United Nations
Office for Outer Space Affairs

POLICY BRIEF

ENSURING RESPONSIBLE AI IN SPACE AND EARTH OBSERVATION

SEPTEMBER 2025

INTRODUCTION

Artificial intelligence (AI) and Earth observation (EO) are advancing rapidly and amplifying each other's impact. New satellite constellations that revisit the same location often, along with radar systems that see through clouds, now deliver near-real-time imagery. Open data sharing and analysis platforms make it possible to handle vast datasets with ease. At the same time, AI is moving from research labs into real missions: on-board processors (small computers on satellites) can filter out unusable images, prioritise valuable data, and coordinate actions more quickly. Large "foundation models" for maps and imagery can adapt to various tasks, but they still require careful testing, clear explanations of their functionality, and responsible energy use. This raises a central question: not only what AI and EO can do, but also how they are governed, who benefits, and how they shape life on our planet.

From its vantage point as the UN's office for space affairs, UNOOSA helps countries transform space data into a public good through the United Nations Platform for Space-based Information for Disaster Management and Emergency Response (UN-SPIDER)'s disaster-risk support and capacity-building, and through Access to Space for All, which opens facilities, training, and missions to emerging space actors.

As governments confront new choices about AI in space and Earth observation, the UN has launched two complementary mechanisms: the Independent International Scientific Panel on AI, to anchor the debate in evidence, and the Global Dialogue on AI Governance, to bring together governments, industry, and civil society. UNOOSA serves as the bridge between these system-wide processes and the space community, convening the United Nations Committee on the Peaceful Uses of Outer Space (COPUOS) members and stakeholders, contributing space-specific use cases and technical insights, channelling lessons from UN-SPIDER and Access to Space for All into the Dialogue and the Panel, and translating their outcomes into practical guidance, benchmarks, and capacity-building. This linkage ensures that AI policy is shaped by real mission needs and delivers benefits for all countries, especially emerging space actors.

This policy brief outlines five objectives to build a trustworthy, inclusive EO+AI ecosystem:

(1) ethical, transparent AI for space operations;

(2) fairness, inclusivity, and global capacity-building;

(3) responsible development and use of geospatial foundation models;

(4) integration of climate resilience and sustainability across the EO+AI life cycle; and

(5) protection of data ownership and integrity, including provenance measures to prevent manipulation.

Taken together and anchored in multilateral processes such as COPUOS, these steps can maximise benefits for all countries while managing key risks from the rising energy demands of AI-enabled infrastructure to the urgent need for durable, content-provenance standards for geospatial products.

SCOPE

This policy brief addresses the civilian and peaceful applications of AI across the space and Earth observation value chain from on-board autonomy and ground processing to analysis, dissemination, and decision support. It proposes practical, near-term actions for governments and space agencies, companies developing spacecraft, data and AI services, as well as universities and research organisations. The recommendations are intended to complement multilateral processes such as COPUOS and to align with emerging UN mechanisms on AI governance.

POLICY OBJECTIVES AND RECOMMENDED ACTIONS

OBJECTIVE 1

Implement Ethical and Transparent AI for Space Operations

Rationale

Space operations are high-stakes and time-critical. When AI systems are opaque, operators cannot fully understand or trust their outputs. Missions, therefore, require AI that is explainable enough to audit, with meaningful human control adapted to context: human-in-the-loop for low-latency operations, and human-on-the-loop with robust safeguards for deep-space missions where real-time intervention is impossible. Any AI capable of changing a spacecraft's state or interacting with other objects must be equipped with dependable fail-safes such as automatic reversion to safe mode or a verified handover to a human operator. Equally, transparency around dual-use capabilities is essential to ensure that civilian activities are not adversely affected.

Recommended actions

UN Member States	Consider developing and adopting an international code of practice for AI in space that mandates: (i) explainability and traceability proportionate to risk; (ii) context-appropriate human oversight (human-in-the-loop or human-on-the-loop); and (iii) guaranteed fail-safe behaviours for critical functions. Promote pre-deployment safety cases, post-incident reporting, and transparency for dual-use AI that could impact civilian operations.
Private sector	Incorporate transparency features into mission AI, such as decision logs and operator-facing visualisations of decision pathways, and publish risk-based safety assessments. Establish internal ethics and safety review processes, including red-teaming for on-board autonomy, and implement robust configuration control, roll-back procedures, and safe-mode fallbacks.
Academia	Develop explainable-AI methods tailored to space-grade hardware and constrained communications. Design and test practical approaches for operationalising human oversight in mission planning and provide targeted training for engineers and operators on responsible AI practices.

OBJECTIVE 2

Ensure Fairness, Inclusivity and Global Capacity Building

Rationale

AI learns from the data it is trained on. If that data is predominantly from the Global North, models can misinterpret conditions elsewhere, leading to weaker results and inequitable outcomes. Access to what makes AI work—high-quality EO data, computing power, and skilled personnel—is also uneven, leaving many countries behind. Yet existing training and support programmes are insufficient to meet global demand. To make EO+AI both fair and effective worldwide, we need more diverse datasets and broader, equitable access to tools and training at the same time.

Recommended actions

UN Member States	Expand UNOOSA's Access to Space for All programme to include AI and EO scholarships and fellowships and create pathways for developing countries to access AI-ready EO datasets and computing resources. Ensure that publicly funded EO missions release their data, models, and benchmarks under open licences to enable global use, with appropriate safeguards for sensitive information.
Private sector	Share anonymised, representative training datasets and engage in structured knowledge-transfer partnerships. Conduct regular fairness audits to detect and mitigate bias in AI models. Support start-ups and research groups in the Global South through mentorship, collaboration, and targeted investment.
Academia	Foster equitable international collaborations that include institutions from under-represented regions. Co-design AI-for-EO projects with local stakeholders and study how socio-cultural factors affect model performance. Develop practical metrics and benchmarks for fairness and inclusivity, and advance privacy-preserving analytics to safeguard sensitive population data.

OBJECTIVE 3

Promote Responsible Development and Use of Geospatial Foundation Models

Rationale

Geospatial foundation models (GFMs) are large AI systems trained on vast, mixed datasets, including satellite imagery, maps, and sometimes text. By learning general patterns, they can be adapted to new tasks with minimal labelled data, which is promising for disaster response and other public-good applications. However, FMs are advancing faster than they are being evaluated, and evidence of their real-world performance and contribution to sustainability goals remains limited. Evaluation should therefore look beyond accuracy to include interpretability (can practitioners understand and audit outputs?); transferability (do results hold across regions, seasons, sensors, and resolutions?); energy efficiency and environmental footprint (how much energy and carbon are used to train and run them?); robustness (how do models perform under noise, outages, or adversarial inputs?); and ethical and social impacts (including bias, misuse, and unintended consequences).

Recommended actions

UN Member States	Develop guidance and best practices for GFMs that establishes minimum standards for documentation and evaluation, including metrics for energy use, robustness, and social impact. Fund open, multi-stakeholder benchmarking initiatives, and require publicly funded models to document training-data provenance, environmental footprint, stated limitations, and intended use cases.
Private sector	Adopt responsible AI principles when developing GFMs: measure and report energy consumption, carbon footprint, and geographic bias; train on diverse, ethically sourced data; clearly communicate limitations; participate in multi-stakeholder benchmarking; and license models to support humanitarian and scientific applications.
Academia	Develop interdisciplinary evaluation frameworks for GFMs that incorporate domain expertise, including perspectives from the Global South. Research methods to reduce energy consumption and improve transferability. Study socio-economic impacts of GFMs and publish best practices for impact-driven deployment.

OBJECTIVE 4

Integrate Climate Resilience and Sustainability into AI for Space and Earth-Observation

Rationale

Climate change is intensifying floods, droughts, and other hazards, while also threatening the long-term reliability of space infrastructure. AI and EO can help countries adapt, reduce emissions, and track progress, but the development and operation of AI for space must itself be sustainable. This means paying attention to energy use and environmental impacts across the whole life cycle. Governance should be aligned with the Paris Agreement and the Sustainable Development Goals (SDGs).

Recommended actions

UN Member States	Embed climate resilience into AI-related treaties and agreements, ensuring that AI applications in space and EO contribute to climate adaptation and mitigation. Support initiatives that integrate space-based data with in situ and socio-economic data to model climate impacts. Promote life cycle assessments of space-AI systems from manufacturing to launch and end-of-life deorbiting to minimise their environmental footprint.
Private sector	Develop AI tools that enable climate-related services such as precision agriculture and disaster-risk reduction and measure their environmental impact. Adopt sustainable practices in satellite manufacturing and AI model training, including the use of renewable energy.
Academia	Advance research methods that integrate AI-driven EO analysis with climate models and socio-economic data. Assess the effectiveness of AI-enabled adaptation measures and identify best practices for leveraging space data to support the SDGs.

OBJECTIVE 5

Safeguard Data Ownership, Integrity, and Prevent Manipulation

Rationale

Questions about satellite data ownership, sharing, and the use of AI-derived products are central to trust. While some actors restrict access for commercial or national advantage, deepening inequalities, overly tight rules can stifle innovation. At the same time, AI makes it easier to modify or misinterpret geospatial products, increasing the risk of misuse and misinformation. To maintain the credibility and fairness of EO+AI, we require clear ownership and licensing, robust end-to-end provenance, and practical trust mechanisms that safeguard sensitive information while enabling its beneficial use.

Recommended actions

UN Member States	Develop international standards for provenance and digital labelling of satellite imagery and AI-derived products (e.g., cryptographic signatures, watermarking, robust metadata) to ensure traceability from source to output. Update data protection, copyright, and export-control frameworks to clarify ownership and permissible uses of geospatial data especially when combined with AI or other sensitive datasets. Facilitate secure data-sharing arrangements (e.g., trusted data spaces or data trusts) that balance national security, commercial interests, and global public benefit.
Private sector	Implement end-to-end data governance: authenticated imagery, rich metadata tagging, and audit trails for all AI processing steps. Publish clear, readable licences and permitted-use policies for both data and models, and monitor downstream use to detect manipulation or misuse. Collaborate with AI and media platforms to counter misinformation that misrepresents geospatial content, and promote verified sources.
Academia	Develop technologies to verify the authenticity and integrity of geospatial data (e.g., watermarking, tamper-evident provenance, cryptographic attestations) and stress-test them against adversarial scenarios. Study the socio-legal dimensions of ownership, consent, and control, and propose governance models (such as data trusts) that empower data providers while enabling public-interest applications. Train the next generation of practitioners in ethical data stewardship, and responsible release and usage practices.

CONCLUSION

AI and EO are already delivering public-good benefits, from faster image analysis and on-orbit autonomy to climate services and disaster support, but governance, evaluation, and access must keep pace. This policy brief outlines five practical objectives to steer EO+AI toward outcomes that are trustworthy, inclusive, and sustainable: (1) ethical, transparent, and human-controlled AI for space operations; (2) fairness, inclusivity, and global capacity-building; (3) responsible development and use of geospatial foundation models; (4) integration of climate resilience and sustainability across the EO+AI life cycle; and (5) safeguarding data ownership and integrity to prevent manipulation.

By acting on these objectives, UNOOSA, together with Member States, the private sector, and academia, will translate principles into practice. Through collaborative actions, such as convening stakeholders, sharing mission-tested lessons, and supporting common benchmarks, they will establish clearer expectations for safety and transparency, strengthen capabilities in developing countries, develop better-evaluated models with lower environmental footprints, and ensure reliable geospatial products.

These objectives align with the long-standing normative framework of COPUOS for the peaceful, responsible, and safe use of outer space, ensuring continuity with existing principles while addressing new technological realities. Ultimately, these steps will ensure that AI in space and Earth observation benefits all countries, advancing the Sustainable Development Goals and upholding the foundational principle that outer space is the province of all humankind.

CONTACT

United Nations Office for Outer Space Affairs (UNOOSA)
Vienna International Centre
oosa@un.org | www.unoosa.org