrictions on who may be included on the research team).
- Involves physical shipment of goods.
- Involves use of equipment controlled by ITAR (defense related technology).
- Some types of Specific Software (e.g. encryption).
- Contains certain types of encryption source code.
- Research which is not intended for publication.
- Research includes work done outside the US.
- Any research which involves "development" as defined in 15 CFR 772 (EAR 2016) defines development as being "related to all states prior to serial production, such as: design, design research, design analyses, design concepts, assembly and testing of prototypes, pilot production schemes, design data, process of transforming design data into a product, configuration design, integration design, layouts.
- Involves any agents or the toxins above the minimum allowable amounts, as detailed by the U.S. Government Policy for Oversight of Life Sciences Dual Use Research of Concern for a well-defined subset of life sciences research that involves 15 agents and toxins and seven categories of experiments.
- Contains confidential technical information received from an outside party, such as a government or industry sponsor.
- Deals with military or dual-use technologies listed within the export control regulations.
- Research results will only be shared at a closed conference instead of an open conference: A conference or gathering is "open" if all technically qualified members of the public are eligible to attend and attendees are permitted to take notes or otherwise make a personal record of the proceedings and presentations."[390]

As discussed, federal and industry sponsored research conducted at universities provides ample scenarios in which AI research involving space applications may trigger these invalidators.

Relying on appropriate institutional administrative and support resources, researchers can navigate these issues by taking actions to ensure openness in research dissemination and avoiding restrictions on foreign national participation. Working closely with institutional export control

---

[390] EAR 15 CFR 734.8; "Export Control - Exclusions, Exemptions, and Exceptions."

experts is the best way to ensure compliance and avoiding potentially criminal penalties and jail time.

4.2.2 Public Domain/Publicly Available Exclusion

ITAR and the EAR exclude information that is published and technology and software that is generally available to the public[391] and these excluded items may be freely shared with foreign nationals.[392]

"This exclusion does not apply to encrypted software, to information if there is reason to believe it may be used for weapons of mass destruction, or where the U.S. government has imposed access or dissemination controls as a condition of funding."[393]

"Public Domain is defined as information that is published and generally accessible to the public:

- through sales at newsstands and bookstores;
- through subscriptions available without restriction to anyone who may want to purchase the published information;
- through second class mailing privileges granted by the U.S. Government;
- at libraries open to the public or from which the public can obtain documents;
- through patents [and patent applications] available at any patent office;
- through unlimited distribution at a conference, meeting, seminar, trade show or exhibition that is generally accessible to the public and is in the United States;
- through public release (i.e., unlimited distribution) in any form (not necessarily published) after approval by the cognizant US government department or agency; and
- through fundamental research."[394]

Specifically for the use of AI in space, it is important to ensure AI software and technology intended for public distribution, dissemination, or other sharing with members of the public does not fall into controlled categories.

Practical steps include reviewing AI technology against export control lists (EAR, ITAR), training researchers on export control regulations, establishing internal compliance programs with regular audits, consulting export control experts before sharing technology, implementing a pre-publication review process, developing technology control plans, seeking necessary clearances or licenses, staying updated with changes in export control regulations, ensuring technology meets criteria for unrestricted access, and maintaining detailed records of all compliance checks and consultations

---

[391] ITAR, 22 CFR §120.11(a); EAR, 15 CFR §734.9; "Export Control - Exclusions, Exemptions, and Exceptions."

[392] "Export Control - Exclusions, Exemptions, and Exceptions."

[393] *Ibid.*

[394] *Ibid.*

4.2.3 Export License Exceptions

Even with a relevant license exception identified, there may still be required approvals and/or documentation. For each type of license exception various compliance actions may need to be taken, such as:

- "GOV Exception (Governments, International Organizations, International Inspections Under the Chemical Weapons Convention, and the International Space Station) requires an Electronic Export Information (EEI) filing of the ECCN and also a Destination Control Statement to be put into the commercial export documents;
- LVS Exception (Shipments of Limited Value) requires proof that the limited value of the yearly shipments has not been exceeded;
- TMP Exception (Temporary Imports, Exports, Re-exports, Transfers) requires items to be securely controlled and returned or destroyed; and
- STA (Strategic Trade Authorization) permits transfer of items to another consignee, but the original and each consecutive consignee must sign an agreement attesting to all the required conditions to allow the transfer. STA also requires an application (via SNAP-R) to be made to determine if items are eligible."[395]

EAR Section 740 describes the documentation, agreements, and records maintenance requirements for the various license exceptions.[396] As an example, an AI component designed for satellite data analysis might qualify for the TMP (Temporary Imports, Exports, Re-exports, Transfers) exception if it is being temporarily exported for testing and will be returned to the original country.

To ensure compliance, thorough documentation is crucial. This includes detailed records of the AI component's purpose, duration of export, and secure handling procedures. Implementing rigorous compliance checks before, during, and after the export process can prevent unauthorized use and ensure adherence to export control regulations.

De Minimis Rule for Spacecraft

> "Integrating U.S. components into a foreign made spacecraft does not cause that spacecraft to fall under U.S. export controls if certain conditions are met. Under the de minimis rule, the foreign-made spacecraft is not subject to U.S. jurisdiction under the EAR as long as: (1) the value of the controlled U.S. content comprises 25% or less of the total value of the item and is not destined for a country subject to U.S. arms embargo; or (2) the value of the controlled U.S. content comprises 10% or less of the total value of the item if it is destined for a country subject to a U.S. embargo."[397]

---

[395] *Introduction to U.S. Export Controls for the Commercial Space Industry.*

[396] *Ibid.*

[397] *Introduction to U.S. Export Controls for the Commercial Space Industry*, *See also* "Guidelines regarding the de minimis rules can be found at:

Consider a scenario wherein a European company collaborates with a US company to develop an AI-powered component for a satellite. If the US content (e.g., AI software or hardware) comprises less than 10% of the total value of the satellite and is not destined for a country subject to a U.S. arms embargo, the de minimis rule may apply.

To ensure compliance, the European company should accurately assess the value of the U.S. content by evaluating the cost of the U.S. AI component relative to the total cost of the satellite. This involves detailed documentation of all components and their respective values, consultations with the U.S. supplier for accurate valuations, and regular compliance audits to verify adherence to the de minimis rule.

Special Considerations for Launch Vehicles

> "By U.S. statute, the launch of a launch vehicle, reentry vehicle, or payload is not considered an export, if the launch occurs from the United States (or U.S. territory). Therefore, a launch vehicle does not require an export license from the Department of State or Department of Commerce when built and launched in the United States. However, the operation of the commercial launch vehicle generally requires a launch license or permit from the Federal Aviation Administration (FAA). FAA licensing requirements are in place to protect public health and safety, property, U.S. national security and foreign policy interests, and international obligations of the U.S.
>
> The launch of a U.S. launch vehicle outside the United States requires ITAR authorization in compliance with MTCR controls. The applicable MTCR controls establish a presumption of denial for exports of U.S. launch vehicles, even to other launching states. A U.S. company wishing to apply for a launch vehicle export license must work with the State Department and the government of the recipient nation to establish a Technology Safeguards Agreement (TSA) covering the export."[398]

This highlights the fact that regulations impact the export and launch of AI-equipped satellites as well as launch vehicles. For example, if an AI-equipped satellite is manufactured in the U.S. but launched from a foreign site, a TSA and export licenses will be necessary due to the ITAR and MTCR controls. Detailed documentation of all AI components and software is needed to verify that they meet de minimis thresholds and do not require additional licensure.

---

www.bis.doc.gov/forms-documents/pdfs/1382-de-minimis-guidance/fle. BIS also provides a tool to assist you in determining whether a non-U.S.-made item located outside the United States is subject to the EAR and can be found at: www.bis.doc.gov/de-minimis-direct-product-rules-decision-tool."

[398] *Introduction to U.S. Export Controls for the Commercial Space Industry.*

Conclusions

Internationally, export control laws are implemented at both national and supra-national levels, such as within the EU. These laws include various exemptions, exceptions, and exclusions. While individual states may have unique export control requirements, international harmonization is achieved through several key multilateral export control regimes, including the Australia Group, the Wassenaar Arrangement, the Nuclear Suppliers Group, and the Missile Technology Control Regime.

AI systems, including those deployed in space, may be subject to export control restrictions and exemptions. Additionally, space technologies often fall under export controls by default, making it critical to consider these controls and exemptions for AI systems intended for space deployment.

To further complement the harmonization of export controls under the Wassenaar Arrangement, member states are increasingly recognizing the need to regulate emerging technologies, such as AI, even though the WA does not currently mandate this. Many have already initiated or are planning to implement stricter export controls on AI technologies that meet specific criteria, reflecting the growing awareness of their strategic importance.

Because Export Controls rely on descriptions of products and technologies, they can be somewhat less effective against highly disruptive technological advancements. Further, due to the fact that disruptive AI developments can be embodied across a spectrum of technologies and products ranging from the hardware to the firmware and software aspects of AI systems.

In the immediate future, it's crucial to conduct a thorough evaluation of AI space systems, focusing on the quality and sensitivity of their training data, to clearly differentiate components that should fall under the military list from those classified as dual-use. Furthermore, assembling teams of AI and space experts would pave the way for smoothly integrating AI-related controls into existing regulatory frameworks, enabling a forward-thinking approach to potential challenges.

Looking ahead, creating a detailed risk matrix for AI space systems will be essential for identifying and addressing potential vulnerabilities. Establishing clear and precise regulations for the export of AI-related hardware, software, and technologies will play a key role in ensuring consistent and effective oversight. These measures will not only bridge current regulatory gaps but also prepare frameworks to evolve alongside advancements in AI and space technologies.

Section 3: Overview of Procurement Laws Regulating the Use of AI for Space Applications

Introduction

This section provides a comprehensive examination of the complexities of how procurement laws and regulations impact the integration of AI in space technologies. This section also addresses the technical challenges faced by companies developing AI systems for space regarding procurement processes relevant to AI systems for outer space applications.

This section further highlights the importance of technical specifications, ethical guidelines, and performance metrics, referencing procurement standards under development by the IEEE and Executive Order 14110 on AI procurement in the United States. As AI continues to revolutionize satellite operations, data analysis, and autonomous decision-making in space missions, it becomes critical to develop international laws and standards that ensure the responsible use of AI in this domain.

The World Trade Organization's (WTO) Agreement on Government Procurement (GPA) and the European Union's Procurement Directives serve as key reference points for developing equitable and transparent procurement standards.

Moving into more nuanced procurement issues, this section also highlights the importance of standardizing procurement practices for AI systems in space.

Special attention is given to the challenges faced by companies in complying with dual-use regulations while navigating the complex landscape of AI-specific procurement processes. This analysis is reinforced by discussions on the role of procurement agreements in shaping international laws, as seen in the GPA's focus on non-discrimination and fair competition, which are critical for ensuring that AI systems in space are acquired ethically and transparently.

In summary, this section aims to analyze the standardization of AI technologies in space by examining the interplay between procurement laws and standards with government and commercial efforts to deploy AI systems in space. In reviewing relevant international agreements and regulations, this section underscores the importance of creating a harmonized legal framework that not only facilitates the advancement of AI in space but also ensures its responsible and secure deployment.

Articles and documents such as the OMB memorandum Advancing Governance, Innovation, and Risk Management for Agency Use of AI are discussed in relation to how procurement strategies can align with emerging AI practices. The section also elaborates on the need for comprehensive technology assessments, both in the European Union and the United States, to ensure that AI components used in space missions meet dual-use regulatory requirements, referencing relevant EU and US export control guidelines and exceptions.

This section draws on a range of articles and documents, including key procurement laws, standards, guidelines, and government memoranda to provide a detailed understanding of how procurement regulations intersect with AI technologies in space exploration.

1. *Navigating Procurement Technical Specifications*

In the field of procurement, various types of technical specifications exist and organizations use different kinds of technical specifications (also referred to simply as "specifications") to ensure procurement of goods or services that will meet the specific needs of the purchaser.[399] They can help buyers evaluate and compare similar products, and they also serve to help protect the respective rights of the purchaser and seller.[400]

Technical specifications detail the required functionalities, performance characteristics, qualities, and other requirements of a given product, system, or service.[401] And they can cover everything from technical hardware specs and firmware and/or software required for system operation, to the fonts, graphics and colors for website user experience.[402]

Specifications should be written in clear and concise language to make them more easily understood by all stakeholders, including purchasing, engineering, compliance and legal teammates.[403] Consequently, the specification must clarify all considerations that purchasers may need to evaluate in making their purchasing decision.[404] This typically includes covering all relevant aspects of "the features, performance requirements, material and finish requirements, as well as any necessary installation or usage instructions."[405]

Specifications can be placed into three general categories: functional, technical, and administrative.[406] Functional specifications define the necessary functions or features of products and systems at a high level.[407]

Administrative specifications (e.g., schedule specifications) describe other things like product packaging and labeling and when the product or system must be completed and delivered.[408] And technical specifications describe in fine detail the required properties and capabilities of a product or system such as its size, weight, operation, performance, components and availability.[409]

Additional types of specifications can include performance, environmental, health, and safety (EHS), interface requirements, i.e., compatibility with existing products or systems, branding, product packaging, look and feel, user interface, user experience, financial, legal, and compliance/ethical.[410]

---

[399] "What Are the Types of Specification in Procurement?," Oboloo FAQ's, Oboloo, November 27, 2023, https://oboloo.com/what-are-the-types-of-specification-in-procurement/.

[400] *Ibid.*

[401] "What Is Technical Specification in Procurement?"

[402] *Ibid.*

[403] *Ibid.*

[404] *Ibid.*

[405] "What Is Technical Specification in Procurement?"

[406] "What Are the Types of Specification in Procurement?"

[407] *Ibid.*

[408] *Ibid.*

[409] *Ibid.*

[410] "What Is Technical Specification in Procurement?"

Functional specifications describe the requirements of a system and/or its subsystem(s), though not from the design perspective.[411]

Functional specifications typically describe the following:

- systems and/or their subsystems;
- necessary features and functions;
- expected user interface and (inter-)operability;
- required and desired performance characteristics; and
- any test protocols.[412]

Interface requirements specifications define interfaces between different systems and help ensure they will be compatible and work together.[413] Whether created during the design or procurement process, interface specifications are most often written in an interface definition language (IDL) format.[414] IDLs enable developers to better understand and test system functionalities by constructing visual depictions of required system actions and interactions.[415]

Contract specifications provide detailed information regarding services offered by contractors and typically define criteria for validating performance as well as requirements for maintaining legal and regulatory compliance.[416]

System specifications describe functional and nonfunctional system requirements, and are sometimes drafted as a requirements document, design document, or test plan.[417] System specification requirements documents describe functional and nonfunctional system requirements according to what the system must do, how it must do it, and when it must do it.[418]

System specification design documents describe the physical structure and layout of a system and often also detail the components, their respective interrelated functions, and how they operate together as in an integrated system.[419] System specification test plans define how to test functional and nonfunctional system requirements and typically details specific tests that must be performed, when to conduct them, and how to execute them.[420]

Documentation specifications provide a formal, detailed description of the procurement process itself and can cover specific requirements for bidding documents, negotiation and performance timelines, as well as procedures for contract administration and dispute resolution.[421]

Documentation specifications enable all parties engaged in a given procurement process to ensure they have clear mutual understandings regarding what needs to be accomplished, when it needs to be completed, and what shall constitute satisfaction of the desired procurement

---

[411] "What Are the Types of Specification in Procurement?"

[412] *Ibid.*

[413] *Ibid.*

[414] *Ibid.*

[415] *Ibid.*

[416] *Ibid.*

[417] *Ibid.*

[418] *Ibid.*

[419] *Ibid.*

[420] *Ibid.*

[421] *Ibid.*

transaction. This enables procurement contract officials to select suppliers and award contracts more efficiently with fewer potential uncertainties and surprises.[422]

Documentation specifications come in several different typical varieties, including requirements specifications, Tender documents, RFQs, and Bid documents, each of which have their own respective advantages and drawbacks that may better fit the specific needs of a given procurement process.[423]

Requirements specifications typically detail the mandatory requirements for a system or product. Tender documents generally outline the kinds of suppliers that should be considered and acceptable price ranges.[424]

In contrast, RFQs enable more creative bidding by specifying desired features and functions only at a high level, rather than defining specific requirements for how the product or system achieves those required features and functions. Bid documents give specific details regarding a proposal submitted by a bidder, including cost quotes and delivery schedules.[425]

Drafters of technical specifications can use many different formats, such as drawings, diagrams, text files, e.g., with embedded images or links, or even as software applications.[426] Irrespective of form, the specification must adequately describe the required features of the product, system, or service to allow sufficiently qualified personnel to build or practice the given product, service, or system.[427]

In order to create a comprehensive technical specification, it is important to have accurate information about the project. This includes knowing the target market, what features are required, and what functionality is desired.[428] Technical specifications will typically include detailed descriptions of product or system requirements for:

- purpose or function;
- inputs and outputs;
- performance characteristics;
- material properties;
- interface protocols;
- hardware and software design; and
- security.[429]

In completing the final draft of the technical specification, it is important to get feedback from all relevant stakeholders, especially prospective purchasers to help ensure satisfaction of their needs. After all stakeholders have agreed to the specification contract preparations can begin,

---

[422] *Ibid.*

[423] *Ibid.*

[424] *Ibid.*

[425] *Ibid.*

[426] "What Is Technical Specification in Procurement?"

[427] *Ibid.*

[428] *Ibid.*

[429] *Ibid.*

which often benefit from the enhanced clarity developed during specification drafting and review.[430]

Following completion of the final draft technical specification and accession by all parties to the document, subsequent steps in the procurement process may include:

- publish the technical specification document;
- commence procurement contract negotiations;
- hire employees and/or contractors to build the product or system or perform the service based on the specification;
- run a proof of concept or pilot to ground-test the specification and update as may be needed;
- ensure the product, system or service meets customer and specification requirements; and
- train employees and/or contractors how to use the product or system and to maintain compliance with appropriate specifications.[431]

Irrespective of the type of specification chosen for a given procurement transaction, ensuring a clear mutual understanding between the parties and all relevant stakeholders of the details included in the specification helps guide the procurement process towards success and the ultimate goal of concluding a contract to satisfy the procurement need.[432] Adhering to standards and guidelines and preparing dedicated specification documents for each stage of the procurement process ultimately helps organizations maximize their efficiencies and minimize costs and other potential problems.[433]

2. *International Procurement Frameworks*

2.1 *World Trade Organization Agreement on Government Procurement*

The GPA is a plurilateral agreement under the overarching WTO treaty regime and not all WTO member states are parties to the GPA.[434] The GPA currently has some 22 parties including 49 WTO members and several other WTO members are actively working through accession negotiations.[435] Another 35 members and observers of the WTO take part in the Committee on Government Procurement as observers.[436]

GPA seeks to open the markets of its member states between them for government procurement of commercial goods and services.[437] Based on multiple previous negotiations, members of GPA have mutually opened over US$ 1.7 trillion to international competition

---

[430] *Ibid.*
[431] *Ibid.*
[432] *Ibid.*
[433] *Ibid.*
[434] "Agreement on Government Procurement," Agreement on Government Procurement, World Trade Organization, 2024, https://www.wto.org/english/tratop_e/gproc_e/gp_gpa_e.htm.
[435] *Ibid.*
[436] *Ibid.*
[437] *Ibid.*

including commercial suppliers from other GPA member states offering goods and services, including construction services.[438]

The GPA 2012 is based on the principles of non-discrimination, transparency and procedural fairness, and contains the following main elements:

- guarantees of national treatment and non-discrimination for the suppliers of parties to the Agreement with respect to procurement of covered goods, services and construction services as set out in each party's schedules;
- provisions regarding accession to the Agreement and the availability of special and differential treatment for developing and least-developed countries;
- detailed procedural requirements regarding the procurement process designed to ensure that covered procurement under the Agreement is carried out in a transparent and competitive manner that does not discriminate against the goods, services or suppliers of other parties, avoids conflicts of interest and prevents corrupt practices;
- additional requirements regarding transparency of procurement-related information (e.g. relevant statutes and regulations) provisions regarding modifications and rectifications of parties' coverage commitments;
- requirements regarding the availability and nature of domestic review procedures for supplier challenges which must be put in place by all parties to the Agreement; and
- provisions regarding the application of the WTO Dispute Settlement Understanding in this area;
- a "built-in agenda" for improvement of the Agreement, extension of coverage and elimination of remaining discriminatory measures through further negotiations.[439]

The two primary components of GPA include the Agreement text and the market access schedules submitted by the parties.[440] The text of the Agreement establishes general principles and detailed procedural requirements GPA parties must adhere to for any covered procurement activities.[441] The Agreement text outlines rules that require member states to ensure fair, open, and transparent competition for government procurements.[442]

Though, the rules set forth in the text do not apply to all government procurement activities conducted by the member states.[443] Instead, the coverage schedules of the respective member states dictate whether a specific procurement activity meets the criteria defined by that member state to be governed by the Agreement.[444]

---

[438] *Ibid.*

[439] "Text of the Agreement," Agreement on Government Procurement, World Trade Organization, 2024, https://www.wto.org/english/tratop_e/gproc_e/gpa_1994_e.htm.

[440] "Agreement on Government Procurement"; "Coverage Schedules," Agreement on Government Procurement, World Trade Organization, 2024, https://www.wto.org/english/tratop_e/gproc_e/gp_app_agree_e.htm.

[441] "Text of the Agreement."

[442] "Agreement on Government Procurement"; "Text of the Agreement."

[443] "Agreement on Government Procurement"; "Coverage Schedules."

[444] *Ibid.*

GPA only regulates government procurement activities conducted by covered entities, e.g., specified government agencies, that seek to purchase listed goods and services, including construction services, of a value surpassing a set threshold amount.[445] Members are empowered to enforce GPA via two main tools, including national level domestic review and WTO dispute settlement at the international level.[446]

The coverage schedules of parties are an integral part of the GPA and are contained in Appendix I to the Agreement.[447] The schedule of each party contains several annexes which define the concerned party's commitment with respect to four dimensions of coverage:

1) the procuring entities covered by the Agreement
2) the goods, services and construction services covered by the Agreement
3) the threshold values above which procurement activities are covered by the Agreement
4) exceptions to the coverage.[448]

Only the procurement activities carried out by a covered entity purchasing covered goods, services or construction services of a contract valued above the relevant threshold, and not specifically exempted in the notes to the schedules, are subject to the rules of the GPA 2012.[449]

Under the GPA 2012, the schedule of each party contains seven annexes:

Annex 1: central government entities
Annex 2: sub-central government entities
Annex 3: other entities
Annex 4: goods
Annex 5: services
Annex 6: construction services
Annex 7: general notes.[450]

To facilitate the handling of individual parties' schedules and to take into account modifications that have been certified since that date, loose-leaf schedules have been prepared. These reflect the parties' current legal binding coverage commitments.[451]

The e-GPA gateway of the WTO provides up-to-date information on coverage, considering certified modifications, including on the applicable thresholds, under the GPA 2012. Parties' coverage schedules under the GPA 2012 are maintained by the WTO.[452]

Under subsection (u) of Article 1 of the GPA:

“technical specification means a tendering requirement that:

---

[445] “Agreement on Government Procurement”; “Coverage Schedules.”
[446] “Agreement on Government Procurement.”
[447] “Coverage Schedules.”
[448] *Ibid.*
[449] “Agreement on Government Procurement”; “Coverage Schedules.”
[450] “Coverage Schedules.”
[451] *Ibid.*
[452] “Coverage Schedules”; “Government Procurement - the Plurilateral Agreement on Government Procurement (GPA),” WTO, n.d., https://www.wto.org/english/tratop_e/gproc_e/gp_app_agree_e.htm.

> i. lays down the characteristics of goods or services to be procured, including quality, performance, safety and dimensions, or the processes and methods for their production or provision; or
>
> ii. addresses terminology, symbols, packaging, marking or labelling requirements, as they apply to a good or service."[453]

GPA Article II Scope and Coverage stipulates the application of the GPA:

> "1. This Agreement applies to any measure regarding covered procurement, whether or not it is conducted exclusively or partially by electronic means.
>
> 2. For the purposes of this Agreement, covered procurement means procurement for governmental purposes:
>
> (a) of goods, services, or any combination thereof:
>
> i. as specified in each Party's annexes to Appendix I; and
>
> ii. not procured with a view to commercial sale or resale, or for use in the production or supply of goods or services for commercial sale or resale;
>
> (b) by any contractual means, including: purchase; lease; and rental or hire purchase, with or without an option to buy;
>
> (c) for which the value, as estimated in accordance with paragraphs 6 through 8, equals or exceeds the relevant threshold specified in a Party's annexes to Appendix I, at the time of publication of a notice in accordance with Article VII;
>
> (d) by a procuring entity; and
>
> (e) that is not otherwise excluded from coverage in paragraph 3 or a Party's annexes to Appendix I.
>
> 3. Except where provided otherwise in a Party's annexes to Appendix I, this Agreement does not apply to:
>
> (a) the acquisition or rental of land, existing buildings or other immovable property or the rights thereon;
>
> (b) non-contractual agreements or any form of assistance that a Party provides, including cooperative agreements, grants, loans, equity infusions, guarantees and fiscal incentives;
>
> (c) the procurement or acquisition of fiscal agency or depository services, liquidation and management services for regulated financial institutions or services related to the sale, redemption and distribution of public debt, including loans and government bonds, notes and other securities;
>
> (d) public employment contracts;
>
> (e) procurement conducted:

---

[453] "Agreement on Government Procurement 2012 and Related WTO Legal Texts," World Trade Organization, 2014, https://www.wto.org/english/docs_e/legal_e/rev-gpr-94_01_e.pdf.

i. for the specific purpose of providing international assistance, including development aid;
ii. under the particular procedure or condition of an international agreement relating to the stationing of troops or relating to the joint implementation by the signatory countries of a project; or
iii. under the particular procedure or condition of an international organization, or funded by international grants, loans or other assistance where the applicable procedure or condition would be inconsistent with this Agreement.

4. Each Party shall specify the following information in its annexes to Appendix I:

(a) in Annex 1, the central government entities whose procurement is covered by this Agreement;
(b) in Annex 2, the sub-central government entities whose procurement is covered by this Agreement;
(c) in Annex 3, all other entities whose procurement is covered by this Agreement;
(d) in Annex 4, the goods covered by this Agreement;
(e) in Annex 5, the services, other than construction services, covered by this Agreement;
(f) in Annex 6, the construction services covered by this Agreement; and
(g) in Annex 7, any General Notes.

5. Where a procuring entity, in the context of covered procurement, requires persons not covered under a Party's annexes to Appendix I to procure in accordance with particular requirements, Article IV shall apply mutatis mutandis to such requirements."[454]

GPA Article III Security and General Exceptions stipulates:

"1. Nothing in this Agreement shall be construed to prevent any Party from taking any action or not disclosing any information that it considers necessary for the protection of its essential security interests relating to the procurement of arms, ammunition or war materials, or to procurement indispensable for national security or for national defense purposes.

2. Subject to the requirement that such measures are not applied in a manner that would constitute a means of arbitrary or unjustifiable discrimination between Parties where the same conditions prevail or a disguised restriction on international trade, nothing in this Agreement shall be construed to prevent any Party from imposing or enforcing measures:

(a) necessary to protect public morals, order or safety;
(b) necessary to protect human, animal or plant life or health;
(c) necessary to protect intellectual property; or

---

[454] "Agreement on Government Procurement 2012 and Related WTO Legal Texts."

(d) relating to goods or services of persons with disabilities, philanthropic institutions or prison labor."[455]

Having examined the fundamental aspects of the WTO Agreement on Government Procurement (GPA), we will now shift our attention to the IEEE draft P3119 standard on AI procurement, the US procurements regime relevant to AI acquisitions, the European Union's Public Procurement Directive (Directive 2014/24/EU), and exemplary contract clauses for AI procurements.

2.2 *IEEE Draft Standard for the Procurement of AI and Automated Decision Systems*

Currently, the Artificial Intelligence Procurement Working Group of the IEEE Social Implications of Technology Standards Committee (SSIT/SC) is working to develop the IEEE P3119 Standard for the Procurement of Artificial Intelligence and Automated Decision Systems for the IEEE Standards Association.[456]

With work starting in 2021 and due to be completed by December of 2025, this standard will create "a uniform set of definitions and a process model for the procurement of Artificial Intelligence (AI) and Automated Decision Systems (ADS)."[457] And it will aim to enable government entities to better serve the public interest by addressing important considerations including socio-techno-economic concerns, as well as responsible and open innovation issues.[458]

The process requirements for developing the standard include grounding procurement in the IEEE Ethically Aligned Design (EAD) principles and guidelines.[459] And further, collaboratively rearticulating the standard stages of procurement to include "problem definition, planning, solicitation, critical evaluation of technology solutions (e.g. impact assessments), and contract execution."[460]

The standard will cover AI procurement generally, and "also government in-house development and hybrid public-private development of AI and ADS as an extension of internal government procurement practices."[461]

3. ***US Procurements Regime Relevant to Acquisitions of AI***

3.1 *US Executive Order 14110 Mandated AI Procurement Standards Development*

OMB issued a request for information (RFI) on the responsible procurement of AI along with their release of the OMB Memorandum titled Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence (the "AI M-memo").[462]

---

[455] *Ibid.*

[456] "Standard for the Procurement of Artificial Intelligence and Automated Decision Systems," IEEE Standards Association, September 23, 2021, https://standards.ieee.org/ieee/3119/10729/.

[457] "Standard for the Procurement of Artificial Intelligence."

[458] *Ibid.*

[459] *Ibid.*

[460] *Ibid.*

[461] "Standard for the Procurement of Artificial Intelligence."

[462] "Request for Information: Responsible Procurement of Artificial Intelligence in Government," Federal Register, March 29, 2024,

Executive Order 14110, Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, charged OMB with creating an "initial means" within 180 days following issuance of the AI M-memo to ensure government procurement contracts for AI systems and services follow its guidance and meet objectives set forth in the Advancing American AI Act.[463]

> "Consistent with Section 7224(d)(1) of the AI Act, this "initial means" [developed by OMB] will at a minimum:
>
> - address protection of privacy, civil rights, and civil liberties;
> - address the ownership and security of data and other information created, used, processed, stored, maintained, disseminated, disclosed, or disposed of by a contractor or subcontractor on behalf of the Federal Government;
> - include considerations for securing the training data, algorithms, and other components of any artificial intelligence system against misuse, unauthorized alteration, degradation, or rendering inoperable; and
> - address any other issue or concern determined to be relevant by the Director to ensure appropriate use and protection of privacy and Government data and other information."[464]

The OMB RFI seeks information within two overarching themes—strengthening the AI marketplace and managing the performance and risks of AI.

> "Strengthening the AI Marketplace:
>
> 1) How may standard practices and strategies of Federal procurement, such as Statements of Objectives, Quality Assurance Surveillance Plans, modular contracts, use of contract incentives, and teaming agreements, as well as innovative procurement practices, such as those in the Periodic Table of Acquisition Innovations, be best used or adapted to reflect emerging practices in AI procurement? Are there additional materials or resources that OMB could provide to vendors or agencies to improve alignment between agency missions and technical requirements?
> 2) How can OMB promote robust competition, attract new entrants, including small businesses, into the Federal marketplace, and avoid vendor lock-in across specific elements of the technology sector, including data collectors and labelers, model developers, infrastructure providers, and AI service providers? Are there ways OMB can address practices that limit competition, such as inappropriate tying, egress fees, and self-preferencing?
> 3) Should the Federal Government standardize assessments for the benefits and trade-offs between in-house AI development, contracted AI development, licensing of AI-enabled software, and use of AI-enabled services? If so, how?

---

https://www.federalregister.gov/documents/2024/03/29/2024-06547/request-for-information-responsible-procurement-of-artificial-intelligence-in-government.

[463] *Ibid.*

[464] *Ibid.*

4) How might metrics be developed and communicated to enable performance-based procurement of AI? What questions should agencies be asking vendors to determine whether AI is already being used in performance-based services contracts?

Managing the Performance and Risks of AI:

5) What access to documentation, data, code, models, software, and other technical components might vendors provide to agencies to demonstrate compliance with the requirements established in the AI M-memo? What contract language would best effectuate this access, and is this best envisioned as a standard clause, or requirements-specific elements in a statement of work?
6) Which elements of testing, evaluation, and impact assessments are best conducted by the vendor, and which responsibilities should remain with the agencies?
7) What if any terms should agencies include in contracts to protect the Federal Government's rights and access to its data, while maintaining protection of a vendor's intellectual property?
8) What if any terms, including terms governing information-sharing among agencies, vendors, and the public, should be included in contracts for AI systems or services to implement the AI M-memo's provisions regarding notice and appeal (sections 5(c)(v)(D) and (E))?
9) How might agencies structure their procurements to reduce the risk that an AI system or service they acquire may produce harmful or illegal content, such as fraudulent or deceptive content, or content that includes child sex abuse material or non-consensual intimate imagery?
10) How might OMB ensure that agencies procure AI systems or services in a way that advances equitable outcomes and mitigates risks to privacy, civil rights, and civil liberties?"[465]

3.2 *The National AI Advisory Committee*

NAIAC was created in April 2022 by the William M. (Mac) Thornberry National Defense Authorization Act and is composed of leading experts across many fields within AI. Representing industry, academia, government, and non-governmental organizations, the group first convened as the committee May 2022.[466]

---

[465] *Ibid*.

[466] "National AI Advisory Committee," AI.gov, The White House, n.d., https://www.ai.gov/naiac/; Miriam Vogel et al., "RECOMMENDATION: Provide Authority and Resources to Promote Responsible Procurement Innovation for AI at Government Agencies," May 2024, https://ai.gov/wp-content/uploads/2024/06/RECOMMENDATION_Provide-Authority-and-Resources-to-Promote-Responsible-Procurement-Innovation-for-AI-at-Government-Agencies.pdf.

NAIAC has five active primary working groups: 1) AI Education & Awareness; 2) International Collaboration; 3) AI Futures Preparedness, Opportunities, and Competitiveness; 4) Safety, Trust, and Rights; 5) AI in Work and the Workforce.[467] The NAIAC Subcommittee on AI and Law Enforcement has another four active working groups: 1) Processes; 2) Identification and Surveillance Set; 3) Accountability AI; and 4) Officer Training.[468]

NAIAC advises the President and the White House National AI Initiative Office (NAIIO) on both the near term and over the long-term intersections with and impacts of AI on various fields including innovation, technology, competition, socio-economic issues, legal and regulatory regimes, international relations, and other domains.[469]

NAIAC guides the U.S. government in adopting and leveraging AI resources and recognizes that greater flexibility in procurement authorities can convey significant advantages. For example:

> “[t]he ability to use a non-FAR contracting environment has enabled agencies like DoD to procure state of the art AI technology and do so quickly. The Chief Data and AI Office (CDAO) at the DoD has been prioritizing experimentation with AI procurement, particularly with non-FAR authorities, for years. Utilizing OTAs, which are not subject to the FAR or any other standard set of regulations, the DoD can pursue agile methodologies for managing the procurement process.
>
> The CDAO’s Tradewind initiative utilizes OTAs to improve the procurement of AI by applying and assessing innovative, agile contracting processes and making adjustments in real-time. The program has been successful at shortening the procurement process, with prototype agreements typically awarded within 30 to 60 days. Tradewind also collects and publishes resources to improve contracting practices, including recommendations for contract language and guides for negotiation. Beyond the more quantifiable benefits of faster delivery, streamlined contracting processes, and better practices, the initiative has critically focused on transforming organizational culture by bringing together a diverse set of stakeholders and experts, including representatives from industry, academia, government, and nonprofits. [...]
>
> Not all agencies have access to OTAs. Only eleven agencies, as of 2016, have received congressional authorization to use other transaction agreements, including NASA, DoD, DOE, HHS, DHS, and DOT among others.24 Nevertheless, more flexible tools for assessment of AI technology are also valuable to civilian agencies. In particular, the ability to learn about performance in the use context and revise evaluation criteria based on performance assessments is critical to judge the performance of AI systems that may degrade

---

[467] “National Artificial Intelligence Advisory Committee (NAIAC),” NIST, n.d., https://ai.gov/wp-content/uploads/2024/06/NAIAC-Impact-May-2024.pdf.

[468] “National Artificial Intelligence Advisory Committee.”

[469] “RECOMMENDATION: Provide Authority and Resources.”

> quickly in new contexts. This is exceedingly difficult, if not impossible, under the FAR."[470]

Consequently, NAIAC has recommended that at least three civilian federal agencies be granted Other Transaction Authority (OTA) to enable them to "develop, document, and disseminate best practices for AI procurement through a procurement or acquisitions innovation lab."[471]

To implement greater procurement flexibility and cultivate its benefits to competition, innovation, equal opportunity and equity, agencies must deploy adequate resources and staff to enable the agency procurement innovation labs to provide the appropriate levels of support necessary at each stage in the procurement process.[472]

To reduce barriers and enable civilian federal agencies to become more proficient at fast and efficient procurement of reliable and trustworthy, state-of-the-art AI, and to level the competitive landscape among AI suppliers, they need to have the freedom, resources, and executive support to test, learn as many best practices, and deploy as many new and improved tools as possible in AI procurement.[473]

OTA and Commercial Solutions Opening (CSO) are agile, adaptable (i.e., negotiable) Non-FAR/DFARS (Federal Acquisition Regulations/Defense FAR Supplement) based procurement mechanisms.[474] Though not predominant, these mechanisms are becoming more prevalent because they give agencies like DoD greater leeway to design and select procurement practices best suited to a given procurement need at hand.[475]

They also empower agency procurement offices to increase supplier diversity, provide greater flexibility for US-based work and US small business preferences, and confer greater ability to work with foreign vendors.[476] However, whenever agency procurement officers have any uncertainty regarding their authority to experiment with procurement practices they typically choose to rely on familiar, albeit potentially much slower procurement procedures that preclude rapid and effective procurement and subsequent adoption of AI.[477]

A possible preferred approach to implementing enhanced AI procurement authorities for civilian federal agencies would be to grant enhanced procurement authorities provisionally to those agencies with existing robust procurement offices that do not currently have OTA authority (e.g., the Department of Veterans Affairs).[478] Well-developed procurement infrastructures already exist at such agencies, and they are consequently better poised to deploy non-FAR/DFARS based

---

[470] *Ibid.*

[471] *Ibid.*

[472] *Ibid.*

[473] *Ibid.*

[474] *Ibid.*

[475] *Ibid.*

[476] *Ibid.*

[477] *Ibid.*

[478] *Ibid.*

contract instruments more quickly and effectively than agencies lacking robust procurement offices.[479]

In fall 2023, NAIAC recommended OMB and other relevant stakeholders to evaluate capabilities of existing procurement offices to help better inform selection of the first ones to provisionally receive OTAs.[480] NAIAC further specified that, before receiving OTA, agencies that have procurement and acquisition innovation labs should create and test comprehensive plans to collect, maintain, and manage contract data, "in accordance with GAO guidance to improve transparency and provide opportunities for evaluation of agencies' OTAs use."[481]

Consequently, the data collection and management plan should include systematically tracking OTA awards and documenting and recording important aspects and provisions of OTA contracts, such as compensation and consideration, lists of parties, technologies involved, strategic focus area, etc.[482]

In accordance with the fall 2023 NAIAC recommendation, procurement or acquisition innovation labs should maintain transparency and collect and share contracting instruments and practices within and across agencies to the greatest possible extent to facilitate the broadest possible dissemination of procurement best practices.[483]

For example, NAIAC also previously recommended procurement or acquisition innovation labs to explore ways to implement several emerging best practices in AI procurement, including Quality Assurance Surveillance Plans (QASPs) and provisions for in-domain evaluation.[484] NAIAC further recommends agencies without an acquisition or procurement innovation lab to form an innovation lab to collect best practices and better capitalize on OTA opportunities and benefits.[485]

Agencies starting out with neither OTA nor a procurement innovation lab will face the longest time horizons and the farthest to go in terms of investments of capital, resources and efforts required to obtain and maximize OTA opportunities.[486]

4. *European Union Public Procurement Directive*

Directive 2014/24/EU,[487] also known as the European Union's Public Procurement Directive, is a crucial legal instrument that seeks to create a standardized regulatory system for public procurement in Member States. The Directive guarantees equal competition and procurement access within the EU, augmenting the GPA's global perspective with a more local

---

[479] *Ibid.*
[480] *Ibid.*
[481] *Ibid.*
[482] *Ibid.*
[483] *Ibid.*
[484] *Ibid.*
[485] *Ibid.*
[486] *Ibid.*
[487] European Union. Directive 2014/24/EU of the European Parliament and of the Council of 26 February 2014 on Public Procurement and Repealing Directive 2004/18/EC. Official Journal of the European Union L 94, March 28, 2014. Accessed September 13, 2024. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32014L0024.

strategy. The Directive aims to achieve market transparency and promote economic integration, while also considering factors such as sustainability, innovation, and social responsibility.

Directive 2014/24/EU sets forth detailed rules for public contracts that exceed specific financial thresholds, adjusted biennially to reflect inflation and other market changes. Starting from January 1, 2024, the limits are set at €5,538,000 for public projects, €143,000 for central authority, €221,000 for municipalities, and €750,000 for social and specific services.[488] These limits align with the EU's responsibilities under the Government Procurement Agreement (GPA) of the World Trade Organization (WTO), aimed at discouraging bias and fostering international trade in public procurement.

The Directive's support for the "most economically advantageous tender" (MEAT) in Article 67 enables contracting authorities to assess bids using a wider range of quality criteria in addition to just the lowest price. As per the text of the Directive, such criteria "may comprise, for instance:

> (a) quality, including technical merit, aesthetic and functional characteristics, accessibility, design for all users, social, environmental and innovative characteristics and trading and its conditions;
> (b) organisation, qualification and experience of staff assigned to performing the contract, where the quality of the staff assigned can have a significant impact on the level of performance of the contract; or
> (c) after-sales service and technical assistance, delivery conditions such as delivery date, delivery process and delivery period or period of completion."[489]

Applied to the space sector, these measurements could involve considering the cost over the entire lifespan, the impact on the environment, and social factors, in line with the European Space Agency's goal of encouraging sustainable and innovative approaches in space exploration and technology advancement. The Directive supports the European Green Deal's goal of promoting environmentally responsible space projects by integrating sustainability factors into the procurement process.

Article 31 introduces the "Innovation Partnership" procedure, promoting collaboration between public authorities and private suppliers to develop innovative technologies. As summarized by the EU Commission, the tendering phase occurs at the outset of the procedure, where the most suitable partners are selected based on their capabilities and capacity to perform the contract, as well as the quality of their tenders.

This process should be employed only under specific conditions, such as when the desired goods, works, or services are innovative, and both the development and procurement phases are included in the procedure, provided they meet agreed performance standards and cost limits.

---

[488] European Union. "Public Procurement." EUR-Lex. Accessed September 12, 2024. https://eur-lex.europa.eu/EN/legal-content/summary/public-procurement.html.
[489] *Ibid.,* Art. 67

Contracts for innovation partnerships are awarded according to the criteria of the best price-quality ratio offered.[490]

Furthermore, Article 46 promotes the breaking down of larger contracts into multiple smaller ones, making it easier for small and medium-sized enterprises (SMEs) to participate. This clause is essential for boosting economic activity and guaranteeing a variety of suppliers in the public procurement sector.

Directive 2014/24/EU includes social and environmental responsibilities in the procurement framework, as required by Articles 18 and 57. Contracting authorities are required to guarantee adherence to key EU labor, social, and environmental regulations. The Directive promotes social justice and sustainability in space initiatives by requiring the exclusion of contractors who fail to meet labor rights and environmental standards, including ILO conventions.

However, when applied to AI technologies, the current EU public procurement framework faces significant challenges. As highlighted by recent studies, Directive 2014/24/EU lacks mandatory provisions to ensure that AI technologies being procured are "trustworthy" or that their use aligns with the protection of individual rights and freedoms.

This gap raises concerns about the compliance of AI procurement with broader digital regulation frameworks, such as the European Declaration on Digital Rights and Principles. Without clear guidelines on the ethical deployment of AI, there is a risk that public procurement processes could inadvertently endorse technologies that do not meet the EU's standards for transparency, accountability, and fairness.[491]

Furthermore, current procurement procedures find it challenging to efficiently handle the intricacies of digital supply chains, which typically involve a significant level of technical and commercial complexity. The integration of AI into public projects poses a heightened risk of cyber vulnerabilities, making cybersecurity a major concern in this scenario.

To sum up, Directive 2014/24/EU has a major impact on public procurement in the European space industry by promoting transparency, competition, and innovation. Its adaptable standards for awarding contracts make it easier to implement cutting-edge technologies like artificial intelligence, crucial for improving efficiency and effectiveness in space missions.

5. *Exemplary Clauses for Use in Contracts on Procurement of AI Systems*

Currently, there are multiple efforts underway to modernize procurement practices with respect to obtaining AI systems. Notably, procurement itself is being transformed through the deployment of AI, but that is an aside to this review.

---

[490] European Commission. Guidance on Innovation Procurement: Innovation Partnership. Brussels: European Commission, 2021. Accessed September 13, 2024. https://single-market-economy.ec.europa.eu/system/files/2021-11/GROW_C2_innovation_partnership_210901.pdf.

[491] Sanchez-Graells, Albert. "Public Procurement of Artificial Intelligence: Recent Developments and Remaining Challenges in EU Law." LTZ (Legal Tech Journal) 2, no. 2024 (January 25, 2024): 122–131. Available at SSRN: https://ssrn.com/abstract=4706400 or http://dx.doi.org/10.2139/ssrn.4706400.

Some of the efforts in progress include the European Commission Proposal for standard contractual clauses for the procurement of Artificial Intelligence (AI) by public organisations,[492] the US OMB Memorandum M-24-18 on Advancing the Responsible Acquisition of Artificial Intelligence in Government,[493] and the IEEE P3119 Draft Standard for the Procurement of Artificial Intelligence and Automated Decision Systems[494] (in development).

A critical best practice in contract negotiations, generally, is to allocate and apportion risk to the parties in the best positions to manage the respective risks.[495] Often in the domain of cutting-edge technologies, such as AI, the party best able to manage risks is the party developing the cutting-edge tech in question.

However, AI is different from many other types of advanced technology in that it remains relatively opaque regarding the precise mechanisms of action that achieve the desired results and how the various AI system components interoperate to achieve those results, and this means AI can be especially rife with unknown and unanticipated risks.[496]

Consequently, appropriate risk allocation is not only key to "the long term viability of the service, but also to achieving best value. Liability for certain areas will reside with the department, particularly around the use and application of the AI-powered solution, and in relation to data access and transfer. Liability may also need to sit with the supplier, including areas focused around technical, security and quality assurance."[497]

Best practices for risk mitigation further demand that the AI system "have an easy way to report any suspected unauthorised behaviour to relevant authorities within or outside the organisation."[498]

The European Commission Proposal for standard contractual clauses for the procurement of Artificial Intelligence (AI) by public organisations provides exemplary provisions related to risk management of AI systems acquired through government procurement. Specifically, Article 2 on risk management systems stipulates that:

> "2.1. The Supplier ensures that, prior to the Delivery of the AI System, a risk management system shall be established and implemented in relation to the AI System.
>
> 2.2. The risk management system shall at least comprise the following steps:

---

[492] European Commission. AI Procurement Clauses Template for High-Risk AI. Public Buyers Community, October 2023, https://public-buyers-community.ec.europa.eu/system/files/2023-10/AI_Procurement_Clauses_template_High_Risk%20EN.pdf.

[493] Executive Office of the President. AI Acquisition Memorandum (M-24-18). The White House, October 2024, https://www.whitehouse.gov/wp-content/uploads/2024/10/M-24-18-AI-Acquisition-Memorandum.pdf.

[494] IEEE Standard 3119-2023: Standard for the Procurement of Artificial Intelligence and Automated Decision Systems, IEEE, 2023, https://standards.ieee.org/ieee/3119/10729/.

[495] Kimberly Munch, "A Comparison Between U.S. Export Controls and EU Export Controls," Export Compliance Training Institute, January 9, 2024, https://www.learnexportcompliance.com/a-comparison-between-u-s-export-controls-and-european-export-controls/.

[496] *Ibid.*

[497] *Ibid.*

[498] *Ibid.*

a. identification, estimation and evaluation of the known and reasonably foreseeable risks to health, safety and fundamental rights of the European Union that are likely to arise in the light of the Intended Purpose of the AI System and Reasonably Foreseeable Misuse;
b. evaluation of other possibly arising risks;
c. adoption of appropriate and targeted risk management measures designed to address the risks identified pursuant to points a and b of this paragraph in accordance with the provisions of the following paragraphs.

2.3. The risk management measures referred to in paragraph 2.2, point (c) shall be such that relevant residual risks associated with each hazard as well as the overall residual risk of the AI system is reasonably judged to be acceptable by the Supplier, provided that the AI System is used in accordance with the Intended Purpose or under conditions of Reasonably Foreseeable Misuse.

2.4. In identifying the most appropriate risk management measures referred to in paragraph 2.2, point (c), the following shall be ensured:

a. elimination or reduction of identified risks as far as technically feasible through adequate design and development of the AI System;
b. where appropriate, implementation of adequate mitigation and control measures in relation to risks that cannot be eliminated;
c. provision of adequate information to the Public Organisation.

2.5. The Supplier ensures that, prior to the Delivery of the AI System, the AI System is tested in order to verify whether the AI System complies with the Clauses and whether the risk management measures referred to in paragraph 2.2, point (c) are effective in light of the Intended Purpose and Reasonably Foreseeable Misuse. If requested by the Public Organisation, the Supplier is obliged to test the AI System in the environment of the Public Organisation.

2.6. All risks identified, measures taken and tests performed in the context of compliance with this article must be documented by the Supplier. The Supplier must make this documentation available to the Public Organisation at least at the time of the Delivery of the AI System. This documentation can be part of the technical documentation and/or instructions for use.

2.7. The risk management system shall consist of a continuous and iterative process run throughout the entire duration of the Agreement. After the Delivery of the AI System the Supplier must therefore:

a. regularly review and update the risk management process, to ensure its continuing effectiveness;
b. keep the documentation described in article 2.6 up to date; and
c. make every new version of the documentation described in article 2.6 available to the Public Organisation without delay.

> 2.8. If reasonably required for the proper execution of the risk management system by the Supplier, the Public Organisation will provide the Supplier, on request, with information insofar as this is not of a confidential nature.
>
> 2.9. <Optional> If the Public Organisation's use of the AI System continues beyond the term of the Agreement, at the end of the term of the Agreement, the Supplier shall provide the Public Organisation with the information necessary to maintain the risk management system by itself."[499]

The EC Proposal for standard contractual clauses for the procurement of Artificial Intelligence (AI) by public organisations further provides exemplary provisions related to indemnification, a contractual mechanism designed to limit responsibility for damages, irrespective of whether direct liability can be effectively limited contractually. Specifically, Article 17 on indemnification stipulates that:

> "17.1. The Supplier will indemnify the Public Organisation against all claims brought by third parties, including supervisors, in respect of any breach of their intellectual property rights, privacy rights or equivalent claims relating to knowledge, unlawful competition and so forth with regards to the Supplier Data Sets and Third Party Data Sets.
>
> 17.2. The Public Organisation will indemnify the Supplier against all claims brought by third parties, including supervisors, in respect of any breach of their intellectual property rights, privacy rights or equivalent claims relating to knowledge, unlawful competition and so forth with regards to Public Organisation Data Sets."[500]

Exemplary clauses used for limitation of liability with respect to the provisioning of AI goods or services can also be sourced from industry. For instance, CEGSOFT, a cloud applications and AI Software as a Service (SaaS) provider has made the following limitation of liability provision available online:

> "Use of Artificial Intelligence and Limitation of Liability
>
> CEGSOFT offers artificial intelligence solutions as part of its services, referred to in this document as 'AI Solutions.' This feature uses natural language processing and machine learning technology and has been trained with CEGSOFT'S proprietary data as well as user feedback through an upvote/downvote system.
>
> Licensee acknowledges and agrees that:
>
> a) The AI Solutions are an assistance tool and should not be considered a substitute for professional, legal, financial, or any other type of advice.

---

[499] Jeroen Naves and Pels Rijcken, "Procurement of AI Community Proposal for Standard Contractual Clauses for the Procurement of Artificial Intelligence (AI) by Public Organisations," draft, 2023, https://public-buyers-community.ec.europa.eu/system/files/2023-10/AI_Procurement_Clauses_template_High_Risk%20EN.pdf.

[500] *Ibid*.

> b) The responses generated by the AI Solutions are the result of automated processes and may not always be accurate, complete, or suitable for all situations.
> c) The content generated by the AI Solutions may vary over time due to the dynamic nature of machine learning and user feedback.
>
> Licensee acknowledges and agrees that CEGSOFT will not be responsible for:
> a) Any decision made by the Licensee based on the information provided by the AI Solutions.
> b) Errors, omissions, or inaccuracies in the responses generated by the AI Solutions.
> c) Any direct, indirect, incidental, special, or consequential damages resulting from the use or inability to use the AI Solutions.
> d) The suitability or adequacy of the AI Solutions' responses for the Licensee's specific purposes.
>
> Licensee agrees to use the AI Solutions responsibly and ethically, and not to rely solely on its responses for making critical or legally binding decisions.
>
> CEGSOFT reserves the right to modify, improve, or discontinue the AI Solutions at any time, without prior notice and without incurring any liability to the Licensee."[501]

A further example, Checksum AI, an AI software applications testing services platform (also SaaS), has made the following limitation of liability provision available online:

> "Limitation of Liability.
>
> 1. Except with respect to liabilities arising out of breaches of Section 10 of this Agreement of a Party's gross negligence or willful misconduct: (A) neither Party will be liable to the other for punitive, incidental, indirect, special, reliance or consequential damages, including lost business, revenue or anticipated profits, regardless of the cause of action and whether or not the Party was advised of the possibility of such loss or damages, and (B) except with respect to Customer's obligation to pay any minimum fees, in no event will a Party's total cumulative liability under this Agreement exceed the greater of (I) amount paid or owed by Customer under this Agreement for the 12 months prior to the date that the cause of action arose and (II) one thousand dollars (USD $1,000.00). These limitations of liability apply even if any remedy specified in this Agreement is found to have failed its essential purpose.
>
> 2. Certain aspects of the services may involve artificial intelligence or machine learning ("AI Functions"). Customer acknowledges that the AI Functions are rapidly evolving field. While Checksum is always working to improve its AI Functions, due to the probabilistic nature of theAI Functions, the services may provide inaccurate output or otherwise not always produce the intended results. As such, Customer acknowledges

---

[501] CEG Software Solutions, End-User License Agreement (EULA), CEGSOFT, 2024, https://cegsoft.com/documents/CEGSOFT_EULA_2024.pdf.

> that no warranties are made by Checksum with respect to (and Checksum will have no liability with respect to) the output (or any Party's use thereof) of Checksum's AI Functions."[502]

As a caveat, inclusion of these sample contract provisions does not constitute an endorsement in any way, and they are included solely for purposes of example. Before entering into a procurement agreement, all parties should consult with appropriate professionals skilled at analyzing the unique circumstances often involved in novel transactions such as the procurement of AI for space applications.

Conclusions

Internationally, export control laws are implemented at both national and supra-national levels, such as within the EU. Adjacent to export controls, and regulating the flow of technology, products, and services in the opposite direction, from the acquisitions side, are procurement laws, regulations, and standards, which are also relevant to AI systems used in the space environment.

Also, of relevance to AI systems designed for use in the space environment, existing and impending procurement standards directly address AI technologies, including the anticipated IEEE P3119 Standard for the Procurement of Artificial Intelligence and Automated Decision Systems.

In parallel, the WTO Agreement on Government Procurement (GPA) plays a pivotal role in fostering open competition in government procurements, ensuring international markets are accessible to a wide range of suppliers, including those providing AI-enabled solutions.

Further, all parties to procurement agreements, especially those pertaining to acquisition of AI systems for space applications, need to consider the potential unforeseen risks and impacts that may arise from deployment of autonomous systems in the outer space environment. Risk management, indemnification, and limitation of liability provisions can provide parties to procurement contracts with flexible mechanisms for apportioning risk, limiting liability, and assigning responsibility for compensation in the event potential damages are realized.

To create international standardization for AI in space based on procurement frameworks, the following steps can be taken:

- Reviewing existing procurement regulations and identify gaps and challenges specific to the acquisition of AI technologies for space purposes.
- Developing standardized guidelines for the procurement of AI systems in space missions, considering factors such as transparency, competition, risk management, evaluation criteria, and the inclusion of ethical and technical requirements.
- Establishing mechanisms for international cooperation and information sharing to promote best practices in AI procurement for space exploration and satellite missions.

---

[502] Checksum AI, Master Subscription Agreement, https://checksum.ai/master-subscription-agreement/.

Section 4: Overview of Telecommunications Frameworks in Regulating the Use of AI for Space Applications

Introduction

This section provides a comprehensive examination of the telecommunications regulatory framework and its impact on the integration of AI in space technologies. Telecommunications and network developments have forced corresponding laws and regulations to focus on the critical role AI plays in spectrum allocation, communication protocols, and security. This section explores the increasing reliance on AI for efficient spectrum management, particularly in dynamically managing non-federal spectrum bands, as discussed in documents from the UK Office of Communications (Ofcom) and the US Federal Communications Commission (FCC).

Examples include AI's application in cognitive radios and dynamic spectrum sharing, with specific reference to Ofcom's and FCC's discussions on AI and spectrum sharing, as well as spectrum allocation and monitoring approaches at the International Telecommunication Union (ITU). Additionally, this section addresses the privacy and security challenges associated with AI-driven communication networks, particularly in deep space missions, citing examples like NASA's Artemis program and the European Union's IRIS2 satellite communication system.

This section draws on a range of articles and documents, including key export control regulations, government memoranda, and AI procurement guidelines, to provide a detailed understanding of how telecommunications regulations intersect with AI technologies in space exploration.

1. *International Agreements: ITU Regulations*

1.1 *ITU Constitution*

The International Telecommunication Union, the United Nations specialized agency for information and communication technologies (ICTs), is the organization of reference in the realm of telecommunications. To that effect, its constituent regulation lays down its main responsibilities and purposes, including serving as a forum to enhance international cooperation among its Member States aimed at improving and utilizing telecommunications of any kind in a rational manner.[503]

Moreover, ITU is entrusted with fostering the development of technical facilities in order to improve the efficiency of telecommunication services by boosting their usefulness, and with bringing the benefits of new technologies to the disposition of the worldwide general public to the extent possible.[504]

With that objective in mind, the second paragraph of Article 1 of the ITU Constitution attributes the following duties to the Union, among which it is remarkable:

---

[503] Article 1(1)(a), *ITU Constitution 1992*.

[504] Articles 1(1)(c) and (d), *ITU Constitution 1992*.

- The allocation, allotment and assignment of the radio-frequency spectrum, and the corresponding registration in the Master International Frequency Register (MIFR), in order to avoid harmful interference between radio stations of different countries.
- The coordination of international efforts to ameliorate the use made of the radio-frequency spectrum for radiocommunication services and of the geostationary-satellite and other satellite orbits.
- The conducting of studies to draft regulations, adopt resolutions, and formulate recommendations and opinions.
- The standardization of telecommunications.

Accordingly, the role of ITU in any standardization efforts in the matter of AI applied to telecommunications is undisputed, given: (1) the mandate of the Union, and (2) the nature of the institution, as one of the organisms under the system of the United Nations, the largest international organization governing international law.

Following Article 4 of the ITU Constitution, the instruments of the Union encompass the Constitution already mentioned, the ITU Convention, and the Administrative Regulations (which include the International Telecommunication Regulations, and the Radio Regulations). All of them are legally binding upon Member States in every of their operative telecommunication offices and stations, either providing international services or potentially disrupting the radio services of other states.[505]

Regarding the general provisions relating to telecommunications (Chapter VI of the ITU Constitution), it is important to reiterate the right of the public to use the international telecommunication service of public correspondence, whose services, charges and safeguards shall remain the same for all.[506] This right includes the international telecommunications carried out by AI systems onboard spacecraft, and would encompass a State obligation to protect the infrastructure through which these communications are performed, in this case, satellite infrastructure.

Similarly, in the event that an AI-based space system conveys a telegram that could potentially pose a threat to national security or that could attempt to undermine the domestic legal foundations, the public order, or decency of a given State, such transmission could be interrupted in observance of their national law. With the exception of circumstances in which such reporting might endanger the security of the State, the prompt notification to the office of origin of the cessation of any such telegram or any part thereof is required to implement the aforementioned stoppage.[507]

Moreover, it is interesting to refer to Article 36 of the ITU Constitution regarding responsibility, in accordance with which "Member States accept no responsibility towards users of the international telecommunication services, particularly as regards claims for damages." When AI-powered satellites are involved, the liability issue becomes more complex due to the

---

[505] Article 6, *ITU Constitution 1992*.

[506] See Article 33, *ITU Constitution 1992*.

[507] See Article 34, *ITU Constitution 1992*.

inherent characteristics to AI technologies and the difficulty in clearly determining the causal link, as it has been previously discussed in other sections.

As space activities, the exploitation of telecommunications satellites is also governed by the space treaties, including the 1967 Outer Space Treaty and its liability for damages clauses. Therefore, depending on the location of the damage, whether in outer space or on Earth, one provision or another of the Liability Convention would apply.

It is true that Article 36 of ITU Constitution enshrines that *"Member States accept no responsibility towards users of the international telecommunication services, particularly as regards claims for damages."* However, this norm is not applicable to damages caused by, for instance, the collision between two telecommunications satellites or one of those and a distinct space object, since it would amount to tort (non-contractual) liability for damages. What appears to be clear then is that, even if the provider of international telecommunications services is a private company, the liability for damages caused by AI-enabled space objects would fall upon the launching State or States, as per the criteria laid down by Article VII of the Outer Space Treaty.

The secrecy of international telecommunications constitutes another relevant provision of the ITU Constitution, following Article 37. In this particular case, cyberattacks against AI systems onboard spacecraft should be referenced as they may lead to a potential leakage of communications, thereby breaching this secrecy duty.

Accordingly, in order to ensure the secrecy of telecommunications and comply with the obligations stated herein, States should take all needed measures to guarantee the respect for communications confidentiality, such as implementing encryption mechanisms and building resilient systems, which are prepared to face this sort of attacks without suffering significant losses and/or data breaches. In the case of autonomous telecommunications satellites, despite the technical complexity for intrusion attempts due to its characteristics (closed systems with encrypted communications), there are potential vulnerabilities identified. For instance, once in orbit, autonomous satellites often receive software updates. If these updates are not secure, they can become an entry point for attacks. This duty not only embraces telecommunications infrastructure, but also telecommunication channels and installations,[508] which could become the potential target for cyberattacks as well.

In turn, two of the cardinal norms contained in the ITU Constitution focus on the use of the radio-frequency spectrum and of the geostationary-satellite and other satellite orbits (Article 44) and on the prevention of harmful interference (Article 45).

The former foresees that "In using frequency bands for radio services, Member States shall bear in mind that radio frequencies and any associated orbits, including the geostationary-satellite orbit, are limited natural resources and that they must be used rationally, efficiently and economically, in conformity with the provisions of the Radio Regulations, so that countries or groups of countries may have equitable access to those orbits and frequencies, taking into account the special needs of the developing countries and the geographical situation of particular countries."

---

[508] See Article 38(3), *ITU Constitution 1992*.

This clause explicitly recognizes the electromagnetic spectrum, concretely radio waves, as a scarce resource which shall be duly administered and utilized observing the criteria of rationality, efficiency and economy, and the right to access and special necessities of developing countries. In this regard, the use of AI technologies can aid in the long-term sustainability of outer space, including radio frequencies, as will be showcased in the pertinent section of this report.

The latter prescribes that "All stations, whatever their purpose, must be established and operated in such a manner as not to cause harmful interference to the radio services or communications of other Member States or of recognized operating agencies, or of other duly authorized operating agencies which carry on a radio service, and which operate in accordance with the provisions of the Radio Regulations."

The non-harmful interference principle should be respected, as well, when an AI-driven satellite is operating in outer space, requiring that no communications channeled through these objects be perturbed. Even if these autonomous space systems may be able to adapt to new scenarios, by predicting potential outcomes based on previous experiences during the learning process, securing communications between the satellite at stake and the ground station is crucial.

Finally, the ITU is encouraged to cooperate with other United Nations agencies and external organizations in the pursuit of similar interests,[509] such as the standardization and development of regulatory and frameworks for the development and deployment of AI systems in space. For that purpose, efforts could be joined with UNCOPUOS in the formulation of international guidelines governing the use of AI in outer space.

1.2 *ITU Radio Regulations*

The ITU Radio Regulations are the international treaty that lays down the regulatory framework governing the use of the radio-frequency spectrum and the GSO and non-GSO satellite orbits.[510] The regulations are subject to periodic review and, if necessary, revision every three to four years at the so-called "World Radiocommunication Conferences" (WRC),[511] convened by the Radiocommunication Sector (ITU-R).[512]

As per Article 7 of the ITU Convention, the agenda may comprise the designation of topics to be examined by the Radiocommunication Assembly and the relevant study groups, along with matters to be considered by the Assembly in relation to future Radiocommunication Conferences (e.g., WRC agenda usually considers recommendations made by previous WRC). In addition, the final draft of the agenda shall be agreed on by the ITU Council two years prior to the conference, with the approval of a majority of Member States, while the general scope of the same should be outlined four to six years in advance.

---

[509] See Article 50, *ITU Constitution 1992*.

[510] https://www.itu.int/wrc-23/about/about-wrcs/

[511] Article 13(1) and (2), *ITU Constitution 1992*.

[512] ITU-R is tasked with the adoption of recommendations on radio communication matters, which is the reason behind the celebration of the WRC, as these latter constitute one of the mechanisms through which ITU-R shall work (according to Article 12(a) of the ITU Constitution).

The most recent WRC corresponds to WRC-23, held in Dubai (United Arab Emirates) from 20 November to 15 December 2023.[513] The agenda[514] included items such as the consideration of regulatory clauses to facilitate the use of radiocommunications for suborbital vehicles (following Resolution 772 (WRC-19)), and the examination of proper regulatory actions to, first, enable the utilization of fixed-satellite service networks for the control and non-payload communications of unmanned aircraft systems (based on the ITU R studies under Resolution 171 (WRC-19)), and, second, provide inter-satellite links in particular frequency bands, or segments thereof. The latter may entail the incorporation of an inter-satellite service allocation where appropriate (on the grounds of ITU R studies conducted in line with Resolution 773 (WRC-19)).

Accordingly, the 2020 Radio Regulations—adopted after WRC-19, in Sharm el-Sheikh (Egypt)—have been, once again, updated: this time, under the name of the 2024 Radio Regulations;[515] recently published and set to take effect on 1 January 2025. If we take a closer look at the content, however, none of its provisions explicitly mentions the component of AI for satellite telecommunications. Still, some Radio Regulations clauses may be applicable to AI-powered satellites, for instance, the prohibition of interference from radio stations (Article 15).[516]

The agenda for WRC-27, the venue and date of which have yet to be determined, has been published by Resolution 813 (WRC-23).[517] Nevertheless, it contains no single reference to AI technologies for space stations on board spacecraft. It would be prudent to eventually address this subject, given its topicality and the need for regulatory provisions that ensure the proper deployment and utilization of these systems for telecommunications purposes.

1.3 *Recommendations for Standardization*

After analyzing the existing telecommunications regulations and standards to identify their applicability to AI-enabled space systems' communication requirements, there appear to be some crucial steps to be taken in order to create international laws for AI in space based on telecommunications frameworks. These include the development of standardized protocols and guidelines for communication, spectrum allocation, and interoperability of AI-driven space vehicles, ensuring efficient and secure communication, and the establishment of international cooperation frameworks to facilitate the coordination of AI-enabled space systems' communication activities, promote spectrum sharing, and address potential interference issues.

As pointed out, AI and machine learning technologies necessitate a more advanced and adaptive spectrum management framework. One critical issue within current ITU protocols is

---

[513] The place and date of WRC-23 were decided upon by means of Decision 623. https://www.itu.int/wrc-23/wp-content/uploads/sites/12/2023/02/S21-CL-C-0096MSW-E.pdf

[514] The agenda for WRC-23 was established by means of Resolution 1399 (C20). https://www.itu.int/md/S20-CL-C-0069/en

[515] International Telecommunication Union. Radio Regulations. 2024. https://www.itu.int/pub/R-REG-RR-2024.

[516] "§ 1 All stations are forbidden to carry out unnecessary transmissions, or the transmission of superfluous signals, or the transmission of false or misleading signals, or the transmission of signals without identification (except as provided for in Article 19)."

[517] https://www.itu.int/dms_pub/itu-r/oth/0c/0a/R0C0A0000110038PDFE.pdf See the ITU-R Preparatory Studies for WRC-27 here: https://www.itu.int/en/ITU-R/study-groups/rcpm/Pages/wrc-27-studies.aspx

assessing whether they adequately support efficient and equitable spectrum use for AI-driven communication systems, particularly in space applications.

Concerning automated spectrum management systems (ASMS), ITU-R SM.1370-2[518] provides comprehensive guidelines by including operational frameworks, database management, and recommendations for secure, efficient data processing and user interface standards. ITU-R SM.1537-1[519] also focuses on automating and integrating spectrum monitoring systems with spectrum management functions, outlines how modern technology, particularly digital signal processing (DSP) and computerized networks, enables efficient spectrum monitoring and management through automation.

These recommendations emphasize the need for automation in managing the growing demand for spectrum and increasing complexity in frequency assignment processes. Nevertheless, AI-driven systems demand adaptive and flexible spectrum use, which may not be fully supported by the current protocols. ITU-R SM.1370-2, for instance, includes aspects of real-time monitoring and interference calculation, which are helpful, but additional protocols specifically addressing dynamic satellite frequency adjustments and decentralized decision-making are critical for AI applications in high-demand scenarios and fully autonomous satellite management.

Additionally, ITU-R SM.1046-3[520] addresses spectrum utilization and efficiency metrics primarily within radio systems and highlights methodologies for evaluating spectrum management effectiveness. These are relevant to assessing the sufficiency of current spectrum management protocols, especially when considering modern, data-intensive applications, such as AI-driven systems.

For AI-driven systems, which demand high data throughput and often operate across multiple frequency bands, the flexibility and specificity provided by the models introduced for assessing spectrum usage can be advantageous. Nevertheless, the document does not directly address AI-specific challenges, such as the dynamic frequency changes or ultra-low latency requirements typical in AI applications. Current protocols may need to incorporate more adaptive, real-time spectrum allocation and interference management capabilities to adequately accommodate the sophisticated demands of AI systems.

Last but not least, ITU-R SM.2039[521] explores the evolving needs and technologies for spectrum monitoring to keep up with the rapid advancements in radiocommunication, such as cognitive radio and software-defined radio (SDR). It discusses improvements needed in spectrum monitoring systems to handle new signal types, higher frequencies, and complex multiplexing techniques.

---

[518] International Telecommunication Union (ITU), *Recommendation ITU-R SM.1370-2 (08/2013)*, Design guidelines for developing automated spectrum management systems. Available at: https://www.itu.int/rec/R-REC-SM.1370-2-201308-I/en

[519] ITU, *Recommendation ITU-R SM.1537-1 (08/2013)*, Automation and integration of spectrum monitoring systems with automated spectrum management. Available at: https://www.itu.int/rec/R-REC-SM.1537-1-201308-I/en

[520] ITU, Recommendation ITU-R SM.1046-3 (09/2017), *Definition of spectrum use and efficiency of a radio system*. Available at: https://www.itu.int/rec/R-REC-SM.1046-3-201709-I/en

[521] ITU, *Recommendation ITU-R SM.2039-0 (08/2013)*, Spectrum monitoring evolution. Available at: https://www.itu.int/rec/R-REC-SM.2039-0-201308-I/en

Some of the key recommendations underscored for evolution in spectrum monitoring include extended monitoring capabilities, enhanced functionalities—such as supporting complex signal separation, which is especially relevant for overlapping AI-driven applications that may utilize multiple domains simultaneously—and user-friendly interface. However, the document's approach may fall short of the flexibility and rapid response required for highly dynamic AI applications, which often require instant access to varying frequencies and robust interference management.

In conclusion, while these protocols provide a foundation for efficient spectrum management, they lack the agility required to support AI-driven systems. Enhanced flexibility in frequency assignment, rapid data processing, and real-time spectrum adaptation will be necessary for effectively managing AI-enabled and autonomous systems within the spectrum. As a result, further modifications or entirely new models might be necessary to handle the unique characteristics and requirements of AI-driven technologies in spectrum utilization and management.

In such a context, a collaborative review by the ITU on the effectiveness of current spectrum allocation standards in meeting the unique needs of AI-driven systems in space would be an essential step toward addressing the emerging challenges outlined in this section. Prioritizing spectrum access for AI systems in space operations could address the growing demand for spectrum in high-density, mission-critical environments. Cooperation among governments, the space industry, and AI experts would help to develop a framework that enhances interoperability, optimizes spectrum use, and promotes equitable, flexible spectrum allocation for future AI-driven space missions.

An essential outcome of this evaluation could be a roadmap for incrementally integrating more flexible, AI-oriented spectrum protocols, which would support rapid technological advancements in AI-driven space systems while maintaining efficient and fair spectrum usage across other services. Such a collaborative evaluation would aid the ITU in better aligning its spectrum management frameworks with the evolving needs of autonomous, AI-powered systems, ensuring that the spectrum remains a robust and equitable resource for future technological advancements in space.

The economic impact of AI in telecommunications is two-folded: for several kinds of commodities, the minimization of production costs and/or the enhancement of the service quality for the customer are likely to occur, due to the anticipated transformation of industry processes. These consequences influence (or should influence) the regulatory decisions to be made, since this 'welfare-enhancing' potential of AI technologies should be considered by legislators if the aim is to set the appropriate incentives to achieve optimal welfare outcomes in the future.[522]

Among the suggested actions for the ITU to undertake in the future, in addition to the collaborative framework suggested to evaluate current protocols, the establishment of a study group on AI in space telecommunications under the ITU-R and the deliberation of possible

---

[522] Balmer, R. E., Levin, S. L., and Schmidt, S. 2020. "Artificial Intelligence Applications in Telecommunications and other network industries." *Telecommunications Policy* 44, p. 8, https://doi.org/10.1016/j.telpol.2020.101977

regulatory provisions for the development and implementation of AI-enabled satellites stand out as the most pressing and suitable avenues for consideration.

By way of example, during the ITU Space Sustainability Forum, held in Geneva on September 11-12, 2024, the term "AI" was invoked on multiple occasions, showcasing the interest of the space industry in these emerging technologies and the potential it holds for space sustainability efforts.

In light of the ITU's mandate regarding the sustainable use of radio-frequency spectrum and associated satellite-orbit resources used by space services,[523] it would be prudent to investigate the extent to which AI systems could be leveraged to achieve the Sustainable Development Goals (SDGs) and to facilitate and support space sustainability initiatives.

2. *Telecommunications: Exploring Communication Protocols, Spectrum Allocation, and Security*

2.1 *Communication Protocols*

In the field of telecommunications, a "communication protocol" (also referred to as "telecommunication protocol," or just "protocol") is typically defined as a set of standardized norms which prescribe the manner in which two or more communication entities[524] should engage in communication and interact. The communication between entities is carried out by exchanging protocol messages, whose nature and format is specified by the protocol beforehand, as well as those rules establishing the requirements to determine when a communication entity may or must send a protocol message to another.[525]

As mentioned, these protocols are standardized –given that, otherwise, the communication entities would be unable to understand each other–, thereby facilitating interoperability, that is, a secure and agile exchange of mutually intelligible data. For instance, interoperability is capable of enabling seamless communication and data sharing, leading to a more efficient and effective service delivery.[526] One of the numerous benefits brought by AI technologies is, precisely, the enhancement of interoperability.

By way of illustration, in view of NASA's preparation of the Artemis mission to return to the Moon,[527] the communication between astronauts and researchers back on Earth can prove

---

[523] On the basis of Resolution 219 (Bucharest, 2022) of the Plenipotentiary Conference and RA-23 Resolution ITU-R 74.

[524] "Devices that communicate in a communication system (a communication network) are referred to as communication entities." See Hercog, D. 2020. Communication Protocols Principles, Methods and Specifications. Cham, Springer, p. 15.

[525] *Ibid*., pp. 15-16.

[526] Tangi, L. et al. 2023. Artificial Intelligence for Interoperability in the European Public Sector An exploratory study, JRC134713, Publications Office of the European Union, Luxembourg, p. 15. doi:10.2760/633646.

[527] Greshko, M. 2022. "How NASA's Artemis program plans to return astronauts to the moon." National Geographic, 22 August 2022. https://www.nationalgeographic.com/science/article/how-nasas-artemis-program-plans-to-return-astronauts-to-the-moon

complex due to extreme distances and harsh conditions, resulting in signal delays and disruption.[528] In this vein, AI could be leveraged to establish and maintain secure space communications networks.

2.2 *Spectrum Allocation*

At the international level, spectrum management is entrusted to the International Telecommunication Union (ITU). This body owns a comprehensive three-step process, namely "allocation" (services), "allotment" (areas or countries), and "assignment" (stations), which collectively facilitate the utilization of radio frequencies by individual radio operators without interference.[529]

The management and use of the wireless spectrum have been conducted over numerous decades through a multifaceted regulatory framework and an assortment of policies. The current methodologies in use to assess spectrum requirements are complex by the burgeoning degree of interdependencies within the spectrum domain, while those utilized for the allocation of spectrum are frequently influenced by small-scale studies, which are usually prone to inherent biases.

Consequently, the resulting policies and usage of the spectrum are often suboptimal and inflexible, which hinders and impedes the efficient use of spectrum. The allocation of spectrum has proven effective thus far; however, with the exponential increase in the number of services and attendant requirements for greater bandwidth, it will become necessary to shift towards more optimal telecommunications management techniques.[530]

In March 2023, the UK Office of Communications (Ofcom) released a discussion paper outlining potential avenues for dynamic or adaptive approaches to spectrum management in the country.[531] The paper prompts the question of whether dynamically managed bands –potentially aided by devices with sensing and/or machine-learning capabilities– could be made available as "top-up" capacity for certain spectrum uses with high projected future demand, including mobile broadband. Some months later, in October 2023, Ofcom published a further discussion paper on flexible access and spectrum sharing,[532] where it analyzed whether technological developments could facilitate the creation of more dynamic spectrum sharing environments.

---

[528] NASA, NASA Develops Advanced Space Communications Process, Glenn Communications, 1 March 2023. https://www.nasa.gov/humans-in-space/nasa-develops-advanced-space-communications-process/

[529] von der Dunk, F. 2015. "Legal aspects of satellite communications." in von der Dunk, F., and Tronchetti, F. 2015. Handbook of Space Law. Cheltenham, Edward Elgar Publishing, p. 465. See Art. 1(2), ITU Constitution 1992: "To this end, the Union shall in particular: a) effect allocation of bands of the radio-frequency spectrum, the allotment of radio frequencies and the registration of radio-frequency assignments and, for space services, of any associated orbital position in the geostationary-satellite orbit or of any associated characteristics of satellites in other orbits, in order to avoid harmful interference between radio stations of different countries".

[530] Telecommunication Engineering Centre. 2021. AI in Spectrum Management, Radio Division, Ministry of Communications and Information Technology, Government of India. https://www.tec.gov.in/pdf/Studypaper/AI_in_Spectrum_management.pdf

[531] See Ofcom. 2023. Opportunities for dynamic or adaptive approaches to managing spectrum in the UK: A discussion paper, 28 March 2023. https://www.ofcom.org.uk/spectrum/innovative-use-of-spectrum/flexible-and-adaptive-spectrum/

[532] *Ibid*.

Concomitantly, in August 2023, the US Federal Communications Commission (FCC) issued a notice of inquiry (NOI), titled "Advancing Understanding of Non-Federal Spectrum Usage,"[533] concerning the potential applications of AI and other emerging technologies in the management of spectrum usage, with a view to determining whether AI techniques might prove useful in managing the use of spectrum between government and non-government entities.

In so doing, the FCC has publicly shown interest in making use of cutting-edge tools, such as AI, to improve spectrum management and improve their understanding of the actual usage of non-federal spectrum bands. The reason lies in the fact that the FCC's spectrum management policies have traditionally leaned on external parties to assess actual spectrum usage. Given the scarcity with which third-party information is made available and the fact that it is generally non-public, being therefore difficult to verify, it would be beneficial for the FCC to conduct its own studies of spectrum usage.[534]

In this line, on July 2023, the FCC, together with the National Science Foundation, hosted a workshop organized a forum with a focus on AI titled "The Opportunities and Challenges of Artificial Intelligence for Communications Networks and Consumers,"[535] where diverse stakeholders gathered to discuss the opportunities of the application of AI in spectrum management and network resiliency, along with the AI-derived challenges in terms of crucial consumer issues like robocalls/robotexts and digital discrimination.

Employing AI in spectrum allocation can enable more efficient spectrum usage, which is important given that radio frequencies and any associated orbits, including GEO, are limited natural resources. Thus, spectrum allocation should be carried out in a rational, efficient and economic manner, paying special attention to the developing countries' needs and the geographical location of particular countries in order to ensure an equitable access for all nations to those orbits and frequencies.[536]

This aim has led, for instance, to the utilization of these novel technologies by software-defined radios, such as cognitive radio, aimed at employing underutilized –yet already licensed– segments of the electromagnetic spectrum in the absence of human intervention. In the case of the FCC, this use of cognitive radios is permitted while the frequency remains unused by its primary user until the latter becomes active again.[537]

In terms of spectrum monitoring, AI can also play a pivotal role. Spectrum monitoring has consistently served as a crucial component of spectrum management processes, providing a comprehensive understanding of the spectrum environment and facilitating the planning,

---

[533] FCC, Advancing Understanding of Non-Federal Spectrum Usage, Notice of Inquiry. 4 August 2023, FCC Record Citation: 38 FCC Rcd 7216 (8). https://www.fcc.gov/document/spectrum-usage-noi

[534] Rosenworcel, J. 2023. August 2023 Open Meeting Agenda. FCC, 13 July 2023. https://www.fcc.gov/news-events/notes/2023/07/13/august-2023-open-meeting-agenda

[535] FCC, The Opportunities and Challenges of Artificial Intelligence for Communications Networks and Consumers. 13 July 2023. https://www.fcc.gov/fcc-nsf-ai-workshop

[536] Article 44(2), ITU Constitution 1992.

[537] Garner, R. 2017. "NASA Explores Artificial Intelligence for Space Communications." NASA News, 8 December 2017. https://www.nasa.gov/directorates/somd/space-communications-navigation-program/nasa-explores-artificial-intelligence-for-space-communications/

optimization, and enforcement of spectrum policies, aside from being key in identifying and mitigating potential interference, ensuring the efficient and lawful use of the spectrum.

AI can be deployed for enhancing the automation of spectrum monitoring, learning, prediction, and allocation tasks, which become increasingly convoluted with the development of new radio technologies. As a case in point, it may prove useful in improving signal recognition, enabling in-situ monitoring of a multitude of automated signaling apparatus and devices, and/or facilitating the identification of sources of interference.[538]

2.3 *Security and Privacy*

As the prevalence of wireless connectivity continues to grow, the necessity for robust security measures to protect users, network equipment, and data from malicious attacks, unauthorized access, and information leakage becomes increasingly apparent.[539] The advent of cognitive technologies represents a significant advancement in the design and architecture of communication systems.

It is posited that these technologies will enhance the efficacy and resilience of communications networks, particularly in the pursuit of missions to deep space. The incorporation of AI and cognitive radios into these systems promises to boost their efficiency, autonomy, and reliability.[540]

Moreover, there is a pressing need for the implementation of a secure data-sharing process for both previously identified and prospective vulnerabilities within the supply chain.[541] For example, the European Union's IRIS² (Infrastructure for Resilience, Interconnectivity, and Security by Satellite) constellation, designed to provide secure communication and high-speed broadband to the EU and its Member States, is exploring the incorporation of AI capabilities into its algorithms as a promising avenue. [542]

Conclusions

Among the various applications of AI in the field of telecommunications, those pertaining to communication protocols, security and privacy, and spectrum allocation are particularly noteworthy. AI has the potential to be harnessed for the establishment and maintenance of secure space communications networks, through the enhancement of communication network efficacy and resilience, notably in the context of deep space missions.

---

[538] ITU. AI will make radiocommunications smarter <-> Radiocommunications will enable AI functioning and connectivity. https://www.itu.int/en/action/ai/emerging-radio-technologies/Pages/default.aspx

[539] Zhang, C., Patras, P., and Haddadi, H. 2019. "Deep Learning in Mobile and Wireless Networking: A Survey," IEEE Communications Surveys & Tutorials 21(3), p. 39. doi: 10.1109/COMST.2019.2904897

[540] Garner. "NASA Explores Artificial… op. cit.

[541] Maguire, P. 2024. "AI at the crossroads of cybersecurity, space and national security in the digital age." SpaceNews, 3 April 2024. https://spacenews.com/ai-crossroads-cybersecurity-space-national-security-digital-age/

[542] European Union Agency for the Space Programme (EUSPA). IRIS².
https://www.euspa.europa.eu/eu-space-programme/secure-satcom/iris2

In particular, the deployment of AI in spectrum allocation can facilitate a more efficient and optimized usage of radio frequencies. This, in turn, may result in the adoption of advanced technologies such as cognitive radio by software-defined radios (SDRs).

Furthermore, AI has the capacity to streamline the automation of complex tasks related to spectrum monitoring, learning, prediction, and allocation, which is increasingly vital in a context where the advent of new radio technologies is rendering traditional approaches to spectrum management increasingly challenging.

Building on a long history of telecommunications regulatory foundation and standard setting, the ITU emerges as a critical player in the standardization of AI in telecommunications. The role of the ITU in standardizing the use of AI in telecommunications is well-established. This is due to the authority vested in the Union by its founding mandate and its status as one of the principal organs of the United Nations system, the preeminent international organization responsible for shaping the norms and standards that govern international law.

Consequently, the ITU is uniquely positioned to contribute to the development of international standards for AI in telecommunications, drawing upon its extensive technical expertise and network of stakeholders.

In this regard, it is noteworthy to highlight several rights enshrined in the ITU Constitution that are pertinent to the discussion of AI-driven telecommunication satellites. These include the right of the public to use the international telecommunication service for public correspondence (art. 33), the responsibility clause (art. 36), the use of the radio-frequency spectrum and of the geostationary-satellite and other satellite orbits (art. 44), and the prevention of harmful interference (art. 45), among others.

In terms of the 2024 Radio Regulations, the ITU's international treaty on the use of the radio-frequency spectrum and satellite orbits, not a single provision makes any reference to the potential role that AI may play in the field of satellite telecommunications. Nevertheless, it is conceivable that specific provisions of the Radio Regulations may have relevance to autonomous satellites (e.g., art. 15 on the prohibition of interference from radio stations).

Likewise, a review of the WRC-27 agenda reveals no mention of AI throughout its entirety. This absence prompts reflection on the potential inclusion of AI as an item in future WRCs. Considering the contemporary salience of this topic and the legal and regulatory gaps concerning the deployment of AI for telecommunications, such addition seems both timely and warranted.

Other recommended courses of action for the ITU to pursue in the forthcoming period are the formation of both a study group on AI in space telecommunications under the ITU-R framework and a collaborative evaluation by ITU on the effectiveness of current spectrum allocations for AI-driven space systems, together with the examination of potential regulatory frameworks for the advancement and deployment of AI-enabled satellites.

PART III: ETHICAL IMPLICATIONS

Section 1: Ethical Concerns in Genetics and Robotics: Case Studies and Their Relevance to the Use of AI in Space Activities

Introduction

The fusion of AI with genetics, genetic engineering (GE), and robotics is leading to transformative advancements in fields like personalized medicine and autonomous systems. AI is being utilized to process vast datasets, optimize complex processes, and automate decision-making in ways that push the boundaries of both biological science and engineering.

This section will delve into case studies that illustrate the impact of AI in two fields that have come about since the advent of human space activities, namely genetic engineering and autonomous robotics (and weapons) systems, where AI has driven considerable advancements, particularly in defense and space operations. In genetics and genetic engineering, AI has enhanced the research and testing of DNA and gene functionality by accelerating genome sequencing, aiding in the design of new genetic therapies, and optimizing tools like CRISPR-Cas9.

These advancements have paved the way for more efficient treatment strategies for complex genetic disorders such as cancer, cystic fibrosis, and Parkinson's disease. AI-driven tools are able to predict the effects of gene editing, helping scientists avoid unintended consequences, such as off-target mutations.

Case studies in this section will explore AI's role in disease biomarker correlations, improvements to gene-editing techniques, and the design of novel proteins through AI-guided synthetic biology. However, while these technological strides offer immense potential, they also bring ethical concerns and strong needs for regulation—ranging from data privacy and discrimination to the risks of destabilizing the human genome through excessive genetic manipulation.

At the same time, AI is playing a key role in the development of autonomous robotic systems, particularly in sensitive areas like military defense and space exploration. Case studies will examine the use of AI in autonomous weapons systems (LAWS), where AI allows for advanced targeting and decision-making with minimal human input, and in space operations, where AI helps satellites navigate and avoid collisions.

These robotic systems, powered by AI, offer enhanced operational capabilities in environments that are either too remote or hazardous for humans, such as space. However, their deployment raises significant questions about accountability, the ethics of delegating life-or-death decisions to machines, and the risk of unintended consequences, such as AI-driven satellites making unauthorized maneuvers or surveillance.

From genome editing to autonomous systems in warfare and space, these case studies will illustrate the importance of thoughtful governance and responsible technological development. By exploring these case studies, along with their particular ethical concerns and regulatory regimes, this section seeks to highlight the dual-edged nature of AI's rapid advancements in both genetics and robotics and provide lessons for prospective regulation of AI in the space domain.

For example, as discussed below with respect to AI, genetics, and space, experts have already proposed that we should genetically engineer humans for better adaptation to the space environment. This is not a question that may only impact one space-faring nation. Even a single astronaut genetically engineered in the right way could pass on engineered traits to their offspring, who could potentially take up residence in any country.

While the potential for groundbreaking discoveries and applications facilitated by AI is undeniable, the risks and ethical dilemmas they pose cannot be ignored. Based on the risks and potential negative outcomes exemplified in both of the case studies, proactive regulation of AI is clearly needed within critical domains such as genetics, autonomous robotics, and outer space.

1. *Genetics Case Studies & Implications for Ethical Use of AI in Space: AI in Human Genetic Testing & Editing*

1.1 *Ethical Dimensions of Genetic AI Applications*

Similarly to outer space technology and capabilities, the fields of AI, genetics and genetic engineering (GE) are advancing very rapidly, and both space and genetics have benefited greatly from steep new improvements in contemporary AI systems.

AI and GE have revolutionized the study and application of genetics, i.e., DNA, genes, their functionalities, as well as genetic therapies and treatment strategies, and have ushered in a new era of enhanced possibilities in biotech and personalized medicine.[543] In large part, this has been powered by recent exponential advancements in accurate predictions and optimizations by AI systems for GE technologies and methods, such as CRISPR-Cas9.[544]

Based on analysis of large datasets of genomes and genetic sequences, AI and machine learning algorithms can inform development of more efficient, effective, and accurate gene editing tools by predicting the most probable unintended consequences, which can result in unwanted damage or mutation to DNA.[545]

Because deploying AI in genome sequencing, analysis, and editing innately relies on machine learning and big data (i.e., very large datasets evaluated using complex neural network and other AI algorithms),[546] AI practitioners in genetics face inherent challenges regarding how to address data minimization mandates and implement appropriate data security measures.

---

[543] Rohit S Vilhekar and Alka Rawekar, “Artificial Intelligence in Genetics,” *Cureus*, January 10, 2024, https://doi.org/10.7759/cureus.52035; https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10856672/.

[544] Vilhekar, “Artificial Intelligence in Genetics.”

[545] Vilhekar, “Artificial Intelligence in Genetics”; N. Peiffer-Smadja et al., “Machine Learning in the Clinical Microbiology Laboratory: Has the Time Come for Routine Practice?,” *Clinical Microbiology and Infection* 26, no. 10 (October 1, 2020): 1300–1309, https://www.sciencedirect.com/science/article/pii/S1198743X20300859.

[546] Shriniket Dixit et al., “Advancing Genome Editing With Artificial Intelligence: Opportunities, Challenges, and Future Directions,” *Frontiers in Bioengineering and Biotechnology* 11 (January 8, 2024), https://doi.org/10.3389/fbioe.2023.1335901.

Deploying AI in genetics and GE can improve and optimize gene editing technologies, rapidly perform genetic sequencing and analysis of genomic data, as well as design new synthetic genes for new proteins and other biomaterials.[547] The following examples demonstrate some ways in which AI can benefit genetics and GE technologies:

- Disease biomarker correlations: AI has demonstrated the ability to analyze patient genomic data to identify correlations between genetic mutations and biomarkers associated with diseases like cancer, diabetes, and Parkinson's;[548]
- Genome editing: AI can propose improvements and optimizations for GE technologies and processes like base, prime, and epigenome editing, all of which can introduce precise alterations to DNA sequences;[549]
- Gene-editing tools: AI-designed gene-editing tools could outperform existing CRISPRs in medical or other applications. For example, large language model and other generative AI systems have demonstrated the ability to rapidly and vastly improve performance when trained and enabled to adaptively learn based on big data and, when deployed within the contexts genetics and GE, these types of capabilities could ultimately empower scientists to analyze, design, and edit genes much more quickly and precisely than current capabilities. This would enable medicines, therapies, and other treatments to be individually tailored to specific patients much faster;[550] and
- Protein design: AI can custom design task-specific proteins and other bio-compounds that, in combination with CRISPR-modified bacteria implemented within a biotech platform, can be used for mass-production of bespoke proteins and other bio-therapeutics. For example, such methods could potentially be employed to manufacture enzymes to degrade plastics into constituent materials for use as concrete fillers/substitutes or to upcycle carbon dioxide and methane into valuable organic feedstocks.[551]

However, despite all the potential incredible benefits and life-saving advantages, the combination of AI and GE also poses serious potential dangers, with underlying concerns ranging from ethics to national security. For example, the terrifying possibility of the emergence of a degraded and unstable genome looms as a potential existential threat to human biology, physiology, and survival, which could possibly arise due to excessive human GE tampering.[552]

---

[547] Vilhekar, "Artificial Intelligence in Genetics"; Dixit, "Advancing Genome Editing With Artificial Intelligence"; Crescenzio Gallo, "Artificial Intelligence for Personalized Genetics and New Drug Development: Benefits and Cautions," *Bioengineering* 10, no. 5 (May 19, 2023): 613, https://doi.org/10.3390/bioengineering10050613.

[548] Vilhekar, "Artificial Intelligence in Genetics."

[549] Dixit, "Advancing Genome Editing With Artificial Intelligence."

[550] Vilhekar, "Artificial Intelligence in Genetics."

[551] The Conversation, "AI and genetic engineering advancements could make designing new proteins possible," Fast Company, June 7, 2024, https://www.fastcompany.com/91137013/ai-genetic-engineering-advancements-could-make-designing-new-proteins-possible.

[552] Matan Arbel-Groissman et al., "The Causes for Genomic Instability and How to Try and Reduce Them Through Rational Design of Synthetic DNA," *Methods in Molecular Biology*, January 1, 2024, 371–92,

Some AI experts have claimed that AI advancements could potentially lead to the extinction of humanity.[553] Supercharging GE tech with generative AI before either technology has fully matured and has appropriate protections in place could potentially accelerate the pace of our own destruction.

Consequently, a deep urgency exists to address policy issues surrounding these emerging technologies, but the intersection of AI and GE technologies has not yet attracted sufficient attention and research from policy experts to develop meaningful perspectives.

GE technologies enable practitioners to make precise changes within DNA sequences of living cells and organisms.[554] The three most advanced GE technologies are zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and CRISPR-Cas-associated nucleases (CRISPR/Cas9), a molecular scissors-like tool that allows practitioners to cut out or remove, i.e., "knock out," and/or insert specific DNA sequences.[555]

CRISPR-based GE technologies like CRISPR/Cas9 enable precise and targeted editing of genetic sequences, providing a transformative innovation to the biotech toolbox.[556] Due to its versatility, effectiveness, and ease of use, CRISPR/Cas9 is the most common GE technology used to both "knock out" or remove defective genes and to insert new genes into cells.[557]

CRISPR-based GE technologies have developed techniques and capabilities for performing different types of gene editing,[558] including base editing,[559] prime editing,[560] and epigenome editing.[561] Each of these methods offers distinct benefits and drawbacks and can be valuable in specific circumstances.

---

https://doi.org/10.1007/978-1-0716-3658-9_21.

[553] Kevin Roose, "A.I. Poses 'Risk of Extinction,' Industry Leaders Warn," The NY Times, May 30, 2023, https://www.nytimes.com/2023/05/30/technology/ai-threat-warning.html; Matt Egan, "AI could pose 'extinction-level' threat to humans and the US must intervene, State Dept.-commissioned report warns," CNN, March 12, 2024, https://www.cnn.com/2024/03/12/business/artificial-intelligence-ai-report-extinction/index.html.

[554] Dixit, "Advancing Genome Editing With Artificial Intelligence"; Dongyuan Ma and Feng Liu, "Genome editing and its applications in model organisms," Genomics, Proteomics & Bioinformatics, January 4, 2016, 336–344, https://academic.oup.com/gpb/article/13/6/336/7224910.

[555] Dixit, "Advancing Genome Editing With Artificial Intelligence"; Thomas Gaj et al., "ZFN, TALEN, and CRISPR/Cas-based Methods for Genome Engineering," *Trends in Biotechnology* 31, no. 7 (July 1, 2013): 397–405, https://doi.org/10.1016/j.tibtech.2013.04.004; Thomas Gaj et al., "Genome-Editing Technologies: Principles and Applications," *Cold Spring Harbor Perspectives in Biology* 8, no. 12 (October 6, 2016): a023754, https://doi.org/10.1101/cshperspect.a023754.

[556] Dixit, "Advancing Genome Editing With Artificial Intelligence"; Swati Tyagi et al., "CRISPR-Cas9 System: A Genome-editing Tool With Endless Possibilities," *Journal of Biotechnology* 319 (August 1, 2020): 36–53, https://doi.org/10.1016/j.jbiotec.2020.05.008.

[557] Dixit, "Advancing Genome Editing With Artificial Intelligence"; Youmin Zhu, "Advances in CRISPR/Cas9," *BioMed Research International* 2022 (September 23, 2022): 1–13, https://doi.org/10.1155/2022/9978571; Leena Arora and Alka Narula, "Gene Editing and Crop Improvement Using CRISPR-Cas9 System," *Frontiers in Plant Science* 8 (November 8, 2017), https://doi.org/10.3389/fpls.2017.01932.

[558] Dixit, "Advancing Genome Editing With Artificial Intelligence."

[559] Nicole M. Gaudelli et al., "Programmable Base Editing of a•T to G•C in Genomic DNA Without DNA Cleavage," *Nature* 551, no. 7681 (November 1, 2017): 464–71, https://doi.org/10.1038/nature24644.

[560] Andrew V. Anzalone et al., "Search-and-replace Genome Editing Without Double-strand Breaks or Donor DNA," *Nature* 576, no. 7785 (October 21, 2019): 149–57, https://doi.org/10.1038/s41586-019-1711-4.

[561] Jacob H. Goell and Isaac B. Hilton, "CRISPR/Cas-Based Epigenome Editing: Advances, Applications, and Clinical Utility," *Trends in Biotechnology* 39, no. 7 (July 1, 2021): 678–91,

This proliferation of GE tools highlights the importance and need for effective AI in the process of making genetic tool design choices specifically tailored to address distinct situations or typical events in the genome editing process.[562]

Integration of AI with CRISPR enhances GE tool chains and pipelines, providing new invaluable insights and understandings, new functionalities and capabilities, and opportunities for intentionally manipulating and adapting genes and other genetic code sequences.

GE technologies can be used to treat human diseases in a number of ways.[563] For example, by using CRISPR/Cas9 to undo genetic mutations responsible for various diseases, such as tumor suppressor genes involved in cancer or targeting certain genes implicated in, e.g., sickle cell anemia, cystic fibrosis, and some cardiovascular diseases.[564]

Further, CRISPR/Cas9 and other GE tech can target genes responsible for neurodegenerative diseases like Parkinson's, Alzheimer’s, and Huntington’s,[565] and, they can even be used to create cells and organisms resistant to viral infections such as HIV and Hepatitis B.[566]

However, CRISPR-Cas tools have also proven to have unintended effects, both on-target and off-target, and to be more susceptible to unintended off-target effects than other conventional gene-editing methods, potentially resulting in unwanted DNA damage or mutations.[567]

CRISPR/Cas9 can cause unintended DNA mutations at unintended locations within the human genome, even when it was not designed to target that specific site. The full potential effects of these off-target effects remain unknown.[568]

Even on-target events can have unintended consequences. For example, CRISPR/Cas9 can cut or cause breaks double-strand DNA at intended locations, but when such DNA breaks are not properly repaired mismatched ends of DNA can join, leading to potentially large-scale genetic rearrangements.[569]

CRISPR/Cas9 can also cause structural abnormalities in genes and chromosomes, potentially leading to cancer-related changes.[570] A study recently published in the journal Cell found that CRISPR/Cas9 can cause serious and potentially fatal side effects in human embryos,

---

https://doi.org/10.1016/j.tibtech.2020.10.012.

[562] Dixit, “Advancing Genome Editing With Artificial Intelligence.”

[563] Dixit, “Advancing Genome Editing With Artificial Intelligence”; Hongyi Li et al., “Applications of Genome Editing Technology in the Targeted Therapy of Human Diseases: Mechanisms, Advances and Prospects,” *Signal Transduction and Targeted Therapy* 5, no. 1 (January 3, 2020), https://doi.org/10.1038/s41392-019-0089-y.

[564] Dixit, “Advancing Genome Editing With Artificial Intelligence.”

[565] *Ibid.*

[566] Dixit, “Advancing Genome Editing With Artificial Intelligence”; Li, “Applications of Genome Editing Technology.”

[567] Wenyi Liu et al., “Applications and Challenges of CRISPR-Cas Gene-editing to Disease Treatment in Clinics,” *Precision Clinical Medicine* 4, no. 3 (July 10, 2021): 179–91, https://doi.org/10.1093/pcmedi/pbab014.

[568] *Ibid.*

[569] *Ibid.*

[570] Li, “Applications of Genome Editing Technology”; Liu, “Applications and Challenges of CRISPR-Cas Gene-editing”; Michael V. Zuccaro et al., “Allele-Specific Chromosome Removal After Cas9 Cleavage in Human Embryos,” *Cell* 183, no. 6 (October 29, 2020): 1650-1664.e15, https://doi.org/10.1016/j.cell.2020.10.025.

including DNA breakage, modified DNA sequences, altered gene and chromosomal structures, and even loss of entire chromosomes.[571]

Nevertheless, progress in CRISPR-based therapies and treatments has led to numerous clinical trials in the rapidly evolving cell and gene therapy sectors. Recent advances in multi-omics technologies generate big data from various sources, such as genes, DNA, RNA, and proteins, and have made AI necessary for analyzing and understanding medical information.[572]

Deep learning (DL) and Machine Learning (ML) models have demonstrated great success in analysis and sense-making on large, complex genetic databases.[573] Such studies could prove invaluable in helping to identify more appropriate architectures and functionalities for AI models, thereby enhancing AI capabilities in predicting editing outcomes such as off-target effects.

For example, in cancer, AI models can utilize genomic data to identify cancer subtypes, and CRISPR-based GE can assist in engineering immune cells that target specific cancer subtypes or disrupt/replace oncogenes.[574]

An ML algorithm known as SPROUT developed and published in 2019 by Leenay et al., demonstrated high accuracy predictions for GE repair outcomes in primary T cells. Trained on a large dataset of CRISPR-Cas9 editing events and outcomes, SPROUT can help design CRISPR experiments to maximize desired editing outcomes and develop new genetic therapies.[575]

Cas9 is not the only CRISPR option for GE, and variants of Cas proteins currently being investigated for similar purposes include CRISPR-Cas12,[576] CRISPR-Cas13,[577] CRISPR-Cas3,[578] and many others. Consequently, with so many tools in the GE toolbox, development of effective

---

[571] Zuccaro et al., "Allele-Specific Chromosome Removal"; So Hyun Park, Mingming Cao, and Gang Bao, "Detection and Quantification of Unintended Large On-Target Gene Modifications due to CRISPR/Cas9 Editing," *Current Opinion in Biomedical Engineering*, no. 28 (December 2023), https://www.sciencedirect.com/science/article/abs/pii/S246845112300034X.

[572] Dixit, "Advancing Genome Editing With Artificial Intelligence"; Pavel Hamet and Johanne Tremblay, "Artificial Intelligence in Medicine," *Metabolism* 69 (April 1, 2017): S36–40, https://www.sciencedirect.com/science/article/abs/pii/S002604951730015X.

[573] Dixit, "Advancing Genome Editing With Artificial Intelligence"; Sameer Quazi, "Artificial Intelligence and Machine Learning in Precision and Genomic Medicine," *Medical Oncology* 39, no. 8 (June 15, 2022), https://doi.org/10.1007/s12032-022-01711-1.

[574] Dixit, "Advancing Genome Editing With Artificial Intelligence"; Alyna Katti et al., "CRISPR in Cancer Biology and Therapy," *Nature Reviews. Cancer* 22, no. 5 (February 22, 2022): 259–79, https://doi.org/10.1038/s41568-022-00441-w.

[575] *Ibid*.

[576] Dixit, "Advancing Genome Editing With Artificial Intelligence"; Renjian Xiao et al., "Structural Basis of Target DNA Recognition by CRISPR-Cas12k for RNA-guided DNA Transposition," *Molecular Cell* 81, no. 21 (November 1, 2021): 4457-4466.e5, https://doi.org/10.1016/j.molcel.2021.07.043; Raghul Senthilnathan et al., "An Update on CRISPR-Cas12 as a Versatile Tool in Genome Editing," *Molecular Biology Reports* 50, no. 3 (January 15, 2023): 2865–81, https://doi.org/10.1007/s11033-023-08239-1.

[577] Dixit, "Advancing Genome Editing With Artificial Intelligence"; Naga Rajitha Kavuri et al., "Applications of CRISPR/Cas13-Based RNA Editing in Plants," *Cells* 11, no. 17 (August 27, 2022): 2665, https://doi.org/10.3390/cells11172665.

[578] Dixit, "Advancing Genome Editing With Artificial Intelligence"; Hiroyuki Morisaka et al., "CRISPR-Cas3 Induces Broad and Unidirectional Genome Editing in Human Cells," *Nature Communications* 10, no. 1 (December 6, 2019), https://doi.org/10.1038/s41467-019-13226-x.

AI approaches is increasingly critical to appropriate determination of the most effective tool for a task at hand.[579]

Additional considerations at the intersection of genetics, AI and outer space, including the ethics involved, are relevant to this discussion. Genetic testing and gene editing can raise ethical concerns in several areas, including informed consent, privacy, confidentiality, data security, equity, discrimination and bias, eugenics, human enhancement, safety, and the stability of the human genetic lineage.

1.1.1 Informed Consent

People have the right to make informed decisions about whether to undergo genetic testing and who may have knowledge of their genetic information. Before undergoing any medical intervention, including genetic testing or gene editing, individuals or their legal guardians should fully understand the known and potential unknown risks, benefits, and potential long-term consequences.

1.1.2 Privacy, Confidentiality, Data Security, and Life Decisions

A key consideration within the domains of privacy, confidentiality, data security, and life decisions is whether family members should be informed of genetic test results, especially in instances of genetic predisposition towards disease.

Presently, depending on country and jurisdiction, genetic information and test results may need to be disclosed to third parties, and there may be a duty to warn at-risk relatives. Some people question whether genetic developments are worth the loss of privacy.

Secure and confidential storage and maintenance of genetic data are also a primary concern for privacy considerations related to genetics as evidenced by multiple incidents of hackers targeting genetic information, e.g., 23andMe, a direct-to-consumer genetics testing company, suffered a security data breach affecting almost 7 million users.[580]

Genetic testing can also lead people to make life-changing decisions, such as not having children, not marrying, or choosing certain careers. And gene editing also comprises a life altering decision, whether to somatic cell lines affecting only the individual or to germ line cells comprising inheritable traits that may affect future members or become a dominant trait of the lineage.

1.1.3 Equity, Discrimination, Bias, and Eugenics

Issues of justice, fairness, and equity can arise in relation to genetic testing. People with genetic disorders may face discrimination from employers, insurance companies, or society. For example, life insurance companies may use genetic testing information to make decisions about coverage and premiums.

---

[579] Dixit, "Advancing Genome Editing With Artificial Intelligence."

[580] Mack Degeurin, "Hackers Got Nearly 7 Million People's Data From 23andMe. The Firm Blamed Users in 'Very Dumb' Move," *The Guardian*, February 26, 2024, https://www.theguardian.com/technology/2024/feb/15/23andme-hack-data-genetic-data-selling-response.

When it comes to gene editing, discrimination and bias can verge upon crimes against humanity when followed through to eugenics. Moreover, how can we curtail and ensure the inclination to augment and enhance humanity—indeed, to create superhumans[581] via genetics and AI—does not become the perpetration of eugenics against our common natural human genetic heritage.

And ultimately, the ethical question of when and where, i.e., whether, to apply ethics also needs to be considered. For instance, it is predicted that by 2030 China will emerge as the world leader in biological and military applications of Human Performance Enhancement Technologies (HPET). These advancements promise incredible outcomes such as eliminating diseases, improved vision, increased strength, and decreased fatigue, all contributing to overall enhanced, 'super' human performance.

China's national strategy of civil-military fusion endeavors a superior biotechnology program focused on gene editing and HPETs to create specific traits, such as high intelligence or reptilian night vision in Chinese soldiers.

However, this remains a complex and delicate undertaking. As these desired features are influenced by many genetic locations with only minute impacts on the specific outcome, reaching their desired result without introducing dangerous side effects remains an ambitious task that lies more than five to ten years away.

Over the next five years, scientists anticipate enhanced precision in editing DNA to repair any damage to a sequence. While challenges may linger, tremendous progress is expected in searching for new genetic repair solutions.[582] And, once super abled humans have arrived, what benefits may they perceive in keeping around all of the congenitally natural humans?

1.1.4 Genetic Editing, Human Enhancement, and Safety

There are many unknowns about the safety of gene editing, including the potential for off-target mutations and multigenerational effects. Gene editing could be used to enhance human traits like height, intelligence, or athletic ability, which could raise concerns about human dignity and accessibility

Some ethical questions surrounding genetics, gene therapies, and genomic editing include:

- "How can "good" and "bad" uses of these technologies be distinguished?
- Who decides which genetic traits are normal and which constitute a disability or disorder?
- Will high costs associated with gene therapy make it available only to the wealthy?

---

[581] "U.S. Official Says China Attempted to Create 'super Soldiers,'" December 4, 2020, https://www.nbcnews.com/politics/national-security/china-has-done-human-testing-create-biologically-enhanced-super-soldiers-n1249914.

[582] "456. China: Leader in Military Application of Biological Human Performance Enhancement by 2030," *Mad Scientist Laboratory* (blog), August 10, 2023, https://madsciblog.tradoc.army.mil/456-china-leader-in-military-application-of-biological-human-performance-enhancement-by-2030/.

- Could the widespread use of gene therapy make society less accepting of people who are different?
- Should people be allowed to use gene therapy to enhance basic human traits such as height, intelligence, or athletic ability?"[583]

1.2 *Outer Space Considerations and Concerns for Genetic Editing and AI*

The idea of human enhancement to better permit living and working in space is simple: because the space environment is hazardous and humans are not adapted by evolution to live there, it makes sense to artificially increase human adaptation to space by biomedical means.

However, the precise nature of enhancement, which may be more or less invasive, reversible or irreversible, and heritable or non-heritable, requires very careful thought and might well be driven by scientific and ethical considerations on Earth. And how we determine what constitutes and who receives equitable access is a crucial question that will need to be answered.

"Genetic engineering, particularly germ-line gene editing, is one of the most controversial forms of bio-enhancement, at least on Earth. However, there are good reasons to assume that in the context of space, there is a stronger rationale for human enhancement than in the terrestrial domain, in which case the ethical analysis should also differ."[584]

With respect to outer space, GE and gene therapies could potentially enhance humans to better live and work in space by making them more resistant and resilient to harmful effects of the space environment. Some desirable traits for astronauts include:

- radiation resistance: space travelers are exposed to increased radiation outside of the Earth atmosphere and magnetic field and providing their DNA with the ability to better withstand and reverse mutations caused by solar, galactic, and cosmic radiation could help them stay healthy in space;
- microgravity resistance: the decreased pull of gravity in space can affect bones and eyes; and
- other high-performance traits: including reduced need for sleep, higher intelligence, stronger bones and muscles, modified metabolism, increased endurance, and reduced pain sensitivity.[585]

> "Modifying the DNA of future astronauts through gene therapies or gene editing to make them more resistant to radiation and other threats from long-term space travel is an ambitious strategy.
>
> - In 2016, geneticist George Church identified more than 40 genes that could be targeted for long-term spaceflight.
> - Another option may involve combining the DNA of other, radiation-hardened species—like microscopic tardigrades—with humans.

---

[583] "What Are the Ethical Issues Surrounding Gene Therapy?" February 28, 2022, https://medlineplus.gov/genetics/understanding/therapy/ethics/.

[584] Szocik, "Future Space Missions and Human Enhancement."

[585] *Ibid.*

- A group at Duke University is trying to tease out the secrets behind the tardigrade's resistance and possibly translate it to other organisms.
- "If that is interesting, and shows promise, then can we put those into rodent models? Can we look into human cell lines—starting to tell the translational story from simple organisms all the way up to people?" Kristin Fabre, of the Translational Research Institute for Space Health, told Axios."[586]

Likely also of special relevance to surviving and thriving in outer space, Alcino J. Silva, a neurobiologist at the University of California, Los Angeles, and team uncovered a major new role of the CCR5 gene in memory and the ability of the brain to form new neural connections.[587]

In 2019, they revealed that the brains of two twin girls born in China in 2018, whose genes were edited by scientists that used CRISPR to make them HIV-immune by removing their CCR5 genes, appear to exhibit changes including enhanced cognition and memory.[588]

Research shows that deletion of the CCR5 gene in mice makes them smarter, and for humans it also makes the brain more resilient and enables improved recovery following stroke and may improve success in school.[589] It is relatively easy to extrapolate that improved resiliency to stroke could also potentially improve resiliency to damage suffered by neurons and the brains of astronauts due to cosmic radiation.

One of the most obvious and realistic body modifications has to do with the human balance system: the vestibular organs of the inner ear and their interactions with other senses to maintain spatial orientation and compensatory reflex responses.[590]

The role of the vestibular system in space exploration has been recognized since the dawn of the aerospace age. It has a central role since one part of the vestibular apparatus is dedicated to measuring linear acceleration and gravity. When gravity is altered, there are consequences: disorientation, nausea, ataxia, and motion sickness.[591]

Travelers in space and on planets with different gravity levels will be faced with challenges related to these factors. Most critical might be problems with manual-control tasks such as piloting, while reduced performance in general can result from the associated malaise.

Vestibular adaptation does occur, but body function can be dangerously deficient in the initial phases of the transition to a new gravity level. However, Earth patients with some types of

---

[586] Miriam Kramer and Bryan Walsh, "Radiation-proofing the human body for long-term space travel," Axios, September 29, 2020, https://www.axios.com/2020/09/29/gene-editing-radiation-space-travel.

[587] Antonio Regalado, "China's CRISPR Twins Might Have Had Their Brains Inadvertently Enhanced," *MIT Technology Review*, April 2, 2020, https://www.technologyreview.com/2019/02/21/137309/the-crispr-twins-had-their-brains-altered/.

[588] *Ibid.*

[589] *Ibid.*

[590] Szocik, "Future Space Missions and Human Enhancement;" Jay Goldberg et al., *The Vestibular System: a Sixth Sense* (Oxford University Press, 2012). https://www.researchgate.net/profile/Jay-Goldberg-4/publication/285931494_The_Vestibular_System_A_Sixth_Sense/links/573240a108aea45ee8364231/The-Vestibular-System-A-Sixth-Sense.pdf.

[591] Szocik, "Future Space Missions and Human Enhancement;" Millard Reschke et al., "Space Flight and Neurovestibular Adaptation," *The Journal of Clinical Pharmacology* 34, no. 6 (June 1, 1994): 609–17, https://doi.org/10.1002/j.1552-4604.1994.tb02014.x.

vestibular pathology face similar issues, and implantable vestibular prostheses are under development which can replace some of the lost function.[592] In addition, the vestibular system could potentially be modified genetically to be better able to adapt to the space environment.

More expansively, safety and performance might be improved if the visual system were to be enhanced to respond to a wider range of wavelengths: infrared, ultraviolet, radio, microwave, and X-ray. This would allow direct perception of things such as heat signatures, permit vision through darkness and dust storms, and provide visual indications of high radiation areas. Specialized retinal implants,[593] though still in the early stages of development, might provide this capability.

For a variety of reasons associated with the multiple interacting stressors of space flight, the gut microbiome (fungus, bacteria, and other intestinal microorganisms) can also be altered.[594] This can cause widespread disruptions because of interactions of the microbiome with many other body systems, including cognition.[595]

Adding to the problem is the fact that in space there will be less regular turnover of the biome as normally results from contact with a wide variety of other organisms. To remedy this, an implantable pump might be used to provide a regular infusion of new microbiome constituents, similar in principle to probiotic supplements or fecal transplants.[596]

While there may be simpler ways to introduce such compounds into the body, a permanent pump could allow for administration of a wider variety of substances and permit constant monitoring and adjustment as needed in order to maintain physiological fitness.

---

[592] Szocik, "Future Space Missions and Human Enhancement;" Justin Golub et al., "Prosthetic Implantation of the Human Vestibular System," *Otology & Neurotology* 35, no. 1 (January 1, 2014): 136–47, https://doi.org/10.1097/mao.0000000000000003.

[593] Szocik, "Future Space Missions and Human Enhancement;" Edward Bloch, Yvonne Luo, and Lyndon Da Cruz, "Advances in Retinal Prosthesis Systems," *Therapeutic Advances in Ophthalmology* 11 (January 1, 2019): 251584141881750, https://doi.org/10.1177/2515841418817501; Yvonne Luo and Lyndon da Cruz, "A Review and Update on the Current Status of Retinal Prostheses (Bionic Eye)," *British Medical Bulletin* 109, no. 1 (February 12, 2014): 31–44, https://doi.org/10.1093/bmb/ldu002.

[594] Szocik, "Future Space Missions and Human Enhancement;" R. Siddiqui et al., "Gut Microbiome and Human Health Under the Space Environment," *Journal of Applied Microbiology* 130, no. 1 (August 10, 2020): 14–24, https://doi.org/10.1111/jam.14789; Alexander A. Voorhies and Hernan A. Lorenzi, "The Challenge of Maintaining a Healthy Microbiome During Long-Duration Space Missions," *Frontiers in Astronomy and Space Sciences* 3 (July 22, 2016), https://doi.org/10.3389/fspas.2016.00023.

[595] Szocik, "Future Space Missions and Human Enhancement;" Mélanie G Gareau, "Microbiota-Gut-Brain Axis and Cognitive Function," *Advances in Experimental Medicine and Biology*, January 1, 2014, 357–71, https://doi.org/10.1007/978-1-4939-0897-4_16; Andrew B Shreiner, John Y Kao, and Vincent B Young, "The Gut Microbiome in Health and in Disease," *Current Opinion in Gastroenterology* 31, no. 1 (January 1, 2015): 69–75, https://doi.org/10.1097/mog.0000000000000139.

[596] Szocik, "Future Space Missions and Human Enhancement;" Silvia Turroni et al., "Gut Microbiome and Space Travelers' Health: State of the Art and Possible Pro/Prebiotic Strategies for Long-Term Space Missions," *Frontiers in Physiology* 11 (September 8, 2020), https://doi.org/10.3389/fphys.2020.553929.

1.3 *Ethical Review*

Due to the myriad capabilities and potential outcomes for genetic testing and GE technologies, developing a thorough understanding of the potential risk-benefit landscape can be challenging, further complicating ethical analyses and the adoption of appropriate cultural norms.

Even if AI were capable of guiding GE technologies to perfect outcomes, what are the appropriate ethics for a world where no one ever dies from natural causes, age-related or otherwise? As in, if we as humanity become capable of providing every individual with a "perfect" genome, i.e., one totally free of all predispositions towards disease, what else could or should we change "for the better"?

For instance, what about super intelligence, super strength, super endurance, super pain tolerance, etc.? According to a 2022 survey conducted by the Pew Research Center, 71% of Americans support use of gene editing to treat serious diseases or health conditions that a person currently has, while 74% oppose using gene editing to enhance physical appearance.[597]

In the U.K. a poll commissioned by the Progress Educational Trust, a fertility and genomics charity, found 53% of respondents supported use of GE and gene therapy to edit human embryonic genomes to prevent the development of serious and life-threatening conditions such as cystic fibrosis.[598]

Whereas a lower proportion supported use of gene editing to prevent milder conditions such as asthma, at 36%. And only 20% supported use of gene editing to create 'designer' embryos. However, personal views on the topic varied greatly with age, and nearly 40% of 16–24-year-old respondents supported use of GE and gene therapy to create designer babies.[599]

According to a 2020 survey conducted by the Pew Research Center, roughly 30% of the global populace supports scientific research on gene editing (56% in India). About 70% support the use of GE and gene therapy to modify embryonic genetic codes to prevent serious diseases or conditions the baby would have at birth (88% in Spain).[600]

Some 60% support GE and gene therapy to modify embryonic DNA to reduce the risk of a serious disease or condition that could occur over the lifetime of the individual (77% in Spain). And only about 14% support use of GE and gene therapy to modify an embryo to make it more intelligent. Views varied greatly across the globe on this topic, with 64% of Indians supporting GE enhanced intelligence.[601]

---

[597] Lee Rainie et al., "7. Americans Are Closely Divided Over Editing A Baby's Genes to Reduce Serious Health Risk," Pew Research Center, July 22, 2024, https://www.pewresearch.org/internet/2022/03/17/americans-are-closely-divided-over-editing-a-babys-genes-to-reduce-serious-health-risk/.

[598] Ian Sample, "Half in UK Back Genome Editing to Prevent Severe Diseases," *The Guardian*, June 22, 2022, https://www.theguardian.com/science/2022/jun/22/half-in-uk-back-genome-editing-to-prevent-severe-diseases.

[599] Sample, "Half in UK Back Genome Editing."

[600] Cary Funk et al., "Biotechnology Research Viewed With Caution Globally, but Most Support Gene Editing for Babies to Treat Disease," Pew Research Center, July 15, 2024, https://www.pewresearch.org/science/2020/12/10/biotechnology-research-viewed-with-caution-globally-but-most-support-gene-editing-for-babies-to-treat-disease/.

[601] Funk, "Biotechnology Research Viewed With Caution Globally."

For now, it might be that, with the nascent state of AI GE technologies, we should err on the side of caution and focus on perfecting our technologies and techniques rather than proceeding too aggressively and potentially making an extinction level event mistake.

Is it also possible that at some point in the evolution of these technologies that an inflection point in the ethical analysis could be achieved where the robustness of AI GE technologies not only allow but demand that we use these technologies to maximize the quality of life for humanity to the greatest extent possible?

What about 100 or 200 years from now, maybe even 25 to 50 years from now, when the technology has been thoroughly debugged and de-risked? If these capabilities truly come within our grasp, would there not be a moral imperative to perfect their applications for the benefit of our descendants and their descendants?

What about the ethical imperative for a bio-diverse community, e.g., genetically and phenotypically? The results from many, many research investigations demonstrate that biological diversity, both intra- and extra-species, is a cornerstone to environmental and individual health.

How can the minds of humans, or even the artificial minds of AI, fully comprehend all the potential consequences and imbalances to natural equilibrium that could arise from the hubris of a few to make irreversible changes within the evolutionary course of humanity and potentially all life? And, at the same time, we have already seen such acts occur and attempts to prohibit rather than regulate would only drive the research underground, precluding ourselves from benefiting from its potential advantages to the many intractable challenges humanity faces.

Beyond potential downsides to mere humans, might it be possible for the deployment of AI GE technologies to undermine the stability of the genetic foundation of all life? Consider for instance, the unknown potential outcomes of the merger of environmental engineering, AI, GE, synthetic biology, hachimoji DNA, human-animal chimeras, and biobots.

1.4 *Policy and Governance*

Several countries have laws related to genetic testing, including mandatory testing, premarital screening, and restrictions on direct-to-consumer (DTC) testing. For instance, many governments in the Gulf Cooperation Council countries require premarital genetic testing for genetic and sexually transmitted diseases. This is due to their high incidence of birth defects and genetically inherited blood disorders, such as sickle cell disease and thalassemia.[602]

Some western countries, such as Switzerland and Germany, have laws founded on protecting personal genetic information. Germany's 2009 Human Genetic Examination Act, for instance, bans direct-to-consumer genetic testing and requires full consent of all parties involved

---

[602] "Bad Blood: Tackling Genetic Disorders in the Gulf," Newsletter, The Center for Strategic and International Studies, July 8, 2010, https://www.csis.org/analysis/bad-blood-tackling-genetic-disorders-gulf; Rahma Al-Kindi et al., "Awareness and Attitude Towards the Premarital Screening Programme Among High School Students in Muscat, Oman," *Sultan Qaboos University Medical Journal* 19, no. 3 (November 5, 2019): 217, https://doi.org/10.18295/squmj.2019.19.03.007.

for doctors to perform genetic tests. France also has laws proscribing private genetic paternity tests, including tests conducted by foreign genetics labs.[603]

The situation is still very much in flux, with respect to global policy and governance of GE technologies. As discussed above, in November 2018, Chinese researcher He Jiankui reported the birth of twin girls whose genomes were edited as embryos using CRISPR/Cas9 to make them impervious to HIV.[604]

Subsequently, "in March 2019 a number of scientists, including Françoise Baylis, Feng Zhang, and Emmanuelle Charpentier, called for a global moratorium on heritable genome editing to allow time for the discussions needed to establish an international framework. An International Commission on the Clinical Use of Human Germline Genome Editing was also convened and has recently released a report in September 2020."[605]

1.4.1 Global Policy and Legal Landscapes

In 2020 researchers evaluated the contemporary policy and legal landscape by reviewing 125 policy documents from 96 countries on human germline and heritable genome editing. Only 40 out of 96 countries had policies and laws specifically addressing germline genome editing, i.e., not for reproduction, with 23 countries prohibiting such research, and 11 explicitly permitting use of genetically modified embryos solely in *in vitro* research.[606] The researchers found no countries that explicitly permitted use of such embryos to initiate pregnancies.[607]

Whereas 78 out of 96 countries had policies and laws addressing heritable genome editing, i.e., for reproduction, with 70 prohibiting GE for human embryos, five (5) countries proscribing it with limited exceptions (Columbia, Panama, Belgium, Italy, and the UAE), and three (3) were indeterminate (Burkina Faso, Singapore and Ukraine). None of the states participating in the survey expressly allow heritable human genome editing.[608]

Yet, with 195 recognized countries in the world, the potential for "ethics dumping," a practice where researchers move their work to countries with no laws prohibiting such research, or those with little to no capability to enforce such laws.[609] Ethics dumping has been observed in

---

[603] L. Kalokairinou et al., "Legislation of Direct-to-consumer Genetic Testing in Europe: A Fragmented Regulatory Landscape," *Journal of Community Genetics* 9, no. 2 (November 18, 2017): 117–32, https://doi.org/10.1007/s12687-017-0344-2.

[604] Farah Qaiser, "Study: There Is No Country Where Heritable Human Genome Editing Is Permitted," *Forbes*, November 2, 2020, https://www.forbes.com/sites/farahqaiser/2020/10/31/study-there-is-no-country-where-heritable-human-genome-editing-is-permitted/.

[605] Qaiser, "Study: There Is No Country Where Heritable Human Genome Editing Is Permitted"; *Heritable Human Genome Editing*, *National Academies Press eBooks*, 2020, https://doi.org/10.17226/25665; "United States: Germline / Embryonic," Global Gene Editing Regulation Tracker, December 31, 2019, https://crispr-gene-editing-regs-tracker.geneticliteracyproject.org/united-states-embryonic-germline-gene-editing/.

[606] Qaiser, "Study: There Is No Country Where Heritable Human Genome Editing Is Permitted."

[607] *Ibid*.

[608] *Ibid*.

[609] *Ibid*.

other fields, including assisted reproductive technologies and unproven, unapproved stem cell therapies.[610]

1.4.2 US Status Quo on Policy and Governance of Genetic Engineering Technologies

For instance, the U.S. Government prohibits use of federal funds for research on germline gene therapy in humans due to ethical concerns. In part, because the people who would be most directly affected by germline GE and gene therapies have yet to be born and, by definition, are thereby precluded from participating in an informed decision on whether to receive such treatment.[611]

However, the U.S. does not have any laws or regulations prohibiting germline GE and gene therapies conducted using private funds.[612] Consequently, it is hypothetically possible to conduct non-clinical, human GE and gene therapy research at a privately funded lab in the U.S.[613]

Some examples of GE and gene therapy research conducted in the U.S. include Columbia University researchers who used CRISPR to remove the gene for retinitis pigmentosa, an inherited form of blindness, from embryos solely for research purposes.[614] And researchers at Oregon Health & Science University used CRISPR to remove a gene that causes cardiomyopathy, a potentially fatal heart condition, from human embryos, also solely for research purposes.[615]

Nevertheless, to commercialize and sell any such therapies in the U.S. market would require US Food & Drug Administration (FDA) approval for clinical studies and marketing.[616] In January 2024, the FDA issued its Guidance for Industry on Human Gene Therapy Products Incorporating Human Genome Editing, in which it provides recommendations on developing human gene therapy products incorporating genome editing (GE) of human somatic cells.[617]

Specifically, it covers information that should be provided in an Investigational New Drug (IND) application in order to assess the safety and quality of the investigational GE product, as required in Title 21 of the Code of Federal Regulations 312.23 (21 CFR 312.23), including information on product design, manufacturing, and testing, nonclinical safety assessment, and clinical trial design.[618]

---

[610] *Ibid*.

[611] "What Are the Ethical Issues Surrounding Gene Therapy?"; "United States: Germline / Embryonic."

[612] "United States: Germline / Embryonic."

[613] *Ibid*.

[614] "United States: Germline / Embryonic"; "CRISPR Used to Repair Blindness-causing Genetic Defect in Patient-derived Stem Cells," Columbia University Irving Medical Center, January 27, 2016, https://www.cuimc.columbia.edu/news/crispr-used-repair-blindness-causing-genetic-defect-patient-derived-stem-cells.

[615] "United States: Germline / Embryonic"; Heidi Ledford, "CRISPR Fixes Disease Gene in Viable Human Embryos," *Nature* 548, no. 7665 (August 1, 2017): 13–14, https://doi.org/10.1038/nature.2017.22382.

[616] "United States: Germline / Embryonic."

[617] "Human Gene Therapy Products Incorporating Human Genome Editing Guidance for Industry," FDA, January 2024, https://www.fda.gov/regulatory-information/search-fda-guidance-documents/human-gene-therapy-products-incorporating-human-genome-editing.

[618] *Ibid*.

However, for germline GE and gene therapies FDA has, since 2016,[619] operated under a consistently-renewed[620] congressional appropriations bill funding restriction against evaluating or responding to requests to approve or otherwise support any clinical trials for therapies that include a human embryo "intentionally created or modified to include a heritable genetic modification."[621]

2. *Robotics Case Study & Implications for Ethical Use of AI in Space: Lethal Autonomous Weapon Systems (LAWS)*

2.1 *Introduction to LAWS*

An autonomous weapon system is pre-programmed to target (in the case of non-lethal autonomous weapons systems)[622] or kill a specific "target profile."[623] The weapon is then deployed into an environment where its AI searches for that "target profile" using sensor data, such as facial recognition. Autonomous systems with adaptive learning capabilities are increasingly being deployed in various fields. However, no commonly agreed definition of LAWS exists.[624]

States are increasingly developing and deploying weapons with autonomous functions. Nevertheless, certain systems incorporating basic autonomous functions have been around for decades.[625]

The most common types of weapons with autonomous functions are defensive systems. This includes systems such as antivehicle and antipersonnel mines, which, once activated, operate autonomously based on trigger mechanisms.

New systems using advanced technologies include missile defense and sentinel systems, which can autonomously detect and engage targets, as well as issue warnings. Other examples include standby munitions, which contain an integrated warhead (munition) and wait (standby) around a predefined area until a target is located by an operator on the ground or by automated sensors on board and then engage the target.

These systems first appeared in the 1980s, but their functionalities have become increasingly sophisticated, allowing, among other things, longer ranges, heavier payloads and the potential incorporation of AI technologies and the use of human protocols in the loop.

---

[619] "United States: Germline / Embryonic"; United States Congress, *CONSOLIDATED APPROPRIATIONS ACT, 2016*, 2015, https://www.congress.gov/114/plaws/publ113/PLAW-114publ113.pdf.

[620] "United States: Germline / Embryonic"; House Republican Conference, "CONSOLIDATED APPROPRIATIONS ACT, 2024," 2024, https://appropriations.house.gov/sites/republicans.appropriations.house.gov/files/First%20FY24%20Package%20-%20Consolidated%20Appropriations%20Act%2C%202024.pdf.

[621] Jocelyn Kaiser, "Update: House spending panel restores U.S. ban on gene-edited babies," Science, June 4, 2019, https://www.science.org/content/article/update-house-spending-panel-restores-us-ban-gene-edited-babies

[622] Alexander Blanchard et al., "Jus in Bello Necessity, The Requirement of Minimal Force, and Autonomous Weapons Systems", *Journal of Military Ethics* 21(3-4), 2022,286-303.

[623] Autonomous weapons, https://autonomousweapons.org/.

[624] Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, 10 March 2023, CCW/GGE.1/2023/CRP.1.

[625] United Nations, Office for Disarmament Affairs, https://disarmament.unoda.org/the-convention-on-certain-conventional-weapons/background-on-laws-in-the-ccw/.

Land and sea vehicles with autonomous capabilities are also increasingly being developed. Those systems are primarily designed for reconnaissance and information gathering but may possess offensive capabilities.

2.2 *The Role of AI in LAWS*

Autonomous weapons systems rely on "autonomy" to perform their functions in the absence of direction or input from a human actor.[626] AI is not a prerequisite for the functioning of autonomous weapons systems, but, when incorporated, AI could further enable such systems. Indeed, not all autonomous weapons systems incorporate AI to execute particular tasks.

Autonomous capabilities can be achieved through predefined tasks or sequences of actions based on specific parameters, or through using AI tools to derive behavior from data, thus allowing the system to make independent decisions or adjust behavior based on changing circumstances. AI can also be used in an assistance role in systems that are directly operated by a human.

AI technologies are also used in space missions, within the context of "traditional" space missions as navigation and communication or in new space programs as mining activities, on-orbit servicing for instance. In fact, AI components can be integrated into autonomous spacecraft for navigation and obstacle avoidance, enabling real-time decisions to avoid collisions with space debris, akin to a satellite employing AI to dynamically navigate around space junk.

Moreover, AI can also assist in space-based surveillance systems, such as satellites equipped with computer vision algorithms to monitor space traffic and detect potential threats. These same technologies could be reconfigured to enable satellite hunter-killer drones.[627]

In addition, AI technologies with facial recognition capabilities could be implemented in satellites for surveillance and reconnaissance purposes. Such satellites could autonomously identify and track objects or individuals in space or on Earth, raising ethical concerns about privacy and the potential for misuse.

Lastly, AI-driven autonomous defense systems could also be developed for space assets to detect and neutralize threats, such as space debris or hostile satellites. These systems would need to make split-second decisions without human intervention, posing significant ethical and safety challenges.

In the same vein, AI could be used in the case of autonomous satellite maintenance and repair missions. AI systems could be programmed to conduct routine checks and repairs on satellites, but the risk of malfunctions or unintended (hostile) actions must be carefully managed. The implementation of robust ethical guidelines and fail-safes would be crucial in such scenarios. These examples illustrate the importance of transparency and accountability in the use of AI technologies in space.

---

[626] *Ibid*.

[627] Matthew Griffin, "US General; 'Swarms of autonomous hunter killers drones will lead marines into battle'", 311 Institute, April 21, 2017, https://www.311institute.com/us-general-announces-swarms-of-hunter-killer-drones-will-lead-marines-into-battle/.

2.3 *Ethical Implications of AI in Warfare and Space Technologies: Recommendations*

2.3.1 Human Responsibility

In practice, the rapid development and scope for the application of AI and machine learning has an important impact on the level of autonomy in weapons systems. Intelligence is related to autonomy in that intelligent systems are capable of determining and performing tasks in more complex environments. The current degree of technological and machine learning development enables the use of AI systems in conjunction with human decision-making.

However, the use of certain technologies during warfare, such as fully autonomous weapons, LAWs or "killer robots," underscore major moral and legal concerns, primarily because they possess the capacity to select and engage their targets without human control.[628]

Significant doubts remain as to whether these weapons ever would be able to imitate the sufficient nuances of human judgment necessary to comply in every circumstance with the legal obligation to distinguish civilian from military targets. These issues are highlighted by the scope for accountability that would arise as a result of unlawful harm caused by fully autonomous weapons.

Moreover, the potential for the development, deployment and use of LAWs highlight serious human rights concerns threatening, for example, the individuals' right to life and to security. Even though these systems are not yet fully operational, but are already deployed, governments and their respective ministries of defense are no doubt undertaking relevant research, given the perception in some quarters that they would potentially enhance military operations in situations where traditional systems may not be able to operate.

However, a broad spectrum of the international community, composed of States, non-governmental organizations and civil society, strongly argue for the regulation, or even total prohibition, of LAWs, due to the complex ethical considerations that they bring to the fore.[629]

From an ethical perspective, it is important to consider the potential future impact of these systems. The use of LAWs risks eroding the moral responsibility of operators. It is crucial that humans stay in the loop, to ensure compliance with the rules of IHL and other relevant and applicable norms of behavior. Nevertheless, any technological achievement may optimize the human-machine interface that arises when hybrid systems, humans and machines are partnered.[630]

Analogous ethical considerations apply to AI in space missions. For example, AI systems used for satellite collision avoidance or space debris management must adhere to ethical guidelines

---

[628] Anne-Sophie Martin, Steven Freeland, "The Advent of Artificial Intelligence in Space Activities: New Legal Challenges", *Space Policy*, 55 (2021) (doi.org/10.1016/j.spacepol.2020.101408).

[629] David Adam, Lethal AI weapons are here: how can we control them?, Nature, April 23, 2024, https://www.nature.com/articles/d41586-024-01029-0; United Nations, First Committee Approves New Resolution on Lethal Autonomous Weapons, as Speaker Warns 'An Algorithm Must Not Be in Full Control of Decisions Involving Killing', November 1st, 2023, https://press.un.org/en/2023/gadis3731.doc.htm; *See* Res. A/RES/78/241, 22 December 2023.

[630] This implicates Brain Computer Interface whereAI has access to a brain computer interface. Thus the question: who is controlling the huma, the person themselves o the AI? See Umer Asgher et al., "Advances in artificial intelligence (AI) in brain computer interface (BCI) and Industry 4.0 for human machine interaction (HMI)', *Frontiers* 17 (2023).

to ensure safety and prevent unintended consequences. Thus, there is a need to keep human oversight in these systems, similar to how human operators must supervise autonomous weapons.

Additionally, consider the ethical implications of using AI in space-based military surveillance, where AI algorithms might need to distinguish between civilian and military activities to avoid accidental targeting or monitoring of non-combatants. There is a need for meaningful human control over decisions in order to ensure that autonomous systems operate under strict oversight and accountability.

2.3.2 Proportionality and Discrimination

The heart of the debate surrounding LAWs involves the legal and ethical considerations associated with the use of systems that may be outside human control. Legal discussions focus primarily on the capacity of autonomous weapons to comply with international law, in particular international humanitarian law (IHL), which involves fundamental principles of distinction which prohibits indiscriminate attacks, that is types of attacks that are of a nature to strike military objectives and civilian objects without distinction; necessity that is the balance between humanitarian considerations and military advantages; and proportionality which prohibits parties to armed conflicts from launching an attack against a military objective which may be expected to cause incidental civilian harm that would be excessive in relation to the concrete and direct military advantage anticipated.[631]

AI systems in space missions must also comply with international laws and ethical standards. For instance, AI algorithms used in space-based surveillance must ensure that data collection and analysis do not violate privacy rights or lead to biased outcomes. Furthermore, one can think of an AI system deployed on a satellite to monitor environmental changes.

It is necessary to ensure that it does not disproportionately target or monitor specific regions unfairly. The main challenge relates to the design and implementation of appropriate human-to-machine and machine-to-machine interactions.

2.3.3 International Norms and Legal Frameworks for LAWS and AI in Warfare

Since 2014, LAWs have been discussed within the context of the CCW.[632] In 2017, these discussions switched from an informal "meeting of experts" to a formal "Group of Governmental Experts" (GGE), with the aim of exploring the technological, military, ethical and legal dimensions of LAWs.[633]

The conclusions of the 2019 GGE session recall some paramount principles applicable to the potential use of lethal weapons systems based on emerging technologies[634]: (i) These systems

---

[631] Human Rights and LAWs, Australian Human Rights Commission, Submission to the Human Rights Council Advisory Committee, November 2023, https://www.ohchr.org/sites/default/files/documents/hrbodies/hrcouncil/advisorycommittee/techmilitarydomain/submissions/1-nhri-australian-hrc.pdf.

[632] *See also Part II Section 2 on CCW and IHL.*

[633] GGE on lethal autonomous weapons systems, https://dig.watch/processes/gge-laws.

[634] *See* CCW/GGE.1/2019/3.

must be conducted in accordance with applicable international law, in particular IHL and its requirements and principles, including inter alia those of distinction, proportionality and precautions in attack; (ii) IHL imposes obligations on States, parties to armed conflict, and individuals, but not machines; (iii) IHL requirements and principles must be applied through a chain of responsible command and control by the human operators and commanders; (iv) Human judgment is essential to ensure that the possible use of lethal weapons systems remains in compliance with international law.

Although there is agreement on the importance of maintaining human control in the use of autonomous weapons systems, further clarification will be needed on the type and degree of human-machine interaction required, including elements of control and judgment in different stages of a weapon's life cycle, to develop shared understandings of these concepts and their applications.

In this context, the GGE has developed some Guiding Principles[635] indicating that:

> (i) IHL continues to apply fully to all weapons systems;
>
> (ii) Human responsibility for decisions on the use of weapons systems must be retained, since accountability cannot be transferred onto machines;
>
> (iii) Accountability for developing, deploying and using any emerging weapons system must be ensured in accordance with applicable international law;
>
> (iv) When developing or acquiring new weapons systems based on emerging technologies in the area of LAWs systems, the risk of acquisition by terrorist groups and the risk of proliferation should all be considered;
>
> (v) Risk assessments and mitigation measures, through appropriate rules of engagement, should be part of the design, development, testing and deployment cycle of emerging technologies in any weapons systems;
>
> (vi) Discussions and any potential policy measures taken within the context of the CCW should not hamper progress in or access to peaceful uses of intelligent autonomous technologies.

These elements are also very relevant for developing policy and legal frameworks for situations where AI components and autonomy are to be incorporated into future space missions.

Given the dual-use nature of the underlying technologies in the area of LAWs, it is important to promote responsible innovation and use of such technologies, even more so with the increasingly prevalent dual-use characteristics of space activities, either within the context of "traditional" space missions such as observation, navigation and communications, or in other novel programs, including mining activity, on-orbit servicing, deep space exploration and space traffic management.

---

[635] Guiding Principles affirmed by the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System, https://www.ccdcoe.org/uploads/2020/02/UN-191213_CCW-MSP-Final-report-Annex-III_Guiding-Principles-affirmed-by-GGE.pdf.

A similar framework could be applied to space technologies. In fact, UNCOPUOS could establish a working group for developing guidelines related to the ethical use of AI in space exploration, similar to the principles established for LAWs.

The guidelines could highlight specific provisions, such as the need for human oversight and accountability in AI applications, which are essential in maintaining control over AI-driven space missions. This could include establishing protocols for AI-driven satellites to ensure they do not perform unauthorized maneuvers or collect data on unintended targets.

3. *Robotics Case Study & Implications for Ethical Use of AI in Space: Autonomous Robots in Border Control*

Technological evolution for border management demonstrated by advancements in biometrics, surveillance and artificial intelligence have reshaped border control.

The term "digital border technologies" includes technologies which are often characterized by the use of "artificial intelligence,"[636] such as drones, robots, and vehicles, equipped with infrared cameras and night vision, and which in some instances may simply operate digitally but in others may already, or in the near future, be 'AI-enabled,' for example, with the capacity to make distinctions between humans and animals.[637]

There is also a growing use of tools reliant on generative AI, including for border enforcement and migration management. For example, the US Department of Homeland Security is investing in robot dogs to patrol remote regions of the US-Mexico border,[638] the EU's Frontex is deploying pseudo-scientific automated AI lie-detection video kiosks for travelers.[639]

The link between the regulation of space activities and border control is emerging in the context of technological surveillance, space-based data and AI. In fact, governments are increasingly turning to satellite technology for border surveillance. These satellites, often equipped with AI-driven imaging and data analysis capabilities, can detect illegal border crossings, trafficking, or environmental violations. In addition, AI-powered satellites can be utilized to

---

[636] United Nations, UN Human Rights Office of theHigh Commissioner, "Digital Border Governance: a Human Rights Based Approach," September 18, 2023, https://www.ohchr.org/sites/default/files/2023-09/Digital-Border-Governance-A-Human-Rights-Based-Approach.pdf.

[637] See Human Rights Council Advisory Committee, "Possible impacts, opportunities and challenges of new and emerging digital technologies with regard to the promotion and protection of human rights: Report of the Human Rights Council Advisory Committee" A/HRC/47/52 (19 May 2021) at §3. For further definitions, see UNESCO, Recommendation on the Ethics of Artificial Intelligence (23 November 2022) at 10; UN Chief Executives Board for Coordination, High-Level Committee on Programmes (HLCP) Inter-Agency Working Group on Artificial Intelligence, Principles for the Ethical Use of Artificial Intelligence in the United Nations System (20 September 2022) at §2.

[638] U.S. Department of Homeland Security, "Feature Article: Robot Dogs Take Another Step Towards Deployment at the Border," February 1st, 2022, https://www.dhs.gov/science-and-technology/news/2022/02/01/feature-article-robot-dogs-take-another-step-towards-deployment.

[639] European Commission, "Smart lie-detection system to tighten EU's busy border," October 24, 2018, https://projects.research-and-innovation.ec.europa.eu/en/projects/success-stories/all/smart-lie-detection-system-tighten-eus-busy-borders.

monitor space activities, ensuring compliance with international regulations similar to border control on Earth. These satellites could prevent unauthorized satellite launches by detecting and reporting any unregistered spacecraft. Additionally, they could identify and track space debris that crosses into different jurisdictions, thus ensuring safe and orderly space traffic management.

3.1 *Human Rights*

The proliferation of surveillance technology amplifies data collection efforts, posing potential threats to individuals' privacy. The extensive gathering of personal information and its storage demand robust data protection measures to mitigate risks. In fact, government proposals for more security might be in conflict with fundamental rights.[640]

Ensuring that surveillance practices do not infringe individual rights to privacy, freedom of movement, and freedom of expression remains a critical challenge. Protecting the rights to life, liberty, privacy and non-discrimination in the face of increasing demands for "security" requires research, cooperation, campaigning and advocacy within and between countries.

In what way does the use of AI and automated technologies for border control impact human rights?

When AI technologies are used in border control, a vast array of human rights violations can occur: facial recognition and algorithmic decision-making, drone surveillance can discriminate; impact people's right to life, liberty, and security of the person. These are just some of the many human rights risks of border control technologies might represent as technology continues to be largely unregulated and non-transparent.

Addressing ethical implications, such as the protection of privacy, is crucial. In space, AI systems could inadvertently collect sensitive data about other nations' space activities. Establishing strict data governance protocols will ensure that privacy is maintained. Furthermore, just as border control technologies on Earth must avoid discriminatory practices, AI in space should be designed to operate without bias. For instance, algorithms used to prioritize collision avoidance maneuvers should treat all satellites equally, regardless of their country of origin.

In the field of space operations, an AI-powered satellite monitoring system could be equipped with advanced machine learning algorithms to analyze patterns in space traffic. For instance, it could identify anomalies, such as unauthorized launches, and automatically notify relevant international bodies to take action.

Similarly, AI-driven space debris detection systems could assess the risk levels of various debris and recommend collision avoidance strategies to satellite operators. Implementing such systems would require collaboration between space-faring nations to develop and adhere to international standards, ensuring that AI technologies are used responsibly and ethically in space missions.

---

[640] *See* 1948 Universal Declaration of Human Rights.

3.2 *Transparency and Accountability*

There is a lack of transparency in the mapping of such technologies and to the assessment and operationalization of human rights protection in the use of digital border technologies.[641] States often link border enforcement and security concerns[642] as well as use justifications of national sovereign control over border enforcement, creating "grey zones" of accountability and oversight.[643]

The use of these technologies for border control is somewhat opaque and, most of the time, their use are revealed after the fact as a result of the work of civil society organizations and human rights observers. When violations of international law occur, the question of accountability is raised.[644]

Direct responsibility holds offenders liable for playing an active role in the commission of a crime. It also creates accountability for the direct perpetrator; it also covers other actors who are directly involved because, for example, they planned or ordered a crime. Robots could not themselves be held responsible for their actions under this doctrine for three reasons.[645]

First although they might commit a criminal act, they could not have the mental state required to perpetrate a crime. Second, international criminal tribunals generally limit their jurisdiction to "natural persons," that is, human beings, because they have the intentionality to commit crimes.

Third, even if this jurisdiction were expanded, on a practical level, autonomous robots could not be punished because they would be machines that could not experience suffering or apprehend or learn from punishment. Autonomous robots would therefore present a new gap in liability: the entity that selects and engages targets, which until now has always been a human being, could not be held directly liable for criminal action resulting from the illegal selection or engagement of targets.[646]

Furthermore, there would be insufficient direct responsibility for a human who deployed or operated a fully autonomous robot that committed a criminal act. A gap could arise because fully autonomous robots would, by definition, have the ability to act autonomously and could therefore conduct independent and unpredictable maneuvers. The operator could only be responsible for deploying the robot, and liability would rest on whether that decision under the circumstances amounted to an intention to commit an attack.

---

[641] United Nations, UN Human Rights Office of theHigh Commissioner, "Digital Border Governance: a Human Rights Based Approach," September 18, 2023, https://www.ohchr.org/sites/default/files/2023-09/Digital-Border-Governance-A-Human-Rights-Based-Approach.pdf.

[642] Gavin Sullivan, The Law of the List: UN Counterterrorism Sanctions and the Politics of Global Security (Cambridge: Cambridge University Press, 2020).

[643] Petra Molnar, "Surveillance sovereignty: migration management technologies and the politics of privatization" in Idil Atak, Graham Hudson (eds), *Migration, Security, and Resistance: Global and Local Perspectives* (UK: Routledge, 2021) 66-82, at 70.

[644] Human Rights Watch, April 9, 2015, https://www.hrw.org/report/2015/04/09/mind-gap/lack-accountability-killer-robots.

[645] *Ibid.*

[646] *Ibid.*

Even if direct liability were legally possible, it would be difficult to prove liability. Robots would have at least two sides supplying the orders: the operator and the programmer (and there would often be several people involved in the programming). Each party could try to attribute fault to the other in order to avoid liability.

With this in mind, one can assume that there is insufficient policy and regulatory framework integrating technology into border governance strategies. There is a similar lack of information regarding how States assess the lawfulness, necessity and proportionality of such deployment, particularly against the potential impacts on human rights, and the availability of less intrusive alternatives, both technological and non-technological.[647]

In the context of space activities, AI algorithms used in satellite navigation or space debris management should be transparent and subject to regular audits to ensure they operate ethically and effectively. Highlighting the need for clear accountability mechanisms, such as detailed logs of AI decision-making processes and regular compliance checks, will underscore the importance of maintaining trust and integrity in the use of AI technologies in space.

3.3 *Non-Discrimination and Bias*

AI creates new challenges for establishing *prima facie* discrimination. Compared to traditional forms of discrimination, automated discrimination is more abstract and unintuitive, subtle, and intangible. These characteristics make them inherently difficult to detect and prove, as victims may never realize that they have been put at a disadvantage.[648]

In the last few years the High-Level Expert Group on AI (AI HLEG) from the European Commission has issued a set of ethics guidelines for achieving trustworthy AI.[649] It is possible to extrapolate such guidelines to intelligent robots:[650] (1) Human Agency and Oversight-robots and robot systems should respect human agency and support oversight of their execution; (2) Technical Robustness and Safety-robots should be robust and safe as they interact with humans and in our society; (3) Privacy and Data Governance-robots should follow the established privacy rules and data governance mechanisms; (4) Transparency: robots should be transparent when making decisions, and about their capabilities, making clear why certain decision is the appropriate; (5) Diversity, Non-discrimination and Fairness-robots should respect not discriminate nor cause discrimination, and guarantee fairness in their decisions; (6) Environmental and Societal well-being-robots should foster societal well-being and contribute to a better society and environment;

---

[647] United Nations, UN Human Rights Office of the High Commissioner, "Digital Border Governance: a Human Rights Based Approach," September 18, 2023, https://www.ohchr.org/sites/default/files/2023-09/Digital-Border-Governance-A-Human-Rights-Based-Approach.pdf.

[648] Sandra Watcher, "Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI", Computer Law&Security Review 41 (2021) 10.

[649] European Commission, Ethics guidelines for trustworthy AI, April 8, 2019, https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

[650] Pedro U. Lima, Ana Paiva, 'Autonomous and Intelligent Robots: Social, Legal and Ethical Issues', in Henrique Sousa Antunes et al., *Multidisciplinary Perspectives on Artificial Intelligence and the Law*, (Springer, 2023) 135-136.

(7) Accountability: a clear accountability process and eco-system should be in place and followed by robot manufacturers, guaranteeing that when problems occur the process can be triggered.[651]

In cases where decisions are based on predetermined criteria or AI, it can be difficult to understand not only which risk models or data are used, but also how this use shapes the outcome of the decision-making process. Challenging bias or discriminatory criteria in these decisions can be even more difficult.

Despite the rapid expansion into border zones and fast uptake by border control agencies, regulations and guidelines for the deployment of AI have been slower to evolve.[652] In April 2024, the European Union released the first legal framework for AI in an attempt to regulate the technology before it becomes even more mainstream.[653]

The proposal for harmonized rules specifically mentions AI systems in migration, asylum, and border control, claiming these processes can affect particularly vulnerable people. It notes that ensuring the accuracy, nondiscriminatory nature, and transparency of AI systems is especially important to ensuring that the rights of vulnerable populations are protected.

The draft regulation therefore classifies the use of AI systems in migration management as "high risk," especially regarding technologies such as risk assessments, document verification, and applications for immigrant status. This approach could mark a turn from previous EU projects such as Roborder and iBorderCtrl.

However, experts have pointed out oversights, including a lack of rules that would impact major technology companies and insufficient focus on people affected by AI systems. AI systems used in space exploration and satellite operations must be designed to prevent bias and ensure fairness. In fact, AI algorithms for satellite image analysis should be trained on diverse datasets to avoid discriminatory outcomes, such as unfairly prioritizing certain geographic areas over others.

Highlighting the importance of ethical guidelines, such as those developed by the European Commission's High-Level Expert Group on AI, will provide a framework for developing fair and unbiased AI systems in space. These guidelines could include specific measures to ensure that AI systems do not disproportionately affect certain regions or populations.

---

[651] Evelien Brouwer, "Challenging Bias and Discrimination in Automated Border Decisions: Ligue des droits humains and the Right to Effective Judicial Protection", VerfBlog, November 5, 2023, https://verfassungsblog.de/pnr-border/ (10.17176/20230511-181734-0).

[652] Hannah Tyler, *The Increasing Use of Artificial Intelligence in Border Zones Prompts Privacy Questions* (Migration Policy Institute, February 2, 2022), https://www.migrationpolicy.org/article/artificial-intelligence-border-zones-privacy.

[653] European Union. Regulation of the European Parliament and of the Council 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). 2021/0106(COD) PE-CONS 24/24. Brussels, 14 May 2024. https://data.consilium.europa.eu/doc/document/PE-24-2024-INIT/en/pdf.

Conclusions

The use of AI in genetics is revolutionizing the field by enhancing the ability to analyze large datasets, accelerate research, and uncover complex genetic patterns that were previously difficult to identify. AI's advanced algorithms enable precise genome sequencing, gene editing, and predictive modeling for genetic disorders, contributing to personalized medicine and new therapeutic strategies.

While AI improves efficiency and discovery, ethical considerations—such as data privacy, bias, and the impact of automation—require careful regulation. Overall, AI's integration into genetics holds transformative potential, but it must be navigated responsibly to ensure its benefits are maximized. The same can also be said of developing AI-based space applications.

In the same vein, the integration of AI in robotics and space programs has brought transformative advancements, enhancing efficiency, precision, and autonomy in exploration, manufacturing, and data analysis. AI-driven robots have been crucial in space missions, enabling remote operations in hostile environments where human intervention is impractical or impossible.

AI enhances navigation, decision-making, and adaptability in uncertain and dynamic conditions, such as those encountered in space exploration. It also optimizes mission planning, spacecraft trajectory, and the management of resources aboard spacecraft, making deep-space exploration more feasible.

However, the growing use of AI in robotics for space also raises important legal, ethical and technical considerations, in particular in the field of responsibility and fundamental rights. Challenges include ensuring the safety and reliability of autonomous systems, protecting against potential malfunctions, and safeguarding against cyber vulnerabilities.

With this in mind, the development of guidelines will support the deployment of AI in space missions. However, the benefits, including enhanced operational capabilities, reduced costs, and expanded reach, underscore AI's pivotal role in advancing space exploration and technology.

AI has become indispensable in the future of genetic, robotics and space programs, enabling breakthroughs that drive human exploration beyond Earth's bounds. As AI technology advances, it will continue to push the frontiers of space science.

Section 2: Recommendations for International Guidelines on Ethical AI Governance Across Emerging Sectors

Introduction

This section examines several key international frameworks that serve as vital guides for the responsible development and deployment of AI, offering foundational principles to address the ethical, legal, and societal challenges associated with AI. Notably, the Council of the Organization for Economic and Development (OECD) Principles on AI, the United Nations Educational, Scientific and Cultural Organization (UNESCO) recommendations, and the EU Artificial Intelligence Act emerge as crucial instruments shaping the future of AI governance, particularly in emerging fields such as outer space activities.

The OECD Principles on AI, first adopted in 2019 and revised in 2023 and 2024, represent the first intergovernmental standards on AI. These recommendations emphasize fostering innovation and trust in AI while promoting transparency, accountability, human rights, and democratic values. While these principles are not legally binding, they serve as a crucial reference for countries in developing their national regulations.

The OECD's focus on AI systems, their lifecycles, and key actors provides a comprehensive framework that supports the ethical use of AI across industries, including space technologies. The guidelines emphasize risk management, traceability, and cross-stakeholder cooperation—elements that are critical for the safe and transparent deployment of AI systems in outer space.

Complementing the OECD's efforts, UNESCO's "Recommendation on the Ethics of Artificial Intelligence," supported by 193 signatories, provides an ethical foundation for AI governance with principles focused on human dignity, transparency, and accountability. Although the recommendations do not explicitly address outer space, their broad acceptance underscores their importance in guiding AI regulation.

These ethical principles can easily be extended to satellite operations and space exploration, where the ethical management of AI is crucial for maintaining transparency and preventing harm to individuals, nations, and the space environment. UNESCO's focus on equity, non-discrimination, and sustainability also aligns with the core values needed to ensure that AI benefits all of humanity, particularly in space applications.

Building on these global standards, the EU AI Act, adopted in 2024, provides a robust legal framework for AI systems within the European Union, with significant implications for space exploration. The Act introduces a risk-based approach to AI governance, categorizing AI applications into minimal, high-risk, and prohibited systems.

This risk assessment framework is essential for space activities, as AI is increasingly used for critical tasks like satellite navigation, Earth observation, and even human-robot interactions during space missions. The extraterritorial application of the EU AI Act ensures that space-faring nations and companies operating within the EU must comply with its ethical and safety standards, promoting human-centric AI that prioritizes safety, rights, and accountability.

Together, these frameworks—the OECD Principles, UNESCO's recommendations, and the EU AI Act—provide a comprehensive foundation for the responsible regulation of AI. They offer the necessary guidance for countries and space-faring organizations to develop national policies that ensure ethical AI applications in space exploration.

By promoting transparency, fostering international collaboration, and ensuring robust data governance and risk management, these frameworks support the safe, ethical, and innovative use of AI in space. As AI continues to play an expanding role in space activities, integrating these principles into international agreements and fostering continuous dialogue between stakeholders will be critical for shaping a sustainable and responsible future for AI-driven space technologies.

1. ***OECD AI Principles: Guiding Responsible Innovation and Governance***

In May 2019, OECD adopted their Recommendation on Artificial Intelligence (AI), revised in 2023 and 2024 to keep up to date with technological development. This set of recommendations constituted the first intergovernmental standards on AI. The recommendation aims to foster innovation and trust in AI by promoting the responsible stewardship of trustworthy AI while ensuring respect for human rights and democratic values.

However, they do not constitute binding laws but rather serve as a voluntary framework that countries may adhere to or draw inspiration from for their national regulations. Indeed, such recommendations are notably important for providing guidance on how AI could be regulated.

The document is of interest for the regulation of AI in outer space activities for mainly two reasons. Firstly, it offers a set of definitions for crucial terms that can sustain a common understanding between different stakeholders. In particular for "AI system," "AI lifecycle," and "AI actors." These definitions include:

- AI system: "An AI system is a machine-based system 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. Different AI systems vary in their levels of autonomy and adaptiveness after deployment."
- AI system lifecycle: "An AI system lifecycle typically involves several phases that include to: plan and design; collect and process data; build model(s) and/or adapt existing model(s) to specific tasks; test, evaluate, verify and validate; make available for use/deploy; operate and monitor; and retire/decommission. These phases often take place in an iterative manner and are not necessarily sequential. The decision to retire an AI system from operation may occur at any point during the operation and monitoring phase."
- AI actors: "AI actors are those who play an active role in the AI system lifecycle, including organizations and individuals that deploy or operate AI."

Secondly, the Recommendation is structured in five high-level values-based principles and five recommendations for national policies and international cooperation.

In particular, the recommendations have a bifocal substantial focus with two sections: principles for responsible stewardship of trustworthy AI, and national policies and international co-operation for trustworthy AI.

1.1 *Principles for Trustworthy AI Stewardship*

Achieving international alignment on AI regulations necessitates a degree of abstraction. To address this, the AI Principles are designed as high-level guidelines and advocate for: 1.1 Inclusive growth, sustainable development and well-being; 1.2 Respect for the rule of law, human rights and democratic values, including fairness and privacy; 1.3. Transparency and explainability; 1.4. Robustness, security and safety; 1.5. Accountability.[654]

The OECD emphasizes that its principles are complementary and should be viewed collectively to promote the responsible stewardship of trustworthy AI. Notable passages relevant to AI in space include the following:

- 1.3. "AI Actors should commit to transparency and responsible disclosure regarding AI systems."
- 1.4. "AI systems should be robust, secure and safe throughout their entire lifecycle so that, in conditions of normal use, foreseeable use or misuse, or other adverse conditions, they function appropriately and do not pose unreasonable safety and/or security risks. Mechanisms should be in place, as appropriate, to ensure that if AI systems risk causing undue harm or exhibit undesired behavior, they can be overridden, repaired, and/or decommissioned safely as needed"
- 1.5. " a) AI actors should be accountable for the proper functioning of AI systems and for the respect of the above principles, based on their roles, the context, and consistent with the state of the art. b) To this end, AI actors should ensure traceability, including in relation to datasets, processes and decisions made during the AI system lifecycle, to enable analysis of the AI system's outputs and responses to inquiry, appropriate to the context and consistent with the state of the art. c) AI actors, should, based on their roles, the context, and their ability to act, apply a systematic risk management approach to each phase of the AI system lifecycle on an ongoing basis and adopt responsible business conduct to address risks related to AI systems, including, as appropriate, via cooperation between different AI actors, suppliers of AI knowledge and AI resources, AI system users, and other stakeholders. Risks include those related to harmful bias, human rights including safety, security, and privacy, as well as labor and intellectual property rights."

---

[654] Tommaso Giardini and Johannes Fritz, 'The Anatomy of AI Rules A systematic comparative analysis of AI rules across the globe' Digital Policy Alert, 2024 <https://digitalpolicyalert.org/ai-rules/the-anatomy-of-AI-rules>.

1.2 *Applying OECD AI Principles to Outer Space Activities*

Contextualizing these principles for the deployment of AI systems in outer space reveals key baseline points essential for ensuring the safety and reliability of AI systems within the space domain.

The elements outlined above underscore the importance of ensuring transparent, reliable, and trustworthy AI in outer space activities. This holds a dual significance: adherence to these principles supports the proper functioning of AI systems in outer space, while their transparency enables traceability throughout the system's lifecycle, particularly in the event of malfunctions or other issues.

Indeed, throughout the AI lifecycle is a critical factor in assigning fault in cases of damage.[655] Section 1.4 highlights the importance of safeguards: "Mechanisms should be in place, as appropriate, to ensure that if AI systems risk causing undue harm or exhibit undesired behavior, they can be overridden, repaired, and/or decommissioned safely as needed."

AI is set to play a pivotal role in enhancing and automating the maintenance and repair of space assets. However, in the event of malfunctions caused by edge AI on spacecraft, a clear intervention plan must be in place to address the issue or safely decommission the satellite.

To prepare for such scenarios and embrace an AI-driven future, space companies should adhere to Recommendation 1.5: "apply a systematic risk management approach to each phase of the AI system lifecycle on an ongoing basis and adopt responsible business conduct to address risks related to AI systems, including, as appropriate, via cooperation between different AI actors, suppliers of AI knowledge and AI resources, AI system users, and other stakeholders."

Effective communication among all stakeholders is essential, as is the development of appropriate standards to ensure not only interoperability between systems but also a unified approach to risk management for autonomous objects in space.

1.3 *National and International Cooperation for Trustworthy AI Governance*

In this section, the OECD provides guidance on operationalizing its recommendations. It emphasizes that adhering states must implement these recommendations in alignment with the principles outlined in Section 1. These efforts should be reflected in national policies and international cooperation, with particular attention to supporting small and medium-sized enterprises (SMEs).

The recommendations focus on five key areas: 2.1. Investing in AI research and development; 2.2. Fostering an inclusive AI-enabling ecosystem; 2.3. Establishing an interoperable governance and policy framework for AI; 2.4. Building human capacity and preparing for labor market transformation; 2.5. International co-operation for trustworthy AI.

For the purposes of this report, space actors and policymakers should take into consideration the recommendations as a whole, but should place their main focus on points 2.2, 2.3, and 2.5. The following points offer key considerations for establishing a suitable regulatory

---

[655] See Part I Section 1 on the Liability Convention

framework for AI in space activities. In particular, they emphasize the importance of the interoperability of different AI systems and advocate for an expansion of the use of standards.

Point 2.5 raises an especially important element that the space community must achieve: "Governments, including developing countries and stakeholders, should actively cooperate to advance these principles and to progress on responsible stewardship of trustworthy AI." This final point will hold an increasingly central role in the success of the AI and space nexus, given the global commons nature of the space environment.

2. *UNESCO's Ethical Guidelines for AI: A Global Perspective*

The "Recommendation on the Ethics of Artificial Intelligence" is a document prepared by the Ad Hoc Expert Group (AHEG) and adopted by UNESCO in November 2021. The instrument is expected to help nations and companies improve their ethical regulatory frameworks based on a universalist vision. UNESCO has 193 member countries.

These principles include transparency and explainability, non-discrimination and equity, respect for human autonomy, prevention of harm, responsibility, privacy and data governance, social benefit, sustainability, accountability, and inclusion. The recommendations are a global regulatory framework that guides countries in creating their legal frameworks to ensure that AI is deployed ethically.

A particular focus of the recommendations, in line with UNESCO's mandate, addresses the ethical implications of AI systems in culture, education, science, information, and communication. Understanding the rationale behind these recommendations is essential, as they can serve as inspiration for different organizations and states to create their own laws, potentially encompassing or influencing the application of AI in outer space.

Indeed, to harness AI's vast potential, approaches to policymaking must support innovation while appropriately managing risk. An analysis of the different principles is provided, connected to their possible impact on space activities.

2.1 *Ensuring Proportionality and Preventing Harm in AI Systems*

UNESCO emphasizes the principle of proportionality, which entails ensuring that AI development and utilization align with their intended purposes without causing undue harm or excess. As the advancement of AI in outer space is largely driven by private entities, it is imperative to foster a culture of proactive safety. National laws can play a crucial role by incorporating provisions that mandate AI features in satellite operations adhere to these principles.

For instance, AI systems used for satellite collision avoidance should prioritize minimizing harm by accurately predicting potential collisions and adjusting satellite trajectories accordingly. Moreover, implementing safety protocols to assess the risks of AI in space missions, such as analyzing the potential impact of AI malfunctions on space debris generation, is essential. These measures would ensure adherence to the principle of proportionality. A thorough assessment of risks and potential side effects associated with AI implementation is also necessary to further support these efforts, ensuring that AI technologies are both safe and effective in their applications

2.2 *Enhancing AI Safety and Security in Space Operations*

While AI can play a vital role in the future of space missions, it is imperative that AI systems themselves ensure durability and resilience. Maintaining system security throughout its lifecycle is essential to prevent, mitigating risks and vulnerabilities that could lead to potential attacks.

In particular, the integration of AI technologies can enhance the security of space assets. Indeed, AI-driven cybersecurity measures can be implemented to protect satellites from hacking attempts by ensuring continuous supervision of network traffic. Any suspicious activities happening in the network would be appropriately detected.

In addition, blockchain technology can be integrated securing data transmission between satellites and ground stations. Therefore, it will be possible to enhance transparency and traceability, ensuring the long-term resilience of AI systems in space.

Indeed, transparency regarding the development process and access to underlying data are an important element to ensure traceability and enhance system security. The integration of such advanced security measures, it can be ensured that AI technologies in space are both safe and reliable, contributing to a more sustainable and secure outer space environment.

2.3 *Promoting Equity and Inclusivity in AI-Driven Space Technologies*

This principle advocates for an inclusive approach to ensure the equitable access to AI technology benefits. In the application of AI in space activities, it is going to be essential to ensure that the benefits of such developments align with the principles outlined in the Outer Space Treaty, benefiting all of humanity.

However, there's a risk associated with AI development competition, where the first operational AI systems could set standards and potentially create silos between nations. To mitigate such risk, an open and interoperable system is imperative given the international nature of the space environment.

For instance, AI algorithms used in Earth observation satellites can ensure that environmental data, such as climate patterns and natural disaster alerts, is accessible to all nations, especially those lacking advanced space capabilities. The development of international partnerships, such as joint space missions and data-sharing agreements, will become instrumental to bridge the technological gap between developed and developing countries.

In particular, UNESCO calls upon technologically advanced nations to uphold a global obligation of solidarity with less developed nations, ensuring equitable distribution of AI benefits and facilitating participation and access to information systems.

These partnerships could take inspiration from initiatives like the UN's Sustainable Development Goals (SDGs), where AI applications in space can be used to address different global challenges. Collaboration and sharing AI-driven space research, will ensure that the advancements in AI benefit all of humanity, promoting a more inclusive and equitable global community.

2.4 *AI for Sustainable Space Operations*

Sustainability is a fundamental consideration in the design and development of AI systems, particularly in the context of space operations. AI has the potential to significantly enhance the sustainability of space activities by optimizing satellite fuel consumption, enabling more efficient orbital maneuvers, and extending the operational lifespans of satellites. Moreover, AI-driven technologies can play a critical role in mitigating space debris by actively removing or repositioning existing debris to prevent potential collisions, thereby contributing to safer and more sustainable space environments.

Additionally, AI can analyze data onboard satellites to prioritize critical information transmission, thereby reducing the energy consumption associated with data downlinking. Indeed, processing data at the edge and only downlinking essential information, AI can help improve the overall environmental footprint of the space industry.

These measures not only enhance the efficiency of space missions but also contribute to the long-term sustainability of downstream and upstream space activities. It will become important that space operations are both technologically advanced and environmentally responsible. Such synergy will pave the way for a more sustainable future in space exploration.

2.5 *Protecting Privacy and Data in Space-Based AI Systems*

UNESCO deeply advocates for the protection of rights that are integral to human dignity, such as the collection, use, and dissemination of personal data. The growing number of satellites and the expected increase in AI system integration may raise concerns about the use of personal data collected by space assets. Integrating privacy principles, such as those outlined in the GDPR—as detailed in Part II, Section 1 of this report—can be beneficial in ensuring the rights of individuals in the face of increasing technological capabilities.

Satellite operators can implement data anonymization techniques to ensure that personal data collected during Earth observation missions is protected. Differential privacy techniques can obfuscate personal data points in satellite imagery, making it impossible to identify individuals while still enabling useful analysis. Another important point will be to establish clear data governance policies that define the permissible use and sharing of satellite data can help safeguard privacy while leveraging AI capabilities.

This includes creating strict access controls and audit trails for satellite data to ensure that only authorized personnel can access sensitive information and setting guidelines on data retention periods to prevent unnecessary storage of personal data. These steps can ensure compliance with data protection laws and protect individual privacy in space-related AI applications.

2.6 *Ensuring Human Oversight in AI-Enabled Space Technologies*

The attribution of responsibility in case of failure or damage of AI systems is a crucial element that requires specific attention. UNESCO emphasizes the importance of developing systems that ensure traceability in the value chain of AI, thereby offering a level of transparency that allows for the identification of responsibility at various stages of data collection, training, and AI application.

Under international space law, it is well acknowledged that the ultimate responsibility lies with the member state granting the license to operate the spacecraft.[656] Therefore, it is incumbent upon the granting state to ensure that technologies used in space provide traceability for potential damage. This can be achieved by dictating specific requirements for the integration of AI in spacecraft.

Human oversight and determination are critical in this context. For example, space agencies can establish protocols requiring human operators to review AI decisions in satellite navigation systems, ensuring accountability in case of anomalies. Additionally, developing traceability mechanisms that log AI decision-making processes can help identify the source of any errors or malfunctions.

These mechanisms facilitate transparent investigations and accountability. By implementing such measures, states can ensure that AI systems used in space missions are both reliable and accountable, aligning with the principles of traceability and responsibility emphasized by UNESCO.

2.7 *AI Transparency and Explainability for Space Applications*

Transparency and explainability of AI systems are important features to sustain the respect of human rights and fundamental freedoms. Transparency ensures that appropriate information is provided when an AI system is deployed, fostering trust in the system itself.[657]

Explainability makes the results of AI systems understandable, including the input, output, operation, and the contribution of each algorithmic building block to the final product.[658] In simple terms, it is the capability of AI to trace back faults in the AI value chain, thanks to the understandability of subprocesses and results.

These elements are vital for the space industry, given the inherent low-risk appetite of any space project. Space projects are generally extremely costly, and the inclusion of autonomous systems with inherent uncertainty has often deterred the use of sophisticated AI models.[659]

These features are deeply intertwined with the capability to ensure human oversight and determination. For instance, developing user interfaces that allow operators to understand and

---

[656] Frans Von Der Dunk, ‘Liability versus Responsibility in Space Law: Misconception or Misconstruction?’, 1991.

[657] Stefan Larsson, and Fredrik Heintz, ‘Transparency in artificial intelligence’, *Internet policy review*, 2020, 9.2.

[658] Andrea Ferrario, and Michele Loi, ‘How explainability contributes to trust in AI’, In: *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. 2022. p. 1457-1466.

[659] Gianluca Furano, Antonis Tavoularis, and Marco Rovatti, ‘AI in space: Applications examples and challenges’, In: *2020 IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFT)*. IEEE, 2020. p. 1-6.

review the decision-making processes of AI-powered satellite systems can enhance transparency. Additionally, implementing explainable AI (XAI) techniques can help demystify complex AI algorithms, making it easier to trace and address faults in the AI value chain.[660]

This approach sustains the correct functioning of national and international liability regimes, which is crucial for the future of space activities, where AI-based systems will increasingly be part of both the downstream and upstream segments of the space industry. By applying these principles, space missions can achieve reliable and accountable operations, aligning with the essential transparency and explainability standards emphasized by UNESCO.

2.8 *Fostering Responsibility and Accountability in AI-Driven Space Operations*

UNESCO stresses that the deployer of an AI system must ensure their ethical duty by assuming responsibility for the decisions and actions of their AI. To guarantee appropriate responsibility, supervision, impact, and evaluation processes should be developed through audits of each process in the AI value chain.[661] Such an approach would strengthen the ethical framework for AI, ensuring a culture of transparency, trust, and responsibility.

In space activities, the duty of supervision lies with the state that grants the license to launch and operate a space object. According to Article VI of the Outer Space Treaty, "[…] activities of non-governmental entities in outer space, including the moon and other celestial bodies, shall require authorization and continuing supervision by the appropriate State Party to the Treaty. […]." Additionally, Article VIII of the Outer Space Treaty elucidates on the retention of jurisdiction and control over a space object through its inclusion in the registry.

International space law imposes an obligation on States to authorize and continuously supervise non-governmental space activities under their jurisdiction to ensure compliance with the Treaty and general international law standards. As suggested by some authors, the state in the registration process can require additional information regarding the involvement of an AI system on-board of a spacecraft.[662]

Indeed, to enhance accountability in AI-based space missions, states can mandate detailed disclosures about the AI technologies used in satellites during the licensing process, including their capabilities, limitations, and risk mitigation strategies. Furthermore, conducting regular audits and compliance checks on AI systems in space can help maintain high standards of accountability and ensure that AI applications adhere to international laws and ethical guidelines.

In addition, some key parts of the Registration Convention, such as Articles II, IV, and VII, would also support this framework. Article II mandates that states register each space object launched into orbit or beyond with the United Nations, including detailed information about onboard AI systems to ensure transparency and accountability.

---

[660] Alejandro Barredo Arrieta, et al. "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Information fusion*, 2020, 58: 82-115.

[661] Gregory Falco, "Governing AI safety through independent audits," *Nature Machine Intelligence*, 2021, 3.7: 566-571.

[662] Anne-Sophie Martin, and Steven Freeland, "The advent of artificial intelligence in space activities: New legal challenges," *Space Policy*, 2021, 55: 101408.

Article IV specifies the information required for registration, including the general function of the space object, suggesting that this should extend to AI technologies to provide comprehensive details about their operational parameters and safety measures. Article VII allows for updates to the registration information, emphasizing the importance of keeping the registry current as AI systems are modified or upgraded.

Therefore, international space law, coupled with these specific measures, ensures that states can effectively supervise and regulate AI-based space systems, establishing a robust framework for responsibility and accountability in the new era of AI in space.

2.9 *Raising Awareness and Education on Ethical AI in Space*

UNESCO emphasizes the importance of cross-sector collaboration, bringing together industries, governments, academia, civil society organizations, and the media. The significance of the space sector and its influence on societal well-being are increasingly recognized by the public. This growing momentum must not be lost. It is essential to leverage the benefits that AI offers, particularly its ability to extract valuable insights from vast amounts of space data.

Space ecosystems depend on enhanced collaboration among diverse stakeholders to ensure that the benefits of space exploration are widely shared and ethically grounded. Encouraging such cooperation helps avoid replicating terrestrial inequalities, where technological advancements often serve only a limited segment of society.

The organization of international workshops and conferences such as the Global Space Exploration Conference (GLEX) or the International Astronautical Congress (IAC) can bring together industry experts, policymakers, and academics to exchange knowledge and best practices regarding AI in space. These events can feature sessions specifically dedicated to AI applications in space, fostering dialogue and collaboration across sectors.

Developing specialized educational programs, such as aerospace engineering degrees with a focus on the ethical use of AI in space exploration, is essential. Collaborating with institutions like the International Space University (ISU) to offer courses or certification programs on responsible AI practices in space missions can equip future space professionals with the knowledge and skills needed to navigate ethical considerations and implement best practices in AI-driven space activities.

Promoting multi-sector collaboration and establishing specialized education on the ethical use of AI in space will enhance awareness and ensure that the benefits of AI advancements are widely shared. This approach will also help ensure that space activities are conducted responsibly and ethically.

2.10 *Adaptive AI Governance and International Collaboration in Space*

In a geopolitically fragmented world, fostering collaboration has become increasingly challenging. Yet, the inherently international nature of the space environment demands a cooperative approach to developing AI for space systems. Global cooperation is vital for fostering synergies and advancing the space industry. While reaching international governmental consensus may prove difficult in the short term, aligning global industry-led standards is essential to ensure interoperability among systems developed across the world.

To achieve this goal, spacefaring nations could establish consortia modeled after the International Space Station (ISS) partnership, but with a specific focus on developing AI technologies for space exploration. Such a consortium should bring together a diverse range of countries and organizations, including the United States, Russia, China, the European Union, the African Union, and other emerging space nations.

Members of the consortium would collaborate on developing AI-driven solutions for challenges such as satellite maintenance, deep space navigation, and space debris management. The partnership's outcomes, including shared research and development efforts and the establishment of standardized testing protocols, would promote interoperability and ensure these technologies adhere to global standards.

Additionally, establishing an international regulatory body for AI in space, similar to the ITU, could harmonize AI policies across nations. This organization would draft guidelines that member states agree to follow, ensuring consistent ethical standards for space missions.

For instance, the body could develop a unified framework for AI applications in autonomous space vehicles, outlining protocols for tasks such as docking with the ISS or navigating through space traffic. Additionally, it could oversee certification schemes to ensure AI systems meet international safety and ethical standards before deployment in space missions.

These efforts would strengthen international cooperation through consortia and regulatory bodies, enabling the space industry to integrate AI advancements effectively and responsibly. A collaborative, standardized approach like this would foster sustainable and ethical space exploration, benefiting all stakeholders.

3. *EU AI Act: Implications for Global AI Governance and Space Activities*

The Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonized 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) 2020/1828l, better known as the Artificial Intelligence Act, was adopted in May 2024.[663]

It is the second high-impact legal text after the GDPR in the information technology industry, having extraterritorial effect according to its scope of application (art. 2). The AI Act represents a legal framework that governs AI systems in the Union, aimed at promoting the uptake of human-centric and trustworthy AI while ensuring a high level of protection for health, safety, and fundamental rights.

---

[663] European Union. Regulation of the European Parliament and of the Council 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). 2021/0106(COD) PE-CONS 24/24. Brussels, 14 May 2024. Accessed June 29, 2024. https://data.consilium.europa.eu/doc/document/PE-24-2024-INIT/en/pdf.

In the context of the space industry, the implications of the EU AI Act are significant. The regulation covers AI systems that are placed on the market or put into service within the EU, regardless of whether the provider or deployer is located inside or outside the EU.[664]

This means that AI systems used in satellites and other space-related technologies fall under the scope of the AI Act if their outputs are used within the EU. Consequently, AI-driven systems used in satellite surveillance and space missions must adhere to the principles of human-centricity and trustworthiness outlined in the EU AI Act if they want to transact in the EU markets.

An important point to mention, given the dual-use nature of space applications, is that AI systems developed for military, defense, or national security purposes are excluded from the scope of the regulation.[665] AI systems developed for public authorities located in a third country or for international organizations are also excluded.[666]

Furthermore, the space industry often involves cross-border collaborations and operations, making the extraterritorial application of the EU AI Act particularly relevant. Space agencies and companies operating within or in partnership with the EU must comply with this regulation, ensuring that their AI systems meet the stringent ethical and safety standards set forth by the Act.

This alignment promotes innovation and trust in AI technologies within the space sector and contributes to the broader goal of responsible AI governance. The following outlines the key principles of the EU AI Act and their connection to space activities.

3.1 *Risk Assessment and Prohibited AI Practices*

As per the Regulation rationales, a proportionate and effective set of binding rules for AI systems cannot be implemented without a clearly defined risk-based approach.[667] The risk-based approach classifies AI systems from minimal to unacceptable risk. There are no particular restrictions associated with the minimal risk AI systems, but there is a defined set of conditions to identify a high-risk AI system.[668]

AI systems having more than a high risk are labeled unacceptable and are therefore prohibited, e.g. discrimination in the use of AI, unauthorized data scraping or surveillance. This is particularly the case for cognitive behavioral manipulation,[669] social scoring,[670] and biometric identification and categorization.[671] In the field of space activities, it is prohibited, for example, to use AI-driven autonomous satellites designed to destroy, disable, or interfere with other satellites.

Although many of the AI applications used in space activities today, such satellite data processing and autonomous navigation, might not be considered high-risk and therefore not be subject to stringent compliance regulations, this position could be challenged by future developments. For example, increased security measures may be necessary for AI systems that

---

[664] *Ibid*., Article 2
[665] *Ibid*., Article 2, para 3
[666] *Ibid*., Article , para 4
[667] *Ibid.*, Introductory Paragraph 26
[668] *Ibid*., Introductory Paragraph 52
[669] *Ibid.*, Introductory Paragraph 28
[670] *Ibid*., Introductory Paragraph 31
[671] *Ibid*., Introductory Paragraph 16; 30.

perform crucial functions like docking to commercial space stations or future "space hotels" especially when they include people.

Companies like SpaceX are already utilizing AI for docking procedures with the International Space Station, setting a precedent for potential regulation of these systems.[672] Moreover, as robots are developed to assist in healthcare or other services in space, strong ethical frameworks will be necessary to guarantee safe and appropriate human-robot interactions, particularly in scenarios involving civilians or clients. Despite the fact that such developments could eventually be governed by contractual risk allocation, they highlight potential future challenges the industry may face as AI technology advances.

Building on the topic of risk allocation, risk assessment is also particularly relevant for space operators when analyzing space data and writing algorithms for space applications. The complexity and critical nature of space missions demands stringent safety and reliability standards to prevent mission failure or harm to individuals.

Additionally, transparency in AI's data collection and processing is conditional on specific tests and reviews, particularly for Earth observation (EO) operators who rely heavily on accurate data for environmental monitoring and resource management. Ensuring transparency through regular evaluations helps build trust in the accuracy and ethical use of AI in space applications, further underscoring the importance of accountability and compliance as AI systems continue to evolve.

3.2 *Ethical Principles for AI*

The EU AI Act reaffirms that AI should be a human-centric technology serving as a tool for people, with the ultimate aim of increasing human well-being.

A foundational basis for drafting codes of conduct under this regulation is the Ethics Guidelines for Trustworthy AI developed by the High-Level Expert Group on Artificial Intelligence (AI HLEG) in 2019. These guidelines, formulated by the independent AI HLEG appointed by the European Commission, outline seven non-binding ethical principles intended to ensure that AI systems are trustworthy and ethically sound. The principles encompass human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination, and fairness; societal and environmental well-being; and accountability.

The EU underscores the importance of these principles and encourages all stakeholders to consider them for the development of voluntary best practices and standards. This inclusive and general approach can serve as a great baseline for the development of the international guidelines on the use of AI systems in space.

---

[672] How SpaceX is Using AI to Advance Its Ambitions." Analytics India Magazine, January 23, 2023. https://analyticsindiamag.com/ai-origins-evolution/how-spacex-is-using-ai-to-advance-its-ambitions/.

3.3 *Data Governance and Management Practices*

Performance of many AI systems relies on the data used for their training. The EU AI Act pays great attention to the datasets used for training, validation, and testing, including the labels, to avoid them becoming potential sources of discrimination. The data should have “high quality,” which, as the Act explains further, is: 1) relevant, 2) sufficiently representative, 3) free of errors to the greatest extent possible, 4) complete in view of the intended purpose of the system, and 5) having the appropriate statistical properties to avoid biases, especially where data outputs influence inputs for future operations (feedback loops).[673]

This concern is particularly relevant in the case of AI-automated data labeling systems, such as those provided by companies like Scale AI.[674] These systems produce labeled data that trains new AI models. If the initial labeling system malfunctions or introduces inaccuracies, it can propagate errors to downstream AI models, creating negative feedback loops that amplify biases or distortions in future operations.

Therefore, ensuring high-quality labeling and robust validation mechanisms is crucial for maintaining the integrity of AI systems over time. To uphold these standards of quality and accountability, it is essential for stakeholders involved in high-risk AI systems to implement appropriate data governance and management practices.[675] According to Art.10 of the Act, those practices shall concern in particular:

- the relevant design choices;
- data collection processes and the origin of data, and in the case of personal data, the original purpose of the data collection;
- relevant data-preparation processing operations, such as annotation, labeling, cleaning, updating, enrichment and aggregation;
- the formulation of assumptions, in particular with respect to the information that the data are supposed to measure and represent;
- an assessment of the availability, quantity and suitability of the data sets that are needed;
- examination in view of possible biases that are likely to affect the health and safety of persons, have a negative impact on fundamental rights or lead to discrimination;
- appropriate measures to detect, prevent and mitigate possible biases
- the identification of relevant data gaps or shortcomings that prevent compliance with the Act.

3.4 *Technical Robustness and Resilience*

Resilience is a key concept in the development of space missions in order to face technical issues in orbit. Article 13.1 of the Regulation states that High-risk AI systems shall be designed

---

[673] *Ibid.*, Introductory Paragraph 67

[674] Scale. "Make the Best Models with the Best Data." Accessed September 29, 2024. https://scale.com/.

[675] *Ibid.* Art. 10

and developed in such a way as to ensure that their operation is sufficiently transparent to enable deployers to interpret a system's output and use it appropriately.

Para.2 specifies that High-risk AI systems shall be accompanied by instructions for use in an appropriate digital format or otherwise that include concise, complete, correct and clear information that is relevant, accessible and comprehensible to deployers, including information concerning the level of accuracy, including its metrics, robustness and cybersecurity against which the high-risk AI system has been tested and validated and which can be expected, and any known and foreseeable circumstances that may have an impact on that expected level of accuracy, robustness and cybersecurity.[676] Those elements are of particular relevance in the field of space activities.

Article 15 specifies that High-risk AI systems shall be designed and developed in such a way that they achieve an appropriate level of accuracy, robustness, and cybersecurity, and that they perform consistently in those respects throughout their lifecycle. In order to assess the level of robustness, a system of benchmarks and measurement methodologies should be established by the Commission (Art.15.2).

The robustness of high-risk AI systems may be achieved through technical redundancy solutions, which may include backup or fail-safe plans. Here, the need for interoperable space systems is of particular importance for future space programs.

Article 59 provides that AI systems shall be developed for safeguarding substantial public interest by a public authority or another natural or legal person in particular in the field of safety and resilience of transport systems and mobility, critical infrastructure and networks.[677] This is particularly relevant in the case of space infrastructure and the necessity of resilience in case of cyberattack for instance or system's malfunction. In addition, resilience is needed for safeguarding critical infrastructures.

3.5 *Governance and Implementation*

The governance and implementation of the EU Artificial Intelligence Act will be supported by the establishment of a Code of Practice,[678] which can serve as an inspiration for enhancing guidelines related to the use of AI in space activities.

Article 57 delineates the establishment of AI regulatory sandboxes designed to facilitate the experimentation and testing of AI systems within a controlled environment. Under the guidance of a regulator, companies can investigate and test innovative products, services, or business models using these sandboxes. These frameworks are particularly advantageous for the space industry, as they enable stakeholders to explore novel AI applications while ensuring compliance with regulatory standards.

Article 58 mandates the European Commission to adopt implementing acts that will detail the arrangements for the establishment, development, implementation, operation, and supervision

---

[676] *Ibid*. Art.13.3.b.II.

[677] *Ibid*. Art. 59.1.a.IV.

[678] *Ibid*. Art. 56.

of these AI regulatory sandboxes. This approach aims to prevent fragmentation across the Union. Member States may also incorporate AI regulations into their national space legislation, aligning their policies with the broader objectives of the AI Act.

Article 62 outlines measures specifically designed for providers and deployers, particularly small and medium-sized enterprises (SMEs) and start-ups, to support their effective implementation of the Regulation. Article 64 introduces the development of an AI Office, while Article 65 establishes the European Artificial Intelligence Board. These bodies will provide essential guidance for the implementation of the Regulation, as detailed in Article 66. The AI Office is empowered to monitor compliance with the Regulation, ensuring that providers of general-purpose AI models adhere to the approved Codes of Practice, as specified in Article 89.

Additionally, the Act calls for the creation of a scientific panel comprising independent experts, as stipulated in Article 69. This panel will contribute to informed decision-making and regulatory oversight.

Member States are required to designate national competent authorities and establish single points of contact, as outlined in Article 70. This framework aims to streamline communication and enhance regulatory efficiency across the Union.

Article 96 further directs the Commission to develop comprehensive guidelines for implementing the Regulation. These guidelines will specifically address:

(a) the application of requirements and obligations for high-risk AI systems;
(b) the prohibited practices defined in Article 5;
(c) the implementation of provisions regarding substantial modification;
(d) the transparency obligations established in Article 50;
(e) the relationship between this Regulation and Union harmonization legislation (Annex I) and other relevant EU laws, ensuring consistency in enforcement;
(f) the definition of an AI system as outlined in Article 3, point (1).

In formulating these guidelines, the Commission will pay particular attention to the needs of SMEs, local public authorities, and sectors most likely to be impacted by the Regulation.

The concepts articulated within the Regulation present valuable opportunities for the space sector, especially through the implementation of established standards.

3.6 *Privacy and Personal Data Protection*

According to the Act, National Data Protection Authorities must have a sufficient number of personnel available on a permanent basis, possessing expertise in areas such as AI technologies, data computation, personal data protection, cybersecurity, fundamental rights, and health and safety risks. This expertise should also encompass knowledge of existing standards and legal requirements.[679]

According to Article 3 (48), the term 'national competent authority' refers to either a notifying authority or a market surveillance authority. In the context of AI systems used or put into service by Union institutions, agencies, offices, and bodies, references to national competent

---

[679] *Ibid.* Art.70.

authorities or market surveillance authorities in this Regulation shall be interpreted as referring to the European Data Protection Supervisor.

Furthermore, the European Data Protection Supervisor has the authority to establish an AI regulatory sandbox for Union institutions, bodies, offices, and agencies, allowing it to exercise the roles and responsibilities of national competent authorities as outlined in Article 57.3.

In addition, the European Data Protection Supervisor (EDPS) has the power to create an AI regulatory sandbox for Union institutions, bodies, offices, and agencies. This enables the EDPS to fulfill the functions and duties usually assigned to national competent authorities, as outlined in Article 57.3 of the suggested Artificial Intelligence Act (AIA).

In the realm of space endeavors, transmitting data like pictures and navigation signals presents distinct regulatory obstacles. Data that is frequently sensitive and crucial for space programs requires supervision to guarantee alignment with data protection and AI regulations.

National authorities are essential in setting rules for AI systems, especially when they are utilized by or integrated into national space agencies. This involves overseeing AI applications in both national and international space programs, ensuring adherence to AI regulations and broader data protection laws like the General Data Protection Regulation (GDPR).

The national governments must weigh the potential for innovation in AI within the space industry against the importance of strict protections to guarantee that data management in space operations complies with privacy and security considerations.

3.7 *AI Literacy Measures*

“AI literacy” means skills, knowledge and understanding that allow providers, deployers and affected persons, taking into account their respective rights and obligations in the context of the Regulation, to make an informed deployment of AI systems, as well as to gain awareness about the opportunities and risks of AI and possible harm it can cause.[680]

In order to obtain the greatest benefits from AI systems while protecting fundamental rights, health and safety and to enable democratic control, AI literacy should equip providers, deployers and affected persons with the necessary notions to make informed decisions regarding AI systems. In the context of the application of the Regulation, AI literacy should provide all relevant actors in the AI value chain with the insights required to ensure the appropriate compliance and its correct enforcement.

The European Artificial Intelligence Board should support the Commission, to promote AI literacy tools, public awareness and understanding of the benefits, risks, safeguards, rights and obligations in relation to the use of AI systems. The Commission and the Member States should facilitate the drawing up of voluntary codes of conduct to advance AI literacy among persons dealing with the development, operation and use of AI.

Providers and deployers of AI systems shall take measures to ensure, to their best extent, a sufficient level of AI literacy of their staff and other persons dealing with the operation and use of AI systems on their behalf, taking into account their technical knowledge, experience, education

---

[680] *Ibid*. Art.3.56.

and training and the context the AI systems are to be used in, and considering the persons or groups of persons on whom the AI systems are to be used.[681]

In order to ensure that all stakeholders are prepared to successfully navigate the complexity of AI in this crucial industry, it is also imperative that AI standards be developed specifically for private and public entities functioning within the space sector.

Conclusions

In conclusion, this section highlights the importance of several key frameworks for AI regulation, including the OECD Principles on AI, UNESCO's recommendations, and the EU AI Act. The OECD's AI Principles, first introduced in 2019 and revised in 2023 and 2024, serve as the first intergovernmental standards on AI and provide a vital foundation for the responsible use of AI, promoting transparency, accountability, and robustness.

While these principles are non-binding, they offer crucial guidance for countries to model their national regulations and are highly relevant for the regulation of AI in outer space activities. By offering common definitions and principles that focus on AI systems, lifecycles, and actors, the OECD framework ensures that AI applications, including those in space, are designed with ethical considerations at the forefront, promoting transparency, risk management, and stakeholder cooperation.

Similarly, UNESCO's recommendations, supported by 193 signatories, provide a broader foundation for the development of future regulatory instruments for AI. While not specific to outer space activities, the widespread acceptance of these recommendations highlights their importance in fostering ethical AI practices, which can be applied to satellite operations and space technologies. These principles, such as transparency and accountability, can serve as a guide for space-faring nations to regulate AI systems in satellites, ensuring alignment with international ethical standards.

The EU AI Act builds on these frameworks by offering a risk-based approach to AI governance. Its extraterritorial reach and focus on human safety, rights, and well-being underscore the need for robust regulations that protect against the misuse of AI in critical sectors like space exploration.

However, the Act also raises concerns about gaps in regulating dual-use AI systems, particularly those designed for military and defense applications. These issues call for ongoing flexibility and adaptability in AI legislation to keep pace with technological advancements, especially in space missions where human-robot interactions are increasingly prevalent.

Together, the OECD Principles, UNESCO recommendations, and the EU AI Act form a comprehensive foundation for ethical AI regulation. As AI continues to evolve and expand into areas like space exploration, the integration of these frameworks into national and international policies will be crucial.

---

[681] *Ibid.* Art.4.

PART IV: ENVIRONMENTAL RESPONSIBILITY

Section 1: AI-Driven Solutions for Space Environment Protection: Enhancing Control and Debris Mitigation

Introduction

The rapid expansion of space exploration and the increasing deployment of satellites have brought forth unprecedented challenges in maintaining the safety and sustainability of space operations. Among these challenges are the need for precise control over space objects, collision avoidance, and mitigating the growing issue of space debris.

AI is emerging as a transformative technology in addressing these challenges, providing powerful tools for the autonomous monitoring, control, and decision-making necessary to protect the space environment. From managing complex spacecraft operations to reducing human intervention in critical moments, AI systems offer a level of precision and efficiency that is becoming indispensable in space missions.

AI has been successfully implemented in controlling and monitoring satellites, performing autonomous docking, and even avoiding collisions with space debris. Real-world applications discussed in this section demonstrate AI's ability to operate with minimal human intervention, executing highly complex tasks with both speed and accuracy. The ability of AI systems to process vast amounts of data in real time makes them essential for tasks such as trajectory estimation, orbit determination, and collision avoidance in increasingly congested space environments.

In addition to improving operational efficiency, AI is playing a critical role in addressing the issue of space debris, which poses a growing threat to both manned and unmanned missions. AI-enabled systems are being developed to enhance collision prediction and automate avoidance maneuvers, significantly reducing the number of false alerts and increasing the accuracy of collision risk assessments. As the number of satellites in low Earth orbit (LEO) continues to rise—fueled by the advent of mega-constellations like SpaceX's Starlink—AI's capacity to process and analyze massive datasets becomes vital for the future of space traffic management.

This section delves deeply into the multifaceted role of AI in space object control, exploring its applications in automating the control of spacecraft, enhancing collision avoidance protocols, and addressing the pressing issue of space debris. By providing detailed case studies and practical examples, the following analysis demonstrates how AI-driven technologies are reshaping space missions and paving the way for a more sustainable and secure space environment.

The growing reliance on AI, coupled with the increasing complexity of space operations, underscores the necessity of developing standards and collaborative frameworks to ensure the responsible use of these technologies in space.

1. *AI-Controlled Space Objects*

AI is being used to monitor and control satellites. AI can also be used to avoid satellite collisions, during take-off and landing of spacecraft to automate engine operations. It is also used to carry out autonomous operations, such as docking with the ISS. Indeed, AI algorithms are used in SpaceX's Dragon spacecraft for docking with the ISS. Another example is the AI system used in the ESA's Sentinel-1 satellite, which performs autonomous monitoring of Earth's surface changes. These real-world examples illustrate the practical use of AI in space missions.

AI is used for autonomous monitoring of space objects.[682] In this case, software systems utilize various data sources such as telemetry data from satellites, radar observations, and optical tracking data. Programs like MATLAB, Python with libraries like NumPy and SciPy, and specialized software developed by space agencies are commonly used for data processing, trajectory estimation, and orbit determination.

Autonomy is generally associated with a system's ability to function without direct human intervention, though it is an area with several levels and gray areas.[683] A system should be able to "sense, think and act" in its environment in order to be considered autonomous, which requires the ability to detect its environment. In addition, an autonomous system must be able to react to non-routine conditions by adapting its behavior to achieve its objective, while remaining safe and secure.

Some autonomous systems in aerospace carry out predetermined acts that do not alter in response to the environment (automatic). Other systems (automated) initiate or modify their behavior or output in response to environmental feedback, while more advanced systems (autonomous) combine environmental feedback with the system's own interpretation of its current situation.

NASA's Eyes on the Solar System tool uses data from various sources for trajectory estimation and orbit determination. Another example is the use of Python libraries in analyzing radar data from the Space Surveillance Network (SSN) to track space debris.

The "Sense-Think-Act" concept[684] means: (i) the sensors (sense) collect data and transfer this data to a format that the software can interpret; (ii) Control ("think") by evaluating sensory data, spacecraft information and expected outcomes before deciding on actions to perform; and (iii) Actuation ("act") by carrying out the operation defined by the control analysis process (without any further human intervention).

In the field of space missions, ESA's Aeolus satellite uses LIDAR sensors (sense) to gather atmospheric data, processes this data to determine optimal flight paths (think), and adjusts its orbit accordingly (act).

It is so relevant to examine the relation between Human-Machine Systems (HMS) which integrate the capabilities of a human operator and a machine. An autonomous system or function is to some extent beyond human control. However, humans can exercise some control during the

---

[682] NASA, Current AI Technology in Space, 2023, https://ntrs.nasa.gov/api/citations/20240001139/downloads/Current%20Technology%20in%20Space%20v4%20Briefing.pdf.

[683] Kathiravan Thangavel et al., "Artificial Intelligence for Trusted Autonomous Satellite Operations," *Progress in Aerospace Science*, 144, January 2024 (doi.org/10.1016/j.paerosci.2023.100960).

[684] *Ibid.*

design and development phase, for example by interrupting the operation of the system. HMS can be managed in different ways:[685]

- Direct control: needs the permanent interaction of a human operator to directly or indirectly control the system's functions, which makes it non-autonomous.
- Divided control: some tasks are managed directly by the human operator, while others are controlled by the machine under the operator's supervision.
- Supervisory control: a system performs tasks autonomously while a human operator supervises and provides advice, intervenes and takes over control if necessary.

On the ISS, robotic arms like Canadarm2 which can be controlled directly by astronauts (direct control) or perform tasks autonomously under human supervision (shared control). Another example is the use of supervisory control in the operation of Mars rovers, where ground teams monitor and adjust the rover's activities based on mission requirements. These examples clarify the role of HMS in space.

AI plays an important role in collision prediction. Current approaches for collision avoidance face serious challenges, mainly[686]: (i) insufficient data and endangered autonomy of action in space; (ii) a high number of false alerts and a large uncertainty; (iii) lack of scalability and automation for an increasing number of assets.

AI algorithms for collision prediction often involve complex mathematical models and machine learning techniques.[687] Programs such as TensorFlow, PyTorch, and scikit-learn in Python are popular for developing and training machine learning models to analyze historical data, identify patterns, and predict future collisions based on orbital dynamics.

Active collision avoidance has become an important task in space operations nowadays, and hundreds of alerts corresponding to close encounters of a satellite and other space objects are typically issued for a satellite in Low Earth Orbit every week. Such alerts are provided in the form of conjunction data messages, and only about two actionable alerts per spacecraft and week remain to be resolved after analyzing all cases. Therefore, building fully automated techniques for predicting the collision risk can help make the process of avoiding collisions less costly, as the number of false positives could be substantially reduced.[688]

ESA is developing a collision avoidance system that will automatically assess the risk and likelihood of in-space collisions, improve the decision making process on whether or not a maneuver is needed, and may even send the orders to at-risk satellites to get out of the way.[689]

---

[685] *Ibid.*

[686] Chiara Manfletti et al., "AI for Space Traffic Management", *Journal of Space Safety Engineering*, 10 (4) 2023, 495-504.

[687] George Choumos et al., "Artificial Intelligence for a Safe Space: Data and Model Development Trends in Orbit Prediction and Collision Avoidance", AIAA SCITECH 2024 Forum, January 2024 (https://doi.org/10.2514/6.2024-2066).

[688] Lukasz Tulczyjew, "Predicting risk of satellite collisions using machine learning", *Journal of Space Safety Engineering* 8(4) 2021, 339-344.

[689] European Space Agency, Automating collision avoidance, October 22, 2019, https://www.esa.int/Space_Safety/Space_Debris/Automating_collision_avoidance.

Indeed, the ESA's Space Debris Office uses AI algorithms to analyze collision risk data and reduce false alerts. In the same vein, NASA uses the Conjunction Assessment Risk Analysis (CARA) system, which incorporates AI to improve the accuracy of collision predictions and automate avoidance maneuvers. These examples illustrate how AI can be used to enhance collision prediction and avoidance in space operations.

Lastly, real-time maneuvering of spacecraft requires fast and efficient control systems capable of processing sensor data and executing commands in milliseconds. Software frameworks like ROS (Robot Operating System), Simulink, and Flight Software (FSW) developed by space agencies provide the infrastructure for implementing real-time control algorithms, guidance systems, and autonomous decision-making routines.

ROS is used in the control systems of NASA's Mars rovers for real-time navigation and decision-making for instance. Another example is the use of Simulink in the development of control algorithms for the European Space Agency's Automated Transfer Vehicle (ATV), which autonomously docks with the ISS. These examples illustrate the practical application of these frameworks in space missions.

2. ***Automated Collision Avoidance Systems***

Over the next few decades, the "New Space" environment will bring about notable changes in the space environment, by the advent of mega-constellations, the proliferation of small satellites and the generalization of low-thrust engines. These changes will render existing safety strategies and collision avoidance procedures inadequate, as the number of conjunction alerts will increase.[690]

Some examples include SpaceX's Starlink mega-constellation—which is expected to consist of thousands of satellites—drastically increasing the number of potential conjunction alerts, the increased deployment of CubeSats by universities and private companies, which also contribute to the congestion of the Low Earth Orbit (LEO). It is also a well-known fact that a number of significant incidents have occurred during the last decades where two space objects have come into close proximity in the outer space environment, in some cases requiring the implementation of collision avoidance maneuvers.[691]

For instance, the February 2009 collision between a U.S. Iridium 33 communications satellite and a Russian Cosmos 2251 communication satellite—resulting in thousands of debris pieces—and the March 2009 RED threshold late notice conjunction threat to the International Space Station (ISS), due to the presence of debris designated as "25090 PAM-D," constitute clear examples underlining the need for improved collision avoidance measures.

---

[690] Sanchez, L., Vasile, M., and Minisci, E. 2020. "AI and Space Safety: Collision Risk Assessment." in Schrogl, K.-U. (ed.). *Handbook of Space Security: Policies, Applications and Programs*. 2nd edition. Cham, Springer, p. 941.

[691] Viikari, L. 2008. *The Environmental Element in Space Law: Assessing the Present and Charting the Future*. Leiden, Brill, p. 39.

In addition, the December 1992 case where the pilot of the Space Shuttle Discovery narrowly avoided a catastrophic collision of the spacecraft with a 10-cm piece of space debris underscores the ongoing risk of space debris to the safety of space operations.

AI appears to be the most promising approach to addressing this new situation, due to its capacity to operate at a faster pace than conventional computer models of physical systems and to make decisions based on a broader variety of parameters than human operators. By way of illustration, AI algorithms can analyze data from NASA's Space Surveillance Network to predict potential collisions more accurately and quickly than traditional methods.

The ESA's use of AI in their Space Debris Office to automate collision risk assessments appears as another practical application. This results in enhanced performance when more data is available, as will be the case in the coming years.[692] In particular, the application of AI could lead to a significant improvement in the process of collision avoidance. This would be achieved by the ruling out of false alarms and the replacement of the current approach, which relies on "just in time" maneuvers, with a more effective strategy.[693]

This subject is of such relevance that, for instance, in 2019, the ESA organized the "Collision Avoidance Challenge," which was a ML competition where participants were asked to build a model to predict the final estimate of the collision risk between a given satellite and a space object (e.g., another spacecraft, space debris, etc.).[694]

2.1 *AI-Driven Analytics for Space Data Processing*

The utilization of AI to enhance data analytics in outer space is one of the flagship applications of this emerging technology in the industry. As a case in point, AI can be employed to process and analyze satellite imagery, which is a considerably more time-efficient process than that of a human operator. Furthermore, it is capable of artfully identifying the valuable data and imagery that is worthy of collection in a manner that is akin to the way in which a human operator would comprehend a wide variety of images.[695]

Likewise, envisaging a future of satellite data being monitored and interpreted in real time is possible, either through the use of deep learning algorithms or through the deployment of 'intelligent' satellites. Consequently, it is anticipated that AI will ultimately prove to be a pivotal tool in the interpretation and analysis of satellite data. Planet Labs and Orbital Insight are just a few examples of space companies currently investigating the potential of using AI techniques to

---

[692] Sanchez, L., Vasile, M., and Minisci, E. "AI and Space… *op. cit.,* p. 941.

[693] Pardini, C., and Anselmo, L. 2021. "Evaluating the impact of space activities in low earth orbit." *Acta Astronautica* 184, p. 21. https://doi.org/10.1016/j.actaastro.2021.03.030

[694] See https://kelvins.esa.int/collision-avoidance-challenge/ The results are available in Uriot, T. et al. 2022. "Spacecraft collision avoidance challenge: Design and results of a machine learning competition." *Astrodynamics* 6, pp. 121-140. https://doi.org/10.1007/s42064-021-0101-5

[695] Martin, A.-S., and Freeland, S. 2021. "The Advent of Artificial Intelligence in Space Activities: New Legal Challenges." *Space Policy* 55, p. 5. https://doi.org/10.1016/j.spacepol.2020.101408

process and analyze satellite imagery[696] for environmental monitoring[697] or to provide geospatial analytics.[698]

In the concrete field of collision avoidance, as will be showcased throughout the section, AI is expected to enhance the data analysis of the different variables for orbit prediction, aiding ground operators in the decision-making process, and ultimately helping in executing collision avoidance maneuvers. To this end, AI can process telemetry data from satellites to predict and prevent collisions, reducing the need for manual analysis.

2.2 *AI-Driven Collision Prediction Models*

One of the principal factors contributing to on-orbit collisions is the current limitation of orbit prediction capabilities.[699] In terms of specific AI techniques used to improve orbit prediction, for example, machine learning models trained on historical orbital data –similar to the way humans utilize past experiences to anticipate future occurrences– can predict future positions of satellites with greater accuracy.[700]

The use of AI in conjunction with the Space Fence radar system can enhance the prediction and tracking of space objects, improving collision prediction capabilities. The current approach to collision avoidance relies on the probability of collision as a primary factor in risk assessment.[701]

Mashiku et al. (2019)[702] present an investigatory approach that employs the most recent advancements in ML to promote the potential for rapid and enhanced close approach predictions, by examining statistical and information theory parameters as opposed to the classical probability of collision computation, aimed at ascertaining the viability and reliability of accurately predicting close approaches.

A further objective is to develop a set of "information parameters" to complement the collision probability in the conjunction assessment process. Additionally, a dataset of conjunction events will be used to train ML algorithms.

---

[696] Andrew Myers, *Will Marshall on Combining Satellite Data and AI for Sustainability (Stanford University Doerr School of Sustainability*, March 25, 2024), https://sustainability.stanford.edu/news/will-marshall-combining-satellite-data-and-ai-sustainability.

[697] https://phys.org/news/2024-02-ai-satellite-imagery-planet.html In this regard, see also White, J. et al. 2024. "Quantifying trade-offs in satellite hardware configurations using a super-resolution framework with realistic image degradation." *Remote Sensing Letters* 15(3): 291-301. https://doi.org/10.1080/2150704X.2024.2318756 With respect to PlanetScope, there are also research studies investigating how to leverage PlanetScope satellite imagery for river water masking, for instance, by implementing AI technologies. See Valman, S. J. et al. 2024. "An AI approach to operationalise global daily PlanetScope satellite imagery for river water masking." *Remote Sensing of Environment* 301: 1-16. https://doi.org/10.1016/j.rse.2023.113932

[698] https://spacenews.com/orbital-insight-raises-9-million-for-ai-imagery-processing/

[699] Peng, H., and Bai, X. 2018. "Improving orbit prediction accuracy through supervised machine learning." *Advances in Space Research* 61(10), p. 2628. https://doi.org/10.1016/j.asr.2018.03.001

[700] Jadala, G., Meedinti, G. N., and Delhibabu, R. 2022. "Satellite Orbit Prediction using a Machine Learning Approach." *Proceedings of the ICAIW 2022: Workshops at the 5th International Conference on Applied Informatics 2022*, 27–29 October 2022, Arequipa, Peru. https://ceur-ws.org/Vol-3282/icaiw_waai_3.pdf

[701] Mashiku, A. et al. 2019. "Predicting satellite close approaches in the context of artificial intelligence." AAS/AIAA Astrodynamics Specialist Conference, 11–15 August 2019, Portland, Estados Unidos, pp. 1-6. https://ntrs.nasa.gov/api/citations/20190029019/downloads/20190029019.pdf

[702] *Ibid.*

Relying also on ML techniques to reduce prediction errors, in Peng and Bai (2018)[703] a computational framework improving orbit prediction accuracy through the ML method is showcased. The prevailing methodologies for orbit prediction are physics-based, thus, their efficacy hinges on a thorough comprehension of the initial state of the space object at the outset of trajectory computation, along with the environmental data (i.e., earth gravity).

The intent information pertaining to the maneuvering objects is crucial, nonetheless, the maneuvering of a spacecraft owned by another State may not be available when performing the orbit prediction. These issues collectively result in the conclusion that current errors in physics-based predictions are too substantial to be employed for the purpose of meaningful action.

These authors opt for the use of ML methodologies instead, which offer a distinct prediction capability in comparison to the physics-based approach, in such a way that the prediction can be conducted without the explicit modeling of space objects, spacecraft maneuvers, and the space environment. In contrast, the models are constructed through the analysis of a vast quantity of observed data, analogous to the manner in which the human mind learns from past experiences in order to predict future events.

In order to capitalize on the insights gained from physics-based models, standing as the current state of the art, the learning process is designed to focus on modeling the prediction errors, rather than the full dynamics.

In this manner, a surrogate modeling approach is used to identify and apprehend the prediction errors that arise. An additional advantage of this methodology lies in the fact that the learning process is not encumbered by the duty to identify and incorporate any incremental corrections to the physics-based prediction, thus minimizing the dimensionality of the learning task.

2.3 *AI-Enabled Decision-Making for Space Operations*

Another area in which AI can play a pivotal role in collision avoidance is by automating operators' tasks. A relevant example is ESA's AI-based collision avoidance system that automates the decision-making process for satellite maneuvers. Additionally, AI can be used to prioritize collision avoidance actions, ensuring that high-risk scenarios are addressed first. The integration of AI in mission control software can provide operators with real-time recommendations, improving response times and decision accuracy.

Such automation *a fortiori* permits operators to allocate a greater proportion of their time and effort to critical decision-making—in particular, to collision risk assessment, collision avoidance maneuvers, and disposal strategies evaluations.[704] Consequently, the time required for decision-taking is reduced.

On this note, Sánchez et al. (2020)[705] have evaluated the use of diverse Machine Learning (ML) techniques to create some intelligent classification systems, associating each of them with a

---

[703] Peng, H., and Bai, X. "Improving orbit prediction… *op. cit.*, pp. 2628-2629.

[704] Sanchez, L., Vasile, M., and Minisci, E. "AI and Space… *op. cit.,* p. 950.

[705] *Ibid.*, p. 951.

recommended action that would be presented to ground operators during the decision-making process, thereby demonstrating the potential of ML for supporting decision-making.[706]

Moreover, the anticipated overcrowding of the space environment will necessitate monitoring and control in terms of satellite numbers, with the volume of information subject to consideration exceeding the operators' capacities. AI systems designed to assist decision-making processes will be capable of managing vast quantities of information and presenting alternative strategies to operators in a significantly shorter time than current methodologies, having regard to a broader array of variables.[707]

In order for an automated decision-making system to actually alleviate the operator's workload, the models developed must ultimately be able to autonomously and reliably reach decisions in borderline scenarios, or at a minimum, provide a trustworthy recommendation mechanism for ground operators.[708]

By way of illustration, ESA is developing a collision avoidance system that will enable the automatic evaluation of risk and the likelihood of in-space collisions, ameliorating the decision-making process regarding the need for a given maneuver, and potentially issuing commands to at-risk satellites to facilitate their avoidance.[709]

2.4 *AI-Powered Execution of Collision Avoidance Maneuvers*

In this context, Vasile et al. (2017)[710] introduce an intelligent-based decision support system designed to aid ground operators when necessarily planning and executing collision avoidance maneuvers. To that end, the authors put forth the suggestion of establishing and running a database of potential predefined maneuvers to be employed in conjunction event scenarios, based on information derived from the analysis of previously applied maneuvers, from both a realistic and a virtual standpoint.

According to these data, an intelligent decision-making system is then enabled to propose alternative courses of action on the grounds of a series of criteria (i.e., the risk of not executing the collision in question compared with the risk associated with future possible collisions).

---

[706] Sánchez Fernández-Mellado, L., and Vasile, M. 2020. "On the Use of Machine Learning and Evidence Theory to Improve Collision Risk Management." 2nd IAA Conference on Space Situational Awareness (ICSSA), 14–16 January 2020, Washington D.C., Estados Unidos, pp. 22-31. https://core.ac.uk/download/pdf/327986972.pdf

[707] In a similar vein, Sanchez et al. (2019) conducted research on machine learning (ML) methodologies to automate and accelerate the evaluation of the collision risk between operational satellites and space debris, utilizing a database with initial parameters. See Sanchez, L., Vasilea, M., and Miniscia, E. 2019. "AI to support decision making in collision risk assessment." IAC–19–A6,IP,20,x53728, 70th International Astronautical Congress (IAC), 21–25 October 2019, Washington D.C., Estados Unidos. https://strathprints.strath.ac.uk/71041/1/Sanchez_etal_IAC_2019_AI_to_support_decision_making_in_collision_risk_assessment.pdf

[708] Ravi, P., Zollo, A., and Fiedler, H. 2023. "AI for Satellite Collision Avoidance — Go/No Go Decision-Making." 2nd International Orbital Debris Conference, 4–7 December 2023, Sugar Land, Estados Unidos. LPI Contribution No. 2852, p. 6043. https://www.hou.usra.edu/meetings/orbitaldebris2023/pdf/6043.pdf

[709] See https://www.esa.int/Space_Safety/Space_Debris/Automating_collision_avoidance

[710] Vasile, M. et al. 2017. "Artificial intelligence in support to space traffic management." IAC–17–A6,7,1,x41479, 68th International Astronautical Congress (IAC), 25–29 September 2017, Adelaide, Australia, pp. 4-12. https://pureportal.strath.ac.uk/en/publications/artificial-intelligence-in-support-to-space-traffic-management

Regarding specific AI techniques used for executing maneuvers, the use of reinforcement learning algorithms can serve as an example of optimizing the timing and direction of satellite maneuvers to avoid collisions while minimizing fuel consumption. The development of AI-based predictive models can also help in simulating different maneuver scenarios, allowing operators to choose the most effective course of action.

3. *AI Applications in Space Debris Mitigation*

In this context, it is interesting to review the Space Debris Mitigation Guidelines (the "Guidelines")[711] published by UNCOPUOS in 2007, and later adopted by the General Assembly,[712] on the basis of guidelines the Inter-Agency Space Debris Coordination Committee had proposed in 2002. Although legally non-binding, they have been adopted by States and implemented in national regulations as *de facto* international standards.[713]

They aim to minimize the creation of space debris for example through modifications of space systems and correct deorbiting. In particular, the Guidelines recommend to:

1) limit debris released during normal operations (for example, dedicated design efforts, prompted by the recognition of the threat posed by such objects, have proved effective in reducing this source of space debris);
2) minimize the potential for break-ups during operational phases (by incorporating potential break-up scenarios in failure mode analysis, the probability of these catastrophic events can be reduced);
3) limit the probability of accidental collisions in space (if available orbital data indicate a potential collision, adjustment of the launch time or an on-orbit avoidance maneuver should be considered);
4) avoid intentional destruction and other harmful activities (when intentional break-ups are necessary, they should be conducted at sufficiently low altitudes to limit the orbital lifetime of resulting fragments);
5) minimize potential for post-mission break-ups resulting from stored energy (all on-board sources of stored energy should be depleted or made safe when they are no longer required for mission operations or post-mission disposal, including residual propellants and compressed fluids and the discharge of electrical storage devices);
6) limit the long-term presence of spacecraft and launch vehicle orbital stages in the LEO region after their mission (spacecraft and launch vehicle orbital stages that have terminated their operational phases in orbits that pass through the LEO region should be removed from orbit in a controlled fashion; if not possible, they should be disposed of in orbits that avoid their long-term presence in the LEO region);

---

[711] A PDF version can be accessed here: Space Debris Mitigation Guidelines of the Committee on the Peaceful Uses of Outer Space (unoosa.org).

[712] UNGA 62/217, pointing out that the Guidelines reflected then current state practice.

[713] Francis Lyall, Paul B. Larsen, *Space Law - A Treatise*, Routledge, Second Edition, 2018, p. 276

7) limit the long-term interference of spacecraft and launch vehicle orbital stages with GEO after the end of their mission (for example, by moving them to an orbit above the GEO region, where they will not interfere with, or return to, the GEO region).[714]

In summary, the Guidelines focus on preventing the creation of additional space debris by modifying the design and planning of space systems and missions before, during and after their operations.

In addition to the 2007 Guidelines, there are other instruments that can be mentioned, such as the European Code of Conduct for Space Debris Mitigation[715] or the ISO Standard on space debris.[716] In 2014, UNCOPUOS published a "Compendium of space debris mitigation standards adopted by States and international organizations" that is regularly being revised and updated.[717]

Furthermore, recent initiatives have been undertaken in the field of space debris mitigation. Indeed, ESA adopted in 2023 the "Zero Debris Charter"[718] with the aim to improve orbital clearance, avoid in-orbit collisions and prevent intentional release of space debris. It is also worth mentioning the World Economic Forum's 2023 phased approach for space debris reduction[719] which deals with post-mission disposal, collision avoidance and maneuverability, as well as data sharing and traffic management in orbit.

At national level for instance, the US adopted in 2019 "Orbital Debris Mitigation Standards"[720] and in 2024 the FCC updated its rules related to "Mitigation of orbital debris in the New Space Age."[721] Interestingly, in September 2024, Italy adopted the *Space Economy Italian Pact*, a product of the *States General of the Space Economy* convened by the Parliamentary Intergroup for the Space Economy. While not legally binding, this strategic document outlines key directions for the future of the space sector. Among its many recommendations, it emphasizes the importance of establishing concrete actions to fully integrate AI into space activities, thereby safeguarding the long-term sustainability of the space environment.[722]

With this in mind, States should adopt phased approaches for debris reduction, starting with soft law guidelines that include AI-driven debris mitigation.

---

[714] For a detailed discussion, *see* Francis Lyall, Paul B. Larsen, *Space Law - A Treatise*, Routledge, Second Edition, 2018. Chapter 10.

[715] Available at: Microsoft Word - CoC_v1.0 [040628].doc (unoosa.org).

[716] The 2023 version can be accessed here: ISO 24113:2023 - Space systems — Space debris mitigation requirements.

[717] The current version can be accessed here: Space Debris Mitigation Standards Compendium - Update (unoosa.org).

[718] European Space Agency, Zero Debris Charter, 2023, https://esoc.esa.int/sites/default/files/Zero_Debris_Charter_EN.pdf.

[719] World Economic Forum, Space Industry Debris Mitigation Recommendations, June 2023, https://www3.weforum.org/docs/WEF_Space_Industry_Debris_Mitigation_Recommendations_2023.pdf.

[720] US Government, Orbital Debris Mitigation Standard Practices, Nov. 2019, https://orbitaldebris.jsc.nasa.gov/library/usg_orbital_debris_mitigation_standard_practices_november_2019.pdf.

[721] US Federal Register, Space Innovation; Mitigation of Orbital Debris in the New Space Age, August 2024, https://www.federalregister.gov/documents/2024/08/09/2024-17093/space-innovation-mitigation-of-orbital-debris-in-the-new-space-age#:~:text=In%20the%20Orbital%20Debris%20Second%20Report%20and%20Order%2C%20the%20Commission,collisions%20that%20would%20create%20debris.

[722] Stati Generali della Space Economy, Space Economy Italian Pact, 2024, https://www.statigeneralidellaspaceeconomy.it/wp-content/uploads/2024/09/MANIFESTO-1.pdf.

Conclusions

The absence of standards governing the use of AI in the space industry, particularly for fostering space sustainability, poses a critical challenge. This issue is magnified in STM, where the lack of standardized protocols and actions during conflict scenarios hinders efforts to address the problem of a congested space environment. While AI-enabled solutions offer significant potential to support operators and automate processes, their effectiveness depends on establishing shared rules and practices among all spacefaring entities. Recommendations for addressing these gaps, with both short-term and long-term goals for STM, will be discussed in detail in Part V, Section 2.

Concrete examples that could inspire standardization efforts include initiatives like the Consortium for Execution of Rendezvous and Servicing Operations (CONFERS). By developing standards for satellite servicing, including AI-enabled docking and maintenance, CONFERS demonstrates how standardized practices can ensure consistency, interoperability, and safety. Such models can guide the creation of AI standards across space operations, fostering a unified framework for sustainable and efficient use of AI in space.

Collaboration among space-faring nations is critical to establishing these standards. For example, UNOOSA could facilitate international workshops and agreements aimed at harmonizing AI standards. This could, for example, take the form of non-binding guidelines, which over time could become binding. This collaboration is crucial to prevent fragmented regulations and to ensure that all space activities adhere to a unified framework, promoting safe and sustainable space operations globally.

Moreover, the guidelines for space debris mitigation, like those adopted by UNOOSA, as well as the ESA’s Charter and the initiatives at national level, highlight the importance of preventing the creation of additional space debris. AI can play a crucial role in implementing these guidelines by enhancing the prediction, detection, and avoidance of debris. For example, AI-driven autonomous systems can aid in real-time monitoring and executing precise avoidance maneuvers. The integration of AI solutions will help mitigate space debris, contributing to safer space operations and promoting the long-term sustainability of space activities.

Section 2: Regulating AI for Environmental Protection in Space: Frameworks for National and International Laws

Introduction

AI's role in space operations—ranging from autonomous satellite management to collision avoidance systems—brings both unprecedented opportunities and significant risks. To address these, the international legal framework governing space, particularly the Outer Space Treaty, requires careful examination.

As discussed at length in Part I of this study, the 1967 Outer Space Treaty provides the foundation for international space law, but it lacks specific provisions that address the unique challenges posed by AI technologies.

Given the difficulties associated with amending the Outer Space Treaty—a process that requires widespread international consensus—alternative solutions must be explored. The development of soft law instruments, such as non-binding guidelines and international standards, could offer a more pragmatic approach in the short term.

These instruments, like the Space Debris Mitigation Guidelines, have proven effective in shaping international norms and fostering responsible behavior in space without the need for formal treaty amendments. Similar soft law mechanisms can be developed to regulate AI applications, particularly in areas such as space traffic management and debris mitigation.

Moreover, integrating AI-specific protocols into the International Code of Conduct for Outer Space Activities (ICOC-OSA) would enhance the regulation of AI-driven space operations. By incorporating guidelines for AI-powered collision avoidance and space debris management, the ICOC could ensure the safe, transparent, and ethical use of AI technologies in space.

However, for these frameworks to be implemented effectively, the establishment of a Working Group on AI Governance in Space within the Legal Subcommittee of UNCOPUOS is essential. This group would be tasked with developing detailed recommendations, international standards, and ethical guidelines for AI deployment in space, ensuring that AI technologies are used responsibly and in line with international humanitarian law and other relevant legal frameworks.

This section delves into the potential amendments to international space law, the role of soft law instruments, and the importance of establishing a multilateral approach through UNCOPUOS to govern the ethical and responsible use of AI in space. By setting forth these recommendations, this study aims to pave the way for the international community to collaboratively address the regulatory challenges posed by deploying AI in space, to ensure the long-term protection of the outer space environment and the related safety and sustainability of space activities.

1. *Amendments to International Space Law for AI Governance*

The Outer Space Treaty provides in Article XV that "[a]ny State Party to the Treaty may propose amendments to this Treaty. Amendments shall enter into force for each State Party to the Treaty accepting the amendments upon their acceptance by a majority of the States Parties to the Treaty and thereafter for each remaining State Party to the Treaty on the date of acceptance by it."

In theory, therefore, the Outer Space Treaty could be amended to include principles for responsible AI deployment in space. Such an inclusion would need to be proposed by one or more State Parties to the Treaty and accepted by a majority of the States Parties to the Treaty.

Given the current difficulties being encountered, however, in efforts to agree on binding international rules, an amendment does not seem to be feasible in the short term. In this context, the development of soft law could serve as an attractive alternative. A look at the example of the Space Debris Mitigation Guidelines set out above, could indicate a way forward.

UNCOPUOS, for example, could work on international standards for responsible AI that could be set out as non-binding guidelines. If acceptable to a large number of States, such guidelines could then be adopted by the United Nations' General Assembly. In order to advance in such a direction, the first step would be to set up a Working Group within the Legal Subcommittee of the United Nations Committee on the Peaceful Uses of Outer Space.

1.1 *Integrating AI Guidelines into the International Code of Conduct for Outer Space Activities (ICOC-OSA)*

Outer space is considered as a "congested, contested and competitive" place. The European Union considered that strengthening the security, safety and sustainability of activities in outer space was an important objective in the context of expanding space activities.[723] The proposal of a draft Code on space activities was conceived as a reply to the United Nations General Assembly resolution 61/75 of 6 December 2006, which called Member States to submit proposals on transparency and confidence building measures (TBCM) within the context of the prevention of an arms race in outer space (PAROS).

In 2007, the EU's Portuguese Presidency prepared a Food for Thought on a Comprehensive Code of Conduct for Space Objects, based on the principles of freedom to use outer space for peaceful purposes, preservation of the security and integrity of space objects in orbit and due consideration for the legitimate security and defense interests of States. The General Assembly reiterated its call in resolution 62/43 "TBCM in Outer Space Activities" of 5th December 2007, and in the similar resolutions adopted in 2008, 2009 and 2010.

---

[723] Sergio Marchisio, The Draft Code of Conduct for Outer Space Activities, UNITED NATIONS/THAILAND WORKSHOP ON SPACE LAW Activities of States in Outer Space in Light of New Developments: Meeting International Responsibilities and Establishing National Legal and Policy Frameworks 16-19 November 2010 Bangkok, Thailand, https://www.unoosa.org/pdf/pres/2010/SLW2010/02-10b.pdf.

The International Code of Conduct for outer space activities (ICoC) was developed after some incidents including the 2009 collision between a US Iridium satellite and a Russian Cosmos satellite underlying the necessity for improved space traffic management and AI guidelines.

Although the process has faced criticism for being non-transparent and non-inclusive,[724] the ICoC offers added value in three key aspects: (i) it covers all dimensions of space operations, civil as well as military; (ii) the commitment to refrain from any action bringing about, directly or indirectly, damage or destruction, of space objects.[725] That corresponded, in effect, to a political commitment to ban the testing of destructive anti-satellite weapons; (iii) the dynamic nature of the Code as Parties would have to revise and update the text in light of the forthcoming developments.

The draft International Code of Conduct proposes several measures on space operations and the mitigation of debris, including: Promoting space safety and sustainability; Pursuing strategic stability; Minimizing the risk of accidents, collisions, and harmful interference in space; Refraining from deliberate damage or destruction of spacecraft, unless in self-defense or to mitigate debris; Taking appropriate measures like prior notification and consultations to minimize collision risks; Improving adherence and implementation of the ITU regulations; Minimizing the creation of long-lived space debris and implementing UNCOPUOS Space Debris Mitigation Guidelines.[726]

A future Code of Conduct could incorporate AI-specific protocols, such as mandatory AI collision avoidance systems in satellites and detailed guidelines for the use of AI in space debris mitigation. These AI systems have already been used successfully in other areas, such as NASA's Autonomous Spacecraft Experiment (ASE),[727] in particular in the field of data collection and procession.

After the 2015 Multilateral Negotiations in New York, the Chair assessed that based on the discussions and considering the importance afforded to the principles of openness, transparency, universality and inclusiveness, the most supported way forward would be the pursuit of negotiations within the framework of the United Nations through a mandate of the General Assembly.

After New York, the EU and its Member States kept open the diplomatic initiative on a non-binding instrument on the security of space activities but the EU did not find, after the suspension of the ICoC initiative, the necessary internal agreement and motivation for pursuing with a convincing way ahead, aimed at convening an open-ended working group of the UN to negotiate a code or a similar instrument.[728] The lessons learned from the ICoC's experience

---

[724] Sergio Marchisio, *The Law of Outer Space Activities* (Roma: Edizione Nuova Cultura, 2022) 340.

[725] *Ibid*., 341.

[726] Chris Johnson, Draft International Code of Conduct for Outer Space Activities, Secure World Foundation, February 2014, https://swfound.org/media/166384/swf_draft_international_code_of_conduct_for_outer_space_activities_fact_sheet_february_2014.pdf.

[727] NASA, Autonomous Sciencecraft Experiment, June 15, 2004, https://www.jpl.nasa.gov/missions/autonomous-sciencecraft-experiment-ase.

[728] Sergio Marchisio, *The Law of Outer Space Activities* (Roma: Edizione Nuova Cultura, 2022) 346.

constitute indeed a precious heritage for the development of the new diplomatic exercise on norms for responsible behavior in outer space.

A future ICOC-OSA Integration of AI Guidelines should integrate the following elements:[729]

- AI components should be clearly identified for transparency purposes and utilization in order to avoid misperception and misunderstanding in the realization of future space missions;
- AI systems should remain under human control whether with a direct human control over the system (*human-in-the-loop*),[730] or a supervised control regime where the device has more autonomy but must be monitored by a human (*human-on-the-loop*);
- AI systems should be inclusive and accessible, and the benefits from the use of this technology should not give rise to unfair discrimination against actors and end users (see UN GA Res. 51/122);[731]
- AI systems should respect and uphold privacy rights and data protection, and ensure the security of data;
- Space programmes with AI components should be reliably operated in accordance with their intended purpose;
- It should be possible to contest the choice of AI use if the relevant space system significantly impacts adversely upon other users (governmental or private) or the space environment. Programmers should be able to detect and avoid unintended consequences, and possess the capability to disengage or deactivate deployed systems that demonstrate unintended behavior;
- Those responsible for the development of different phases of the AI system should be identifiable and accountable for the outcomes of the AI systems, and human oversight of AI systems should be maintained.

Future guidelines could also integrate the following aspects: AI components should be registered with the international body, and their capabilities and limitations should be documented and shared with all stakeholders; creation of an international AI monitoring body that oversees the deployment and operation of AI in space to ensure compliance with ethical and safety standards.

Indeed, a concrete implementation could be the creation of an international AI monitoring body to oversee compliance with these guidelines and ensure transparency and accountability in AI deployments in space.

---

[729] A.S. Martin, S. Freeland, "Artificial Intelligence –A Challenging Realm for Regulating Space Activities", *Annals of Air and Space Law*, XLV, 2020, 275-306.

[730] See Ge Wang, "Humans in the Loop: the Design of Interactive AI Systems" Stanford University, October 20, 2019, https://hai.stanford.edu/news/humans-loop-design-interactive-ai-systems; see also Fabio Massimo Zanzotto, "Viewpoint: Human-in-the-loop Artificial Intelligence, *J Artificial Intelligence Research* 64 (2019) 243-252.

[731] *Declaration on International Cooperation in the Exploration and Use of Outer Space for the Benefit and in the Interest of All States, Taking into Particular Account the Needs of Developing Countries*, GA Res 51/122, UNGAOR, 1996, Supp No 49, UN Doc A/51/49 para 3.

Many of these elements for future AI governance could potentially be incorporated within Transparency and Confidence Building Measures (TCBMs), which are increasingly regarded as significant tools to create a trustworthy environment for the carrying out of outer space activities by various actors, public and private.[732]

Indeed, TCBMs allow for the clarification of States' intentions by determining norms of behavior in outer space, albeit in a non-binding framework.[733] They are essential to reinforce mutual trust and build confidence among space actors and have been an increasingly common modality within the agreed "rules of the road" for the use of space.[734]

1.2 *United Nations Resolution on AI Governance in Space (UNR-AIGS)*

Another element to consider is the adoption of a resolution calling for the regulation of AI in space. This resolution could refer to the recent one adopted by the UNGA in March 2024 on AI— "Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development."[735]

A future resolution on AI in space should recognize and integrate the need for safe, secure and trustworthy artificial intelligence systems (civil/military domain) whose life cycle includes the following stages: pre-design, design, development, evaluation, testing, deployment, use, sale, procurement, operation and decommissioning. Furthermore, AI systems should be reliable, explainable, ethical, inclusive, and developed in respect of human rights, fundamental freedoms and international law, with sustainable development oriented, and responsible.

The rapid acceleration of the design, development, deployment and use of artificial intelligence systems and rapid technological change imply to facilitate international cooperation to formulate and use effective, internationally interoperable safeguards, practices and standards that promote innovation and prevent the fragmentation of the governance.

The governance of AI systems is an evolving area and the need for continued discussions on possible governance approaches that are appropriate, based on international law, interoperable, agile, adaptable, inclusive, responsive to the different needs and capacities of developed and developing countries alike and for the benefit of all.

States should (i) promote the development and implement domestic regulatory and governance approaches and frameworks to support responsible and inclusive artificial intelligence innovation and investment for sustainable development; (ii) encourage the development of effective measures, that promote innovation for the internationally interoperable identification, classification, evaluation, testing, prevention and mitigation of vulnerabilities and risks during the

---

[732] See e.g., Cassandra Steer, "Sources and Law-Making Processes Relating to Space Activities" in Ram S Jakhu, Paul Stephen Dempsey, eds, Routledge *Handbook of Space Law (*New York: Routledge Taylor & Francis Group, 2017) 22.

[733] See Peter Martinez et al, "Criteria for developing and testing Transparency and Confidence-Building Measures (TCBMs) for outer space activities", *Space Policy* 30, 2 (2014) 1-7.

[734] See generally Steven Freeland, "The Role of "Soft Law" in Public International Law and its Relevance to the International Legal Regulation of Outer Space" in Irmgard Marboe, ed, *Soft Law in Outer Space: The Function of Non-binding Norms in International Space Law* (Austria: Bohlau Publishing, 2012) 9–30.

[735] Resolution A/78/L.49 adopted on 11 March 2024 on AI.

design and development and prior to the deployment and use of artificial intelligence space systems; (iii) encourage the incorporation of feedback mechanisms and a mechanism of lessons learnt considering misuses and incidents of artificial intelligence systems; (iv) foster the development of mechanisms for risks monitoring and securing data; (v) strengthen investment in developing and implementing effective safeguards for artificial intelligence systems security; (vi) encourage the development and deployment of effective, accessible, adaptable, internationally interoperable technical tools, standards and practices in the use of artificial intelligence in space activities; (vii) facilitate the development and implementation of effective, internationally interoperable frameworks, practices and standards; and (viii) promote transparency, predictability, reliability and understandability throughout the life cycle of artificial intelligence space systems.

Practical implementation of such resolution could be through public awareness, involving educational campaigns and workshops for stakeholders in the space sector. In addition, another relevant element to consider is the development of interoperable tools such as the use of blockchain for secure data sharing among space agencies.

2. *Establishing a Working Group on AI Governance in Space*

A Working Group on AI governance in space could be established within the Legal Subcommittee of UNCOPUOS as it has been the case for the WG on space resources as UNCOPUOS deals with new technology and new space activities in its Agenda.

Maybe first a new agenda item on AI governance and then a working group could be established as UNCOPUOS represents a diplomatic body with its own timeline.

In recent years, there has been significant interest in the potential of AI to transform future space activities. This growing enthusiasm has prompted several countries to introduce policies, strategies, and draft national legislation aimed at promoting AI-driven advancements in the space sector.

A first step could be the introduction of a single agenda item on AI in space, thus allowing a more formal and dedicated process by which Member States could express their particular views on the issues.

It becomes increasingly clear that this is a topic of significant interest for many Member States, so a Working Group could be established on the issue. First, it could be established a "scheduled informal consultations" (SIC), to take place in addition to the discussions under the single agenda item (which could be retained each year) with co-Moderators for consultations and, then during LSC for instance, at the conclusion of SIC, the Member States could agree to establish a formal Working Group.

Member States' WG agreed the specifics of the Mandate, Terms of Reference and Work Plan/Methods of Work for the Working Group.

It demonstrates the importance of the multilateral process through UNCOPUOS and is indicative of the good faith and widespread flexibility and willingness of Member States to find a common path forward in relation to the "big" issues regarding the peaceful exploration and use of outer space.

Any consideration of possible future AI space activities will necessarily involve not only legal and governance considerations, but also very significant technical, economic, political, cultural, scientific and other factors.

In this regard, the Member States would have to determine a clear methodology for the Working Group to collect relevant information and, in conformity with the established practice of UNCOPUOS, also to take into account inputs from a broad range of stakeholders and permanent observers of UNCOPUOS. This will facilitate an inclusive and informed discussion among Member States at the meetings of the Working Group.

The mandate of the SRWG foresees a step-by-step approach hopefully leading to the development of "initial principles" for the safe, sustainable, rational and peaceful conduct of space activities with AI. It is necessary to involve non-governmental stakeholders like companies or academia, and permanent observers for inputs to ensure that the practice of UNCOPUOS is not departed from. Another idea could be that the WG relies on dedicated "expert groups" to support it in the identification of emerging issues, and the development of guidelines for AI governance in space activities.

It is usually said that space diplomacy is too slow compared to the pace of business and technological development. The issue of the use of AI space is a very significant issue for humanity.

It is crucial that whatever outcomes the global community determines—through the multilateral process at UNCOPUOS—are appropriate not just for the short term, but for many years. Short-term solutions regarding such a complex issue would not appropriately serve the global community.

A Working Group on AI as an important step so as to ensure that a logical, comprehensive, inclusive and open multilateral process is undertaken to determine what is required. For example, the WG could start by developing guidelines on the ethical use of AI in space missions, ensuring data privacy, and setting safety standards for AI systems.

The WG could propose specific articles or clauses that could be included in the new guidelines to address transparency, such as mandatory reporting on AI system capabilities and limitations. Additionally, some subcommittees could be introduced within the WG, each tasked with a specific area such as AI in satellite communications, AI in space debris management, and AI in planetary exploration.

National licensing and AI regulation. States could introduce AI requirements in their national space law as it has been done with the Space Debris Mitigation Guidelines. Building on the approach used for the Space Debris Mitigation Guidelines, this method could provide an interim solution while the multilateral process at UNCOPUOS is underway. Nations could adopt national licensing requirements that mandate compliance with AI safety and ethical guidelines.

For example, countries could require that AI systems on spacecraft undergo rigorous testing and validation to ensure they meet international safety standards. This would allow for a faster implementation of essential regulations, ensuring that AI technologies in space are used responsibly and safely while broader international agreements are being negotiated.

Moreover, these national requirements could serve as a model for future international standards, providing a foundation upon which UNCOPUOS Working Group can build. This approach would not only expedite the regulation process but also promote consistency and cooperation among space-faring nations.

Conclusions

The development of a Code of Conduct and the establishment of a Working Group on AI in space are critical steps toward ensuring the responsible and ethical use of artificial intelligence in space exploration and related activities. As AI becomes increasingly integrated into space missions, from autonomous spacecraft and robotic systems to data analysis and mission planning, there is a growing need for comprehensive guidelines to govern its use.

A Code of Conduct would provide a framework for safe, transparent, and ethical AI applications in space. It would address key concerns such as safety, accountability, data privacy, and the prevention of harmful interference in space operations. This is particularly important given the complexity and potential risks of AI-driven autonomous systems in space environments, where malfunctions or unintended behavior could have serious consequences.

The establishment of a UNCOPUOS Working Group on AI in space would enable continuous collaboration among international stakeholders. Such a group would foster the exchange of knowledge, monitor AI advancements, and ensure compliance with ethical standards, promoting international cooperation in space exploration.

Those elements could also be introduced in a future UNGA Resolution. These initiatives would help mitigate risks, ensure ethical development, and lay the groundwork for a future where AI significantly contributes to humanity's space exploration.

PART V: VOIDS AND RECOMMENDATIONS

Section 1: Recommendations to Lead Standardization to Soft Institutional Laws

Introduction

Given the geopolitical complexities and competing economic interests among nations, formal, binding treaties for outer space activities have been difficult to achieve since the late 1970s. In response, states have turned to the development of "soft law"—non-binding principles and guidelines that reflect a growing consensus on responsible practices without the weight of enforceable legal obligations. Over time, these soft law instruments can evolve into widely accepted norms, influencing both national regulations and customary international law.

This section explores how soft law and international technical standards are playing a pivotal role in shaping the regulation of AI in outer space. Soft law frameworks, such as the Space Debris Mitigation Guidelines and the Guidelines for the Long-Term Sustainability of Outer Space Activities, have already demonstrated their effectiveness in fostering international cooperation and guiding state behavior.

Though not legally binding, they offer a flexible and pragmatic approach to navigating the complex international environment, providing a foundation for responsible AI use in space until binding agreements can be established.

International standards-setting organizations, including the International Organization for Standardization (ISO), the International Telecommunication Union (ITU), and the International Electrotechnical Commission (IEC), contribute to this effort by creating technical guidelines that shape AI development, deployment, and operational practices. These technical standards are critical for ensuring interoperability, safety, and transparency in the use of AI systems in space.

While not legally binding, they serve as an essential component of the broader governance framework, particularly when there are no formal treaties to regulate AI in this domain. The combination of soft law and technical standards provides a crucial framework for addressing immediate regulatory needs.

By establishing common guidelines and expectations, soft law mechanisms help mitigate risks and foster trust among states and private actors. Similarly, technical standards ensure that AI systems are designed and deployed in a manner that promotes safety, reliability, and accountability.

Together, these mechanisms offer a flexible yet robust approach to regulating AI in space, enabling international cooperation while laying the groundwork for future, binding regulations. This section will explore these frameworks in detail, focusing on their role in shaping the future of AI governance in outer space.

1. *Soft Law in International Space Regulation*

In the context of competing economic interests and geopolitical tensions, international lawmaking is facing significant challenges. With regard to outer space, since the late 1970s, it has been impossible to agree on a new binding treaty. States are addressing this situation by working on so-called soft law, meaning principles and guidance that express a certain consensus, without, however, having a binding legal character.

Thus, the General Assembly of the United Nations adopted several declarations and principles in the 1980s and 1990s. These include for example Space Debris Mitigation Guidelines, a Safety Framework for Nuclear Power Source Applications, and most recently, Guidelines for Long-Term Sustainability of Outer Space Activities.

While non-binding, these guidelines can over time acquire a quasi-mandatory character. Soft law has been defined as a continuum, or spectrum, running between fully binding treaties and fully political positions,[736] and as those non-binding rules or instruments that interpret or inform our understanding of binding legal rules or represent promises that in turn create expectations about future conduct.[737]

When soft law works well, it produces regularity of practice and reliability, a common understanding of the concepts and a mutual acceptance leading to trust.[738] Over time, they can even become binding on States as customary international law.

Soft law may evidence the formation of customary international law, guide the interpretation of treaties, authorize action by international organizations, and give rise to duties of good faith such as a duty to consider.[739] According to Article 38(1)(2) of the International Court of Justice, customary international law consists of "general practice accepted as law." State practice is not sufficient, to acquire a binding character, States need to comply with soft law with the understanding that such compliance is required as law.

Despite being officially non-binding, the 2007 Space Debris Mitigation Guidelines have for example been particularly successful. The United Nations General Assembly Resolution 62/217 recognized they reflected then current State practice, and other States have since implemented them into their national regulations as international standards.

While it will not be until States consider them binding that they will become so as customary international law, this example shows that soft law over time can play an important role in shaping international behavior. In addition, even if not yet officially recognized as customary international law, soft law, which is commonly accepted as reflective of international standards, could be taken into account by courts and tribunals having to determine a State's due diligence efforts.

The Liability Convention stipulates in Article III that "[i[n the event of damage being caused elsewhere than on the surface of the earth to a space object of one launching State or to persons or property on board such a space object by a space object of another launching State, the latter shall be liable only if the damage is due to its fault or the fault of persons for whom it is

---

[736] Andrew T. Guzman and Timothy L. Meyer, "International Soft Law," *Oxford Journal of Legal Analysis*, Spring, 2010: Volume 2, Number 1, pp. 171-225 at 173.

[737] *Ibid.* at 174.

[738] Francis Lyall and Paul B. Larsen, *Space Law - A Treatise* (Routledge, 2018) at 47.

[739] Matthias Goldmann, "We Need to Cut Off the Head of the King: Past, Present, and Future Approaches to International Soft Law," *Leiden Journal of International Law* (2012), 25, pp. 335-368 at 336.

responsible." The Liability Convention does not determine how "fault" would be established. A court or arbitral tribunal facing this question might thus well look at soft law, which could have informed State behavior and if followed could have avoided the damage caused.

With respect to the regulation of AI in space, the elaboration of soft law on the use of AI, for example in the form of non-binding guidelines, which would reflect a certain consensus or standard practice, would help State and private actors in determining which uses are accepted and which should be avoided or addressed with additional caution. Until the political climate is ripe for the enactment of binding regulation for AI use in space, soft law could provide a useful alternative.

2. *The Role of International Standards in AI and Outer Space Applications*

International technical standards hold a somewhat ambiguous position under international law, and "can be considered as part of international law [...] lay[ing] somewhere between hard law and soft law."[740] Importantly, technical standards and their definitions offer valuable resources for building common technical understandings of products, processes, and practices relevant to both private and public sectors.

Many international standards setting organizations (SSOs) develop and publish international technical standards, covering everything from agricultural practices and food production to computers, network routers, and communications protocols. International standardization bodies such as the ISO, the ITU, and the IEC develop technical standards and guidelines, some of which relate to spacecraft and AI.

Technical standards serve to establish uniform norms or requirements directed to subjects including corporate management practices, such as procurement practices and International Financial Reporting Standards, manufacturing processes, production methods, use of products, product and equipment engineering and technical specifications, such as operational processes, functional criteria, performance characteristics, and various other rules, guidelines, conditions, methods, practices, and procedures.[741]

These technical standards can cover various aspects of spacecraft and AI (e.g., avionics, life support, algorithmic transparency, interoperability, safety and ethical criteria, etc.), and they provide a common framework for organizations and stakeholders to adhere to when developing, deploying, and using AI systems.

---

[740] Andrea Barrios Villarreal, "International Standards as Part of International Law," in *Cambridge University Press eBooks*, 2018, 58–77, https://doi.org/10.1017/9781108591348.003.

[741] Villarreal, "International Standards as Part of International Law;" Mark Sharron, "Technical Standard," ISMS.online, March 8, 2024, https://www.isms.online/glossary/technical-standard/; National Research Council, *Standards, Conformity Assessment, and Trade*, *National Academies Press eBooks*, 1995, https://doi.org/10.17226/4921; Jennifer Fong, "Why Technical Standards Are Essential in Product Development," IEEE Innovation at Work, March 29, 2018, https://innovationatwork.ieee.org/why-technical-standards-are-essential-in-product-development/; "Engineering Standards," Georgia Southern University, October 25, 2021, https://www.georgiasouthern.edu/cec/ece/engineering-standards.

US OMB Revised Circular A-119 defines standards and technical standards in accordance with the following:

> "Revised OMB Circular A-119 establishes policies on Federal use and development of voluntary consensus standards and on conformity assessment activities. Pub. L. 104-113, the 'National Technology Transfer and Advancement Act of 1995,' codified existing policies in A-119, established reporting requirements, and authorized the National Institute of Standards and Technology to coordinate conformity assessment activities of the agencies. [...]
>
> 3. What Is A Standard?
>
> a. The term 'standard,' or 'technical standard' as cited in the Act, includes all of the following:
>
> (1) Common and repeated use of rules, conditions, guidelines or characteristics for products or related processes and production methods, and related management systems practices.
>
> (2) The definition of terms; classification of components; delineation of procedures; specification of dimensions, materials, performance, designs, or operations; measurement of quality and quantity in describing materials, processes, products, systems, services, or practices; test methods and sampling procedures; or descriptions of fit and measurements of size or strength.
>
> b. The term 'standard' does not include the following:
>
> (1) Professional standards of personal conduct.
>
> (2) Institutional codes of ethics.
>
> c. 'Performance standard' is a standard as defined above that states requirements in terms of required results with criteria for verifying compliance but without stating the methods for achieving required results. A performance standard may define the functional requirements for the item, operational requirements, and/or interface and interchangeability characteristics. A performance standard may be viewed in juxtaposition to a prescriptive standard which may specify design requirements, such as materials to be used, how a requirement is to be achieved, or how an item is to be fabricated or constructed."[742]

Technical standards may be developed independently or collaboratively, for example individually or collectively by companies, trade unions, trade associations, regulatory bodies, government agencies, and militaries.[743] SSOs frequently take into account more collaborative and diverse input and typically result in the establishment of voluntary standards, which may become

---

[742] "CIRCULAR NO. A-119 Revised," The White House, n.d., https://obamawhitehouse.archives.gov/omb/circulars_a119/#3.

[743] Sharron, "Technical Standard;" National Research Council, *Standards, Conformity Assessment, and Trade*; "Setting the Standards: Strengthening U.S. Leadership in Technical Standards," NIST, March 17, 2022, https://www.nist.gov/speech-testimony/setting-standards-strengthening-us-leadership-technical-standards.

mandatory if adopted by a government, e.g., via legislation, or under a business contract, by mutual agreement of the parties, etc.

And, unless required for some specific group of participants, e.g., for gear, operations, or contractors under a military standard, creation of a technical standard by an SSO usually does not prohibit or prevent the adoption and use of other existing standards, including those established by different SSOs.

Three of the oldest and largest SSOs are the ITU, the IEC, and the ISO, all based in Geneva, Switzerland and respectively founded in 1865, 1906, and 1947.[744] Together these three SSOs form the World Standards Cooperation (WSC) alliance and have technical standards addressing nearly every conceivable topic and numbering in the tens of thousands.[745]

2.1 *The ISO, IEC, and ITU Standards Setting Organizations*

ISO and IEC are private international organizations formed by constituent members and not created by any international treaty.[746] Their members comprise national standards bodies (NSBs) or committees, one per country or member economy.[747] Like IEC, "ISO is [also] a multinational forum that enables the development and publication of international standards through its members by bringing together experts to share knowledge and develop voluntary, consensus-based, market relevant international standards."[748]

Conversely, the ITU is a treaty-based organization and a permanent agency of the United Nations, in which governments are the primary members.[749] Further, companies and other non-governmental organizations can also participate in ITU membership, which includes 193 Member States and more than 1,000 companies, universities, research institutes and international organizations.[750]

The structure of the ITU is comprised, among others, of three sectors of activity: (1) the Radiocommunication Sector, including world and regional radiocommunication conferences, radiocommunication assemblies and the Radio Regulations Board; (2) the Telecommunication

---

[744] "About ISO," ISO, n.d., https://www.iso.org/about; "About International Telecommunications Union (ITU)," ITU, n.d., https://www.itu.int/en/about; "History," IEC, n.d., https://www.iec.ch/history; "About Us - What Does IEC Do?," IEC, n.d., https://www.iec.ch/about-us.

[745] "World Standards Cooperation," WSC, n.d., https://www.worldstandardscooperation.org/; Ulrich Harmes-Liedtke, "Data on International Standards," Quality Infrastructure for Development, January 31, 2022, https://qi4d.org/2022/01/24/data-on-international-standards/.

[746] *Using and Referencing ISO and IEC Standards to Support Public Policy*, International Organization for Standardization and International Electrotechnical Commission, 2015, https://www.iso.org/files/live/sites/isoorg/files/store/en/PUB100358.pdf; "National Standards Bodies" NIST, August 25, 2016, https://www.nist.gov/iaao/national-standards-bodies.

[747] *Ibid.*; "Members," ISO, n.d., https://www.iso.org/about/members; "National Committees," IEC, n.d., https://www.iec.ch/national-committees.

[748] "Space Industry Technical Standards," NOAA, 2024, https://www.space.commerce.gov/space-industry-technical-standards/.

[749] "National Committees," IEC; "About International Telecommunications Union (ITU)," ITU; "National Standards Bodies" NIST.

[750] "Membership," ITU, September 19, 2024, https://www.itu.int/hub/membership/; "About International Telecommunications Union (ITU)," ITU.

Standardization Sector, including world telecommunication standardization assemblies; and (c) the Telecommunication Development Sector, including world and regional telecommunication development conferences.[751]

In terms of standardization, as its own name indicates, the Telecommunication Standardization Sector is the body of reference. The ITU Telecommunication Standardization Sector (ITU-T) is a standing body of the ITU responsible for studying and making recommendations on technical, operational and tariff issues for the purpose of standardizing global telecommunications.

In addition, many other independent organizations such as the American Society of Mechanical Engineers (ASME), the American Society for Testing and Materials (ASTM International), the International Air Transport Association (IATA), the International Civil Aviation Organization (ICAO), the Institute of Electrical and Electronics Engineers (IEEE), the Internet Engineering Task Force (IETF), the International Maritime Organization (IMO), SAE International (formerly the Society of Automotive Engineers), and the World Wide Web Consortium (W3C) develop and publish a vast number of technical standards.[752]

As an example of collaborative technical standards establishment, SAE International formed the Airlines Electronic Engineering Committee (AEEC) in 1949 to develop engineering standards for avionics systems in cooperation with other aviation organizations, including ICAO.[753]

Membership in these international standards organizations is often open to parties interested in joining and agreeing to the respective organizational by-laws. In many cases, these organizations include a mix of members such as companies, NGOs, and individual technicians and experts. SSOs have historically established technical standards related to outer space technologies and activities. More recently, SSOs have started publishing technical standards for AI hardware and software systems, and processes, practices, or actions implemented by those systems.

Although, as of the drafting of this manuscript, the authors have found no technical standards specifically related to applications of AI in outer space, the separate respective technical standards established for space and AI systems and practices can be invaluable to building a full

---

[751] See Article 7, *ITU Constitution 1992*.

[752] “About ASME Standards and Certification,” ASME, n.d., https://www.asme.org/codes-standards/about-standards; “ASTM International - Standards Worldwide,”ASTM International, n.d., https://www.astm.org/; “Manuals, Standards & Regulations,” IATA, n.d., https://www.iata.org/en/publications/manuals/; “SARPs - Standards and Recommended Practices,” ICAO, n.d., https://www.icao.int/safety/safetymanagement/pages/sarps.aspx; IEEE SA, “Home - IEEE Standards Association,” IEEE Standards Association, September 10, 2024, https://standards.ieee.org/; “Introduction to the IETF,” IETF, n.d., https://www.ietf.org/about/introduction/; “Introduction to IMO,” IMO, n.d., https://www.imo.org/en/About/Pages/Default.aspx; “SAE Standards for Mobility Knowledge and Solutions,” SAE, n.d., https://www.sae.org/standards; “Web Standards,” W3C, n.d., https://www.w3.org/standards/.

[753] “ARINC Industry Activities,” SAE ITC, n.d., https://aviation-ia.sae-itc.com/activities/aeec; “ARINC Industry Activities,” SAE ITC, n.d., https://aviation-ia.sae-itc.com/news-articles/airlines-electronic-engineering-committee-aeec-arinc-ia-sets-standards-avionics-aircraft-worldwide.

understanding and appreciation of at least some of the potential regulations and implications for designing, deploying, and maintaining space-based AI solutions.

For instance, consider a scenario in which an AI subsystem controlling the navigation of a satellite malfunctions, leading to a collision with another satellite. In such a case, technical standard guidelines on AI safety and reliability, such as those proposed in international forums or adopted by leading spacefaring nations, could be crucial in determining fault.

Another hypothetical scenario could involve an AI-driven satellite tasked with space debris mitigation inadvertently causing damage to another AI-driven spacecraft. Technical standard guidelines on ethical use and transparent, accountable decision-making for AI systems employed in space operations, including space debris management, could influence tribunal judgment on whether the satellite operator acted with sufficient due diligence.

2.2 *ISO Standards for AI*

There are several examples of specific standards that can impact space-related AI technologies and illustrate the practical application of these standards in the space sector. Among others, these include ISO/IEC 22989:2022 on AI concepts and terminology, which standardizes the language and verbiage used to describe and define AI systems, e.g., for satellite sensing, communications, and processing. Moreover, ISO/IEC 23053:2022 on AI systems using machine learning may apply to machine learning models used in analyzing space data.

ISO/IEC 22989:2022, on AI concepts and terminology was published on July 19, 2022, and "establishes terminology for AI and describes concepts in the field of AI".[754] It applies to "all types of organizations (e.g. commercial enterprises, government agencies, not-for-profit organizations)" and can be used to provide a common framework for communications between stakeholders and in developing new standards.[755]

For example, consistent terminology can improve collaboration between international space agencies and private companies working on AI-driven satellite navigation systems, ensuring that all parties have a clear understanding of the technical terms and concepts being used.

ISO/IEC 23053:2022, on Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML) was published on June 20, 2022, "establishes an Artificial Intelligence (AI) and Machine Learning (ML) framework for describing a generic AI system using ML technology."[756] The framework applies to "all types and sizes of organizations implementing or using AI systems, including public and private companies, government entities, and not-for-profit organizations," and "describes the system components and their functions in the AI ecosystem."[757]

For instance, an AI system designed for autonomous navigation of spacecraft can benefit from a standardized framework, ensuring that its components and functions are well-defined and interoperable with other systems used by international space agencies.

---

[754] "ISO/IEC 22989:2022," ISO, n.d., https://www.iso.org/standard/74296.html.

[755] *Ibid*.

[756] "ISO/IEC 23053:2022," ISO, n.d., https://www.iso.org/standard/74438.html.

[757] *Ibid*.

ISO/IEC 23894:2023, on guidance for AI risk management was published on February 6, 2022, and "provides guidance on how organizations that develop, produce, deploy or use products, systems and services that utilize artificial intelligence (AI) can manage risk specifically related to AI. The guidance also aims to assist organizations to integrate risk management into their AI-related activities and functions."[758] It "describes processes for the effective implementation and integration of AI risk management" that "can be customized to any organization and its context."[759]

Using this guidance, risk management standards could be applied to AI technologies used in space exploration. For example, these guidelines can help manage risks associated with deploying AI for use in satellite collision avoidance systems by establishing protocols for detecting potential collisions and automatically adjusting satellite trajectories.

Additionally, for automated data analysis from space missions, these guidelines could ensure that the AI systems are regularly tested and validated to identify and mitigate any biases or errors that could impact mission outcomes. Specific risk management practices such as continuous monitoring, incident response plans, and stakeholder communication strategies can help ensure the safe and reliable operation of AI technologies in the space environment.

ISO/IEC 42001:2023, on AI management systems was published on December 18, 2023, "specifies requirements for establishing, implementing, maintaining and continually improving an [AI] management system within organizations."[760] It is intended for use by any organization, regardless of size, type and nature, that provides or uses AI-based products or services, to help them develop, provide or use those AI systems responsibly in pursuing their objectives and meet applicable requirements and obligations.[761]

Examples of potential implementation of AI management systems in space missions may include use by a space agency to manage AI applications for satellite operations by establishing a structured framework that includes protocols for regular system checks, data validation, and anomaly detection. This would help ensure all AI systems onboard satellites adhere to safety protocols and regulatory requirements.

Furthermore, this standard could potentially be used to guide the management of AI-driven spacecraft navigation systems, ensuring they operate reliably and can be audited for compliance with mission-specific safety standards. This standard can help maintain high levels of accountability and efficiency in the operation of AI systems in outer space.

For example, an AI management system could include detailed documentation and logging practices to track decision-making processes and system performance, which is crucial for post-mission analysis and continuous improvement of AI technologies used in space.

---

[758] "ISO/IEC 23894:2023," ISO, n.d., https://www.iso.org/standard/77304.html.

[759] *Ibid.*

[760] "ISO/IEC 42001:2023," ISO, n.d., https://www.iso.org/standard/81230.html.

[761] *Ibid.*

Technical standard ISO/IEC TR 5469:2024, on functional safety and AI systems was published January 8, 2024, and "describes the properties, related risk factors, available methods and processes relating to:

- use of AI inside a safety related function to realize the functionality;
- use of non-AI safety related functions to ensure safety for an AI controlled equipment;
- use of AI systems to design and develop safety related functions."[762]

This standard outlines best practices for the use of AI in design, development, deployment, and maintenance of AI-enabled safety functions, including the requirement for non-AI redundancies in the provision of such safety functions.[763] Given that nearly every aspect of space exploration must inherently consider safety issues, including potentially conflicting safety concerns, this standard promises to provide a great resource for design and use of AI systems to provide safety related functions in space.

If the AI system in question failed to meet these standards, a court could reference this guideline to assess whether due diligence was exercised. For instance, the court might examine whether the developers conducted thorough testing to ensure reliability of the AI-enabled safety function under various space conditions, or whether adequate fail-safes were implemented to prevent malfunctions.

2.3 *ITU Standards for Satellite and Space-Based Telecommunications*

ITU has established several standards applicable to telecommunications networks and systems that also implicate satellites and space-based telecoms operations.[764] Because AI and Machine Learning have become integral to solutions and applications across industry sectors, ITU has cultivated a global dialogue on the implications of AI for the future of society, anchored by the AI for Good Global Summit.[765]

This AI for Good Global Summit has helped establish motivations underpinning recent ITU Focus Groups, which accelerate studies in fields of growing strategic relevance to ITU membership.[766] For example, ITU Focus Groups related to AI and machine learning include Focus Groups on Machine Learning for Future Networks including 5G, AI for autonomous and assisted driving, and Environmental Efficiency for AI and other Emerging Technologies.[767] Open to all interested parties, ITU Focus Groups prepare a basis for related standardization work in ITU-T Study Groups.[768]

> "ICT companies in the networking business are introducing AI and Machine Learning as part of their innovations to optimize network operations and increase

---

[762] "ISO/IEC TR 5469:2024," ISO, n.d., https://www.iso.org/standard/81283.html.

[763] *Ibid.*

[764] "International Standards for an AI Enabled Future," AI For Good, September 2, 2021, https://aiforgood.itu.int/international-standards-for-an-ai-enabled-future/.

[765] *Ibid.*

[766] *Ibid.*

[767] *Ibid.*

[768] *Ibid.*

> energy and cost efficiency. New ITU standards provide an architectural framework for the integration of machine learning into 5G and future networks (ITU Y.3172), a framework to evaluate intelligence levels across different parts of the network (ITU Y.3173), and a framework for data handling in support of machine learning (ITU Y.3174)."[769]

The transition to 5G is a transformative shift in telecommunications and promises unprecedented speeds, ultra-low latency, and the capability to support millions, even billions, of connected devices. However, this increase in complex networks gives rise to significant challenges. Traditional telecommunication management methods, which relied on manual interventions and predefined rules, are no longer entirely sufficient. ML offers an attractive solution by enhancing network operation and optimization.

By way of example, the ITU-T Recommendation Y.3172 provides a good framework for integrating ML into new networks by outlining a framework for achieving greater automation, efficiency, and performance. Network management originally used to be quite straightforward. Operators relied on predefined rules and manual adjustments to maintain network stability.

However, with the rapid increase in connected devices, the rise of the Internet of Things (IoT), and the ever-increasing demand for more data-intensive applications, such as virtual reality applications and autonomous vehicles, networks have become much more complex. Nowadays, networks must dynamically adapt to fluctuating conditions, anticipate user demands, and respond to potential threats in real-time within milliseconds.

Old management techniques are struggling to keep up with these demands, which results in network inefficiencies and vulnerabilities. Machine learning is capable of addressing these problems by enabling networks to learn from data, predict future trends, and make intelligent decisions autonomously. Consequently, network operations can be significantly enhanced.

The Y.3172 establishes a framework, which is designed to address the unique challenges associated with the integration of ML into modern networks. It adopts a modular approach, allowing ML components to be added, removed, or updated without disrupting the network. At the core of the framework established by Y.3172 is the so-called "ML pipeline". This ML pipeline covers the entire ML process from data collection to feedback.

The ML pipeline begins with data collection. It gathers data from various network sources such as user devices, sensors or network nodes. This data then forms the foundation for training the ML models. This is followed by data preprocessing, which involves cleaning and transforming raw data into something the ML algorithm can use. Afterwards, model training uses the preprocessed data to create ML models, which can be trained through techniques like supervised, unsupervised, or reinforcement learning.

Once trained, the ML models are used in the network to make real-time decisions, such as predicting traffic congestion, detecting anomalies, or optimizing resource allocation. A key feature of the ML pipeline is the feedback loop, where the outcomes of the inference process are analyzed

---

[769] *Ibid*.

to further refine the models. This continuous learning approach ensures that the models remain accurate and relevant over time.

The framework also includes mechanisms to manage data sources and repositories. Effective ML is dependent on access to high-quality data. Thus, data is collected from various sources and stored in secure, scalable repositories designed to handle the volume and variety of data encountered in modern networks. Additionally, the ML model repository serves as a centralized storage location for all models used within the network.

The inference engine executes ML models in real-time and processes incoming data streams to generate actionable insights. These insights can then be utilized to adjust network parameters, allocate resources, trigger alarms and detect anomalies. This inference engine must operate efficiently to meet the stringent latency requirements of 5G networks. The orchestrator, the brain of the ML architecture, manages the entire lifecycle of ML models and it ensures that models are consistently aligned with the network's operational goals.

Integrating into and adding ML to networks requires data management and model training and deployment. The former is concerned with collecting, preprocessing, and storing data. This data must be accurate, complete and available. The latter deals with the training and deployment, which require significant computational resources, especially for large datasets. Once a model is trained, it must be seamlessly integrated into the network. Security and privacy are essential in this process and Y.3172 recommends measures to protect models.

Deploying ML in 5G networks presents several challenges that must be managed carefully. Scalability is a key concern due to the massive amounts of data generated by 5G networks. The architecture must be capable of processing and storing large volumes of data while maintaining performance. Latency is but another critical factor, particularly for applications such as autonomous driving and real-time gaming.

The architecture must minimize delays in data processing and model inference to ensure timely, and most importantly safe decision-making. Finally, model lifecycle management is also crucial because ML models require continual updates and refinements to remain effective and safe. The architecture recommended by the ITU supports the full lifecycle of ML models.

The potential applications of ML in 5G networks are vast. For example, ML can be used for orbital traffic prediction and management. It can analyze historical as well as real-time data to predict future orbital congestion or potential collision hazards. This allows satellite operators to manage trajectories efficiently and helps them avoid collision.

In terms of security, ML excels in anomaly detection and can identify unusual patterns that may indicate threats or faults. Early detection allows for quick responses and prevents potential disruptions.

ML continues to advance and its applications in network management will evolve. Future developments, such as deep learning and reinforcement learning, could lead to even more sophisticated models capable of handling increasingly complex network environments. However, these advancements will need even more powerful computational resources and addressing ethical considerations related to AI decision-making.

The integration of ML into 5G networks represents a significant leap forward in telecommunications. The ITU-T Y.3172 framework provides a detailed guide for deploying ML in these networks. By leveraging ML to enhance both automation and performance, the future of telecommunications will further transform, and ML will play a central role in shaping the next generation of network management.

The ITU-T Y.3172, ITU-T Y.3173, and ITU-T Y.3174 "'Machine Learning for 5G' standards are also guiding contributions to a new ITU Global Challenge on AI and Machine Learning in 5G."[770]

> "AI and Machine Learning are widely used in developing models to assess the quality of speech, audio and video, for example in ITU standards for the quality assessment of audiovisual streaming, in particular ITU P.1203 (progressive-download and adaptive-bitrate AV) and ITU P.1204 (video streaming services up to 4K). New ITU standards address intelligent network analytics and diagnostics (ITU E.475) and the creation and performance testing of Machine Learning-based models to assess the impact of the transmission network on speech quality for 4G voice services (ITU P.565). Other notable new ITU standards relevant to AI and Machine Learning address environmental sustainability [...] and operational aspects of service provision and telecom management."[771]
>
> "AI is one of the five characteristics of a new ITU framework to support smart service operation, network management and infrastructure maintenance (ITU M.3041). New ITU standards under development in this domain will address AI-enhanced telecom operation and management, energy saving for 5G Radio Access Networks with AI, and robot-based smart patrols of telecoms networks."[772]

These ITU standards and efforts related to 5G, 6G, and future networks are directly relevant to satellites, spacecraft and other space-based operations, with the U.S. and other militaries currently integrating 5G into their satellite-based networks and commercial deployment expected soon thereafter.

2.4 *International Standards for Outer Space*

International standards play a critical role in the development and operation of space systems, ensuring that equipment and technologies used in space are safe, reliable, and interoperable. These standards provide detailed specifications and guidelines for the design, operation, and maintenance of space systems, fostering global collaboration and compatibility across space missions. This section explores the organizations and frameworks responsible for developing these standards, including ISO, CCSDS, and government agencies, as well as their relevance to emerging technologies like AI and machine learning in space activities.

---

[770] "International Standards for an AI Enabled Future," AI For Good.

[771] *Ibid.*

[772] *Ibid.*

> "Space standards are guidelines, best practices, and recommendations that describe the specifications, dimensions, and requirements for designing and operating equipment, and systems in space. These help ensure the safety, reliability, and compatibility of space missions and activities within and across organizations, as well as facilitate international cooperation and coordination.
>
> Space standards are vital for the design, operation, and success of space missions and activities. They provide a common language and framework for space agencies and organizations across the globe, enabling them to work together and achieve their goals. By following standards, space professionals can ensure the safety, reliability, and compatibility of space systems in a safe and sustainable manner.
>
> Space standards are often developed via consensus building on what the standard addresses, most often through international standards developing organizations, such as ISO, the Consultative Committee for Space Data Systems (CCSDS), or ASTM International. Standards from these organizations are voluntary consensus standards and are usually appropriate or adaptable for the government's purposes. The CCSDS is a multi-national organization of international space agencies and develops open communications and data standards for space systems. CCSDS has multiple working groups developing and publishing standards. The Navigation Working Group family of space data messages are most applicable to be used by space launch operators, spacecraft operators, SSA (space situational awareness) service data providers, analysts, and message exchange partners and are freely accessible at the CCSDS website.
>
> [... ISO] Technical Committee (TC) 20, Aircraft and Space Vehicles has subcommittees (SC) focused on the standardization of materials, components and equipment for construction and operation of aircraft and space vehicles as well as equipment used in the servicing and maintenance of these vehicles. TC 20 has to-date published 682 ISO standards with 17 participating member countries and 28 observing member countries. There are two subcommittees supporting space data messaging and SSA: SC 13 Space data and information transfer systems and SC 14 Space systems and operations. Products from SC 13 are identical to products from the CCSDS via formal arrangements between ISO and CCSDS.
>
> In addition, government agencies, such as NASA, National Institute of Standards and Technology (NIST), and European Space Agency (ESA), and professional and trade organizations, such as SAE International, develop standards specific to technical disciplines as required for their needs and trade, usually through consensus building processes with their stakeholders."[773]

Space Standards identified by the NOAA Office of Space Commerce as relevant to AI and machine learning include:

---

[773] *Ibid.*

- JA7496: 202206 Cyber Physical Systems Security Engineering Plan (SAE Technical Committee G-32 Artificial Intelligence in Aviation), published in June, 2022,
- JA6678: Guidelines for Establishing and Maintaining Cyber-Physical-Systems' Cyber-Resilience (SAE Technical Committee G-32 Artificial Intelligence in Aviation), currently in development, and
- ARP6983: Process Standard for Qualification of Aeronautical Systems Implementing AI: Development Standard (SAE Technical Committee G-34 Artificial Intelligence in Aviation), also in development.[774]

> "Small Spacecraft Avionics (SSA) consist of all the electronic subsystems, components, instruments, and functional elements of the spacecraft platform, including the primary flight sub-elements Command and Data Handling (CDH) and Flight Software (FSW), as well as other critical flight subsystems such as Payload and Subsystems Avionics (PSA). All [these components] must be configurable into specific mission platforms, architectures, and protocols, and be governed by appropriate operations concepts, development environments, standards, and tools.[775] The CDH and FSW are the brain and nervous system of the integrated avionics system, and generally provide command, control, communication, and data management interfaces with all other subsystems in some manner, whether in a direct point-to-point, distributed, integrated, or hybrid computing mode. The avionics system is essentially the foundation for all components integrated on the spacecraft and their functions. As the nature of the mission influences the avionics architecture design, there is a large degree of variability in avionics systems."[776]

Of further relevance, the International Avionics System Interoperability Standards (IASIS) is a collaborative outcome of the International Space Station (ISS) partners efforts to establish standards for interoperable interfaces and corresponding terminology to facilitate collaborative space cislunar and deep space exploration.[777]

> "NASA will maintain the IASIS under Human Exploration and Operations Mission Directorate (HEOMD)," and Configuration Management is under the responsibility and control of the Multilateral Coordination Board (MCB), which must approve any revisions to the document.[778]

The IASIS standards are available for international and commercial partnerships, "to enable academia, industry, government agencies and international entities to independently

---

[774] "Space Industry Technical Standards," NOAA.

[775] "8.0 Small Spacecraft Avionics," NASA, n.d., https://www.nasa.gov/smallsat-institute/sst-soa/small-spacecraft-avionics/.

[776] "8.0 Small Spacecraft Avionics," NASA.

[777] William H. Gerstenmaier et al., "International Avionics System Interoperability Standards (IASIS) Baseline," February 23, 2024, https://internationaldeepspacestandards.com/wp-content/uploads/2024/02/avionics_baseline_final_3-2019.pdf.

[778] *Ibid.*

develop systems for deep space exploration that would be compatible aboard any spacecraft, irrelevant of the spacecraft developer."[779]

IASIS standards "are not intended to specify system details needed for implementation nor do they dictate design features behind the interface. Specific requirements are defined in unique documents." Rather, IASIS integrates established and internationally recognized standards, selected to enable a larger diversity of providers.

> "Increasing commonality among providers while decreasing unique configurations has the potential to reduce the traditional barriers in space exploration: overall mass and volume required to execute a mission. Standardizing interfaces has the benefit of reducing the scope of [necessary] development efforts."

IASIS presents "a set of parameters, which if accommodated in the system architecture support greater efficiencies, promote cost savings, and increase the probability of mission success."

> "The purpose of the IASIS is to provide a starting framework that allows developers to independently design compatible Avionics systems for human exploration and associated interfaces in cislunar and deep space environments."[780]
>
> "The first application of these standards was to the Gateway element of the Artemis program and is being applied, as appropriate, to other Artemis elements. However, these standards are meant to be applicable to the LEO, cis-lunar, lunar surface, cis-Mars, Martian surface, and environments between the Earth, Moon and Mars."[781]
>
> "This document specifies data link protocols and physical layer standards options to be used to architect the interfaces between spacecraft elements, including docking adapters connecting the elements as well as element services in support of items that are expected to be common among elements, such as Low Profile Grapple Fixtures for Robotics.
>
> Although these standards focus on interoperability between elements as well as common item data interfaces among elements, the standards are not exclusive to any element's network selection or other network standards. It is understood that other data network protocols will be needed and utilized within an element."[782]

International standards are essential for ensuring interoperability, reliability, and efficiency in space exploration. As AI and machine learning transform the field, frameworks like IASIS should be continuously updated to keep pace with technological advancements.

---

[779] "FAQ," International Deep Space Interoperability Standards, n.d., https://internationaldeepspacestandards.com/faq/; Gerstenmaier, "International Avionics System Interoperability Standards (IASIS) Baseline."

[780] Gerstenmaier, "International Avionics System Interoperability Standards (IASIS) Baseline."

[781] "FAQ," International Deep Space Interoperability Standards.

[782] Gerstenmaier, "International Avionics System Interoperability Standards (IASIS) Baseline."

These updates are vital for promoting collaboration, reducing costs, and enabling sustainable missions, ultimately driving global cooperation and the future of space exploration.

Conclusions

The analysis in this section highlights the crucial role of soft law and international standards in addressing the regulatory gaps surrounding the use of AI in space. As the complexities of space exploration and AI technologies outpace the creation of binding international treaties, soft law mechanisms offer a viable and flexible framework for guiding state and private sector behavior.

These non-binding instruments, such as the Space Debris Mitigation Guidelines and the Guidelines for Long-Term Sustainability of Outer Space Activities, have already proven effective in shaping international conduct. They provide a foundation for the standardization of AI technologies, even without the rigidity of formal legal obligations. Over time, such guidelines can evolve into customary international law, especially when adopted by states in practice and through national regulations.

The evolution of soft law into binding customary international law, as discussed, hinges on states' continued adherence to these non-binding frameworks with the understanding that compliance reflects a collective commitment to maintaining order and responsibility in space activities.

The successful implementation of the 2007 Space Debris Mitigation Guidelines serves as a precedent, illustrating how soft law can influence national legislation and, over time, acquire the force of law. In the context of AI in space, similar guidelines could play a pivotal role in defining the responsible use of AI systems, providing a clear distinction between acceptable uses and those that require more stringent safeguards.

Equally important is the role of international standards-setting organizations such as the ISO, IEC, and ITU, whose technical standards provide a common framework for the design, deployment, and regulation of AI technologies. These standards, while technical in nature, are a critical element of the broader governance framework. They offer a shared language and set of practices that foster interoperability, ethical use, and safety across the public and private sectors involved in space activities. For instance, standards like ISO/IEC 22989:2022 on AI concepts and terminology or ISO/IEC 23053:2022 on AI systems using machine learning are instrumental in ensuring that AI systems used in space are both technically sound and ethically designed.

Furthermore, the establishment of risk management standards such as ISO/IEC 23894:2023 and AI management systems like ISO/IEC 42001:2023 ensures that AI-driven systems in space operations are subject to rigorous testing, continuous monitoring, and compliance with safety protocols. These standards help mitigate risks associated with AI malfunction or unintended consequences, such as satellite collisions or data inaccuracies, thereby fostering greater accountability.

In conclusion, this section demonstrates that soft law and international technical standards are integral to the standardization of AI in space. Together, they form a flexible yet robust regulatory framework that not only addresses immediate governance needs but also lays the groundwork for future, binding legal agreements.

As states and private actors continue to adopt and adhere to these frameworks, the potential for these norms to evolve into customary international law increases, thereby providing a more structured and enforceable governance regime for AI in space.

The strategic use of soft law and technical standards is not only a pragmatic approach in the face of current geopolitical challenges but also a necessary one to ensure the responsible, ethical, and sustainable development of AI technologies in space.

Section 2: Frameworks from Air and Maritime Law for Standardizing AI in Space

Introduction

In the realm of space operations, the integration of AI necessitates a robust framework of standardization and regulation to ensure its responsible and ethical deployment. This section provides an in-depth analysis of the voids in current space regulations and presents comprehensive recommendations for leading the standardization of AI in space.

By emphasizing the adoption and implementation of international initiatives and guidelines, alignment with universal human rights and civil rights conventions, and the establishment of norms and best practices, this section draws on the established frameworks in air and maritime regulations to inform the development of AI regulations in space.

First, an overview of international regulations highlights the roles of the International Civil Aviation Organization (ICAO) and the International Maritime Organization (IMO) in setting standards for safety, security, and efficiency in aviation and maritime operations. The principles and guidelines from ICAO and IMO offer valuable insights for creating international laws for AI use in space, particularly in the areas of safety management, operational security, navigation, and communication.

Next, the section delves into specific recommendations for enhancing the safety and operational security of AI-driven systems in space. This includes implementing safety management systems tailored for AI components, enhancing redundancy and failover mechanisms, enforcing regular software updates and maintenance, defining standards for AI hardware resilience, deploying comprehensive security measures, and developing detailed security plans and access controls.

Furthermore, the discussion extends to navigation and communications, proposing the adoption of standardized communication protocols, interoperability standards, and advanced surveillance systems for effective Space Traffic Management (STM). The development of a comprehensive regulatory framework encompassing communication protocols, interoperability requirements, surveillance mandates, compliance and certification processes, and sophisticated AI algorithms for collision avoidance is essential for the safe and efficient operation of AI-driven spacecraft.

In addressing liability and insurance, the section draws on frameworks such as the ICAO's Montreal Convention and the IMO's HNS Convention. Recommendations include establishing clear liability tiers, mandatory insurance requirements, a broad scope of liability, and addressing ownership and control changes. Including AI in the definition of space objects and developing new legal instruments specifically for AI-related challenges in space are crucial steps in creating a robust regulatory environment.

We propose to expressly clarify that AI systems should be regarded as a component of a “space object”—not to consider AI software as an independent space object. The definition of "product" established in the new EU's Product Liability Directive (PDL),[783] whose latest version includes a novel definition of the concept 'product' in the following terms: *"'product' means all movables, even if integrated into, or inter-connected with, another movable or an immovable; it includes electricity, digital manufacturing files, raw materials and software"* (art. 4(1)). Similarly, our suggestion is that the definitional interpretation of "space object" includes AI software, analogously to the PDL definition. However, as AI technology evolves, particularly in cases where software systems perform autonomous functions beyond the control of the parent space object, this framework may need to be revisited to address emerging challenges and ensure regulatory relevance.

Finally, the significance of international cooperation is emphasized through short-term and long-term recommendations for STM. These recommendations balance immediate actions with a phased approach, ultimately aiming to establish a framework that fosters international collaboration. For the long term, both centralized and decentralized governance models are proposed, providing flexibility while maintaining the overarching goal of global coordination. By aligning innovation with regulatory coherence, these measures seek to ensure the safe, secure, and efficient integration of AI-driven technologies in space.

1. *Leveraging Air and Maritime Standards for Safe AI Operations in Space*

The ICAO, a specialized agency of the United Nations, plays a pivotal role in developing standards and regulations for international civil aviation. Although space operations differ significantly from aviation, the ICAO's guidelines offer valuable insights into creating international laws for AI use in space.

Key areas of consideration include safety and security, where the ICAO sets standards for safety management systems, airworthiness, and aviation security.[784] These principles can be adapted to ensure the safe and secure use of AI systems in space operations.

Additionally, the ICAO's regulations for air traffic management, encompassing airspace management, communication, navigation, and surveillance, provide a foundation for similar principles in STM, regulating the movement and coordination of AI-powered space vehicles.[785]

Similarly, the IMO is responsible for regulating international maritime shipping.[786] While space operations diverge from maritime activities, aspects of IMO regulations can serve as a basis for developing guidelines for AI use in space.

---

[783] Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on liability for defective products and repealing Council Directive 85/374/EEC.

[784] See Lapesa Barrera, David. "The Nature and Role of ICAO." In *Aircraft Maintenance Programs*, Springer Series in Reliability Engineering. Springer, Cham, 2022. https://doi.org/10.1007/978-3-030-90263-6_1.

[785] See McClintock, Bruce, et al. *The Department of the Air Force's Plan to Field Advanced Battle Management System Capabilities: Accelerating Decision Making in Air, Space, and Cyberspace*. Research Report, RAND Corporation, 2023. https://www.rand.org/content/dam/rand/pubs/research_reports/RRA1900/RRA1949-1/RAND_RRA1949-1.pdf.

[786] See Lapesa Barrera, David. "IMO Institutional Structure and Law-Making Process." In *Aircraft Maintenance Programs*, Springer Series in Reliability Engineering. Springer, Cham, 2022. https://doi.org/10.1007/978-3-030-90263-6_1.

In navigation and communication, the IMO establishes standards for safety and collision avoidance, setting guidelines for communication systems and equipment.[787] These principles can inform the development of regulations for AI-driven space vehicles.

Furthermore, the IMO's coordination of global search and rescue operations at sea provides a model for establishing frameworks for AI-enabled space systems, addressing emergency situations, response protocols, and international cooperation.[788]

1.1 *Space Safety and Security: Applying Air and Maritime Standards to AI*

In the realm of international regulations, both the ICAO and the IMO play pivotal roles in establishing frameworks that ensure safety, security, and efficiency in aviation and maritime operations.

ICAO's Annex 6, "Operation of Aircraft," sets forth rigorous standards for aircraft operations, focusing on safety management systems (SMS), equipment standards, and operational procedures aimed at mitigating risks and ensuring the safe conduct of air transport operations.[789] Complementing this, ICAO's Annex 17, "Security," mandates comprehensive security protocols, encompassing threat assessments and contingency planning, to safeguard aviation infrastructure and uphold operational integrity.[790]

Similarly, the IMO's International Ship and Port Facility Security (ISPS) Code provides a globally recognized framework for enhancing security measures across ships and port facilities, emphasizing risk assessments and tailored security plans.

Drawing upon these established frameworks from ICAO and IMO, the following recommendations could be refined for standardization to bolster the safety and security of AI-driven systems in space operations:

- Implementing Safety Management Systems (SMS) for AI components: Implementing SMS tailored for AI systems in space missions, akin to those in aviation under ICAO's Annex 6, and detailed in Doc 9859 - Safety Management Manual, is critical.[791] These systems should integrate real-time diagnostic tools to continuously monitor and evaluate AI algorithms used in navigation, communication, and data processing. This approach ensures proactive management of AI system performance, identifying anomalies promptly to maintain mission safety and reliability.

---

[787] *Ibid.*

[788] *Ibid.*

[789] See International Civil Aviation Organization. *Annex 6 to the Convention on International Civil Aviation: Operation of Aircraft, Part I: International Commercial Air Transport – Aeroplanes.* 11th ed. Montreal: ICAO, 2018. https://applications.icao.int/postalhistory/annex_6_operation_of_aircraft.htm.

[790] See International Civil Aviation Organization. *Annex 17 to the Convention on International Civil Aviation: Security – Safeguarding International Civil Aviation Against Acts of Unlawful Interference.* 11th ed. Montreal: ICAO, 2020. https://applications.icao.int/postalhistory/annex_17_security.htm.

[791] See ICAO, Annex 6, 2018; *See also* International Civil Aviation Organization. *Safety Management Manual (SMM)*. Doc 9859. 4th ed. Montreal: ICAO, 2018.

- Enhancing redundancy and failover mechanisms: Enhancing redundancy and failover mechanisms, inspired by aviation standards, involves integrating dual AI processors for critical systems. This facilitates seamless transitions and maintains operational continuity in the event of primary processor failure. Additional redundancy measures should cover power supplies, communication modules, and data storage to minimize downtime and uphold mission integrity during adverse conditions.
- Enforcing regular software updates and maintenance: Regular software updates and maintenance, akin to aviation practices under ICAO, are imperative for AI systems in space. Establishing a structured regimen ensures timely deployment of security patches, performance optimizations, and feature enhancements for AI algorithms and systems. Automated update processes should be implemented to streamline software management, reducing vulnerability exposure and bolstering cybersecurity resilience in the dynamic space environment.
- Defining standards for AI hardware resilience: Defining robust standards for AI hardware resilience, similar to aviation equipment standards under ICAO, involves specifying radiation-hardened components and employing materials capable of withstanding high levels of radiation and extreme environmental conditions in space. Advanced thermal management solutions, including heat-resistant materials and active cooling systems, could be implemented to maintain optimal operating temperatures for AI processors and sensors.
- Deploying comprehensive security measures for AI systems: Deploying comprehensive security measures, drawing from IMO's ISPS Code principles for maritime security, is essential to safeguard AI systems in space from cyber threats and physical attacks.[792] Implementing end-to-end encryption protocols, such as the Advanced Encryption Standard (AES), ensures secure data transmissions between AI systems and ground control stations.[793] Robust intrusion detection systems (IDS) should be deployed to monitor network traffic and promptly detect and mitigate potential cybersecurity threats.[794] Strict access controls and authentication mechanisms should be enforced to prevent unauthorized access and modifications to critical AI systems.
- Developing detailed security plans and access controls: Developing detailed security plans, inspired by IMO's ISPS Code security plans for maritime operations, tailored to specific threats faced by AI systems in space operations is

---

[792] See International Maritime Organization. *International Ship and Port Facility Security (ISPS) Code*. Adopted December 12, 2002. https://www.imo.org/en/OurWork/Security/Pages/SOLAS-XI-2%20ISPS%20Code.aspx.

[793] See National Institute of Standards and Technology (NIST). *Announcing the Advanced Encryption Standard (AES)*. Federal Information Processing Standards Publication 197, November 26, 2001. https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.197.pdf.

[794] See Khraisat, Ahmad, et al. "A Critical Review of Intrusion Detection Systems in the Internet of Things: Techniques, Deployment Strategy, Validation Strategy, Attacks, Public Datasets and Challenges." *Cybersecurity* (2020). https://cybersecurity.springeropen.com/articles/10.1186/s42400-020-00008-y.

essential.[795] Conducting regular security audits and vulnerability assessments enables proactive identification and mitigation of security risks. Implementing role-based access controls (RBAC) and multi-factor authentication (MFA) ensures that only authorized personnel can access and modify AI configurations and operational parameters, minimizing the risk of unauthorized system access and compromise.[796]

- Establishing operational guidelines for AI systems: Drawing from ICAO's Doc 10037 - Manual on Remotely Piloted Aircraft Systems (RPAS), establishing standardized operational procedures for AI-driven space systems is crucial.[797] This includes robust communication protocols, secure data transmission methods, and predefined operational guidelines to ensure AI systems in space can operate harmoniously, manage data securely, and respond swiftly to changing conditions. These guidelines should also cover emergency procedures and contingency planning to address potential system failures or unexpected anomalies.

By integrating these specific recommendations derived from ICAO's Annex 6 and Annex 17, along with IMO's ISPS Code, space agencies and operators can establish a robust framework for enhancing the safety, security, and operational resilience of AI-driven systems in space operations.

Implementation of these measures could be facilitated through international consultations within existing bodies such as UNCOPUOS or through dedicated working groups that focus on developing and coordinating standards for space safety and AI technologies. These measures are crucial for mitigating risks, maintaining mission integrity, and ensuring the reliability of AI technologies in the challenging and dynamic space environment.

1.2 *AI Navigation and Communication in Space: Drawing from Air and Maritime Protocols*

Navigating and communicating effectively in airspace and maritime environments is governed by robust international frameworks established by organizations such as the ICAO and regulations like Chapter V of the International Convention for the Safety of Life at Sea (SOLAS).

These frameworks ensure safety, efficiency, and reliability through standardized communication, navigation, and surveillance systems. ICAO's Annex 10, "Aeronautical Telecommunications," and SOLAS Chapter V set forth comprehensive guidelines for air and maritime traffic management, respectively.

ICAO's Annex 10 addresses communication, navigation, and surveillance systems in air traffic management, emphasizing data exchange, interoperability, and communication protocols.

---

[795] See International Maritime Organization, ISPS Code, 2002.

[796] See National Institute of Standards and Technology (NIST). *Guide to Enterprise Telework, Remote Access, and Bring Your Own Device (BYOD) Security.* Special Publication 800-46, Revision 2, July 2016. https://doi.org/10.6028/NIST.SP.800-46r2.

[797] See International Civil Aviation Organization. *Manual on Remotely Piloted Aircraft Systems (RPAS)*. Doc 10037. Montreal: ICAO, 2015. https://skybrary.aero/sites/default/files/bookshelf/4053.pdf

Translating these principles to STM can guide the development of standards for AI-driven spacecraft.[798]

This includes specifying communication protocols for telemetry data exchange, ensuring interoperability between AI systems on different spacecraft, and establishing real-time surveillance systems to monitor AI-driven spacecraft positions and statuses.[799]

Similarly, SOLAS Chapter V provides a framework for maritime safety through navigation systems, equipment standards, and procedural guidelines.[800] Applying these principles to space operations involves leveraging advanced AI technologies to manage orbital traffic effectively and enhance safety amidst the complexities of space environments:

- Standardized communication protocols and data exchange: Implementing Consultative Committee for Space Data Systems (CCSDS)-recommended communication protocols for telemetry data exchange between AI-driven spacecraft and ground stations could support secure, reliable, and real-time data transmission to facilitate coordinated maneuvers and operational planning across diverse space missions.[801]
- Interoperability standards for AI systems: Establishing interoperability standards that specify common interfaces (e.g., ASTERICS) and data formats for AI systems on different spacecraft could ensure compatibility to facilitate seamless communication and cooperative maneuvers, thereby enhancing overall STM efficiency.[802]
- Deployment of advanced surveillance systems: Deploying advanced surveillance systems that leverage Space Situational Awareness (SSA) data and tracking technologies (e.g., S-track) for real-time monitoring of AI-driven spacecraft positions, trajectories, and operational status could enable effective alert systems for collision avoidance and proactive mission management.[803]

---

[798] See International Civil Aviation Organization. *Annex 10 to the Convention on International Civil Aviation: Aeronautical Telecommunications.* Volume II: Communication Procedures including those with PANS status. 7th ed. Montreal: ICAO, 2016.
https://www.iacm.gov.mz/app/uploads/2018/12/an_10_v1_Aeronautical-Telecommunications_7ed._2018_rev.91_01.07.18.pdf

[799] *Ibid.*

[800] See International Maritime Organization. *International Convention for the Safety of Life at Sea (SOLAS), 1974, as Amended.* Chapter V: Safety of Navigation. London: IMO, 2004.

[801] See von der Dunk, Frans G. "Space Traffic Management: Legal Aspects." *Nebraska Law Review* 87, no. 2 (2015). https://digitalcommons.unl.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1091&context=spacelaw.

[802] See Sadat, Mir, and Julia Siegel. "Space Traffic Management: Time for Action." *Atlantic Council*, August 2021. https://www.atlanticcouncil.org/in-depth-research-reports/issue-brief/space-traffic-management-time-for-action/.

[803] Murakami, David, et al. "Space Traffic Management: A Perspective from Developing Countries." *Journal of Space Safety Engineering* 7, no. 2 (2019): 1-12. https://core.ac.uk/download/pdf/227726191.pdf.

- Development of a comprehensive regulatory framework: Developing a comprehensive regulatory framework could include several key components:[804]
    - Communication protocols: Mandating adherence to CCSDS standards for telemetry data exchange, ensuring compatibility and reliability across missions.
    - Interoperability requirements: Defining mandatory interfaces and data formats (e.g., ASTERICS) for AI systems to ensure seamless communication and cooperative maneuvers.
    - Surveillance mandates: Requiring the deployment of advanced surveillance systems (e.g., S-track) for real-time monitoring of spacecraft positions and operational status.
    - Compliance and certification: Establishing international compliance requirements and certification processes to verify adherence to regulatory standards, ensuring uniformity and safety across global space operations.
- AI algorithms for collision avoidance: Developing and standardizing sophisticated AI algorithms for collision avoidance is paramount. These algorithms could integrate data from SSA systems to predict and preempt potential collisions between spacecraft and space debris.[805]
- Integration of reporting systems: Establishing robust reporting systems akin to maritime Automatic Identification Systems (AIS) is essential for effective STM.[806] AI-driven spacecraft should regularly report their positions, velocities, and operational intentions to a centralized STM authority. Standardizing data formats and reporting protocols facilitates seamless information exchange, enhancing coordination among space operators and regulatory bodies.[807]
- Integration of Explainable AI (XAI) for transparency and compliance: Integrating explainable AI (XAI) techniques into STM systems could ensure transparency and regulatory compliance.[808] Methods such as "Local Interpretable Model-agnostic Explanations" (LIME) and "Shapley Additive exPlanations" (SHAP) could provide interpretable insights into AI-driven decisions regarding navigation, collision

---

[804] *See* Larsen, Paul B. "Space Traffic Management Standards." *Journal of Air Law and Commerce* 83, no. 3 (2018): 359-387. https://scholar.smu.edu/cgi/viewcontent.cgi?article=4087&context=jalc.

[805] *See* Frandsen, Hjalte Osborn. "Current Developments in Space Traffic Management." *Space Policy* 9, no. 2 (2021): 231-238. https://www.sciencedirect.com/science/article/pii/S2468896722000064; *See also* Schrogl, Kai-Uwe, et al. Space Traffic Management: Towards a Roadmap for Implementation. *International Academy of Astronautics*, 2018. https://www.black-holes.eu/resources/IAA_spacetrafficmanagement.pdf.

[806] See Patel, Neel V. "Why We Need a Traffic Cop for Space." *MIT Technology Review*, August 23, 2021. https://www.technologyreview.com/2021/08/23/1032386/space-traffic-maritime-law-ruth-stilwell/.

[807] See Stilwell, Ruth. Diplomacy and Space Traffic Management. *United Nations Office for Outer Space Affairs*, 2020. https://www.unoosa.org/documents/pdf/WSF/Posters/Diplomacy/Diplomacy_-_Ruth_Stilwell_-_Stilwell_Poster_WSF.pdf.

[808] See A. Adadi and M. Berrada, "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," in IEEE Access, vol. 6, pp. 52138-52160, 2018, doi: 10.1109/ACCESS.2018.2870052.

avoidance, and operational planning, fostering stakeholder confidence and regulatory oversight.[809]

- Continuous Improvement through Iterative Updates: Promoting continuous improvement by instituting a framework for iterative updates of AI-driven STM systems could establish feedback mechanisms to incorporate technological advancements and operational insights. This would foster collaboration among space agencies, industry stakeholders, and regulatory bodies to optimize operational efficiency in dynamic space environments.

By adopting the principles outlined in ICAO's Annex 10 and SOLAS Chapter V, STM can develop robust standards for AI-driven spacecraft. These standards encompass communication protocols, interoperability requirements, surveillance mandates, regulatory frameworks, and integration of advanced AI for collision avoidance and operational transparency. Such initiatives are essential for enhancing safety, efficiency, and sustainability in orbital environments while fostering international collaboration and regulatory alignment among spacefaring nations and commercial entities.

2. *Space Liability and Insurance: Adapting Models from Aviation and Maritime Law*

Liability and insurance are pivotal aspects in regulating AI systems within the realm of space. The ICAO's Montreal Convention sets forth liability and compensation guidelines for international air carriage, encompassing personal injury, death, baggage, and cargo.

These principles serve as a foundational framework for establishing liability rules tailored to AI systems in space, ensuring accountability for operators in cases involving malfunction or erroneous decisions by AI technologies. Clear guidelines for liability determination and the implementation of insurance schemes are essential components to provide compensation for any resulting damages.

Similarly, the IMO's International Convention on Liability and Compensation for Damage in Connection with the Carriage of Hazardous and Noxious Substances by Sea (HNS Convention) offers insights into managing liability complexities for hazardous substances, offering a precedent for addressing similar challenges posed by AI systems in space.

For instance, spacecraft powered by AI that transport hazardous payloads may be subject to stringent liability rules and mandated to carry adequate insurance coverage to mitigate potential damages.

2.1 *Montreal Convention*

The Convention for the Unification of Certain Rules for International Carriage by Air (Montreal Convention 1999) was conceived as a private international air law instrument to harmonize and create uniformity across different jurisdictions resolving disputes arising against

---

[809] Linardatos, Pantelis, Vasilis Papastefanopoulos, and Sotiris Kotsiantis. "Explainable AI: A Review of Machine Learning Interpretability Methods." *Entropy* 23, no. 1 (2021): 18. https://doi.org/10.3390/e23010018.

air carriers in international travel involving passenger delay, injury or death, as well as loss, damage or delay of cargo.[810] It was meant to replace the existing private international legal regime pertaining to disputes against air carriers established by the Warsaw Convention and its related agreements.[811]

While the Montreal Convention applies to "international carriage of persons, baggage or cargo performed by aircraft for reward"[812] or gratuitously, it can provide us with some valuable insights into a possible solution for liability arising in space travel and transport involving artificial intelligence. This is mainly due to the wording of the liability provisions of the Montreal Convention and how they are framed. A short, more detailed survey of these provisions provide a unique insight into how similar rules could be adopted for the space industry.

The liability provisions of the Montreal Convention are two-fold: they provide compensation for damage due to (1) delay, death or bodily injury of a passenger as well as delay, damage and loss of checked and unchecked baggage, and (2) loss, damage or delay to cargo. The liability provisions of the Montreal Convention are strict, and fault is not relevant to assess the liability of the air carrier.[813]

The onus of proving the fault of the plaintiff remains with the air carrier (see below). In its entirety Article 17 paras 1 and 2 state:

> "1. The air carrier is liable for damage sustained in case of the death or bodily injury of a passenger upon condition only that the accident which caused the death of or injury took place on board of the aircraft or in the course of any of the operations of embarking or disembarking.
>
> 2. The carrier is liable for damage sustained in case of destruction or loss of, or of damage to, checked baggage upon condition only that the event which caused the destruction, loss or damage took place on board the aircraft or during any period within which the checked baggage was in the charge of the carrier. However, the carrier is not liable if and to the extent that the damage resulted from the inherent defect, quality or vice of the baggage. In the case of unchecked baggage, including personal items, the carrier is liable if the damage resulted from its fault or that of its servants or agents."

The wording of the provision using the term "accident which caused" and the "event which caused" means that the plaintiff will only need to prove: (1) the existence of the accident or event, (2) the causal connection between the accident or event and the damage sustained, (3) the passenger injury or death took place on board the aircraft or during embarkation and disembarkation. This amounts to a strict liability and, crucially, fault is not needed.

---

[810] Paul S. Dempsey, "Origins of the Montreal Convention 1999," in The Montreal Convention, ed. Dimitrios Leloudas, Paul S. Dempsey, and Chassot (Cheltenham: Edward Elgar Publishing Ltd, 2023), 1, 2.

[811] These include the Hague Protocol 1955, the Guadalajara Convention 1961, Guatemala City Protocol 1971 and the Montreal Protocols 1975. Together with the Warsaw Convention these are often referred to as the Warsaw System.

[812] Convention for the Unification of Certain Rules for International Carriage by Air, May 28, 1999, 2242 U.N.T.S. 309, entered into force November 4, 2003, art. 1.

[813] The only exception is the exoneration provision found in article 20 which requires "negligence or other wrongful act or omission."

In a similar fashion, for cargo the Montreal Convention stipulates in Article 18 that "the carrier is liable for damage sustained in the event of the destruction or loss of, or damage to, cargo upon condition only that the event which caused the damage so sustained took place during the carriage by air".[814] Carriage by air within the meaning of Article 18 is defined as the time period in which the air carrier is in charge of the cargo.[815]

Finally, the air carrier is also liable for any damage caused by delay in the carriage by air of passengers, baggage or cargo unless the carrier proves that it and its servants and agents took all measures that could reasonably be required to avoid the damage, or it was impossible to take such measures.[816]

The ability for air carriers to escape liability under the Montreal Convention 1999 is very limited. According to Article 20 the air carrier can be wholly or partly exonerated from liability when it proves that the damage was caused or contributed to by the negligence or other wrongful act or omission of the person claiming compensation, or the person from whom they derive their rights. This reverse onus clause places the burden of proof on the air carrier and constitutes the only provision that contains reference to fault.

A reverse onus clause, similar to that found in the Montreal Convention 1999 would be suitable for space travel to exonerate the space carrier. All of this makes the liability regime established in the Montreal Convention quite a useful template for a possible future regulation pertaining to the space sector.

First, there has been a remarkable shift in space activities from public to private entities. While outer space used to be solely a state-dominated area, the majority of space applications are conducted by private entities nowadays, with many States utilizing private companies for space missions.[817] While passenger transport in and into outer space, apart from suborbital tourist flights, does not yet exist, cargo transport is already being routinely used.

NASA for example tasked SpaceX with resupplying the ISS.[818] There seems therefore to be a development towards a "space carrier." For the moment, should damage occur, it will primarily be governed by the Outer Space Treaty and the Liability Convention and, once passenger and cargo transport become more widespread, the need for dedicated liability provisions will arise and the provisions in the Montreal Convention will provide a good template that can be adopted for specific regulation of space activities.

---

[814] Article 18 contains an escape clause. According to para 2 the air carrier is not liable "if and to the extent it proves that the destruction, or loss of, or damage to, the cargo resulted from one or more of the following:
(a) inherent defect, quality or vice of that cargo;
(b) defective packing of that cargo performed by a person other than the carrier or its servants or agents;
(c) an act of war or an armed conflict;
(d) an act of public authority carried out in connection with the entry, exit or transit of the cargo."

[815] Article 18/3 Montreal Convention 1999

[816] Article 19 Montreal Convention 1999

[817] Hamza Hameed, "The Concept of Launching State in Democratized NewSpace," in Proceedings of the International Institute of Space Law 2018, ed. P. J. Blount et al. (The Hague: Eleven International Publishing, 2018). Dave Baiocchi and William IV Welser, "The Democratization of Space," Foreign Affairs 94, no. 3 (2015): 98.

[818] NASA, "Commercial Resupply Services Overview," NASA, accessed on September 27 2024, https://www.nasa.gov/commercial-resupply-services-overview/.

Second, the introduction of AI to the outer space environment adds a myriad of "new" problems to the legal arena. This is perhaps most pressing in the realm of liability. As this report has highlighted, the introduction of AI will pose a challenge to our existing understanding of "fault" as it creates difficulty in establishing negligence or intent on which a fault analysis is ordinarily built.

Dedicated liability provisions that remove the requirement of fault are recommended to adequately account for AI. The provisions of the Montreal Convention 1999 provide such a legal framework, and we advise to adopt a similar framework for the transport of cargo, baggage and passengers in outer space, particularly in light of the growing use of AI.

A strict liability regime where the plaintiff would only need to prove the existence of an accident and the causal connection between the damage and the accident would circumvent the problems that AI typically causes in the fault analysis. Perhaps, to provide even better clarity, a provision could be adopted that either expressly classifies AI as an agent of the space carrier or alternatively makes the AI carrier more generally liable for the use of AI in its operations.

Third, a dedicated regime for space transportation of cargo, passengers and baggage modeled after the Montreal Convention would provide clarity, not only for regulators but also for private entities which can obtain adequate insurance coverage and sufficiently prepare for their obligations. Further, such a regime would come with the distinct advantage of being able to be applied to AI and non-AI enabled space applications alike.

For the space-related context, the wording of the Montreal Convention could be adopted with only minor changes. For example, a liability provision in a space context could state that "*the space carrier is liable for damage sustained in the event of the destruction or loss of, or damage to, cargo upon condition only that the event which caused the damage so sustained took place during space carriage.*"

Although advisable and recommended, regulation of this kind would not come without its challenges.

First, regulators would need to determine a definition of "space carriage," "space carrier" and carriage performed by a "space object." To date, there is neither an accepted definition of "space object" nor agreement on when airspace ends, and outer space begins. These questions will be of importance to make sure the system is applied correctly.

Space carrier could be classified as "any entity that provides transportation to, from and in outer space" and space carriage could be defined as "all carriage of passengers, baggage and cargo performed by a space carrier to, from and in outer space gratuitously or for reward. To provide clarity as to where airspace ends, such a dedicated regime could clarify that outer space means the environment beyond the Van Kármán Line. These definitions would ensure clear delimitation and provide a clear interpretation.

Second, there needs to be a clear understanding as to whom the law considers crew and whom a passenger, the former being exempt from bringing claims under such a regime. Currently, the space law regime refers to the term 'astronaut' (which in turn is also not clearly defined).

We recommend making a clear distinction between "astronaut" and "non-astronaut." US legislation provides guidance on how this could be achieved. The Federal Aviation Administration, for example, refers to "space flight participants" and defines them as "an individual, who is not crew, carried aboard a launch vehicle or reentry vehicle."[819] Such a definition is important if we regulate bodily injury, death or delay of a passenger (i.e., space flight participant).

We recommend the same for a dedicated liability regime for the transport of passengers, baggage and cargo by space carriers. An astronaut could encompass any personnel of the air carrier that is involved in the operation of the space transport, and conversely a space flight participant would be any person who is not involved in the operation of the space transport.

Finally, the scope of application must be clearly determined. Staying close to the treaty text of the Montreal Convention, we recommend that such liability provisions apply to accidents on board of a space object—such terminology would cover all phases including launch, transport and landing procedures—and during the embarking and disembarking of passengers and to cargo the moment it is in the charge of the "space carrier."

It is recommended that a clear temporal and physical scope be imposed on the terms embarkation and disembarkation.[820] Additionally, the period of carriage must be determined. It is not hard to imagine that the space industry will employ a multi-modal transport system.

Cargo could be transported via land, water or air to the launch site before it is loaded onto the space object. The definition of the Montreal Convention provides some good guidance, and the same principles could be adopted for a dedicated regime for space activities.

Accordingly, any carriage performed by air, land, sea or inland waterways should fall outside the scope of such a dedicated regime. For cargo and baggage, this could be the moment the cargo or baggage was given into the charge of the space carrier on the space board, and for passengers as soon as the boarding procedures have commenced.

Regarding a possible exoneration provision, we view the exoneration provision found in Article 20 of the Montreal Convention to be suitable for a potential space law regime. Article 20 of the Montreal Convention provides for the exoneration of the air carrier if it can prove negligence or other wrongful omission by the plaintiff. This provision balances the right for the air carrier to escape liability with the interests of the customer (i.e., passenger and shipper).

For example, if a passenger onboard a space carrier willfully ignores instructions from the crew which caused the passenger bodily harm, then the space carrier could be exonerated from liability based on this. Similarly, where should the shipper themselves take over the loading of cargo, and negligently fail to follow proper procedures, then the space carrier should be able to escape liability.

---

[819] *Federal Aviation Administration Commercial Space Transportation Regulations, 14 C.F.R. § 401.5 (2023).*

[820] A drawback of the Montreal Convention, indeed the whole Warsaw system, is that there is no definition provided for the terms "embarkation" and "disembarkation" which has led to a substantial amount of literature and case law being devoted to interpret these terms. See e.g. Andrew J. Harakas and Robert Lawson KC, "Death and Injury of Passengers: Damage to Baggage," in The Montreal Convention, ed. Dimitrios Leloudas, Paul S. Dempsey, and Chassot (Cheltenham: Edward Elgar Publishing Ltd, 2023), 160, 179ff

Additionally, clauses could be adopted that would take into consideration the unique environment of outer space. For example, the space carrier could be exonerated for any natural disasters that are outside the control of the carrier (e.g., a solar flare) or bodily injury or damage caused by the heightened radiation levels in outer space.

While, in the context of current space activities, it could be argued that space travel is inherently more dangerous than conventional air travel—thus passengers and shippers not needing a higher degree of protection—this will likely change in the future once space transport becomes more common.[821]

2.2 *Insights from the HNS Convention and the Space Liability Convention*

The IMO's International Convention on Liability and Compensation for Damage in Connection with the Carriage of Hazardous and Noxious Substances by Sea (HNS Convention) provides a valuable framework for managing liability complexities associated with hazardous substances. This convention, along with the principles expressed in the Liability Convention (discussed in the earlier sections), offers a robust foundation for addressing the unique challenges posed by AI systems in space.

The HNS Convention's two-tier system for liability and compensation, where initial compensation is provided by the shipowner and supplemented by an international fund for excess damages, could be adapted for AI systems in space.[822] Operators of AI-powered spacecraft would bear initial liability for any damages caused by their operations.

Establishing an international space liability fund, akin to the International Oil Pollution Compensation Funds (IOPC Funds), could provide a secondary layer of financial security.[823] This fund would ensure adequate compensation beyond the operators' coverage, reflecting Article 13 of the HNS Convention.[824] This tiered approach provides comprehensive coverage and encourages operators to maintain high safety standards.

Mandatory insurance requirements, as outlined in Article 12 of the HNS Convention, could be similarly mandated for operators of AI spacecraft.[825] This ensures operators have the necessary financial resources to address damages arising from AI malfunctions or hazardous payloads. Embedding such a mandate within space liability regulations would mirror the insurance provisions in the HNS Convention, ensuring robust financial preparedness.

---

[821] While this development might take a bit longer for passenger travel, cargo is already routinely being transported.

[822] See International Maritime Organization. *Protocol of 2010 to the International Convention on Liability and Compensation for Damage in Connection with the Carriage of Hazardous and Noxious Substances by Sea, 1996 (HNS Convention).* London: IMO, 2010. https://www.hnsconvention.org/wp-content/uploads/2019/05/2010-HNS-Convention-English.pdf.

[823] See International Maritime Organization. *International Convention on the Establishment of an International Fund for Compensation for Oil Pollution Damage (FUND).* London: IMO, 1992. https://www.imo.org/en/About/Conventions/Pages/International-Convention-on-the-Establishment-of-an-International-Fund-for-Compensation-for-Oil-Pollution-Damage-(FUND).aspx

[824] See IMO, *HNS Convention*, Article 13.

[825] *Ibid*., Article 12.

The scope of liability under the HNS Convention includes personal injury, property damage, and environmental harm.[826] For AI systems in space, liability similarly encompasses personal injury, property damage, environmental damage resulting from erroneous AI decisions. If an AI system on a space object makes autonomous decisions leading to damage, liability must be clearly assigned. For a comprehensive analysis of these topics, refer to Part I, Section 1, which discusses the Liability Convention in detail, and Part II, Section 3, which examines procurement.

To address complexities, the following recommendations are proposed:

- Establishing clear liability tiers: The primary liability should rest with the operator of the spacecraft utilizing AI systems, similar to the HNS Convention's principle where the shipowner bears initial liability. An international space liability fund could be created to provide additional compensation if damages exceed the operator's liability coverage. This fund would act as a financial safety net, ensuring comprehensive compensation mechanisms are in place.
- Implementing mandatory insurance requirements: Operators of AI spacecraft should be required to maintain insurance or other financial security sufficient to cover their potential liabilities. This can be legislated through a specific article that mirrors Article 12 of the HNS Convention.[827] Ensuring that operators are financially prepared to handle damages resulting from AI malfunctions or hazardous payloads will promote accountability and risk mitigation.
- Broadening the scope of liability: The liability scope should be broadened to include not just personal injury and property damage, but also environmental harm and any other adverse effects caused by AI decisions. This comprehensive liability scope, akin to the provisions of the HNS Convention, ensures that all potential risks are covered, creating a robust framework for accountability.
- Addressing ownership and control changes: Regulations and contracts should clearly define liability in scenarios where space objects utilizing AI systems change ownership or control mid-mission. For instance, if a satellite with integrated AI systems is sold to another operator while in orbit, the liability for any subsequent damages caused by the AI system must be clearly assigned. Liability should transfer with ownership to ensure that the responsible party is held accountable.
- Including AI in the definition of space objects: To clarify liability issues, AI systems should be explicitly included in the definition of space objects. This inclusion ensures that all aspects of AI technology are covered under existing frameworks, providing clear guidelines for operators and regulatory bodies.

Adapting principles from the HNS Convention and the Liability Convention of Space to AI systems in space involves establishing clear liability tiers, mandatory insurance requirements, and a broad scope of liability. Integrating these principles into space liability regulations ensures that the space industry has robust mechanisms to address the unique challenges posed by AI.

---

[826] See IMO, *HNS Convention*, 2010.
[827] See IMO, *HNS Convention*, Article 12.

This comprehensive approach promotes international cooperation, responsible AI usage, and adherence to established global standards and legal requirements. By leveraging existing regulatory frameworks and creating new legal instruments, the space industry can ensure the safe, secure, and efficient utilization of AI in space operations, mitigating risks and fostering technological advancements.

3. *Coordinated Approaches to STM and AI Regulation*

AI technologies play a critical role in STM by enabling real-time decision-making, predictive analytics, and autonomous operations that improve collision avoidance, debris tracking, and overall traffic coordination in increasingly crowded orbital environments. Building on the principles established in the Chicago Convention, the IMO offers a complementary framework that can be adapted to regulate AI technologies in space. The International Convention on Maritime Search and Rescue (SAR Convention), adopted in 1979, establishes a comprehensive system for maritime search and rescue operations, ensuring efficient and effective international collaboration.[828] This convention's structure and principles provide valuable insights for managing AI-driven space activities, particularly in terms of international cooperation and operational coordination.

Regarding legal force and implementation, it's important to distinguish that while ICAO annexes lack the same binding force as convention articles, they enforce compliance through audits of member states' adherence to Standards and Recommended Practices (SARPs).[829] This auditing mechanism ensures accountability and encourages member states to uphold established standards within their jurisdictions. Conversely, the IMO adopts enforceable legislation, where governments integrate conventions into national law for implementation and enforcement.[830]

This model provides a pathway for evolving AI regulation in space. Through the development of an international convention specifically for AI in space, signatory nations can adopt these regulations into their national legal frameworks. A dedicated regulatory body can oversee the implementation and enforcement of these standards, ensuring global adherence and alignment with best practices.

This approach fosters international cooperation and promotes responsible AI usage in space, aligned with established global standards and legal requirements. By integrating these international regulations and standards, the space industry can establish a comprehensive framework that ensures the safe, secure, and efficient utilization of AI in space operations. This holistic approach not only mitigates risks associated with AI in space but also leverages existing regulatory frameworks to create a robust system adaptable to technological advancements.

---

[828] See International Maritime Organization. *International Convention on Maritime Search and Rescue (SAR Convention)*. Adopted April 27, 1979. https://www.imo.org/en/About/Conventions/Pages/International-Convention-on-Maritime-Search-and-Rescue-(SAR).aspx.

[829] See International Civil Aviation Organization. *Annex 19 to the Convention on International Civil Aviation: Safety Management*. 2nd ed. Montreal: ICAO, 2016. https://www.icao.int/safety/safetymanagement/pages/sarps.aspx.

[830] Lapesa Barrera, "IMO Institutional Structure and Law-Making Process," 2022.

3.1 *Chicago Convention*

The Chicago Convention, formally known as the Convention on International Civil Aviation, was signed in 1944 and established the ICAO.[831] This Convention laid the foundational framework for international aviation law, promoting the safe and orderly growth of international civil aviation and establishing uniform standards and regulations.[832] Although primarily focused on aviation, the principles and structure of the Chicago Convention can offer valuable guidance for developing AI regulations in space.

The Chicago Convention comprises 96 articles covering a wide range of issues from the sovereignty of airspace to the facilitation of international air navigation.[833] Key aspects relevant to AI regulation in space include the establishment of international standards and recommended practices (SARPs), the emphasis on safety and security, and the promotion of international cooperation.[834]

3.1.1 Article 12: Rules of the Air

As ever larger numbers of spacecraft seek to utilize Earth's limited orbital volume in increasingly dense regimes, enhanced coordination is essential to ensure these spacecrafts operate safely while avoiding physical collisions, radio-frequency interference, and other hazards. While efforts to date have focused on improving SSA and enabling operator-to-operator coordination, there is growing recognition that a comprehensive system for STM is necessary.

The 2018 International Academy of Astronautics (IAA) study on STM has laid important groundwork in this field.[835] This study identifies two key dimensions of space traffic: the scientific-technical area and the regulatory field.[836] Additionally, the study outlines three phases of STM: the launch phase, in-orbit operation phase, and re-entry phase.[837] Securing the information needed for SSA involves defining necessary data, establishing rules for data provision and management, and setting up an information service on space weather.[838]

The STM architecture forms the framework for an STM ecosystem, facilitating the inclusion of third parties that can identify and fill niches by providing new, valuable services.[839] By making STM functions available as services, the architecture minimizes the internal expertise required within individual organizations, thereby lowering the barriers to operating in space and equipping participants with the necessary information to behave responsibly. Operational support

---

[831] International Civil Aviation Organization. *Convention on International Civil Aviation. Signed at Chicago*, December 7, 1944. https://www.icao.int/publications/Documents/7300_orig.pdf.

[832] *Ibid.*

[833] *Ibid.*

[834] *Ibid.*

[835] See Schrogl, Kai-Uwe, ed. Space Traffic Management: Towards a Roadmap for Implementation. *International Academy of Astronautics (IAA)*.

[836] *Ibid.*

[837] *Ibid.*

[838] *Ibid.*

[839] See NASA. *System and Method for Autonomous Navigation of a Vehicle and Related Vehicle System.* https://technology.nasa.gov/patent/TOP2-294.

for collision avoidance and separation is managed through a decentralized architecture, rather than a single centralized government-administered system.

The STM system relies on standardized Application Programming Interfaces (API) to allow easier interconnection and conceptual definition of roles, thus enabling suppliers with diverse capabilities to add value to the ecosystem.[840] The architecture supports essential functions such as registration, discovery, authentication of participants, and auditable tracking of data provenance and integrity. This technology can integrate data from multiple sources, enhancing the overall safety and efficiency of space operations.

Drawing parallels to Article 12 of the ICAO's Rules of the Air, which mandates that every aircraft comply with the rules and regulations of the state over which it is flying and requires ICAO to establish uniform rules and regulations for international aviation, ensuring adherence by member states, there is a clear assignment of responsibility to aircraft operators.[841] This ensures a high degree of compliance and standardization. However, applying this principle to AI in space requires adaptation due to the distinct governance structures between airspace and outer space.

For AI systems in space, the principle of adhering to the governing body's regulations can be similarly applied. Establishing standardized rules for autonomous navigation is crucial to ensure that AI-driven spacecraft can navigate safely without human intervention. This would involve creating protocols that mandate AI systems use predefined safe corridors in space, akin to air traffic corridors, to reduce the risk of collisions with other space objects.

Similar to the IAA working group proposals, the management of space traffic should include defining the necessary data, establishing rules for data provision and management, and setting up an information service on space weather.[842] Additionally, implementing a comprehensive notification system is crucial. This system should specify parameters for the notification of launches, operation of space objects, orbital maneuvers, possible encounters, and re-entry, as well as provisions for the end-of-lifetime of space objects.[843]

The IAA working group also emphasized that the design characteristics of space objects should include materials designed for demise, avoidance of electromagnetic interference, and the use of green technologies.[844] Traffic rules proposed by the group should cover safety provisions for launches, specific regimes for the space between airspace and outer space, zoning (selection of orbits), right of way rules for in-orbit phases, prioritization regarding maneuvers, security rules for human spaceflight, specific rules for geostationary orbits (GSO), Lagrange points (LG), polar orbits, and low Earth orbit (LEO) satellite constellations. Additionally, there should be debris mitigation regulations, safety rules for re-entry (including descent corridors), environmental provisions (such as pollution of the atmosphere/troposphere), and rules for radiofrequency use and avoidance of interference.[845]

---

[840] *Ibid.*

[841] See ICAO, *Annex 2*, Article 12.

[842] See International Academy of Astronautics, *IAA at European STM Conference Hearing*, 2021.

[843] *Ibid.*

[844] *Ibid.*

[845] *Ibid.*

Mechanisms for implementation and control suggested by the IAA working group should include harmonized national licensing mechanisms, enforcement and arbitration mechanisms (such as policing in outer space and renouncement of access to information or frequencies), operative oversight, and clearly defined civilian-military coordination and cooperation.[846]

The group proposed an incremental bottom-up approach for STM, allowing the coexistence of regulatory instruments of different natures and purposes, enabling individual solutions at the domestic level, and providing flexibility to tackle key issues in a timely manner, albeit with the risk of fragmentation.[847]

Alternatively, the IAA working group recommended a top-down approach, involving the creation of a comprehensive and inclusive STM regime, combining legal norms (evolving existing space law) and institutional management.[848]

They suggested an "ITU approach" as a possible model, with three levels: Level 1, the Outer Space Convention (OSC), comparable to the ITU Constitution and Convention and rarely updated; Level 2, the Outer Space Traffic Rules (OSTR), comparable to the ITU Administrative Regulations and regularly reviewed and updated; and Level 3, the Outer Space Traffic Technical Standards (OSTTS), comparable to the ITU Standards and regularly reviewed and updated with the involvement of non-governmental stakeholders.[849]

3.1.2 Article 28: Air Navigation Facilities and Standard Systems

Article 28 emphasizes the need for states to provide air navigation facilities and ensure that these systems are standardized and interoperable.[850] This article promotes international collaboration to maintain high standards of air navigation systems, ensuring safety and efficiency in global air traffic.[851] However, the principles of air navigation require significant adaptation for STM.

In translating Article 28 to space operations, states could be required to provide and maintain STM facilities. Developing and implementing STM systems that track and manage the movements of AI-driven spacecraft is crucial. These systems would use data from multiple sensors and satellites to provide real-time tracking and management, ensuring safe distances are maintained.

For instance, a global STM network could be established, where AI algorithms process data from various sources to predict and prevent potential collisions. This system could function similarly to air traffic control systems, providing real-time updates and guidance to AI spacecraft.

Creating standardized data exchange protocols is also critical to ensure interoperability between different nations' STM systems and AI spacecraft. This would involve establishing common data formats, communication standards, and protocols for data sharing. For example, AI

---

[846] *Ibid.*
[847] *Ibid.*
[848] *Ibid.*
[849] *See* ICAO, Annex 2, Article 12.
[850] *Ibid.*
[851] *Ibid.*

systems on different satellites could use standardized communication protocols to share their position and velocity data, allowing them to coordinate movements and avoid collisions.

Additionally, establishing ground-based control centers that use real-time data to monitor and manage AI spacecraft movements can significantly reduce the risk of collisions and enhance operational safety. These centers could utilize AI to predict potential conflicts and take preemptive actions to avoid them, ensuring a robust and adaptive STM system.

3.1.3 Article 37: Adoption of International Standards and Procedures

Article 37 requires ICAO to adopt and amend international standards and recommended practices (SARPs) related to aviation safety, regularity, and efficiency.[852] This article's strength lies in its provision for ongoing development and updating of standards to keep pace with technological advancements and emerging safety issues. Specifically, Article 37 covers the following areas[853]:

(a) Communications systems and air navigation aids: In aviation, this ensures reliable communication and navigation. For space, AI systems require robust communication protocols and navigation aids to operate autonomously and interact with other spacecraft and ground stations. Developing standardized communication protocols for AI systems to ensure interoperability between different nations' space assets is essential. Implementing navigation aids, such as AI-driven positioning systems, will provide real-time location data to spacecraft, ensuring safe and precise maneuvering.

(b) Characteristics of airports and landing areas: This pertains to the physical and operational standards of airports. For space, the equivalent would be launch and landing sites. Establishing international standards for spaceports and landing sites used by spacecraft equipped with AI systems is necessary. These standards should encompass safety measures, operational procedures, and technical specifications to ensure safe and efficient launches and landings.

(c) Rules of the air and air traffic control practices: These rules ensure orderly and safe aircraft operations. For space, similar rules are necessary for effective STM. Developing guidelines for the operation of spacecraft that use AI systems, including right-of-way protocols, speed limits, and communication requirements, is critical. Implementing space traffic control practices to monitor and manage spacecraft movements will significantly reduce the risk of collisions.

(d) Licensing of operating and mechanical personnel: This ensures that personnel are qualified and competent. For space, this would include operators of AI systems and those involved in their maintenance. Establishing certification processes for individuals responsible for developing, operating, and maintaining AI systems in

---

[852] *See* ICAO, *Annex 2*, Article 37.

[853] *Ibid*.

space ensures that personnel have the necessary skills and knowledge to handle AI technologies safely and effectively.

(e) Airworthiness of aircraft: This ensures that aircraft are safe for flight. For space, it translates to the reliability and safety of AI systems. Creating standards for the design, testing, and certification of AI systems used in space operations is essential. This would include rigorous testing protocols to ensure AI algorithms are fail-safe and capable of handling unexpected situations.

(f) Registration and identification of aircraft: This facilitates tracking and identification. For space, while a global registry of space objects already exists under the Convention on Registration of Objects Launched into Outer Space (1976), it primarily tracks high-level details of space objects. Enhancing this registry to include specific information about spacecraft utilizing AI systems—such as descriptions of AI functionalities, operational constraints, and points of contact—would enable improved tracking, management, and accountability of space assets.[854]

(g) Collection and exchange of meteorological information: This ensures safe flight operations by providing weather data. For space, similar data exchange is necessary for space weather and situational awareness. Implementing systems for the collection and exchange of space weather data, including solar activity and space debris monitoring, is vital. AI systems could use this data to make informed decisions about navigation and operations.

(h) Logbooks: Logbooks track aircraft operations and maintenance. For space, spacecraft equipped with AI systems similarly require operational logs. Mandating the use of digital logbooks for AI systems to record all operations, decisions, and maintenance activities is essential. These logs would play a critical role in troubleshooting, ensuring transparency, and facilitating post-mission analysis.

(i) Aeronautical maps and charts: These provide essential navigation information. For space, comparable charts are necessary for orbital navigation. Developing detailed and regularly updated maps of orbital paths, space objects, and potential hazards—accessible to spacecraft utilizing AI systems—would be crucial for ensuring safe and precise navigation in space.

(j) Customs and immigration procedures: This ensures orderly and secure international travel. For space, similar procedures may be necessary for interplanetary travel and missions. Establishing protocols for customs and immigration-like procedures for interplanetary missions ensures that AI systems adhere to international agreements and regulations during such missions.

---

[854] *See* United Nations, Convention on Registration of Objects Launched into Outer Space, adopted November 12, 1974, entered into force September 15, 1976, United Nations Treaty Series, vol. 1023, https://www.unoosa.org/oosa/en/ourwork/spacelaw/treaties/registration-convention.html.

(k) Aircraft in distress and investigation of accidents: This ensures effective responses to emergencies and thorough investigations. For space, similar protocols are necessary for spacecraft utilizing AI systems. Developing guidelines for AI systems to manage distress situations, including automated emergency responses and real-time communication with ground control, is essential. Equally important is establishing robust procedures for investigating accidents involving AI systems to determine causes and implement measures to prevent future incidents.

(l) Other matters concerning safety, regularity, and efficiency: This encompasses additional aspects that ensure the safe, regular, and efficient operation of air navigation. For space, this principle can be applied to create comprehensive safety and operational standards for AI systems, covering all aspects of space missions.

3.1.4 Article 44: Objectives of ICAO

Article 44 of the ICAO outlines its objectives, which include fostering the planning and development of international air transport and promoting safety and security.[855] This article's strength lies in its comprehensive approach to promoting international cooperation and ensuring the safety and security of global air transport. Specifically, Article 44 states:

"The aims and objectives of the Organization are to develop the principles and techniques of international air navigation and to foster the planning and development of international air transport so as to: (a) Insure the safe and orderly growth of international civil aviation throughout the world; (b) Encourage the arts of aircraft design and operation for peaceful purposes; (c) Encourage the development of airways, airports, and air navigation facilities for international civil aviation; (d) Meet the needs of the peoples of the world for safe, regular, efficient, and economical air transport; (e) Prevent economic waste caused by unreasonable competition; (f) Insure that the rights of contracting States are fully respected and that every contracting State has a fair opportunity to operate international airlines; (g) Avoid discrimination between contracting States; (h) Promote safety of flight in international air navigation; (i) Promote generally the development of all aspects of international civil aeronautics."[856]

For AI in space, Article 44's objectives can guide the establishment of an international regulatory body dedicated to overseeing the development and implementation of AI technologies in space activities.[857]

3.2 *Strategic Phases for Space Traffic Management: Short and Long-Term Approaches*

The increasing volume and variety of satellites orbiting Earth have significantly heightened the potential for overcrowding, debris creation, and collisions as the most useful orbital altitudes

---

[855] *See* ICAO, *Annex 2*, Article 44.
[856] *Ibid*.
[857] *Ibid*.

approach their carrying capacities.[858] Space operators must maneuver their satellites to avoid potential collisions, which imposes additional fuel costs and shortens the lifespan of the satellites.

These shortened lifespans not only increase costs but also result in added debris if the defunct satellites cannot be disposed of sustainably.[859] Studies estimate that tens of thousands of additional satellites will likely be launched into low Earth orbit by 2030, further increasing the risk of collisions and threatening the sustainable use of Earth's orbits.[860]

At present, there is only limited STM in place—that provided by the ITU which regulates the orbital slots for communication satellites.[861] Interestingly, this management deals only with frequencies used by satellites, not the physical locations of satellites.[862]

As a result, satellites broadcasting in different frequencies can occupy essentially the same physical space.[863] For example, at some longitude positions, nearly 60 satellites in geosynchronous equatorial orbits (GEOs) share a common region of space.[864]

RAND researchers have highlighted that managing space traffic is primarily a governance challenge rather than a technical one. Effective STM requires operators to coordinate, communicate, exchange data, enable situational awareness, avoid conflicts, and define processes to adjudicate maneuvers.[865]

Currently, the management of space objects is an informal, ad hoc, and often ill-coordinated process. However, the increased complexity and danger of orbital activities have led to calls for improved international governance of space traffic to ensure continued safety and sustainability.[866]

This urgency has escalated global space traffic debates from academic circles to the highest levels of government, with STM becoming an annual topic at UNCOPUOS Legal Subcommittee proceedings.[867] Despite this, there is little agreement on what structure international STM should eventually take or what new steps should be taken in the short term. The following consider possible short-term and long-term actionable steps for STM.

3.2.1 Short-Term Approaches for Effective STM

The rapid growth of satellite launches and the increasing congestion in Earth’s orbits necessitate immediate action to mitigate risks such as collisions and debris generation. Short-term

---

[858] See McClintock, Bruce, et al. "The Department of the Air Force's Plan to Field Advanced Battle Management System Capabilities: Implications for Future Force Design." *RAND Corporation*, June 5, 2023. https://www.rand.org/pubs/research_briefs/RBA1949-1.html.

[859] *Ibid*.

[860] *Ibid*.

[861] See Ailor, William H. "The Emerging Threat of Space Debris." Acta Astronautica 58, no. 5 (March 2006): 279-286. https://doi.org/10.1016/j.actaastro.2006.01.008.

[862] *Ibid*.

[863] *Ibid*.

[864] *Ibid*.

[865] See McClintock et al., "The Department of the Air Force's Plan," 2023.

[866] *Ibid*.

[867] *Ibid*.

measures for effective STM must focus on actionable frameworks for data sharing and the establishment of collaborative governance models involving multiple stakeholders.

A vital initial step is the creation of standardized protocols for data sharing among satellite operators to ensure consistent and reliable communication. Drawing from existing systems like the Global Navigation Satellite Systems (GNSS) and the ICAO Standards and Recommended Practices (SARPs), these protocols can address the exchange of information regarding satellite positions, trajectories, and planned maneuvers. GNSS, as a system, consists of a constellation of satellites providing global positioning and navigation services.[868] It operates through an intricate structure of ground control stations, satellites, and user devices, coordinated to ensure accurate and reliable data transmission. A similar structure could be adapted for STM by establishing centralized coordination nodes responsible for collecting and disseminating data related to satellite positioning and maneuver plans. This would allow operators to rely on a unified and trusted source for critical information, enhancing situational awareness and reducing risks of collision.

Standardizing data formats and terminologies will facilitate interoperability among diverse operators, while real-time data sharing will enhance situational awareness and improve collision avoidance strategies. Additionally, the integration of robust safeguards to protect proprietary information and national security concerns is essential for fostering transparent communication. Advances in machine learning and predictive analytics can also be leveraged to process shared data and provide actionable insights for operational decisions. The development and implementation of such a framework require international consensus and oversight to ensure compliance and address potential disputes. This initiative would transition the current informal and fragmented STM efforts into a structured and reliable system.

Fostering a collaborative environment through a multi-stakeholder engagement model is equally crucial. A model like the Joint Authorities for Rulemaking on Unmanned Systems (JARUS) can provide a blueprint for STM governance. JARUS comprises a group of national aviation authorities and international organizations working collectively to develop recommendations for the safe and efficient integration of unmanned systems into airspace.[869] Its structure includes working groups focused on specific technical, operational, and regulatory aspects, with representatives from diverse stakeholder groups. For STM, a similar approach could involve creating specialized committees tasked with addressing issues such as collision avoidance, debris mitigation, and operational standards. These committees would include representatives from governments, private satellite operators, and academic institutions to ensure comprehensive input and balanced decision-making.

Inclusive representation of emerging space nations, commercial operators, and research institutions ensures diverse perspectives and expertise are considered. Collaborative policy formulation enables the co-creation of standards and guidelines that balance innovation with the need for sustainability and safety. Establishing mechanisms for regular review and refinement of

---

[868] European Global Navigation Satellite Systems Agency. *What is GNSS?*. https://www.gsa.europa.eu/segment/gnss.

[869] Joint Authorities for Rulemaking on Unmanned Systems (JARUS). *About JARUS*. https://www.jarus-rpas.org/content/about-jarus.

policies will address technological advancements and emerging challenges. Furthermore, promoting training programs, research collaborations, and workshops will foster global expertise in STM.

By combining actionable data-sharing frameworks with an inclusive and adaptive governance model inspired by the structures of GNSS and JARUS, the international community can effectively address immediate STM needs while laying the groundwork for a sustainable future in space operations.

3.2.2 Long-Term Approaches for Effective STM

The long-term implementation of STM can take either a centralized or decentralized approach, similar to the top-down and bottom-up methodologies highlighted in the IAA study. Establishing an international Space Traffic Management Authority (STMA) would address governance challenges by adopting a cooperative, collaborative, and inclusive organizational framework. Drawing inspiration from models like ICAO and IMO, STMA could ensure effective regulation of space traffic while integrating AI-driven technologies into its operations.

The first phase of a long-term strategy involves the development of national-level STM standards. Each nation would create regulations that incorporate AI-driven models for satellite operations, space debris mitigation, and collision avoidance. These national regulations would be guided by the Legal Subcommittee of the STMA to ensure alignment with international standards and frameworks. This approach not only fosters consistency but also lays the groundwork for seamless global coordination.

Building on national frameworks, the second phase would focus on promoting bilateral agreements between nations. These agreements would facilitate the exchange of STM data and the creation of common operational rules. By prioritizing legal compatibility across jurisdictions, bilateral agreements would establish a foundation for global cooperation, addressing potential inconsistencies and fostering mutual trust among space-faring nations. Such agreements would also serve as a steppingstone toward broader multilateral arrangements, fostering a culture of collaboration and shared responsibility in space governance.

Following the bilateral agreements, the third phase would involve the International Organization for Standardization (ISO) developing technical standards. ISO's role in creating universally recognized standards would be pivotal in ensuring that all space-faring nations follow consistent guidelines for STM. By leveraging its expertise in standardization across diverse sectors, ISO would draft detailed technical specifications for collision avoidance, space debris mitigation, and AI system integration into satellite operations. These standards would provide the foundation for interoperability among different national and regional systems, ensuring that technical processes align globally and reduce the risks of miscommunication or conflicting methodologies.

To implement ISO's standards effectively, international workshops and collaborative research initiatives could be established, allowing stakeholders to contribute technical insights and feedback. These processes would institutionalize the standards into operational frameworks, ensuring their enforceability and actionability through collaboration with national and regional regulatory bodies.

The culmination of these efforts would be the establishment of a stand-alone STMA as an independent international agency. This agency would be modeled after ICAO and IMO, with a dual focus on regulatory oversight and technical standardization. The STMA's structure would encompass two core subcommittees. The Technical Subcommittee, in collaboration with ISO, would focus on areas such as refining and operationalizing technical standards for collision avoidance, debris mitigation, and satellite maneuvering, as well as addressing transparency protocols, detecting dual-use risks, and preventing misuse of autonomous systems to tackle military and security challenges in STM. Meanwhile, the Legal Subcommittee would draft and enforce global legal frameworks governing space traffic, including liability, regulatory compliance, and dispute resolution mechanisms. Together, these subcommittees would ensure that technical standards are legally binding and globally enforced.

To establish the STMA, international discussions would need to be initiated at global space forums or summits, akin to how ICAO was established through the Chicago Convention. These discussions could take place at UNCOPUOS or through dedicated space cooperation agreements. A multilateral treaty or agreement would then be drafted to formalize the creation of the STMA, granting it binding authority to regulate global space traffic.

The centralized approach, while comprehensive, is also particularly challenging in the current geopolitical climate. Differences in national interests, varying levels of technological capability, and reluctance to cede sovereignty to an international body pose significant obstacles. Negotiating and achieving consensus on such a framework would require extensive diplomatic efforts and substantial trust-building among nations. Moreover, enforcement mechanisms under a centralized model could face resistance, particularly from major space-faring nations that prioritize autonomy in their operations.

In contrast, a decentralized approach to long-term STM governance, as described in the 2018 IAA STM study by Kai-Uwe Schrogl, offers a more immediately feasible alternative in this geopolitical environment. Nations would develop their own STM standards, emphasizing AI for satellite operations and collision avoidance. Coordination between countries would rely on bilateral or multilateral agreements, resulting in a fragmented landscape of sector-specific regulations. This approach allows for greater flexibility and faster implementation, as countries retain control over their own policies without requiring extensive international negotiations.

However, the decentralized approach also has significant drawbacks. Regional bodies and sector-specific agreements could lead to variability and inconsistencies in STM practices, creating challenges for operators navigating multiple regulatory regimes. Adapting to future technologies would further complicate matters, as independent regulatory evolution across sectors and regions could result in misalignments and conflicts. Continuous coordination and harmonization efforts by international organizations and legal bodies would be necessary to mitigate these risks.

Both centralized and decentralized approaches present viable paths for addressing the governance challenges of STM. The centralized model, despite its difficulties, offers a more robust and comprehensive solution for ensuring the long-term safety and sustainability of space operations. Conversely, the decentralized model provides a pragmatic option in the short term, enabling incremental progress while addressing the current geopolitical realities. Ultimately, the choice between these approaches must balance the need for global coordination with the practicalities of international collaboration in a fragmented political landscape.

Conclusions

The integration of AI into space operations necessitates a comprehensive international regulatory framework to ensure safety, security, and operational efficiency. Drawing insights from existing international regulations in aviation and maritime domains, particularly those governed by the ICAO and the IMO, offers a valuable foundation for managing AI systems in space.

Implementing robust safety and security measures is critical. By adopting principles from ICAO's safety management systems and IMO's security protocols, space missions can incorporate SMS tailored for AI components, enhance redundancy and failover mechanisms, and enforce regular software updates and maintenance. Comprehensive security measures must be deployed to protect AI systems from cyber threats and physical attacks, ensuring the integrity and reliability of space operations.

Standardized communication protocols and data exchange methods are vital for ensuring interoperability and coordination of AI-driven spacecraft. Drawing from ICAO's Annex 10 and SOLAS Chapter V, advanced surveillance systems and defined interoperability standards will facilitate real-time monitoring and coordination, reducing collision risks and enhancing overall mission safety.

Addressing liability and insurance for AI systems in space is another critical aspect. By adapting principles from the Montreal Convention and the HNS Convention, a clear liability framework should be established. This includes mandatory insurance requirements and the creation of an international space liability fund to ensure comprehensive coverage and financial preparedness. Liability provisions should cover personal injury, property damage, and environmental harm caused by AI systems, ensuring accountability and compensation.

Finally, the significance of international cooperation is highlighted, advocating for a phased approach to STM that evolves from short-term measures into a structured long-term governance framework. Short-term efforts could focus on developing standardized data-sharing protocols among satellite operators, fostering multi-stakeholder collaboration, and addressing immediate risks such as collisions and debris generation. Inspired by models like GNSS and JARUS, these measures would enhance situational awareness, interoperability, and transparency, laying the foundation for reliable STM systems.

In the long term, STM could progress through several phases. Initially, nations could establish national-level regulations tailored to their unique operational contexts. These national frameworks would then facilitate bilateral agreements to promote cross-border cooperation and the exchange of STM data. As collaboration strengthens, technical standards could be developed at the international level by organizations such as ISO, ensuring consistency across space-faring nations and reducing risks of miscommunication or conflicting methodologies. These phases lay the groundwork for establishing a Space Traffic Management Authority (STMA) as a central body to oversee global coordination and regulate AI-driven technologies specific to STM. Its responsibilities could include collaborating with ISO to refine and implement technical standards for collision avoidance, debris mitigation, and satellite maneuvering; addressing transparency protocols; detecting dual-use risks; preventing the misuse of autonomous systems; and tackling military and security challenges in STM, while enforcing compliance through technical and legal mechanisms.

Two governance models are proposed for long-term STM: a centralized model involving a stand-alone international regulatory authority like the STMA, modeled on ICAO or IMO, and a decentralized model relying on regional or sector-specific agreements. While decentralization offers flexibility and quicker implementation, centralization provides a more cohesive global framework. Combining short-term actions with a phased strategy balances innovation, collaboration, and regulatory consistency, ensuring the safe and efficient integration of AI technologies in space.

In conclusion, by integrating these established principles and developing new legal instruments tailored to the unique challenges of AI in space, we can create a robust regulatory framework. This framework will ensure the safe and efficient utilization of AI in space operations, foster international cooperation, promote technological advancements, and uphold the sustainability of space activities. Establishing a centralized international regulatory authority, with the STMA under its aegis, is pivotal in navigating the complexities of AI in space, ensuring a secure and prosperous future for space activities and utilization.

Final Remarks

The integration of AI into space activities presents unprecedented opportunities for innovation and advancement, while also raising critical challenges in governance, ethics, and sustainability. This study highlights the multifaceted legal, regulatory, and ethical frameworks necessary to address the complexities of AI-driven technologies in space, offering a roadmap for responsible development and use.

The legal discussions emphasize the adaptability of foundational international treaties such as the Outer Space Treaty and the Liability Convention to the AI era, while advocating for enhanced clarity on liability attribution and dual-use concerns. Insights from International Humanitarian Law underscore the need for strict adherence to principles of distinction, proportionality, and accountability in AI's military applications, ensuring compliance with global norms.

Regulatory considerations, including GDPR applications, export controls, and telecommunications frameworks, reveal the urgency of harmonizing international standards to manage risks associated with data privacy, dual-use technologies, and secure communications. Drawing lessons from aviation and maritime sectors, the study proposes adaptive and phased approaches to standardize AI technologies for safe and efficient space operations.

Ethical dimensions, explored through parallels in genetics, robotics, and border control, provide a foundation for addressing fairness, transparency, and human oversight in AI deployment. The role of soft law and international standards, complemented by insights from air and maritime law, demonstrates the importance of flexible yet enforceable frameworks to manage evolving technologies.

For STM, the proposed centralized model, led by a Space Traffic Management Authority (STMA), offers a cohesive approach to ensure global consistency. Alternatively, a decentralized model could allow flexibility but requires careful alignment to mitigate fragmentation. Both strategies highlight the need for short-term measures, such as data-sharing protocols, and long-term visions to sustain safe and collaborative orbital environments.

Ultimately, this study calls for international collaboration, iterative governance structures, and continuous updates to regulatory frameworks to balance innovation with accountability. By addressing the voids in existing policies and fostering global cooperation, AI technologies can be harnessed to advance space exploration ethically, sustainably, and equitably, ensuring the shared benefits of outer space for humanity.