PHYSICAL AI FRONTIERS

EDITION 01 · JULY 2026

# AI-Enabled Spacecraft Health Management

Telemetry analytics, digital twins, and bounded autonomy in the risk architecture of a satellite fleet.

**~13,000**
active satellites in orbit

**~300**
reported to carry in-orbit cover

**~$500M**
estimated value of Intelsat-33e, uninsured

**355,000+**
avoidance maneuvers in 12 months

**Lise DELOBEL**
lisedelobel.fr · Lise.delobel@me.com

01

# Executive Summary

The satellite economy has scaled past the operating model it was built on. Nearly 13,000 active satellites now orbit Earth, yet industry reporting indicates only around 300 carry comprehensive in-orbit insurance. Intelsat-33e — commonly reported as a roughly $500 million spacecraft — was uninsured against in-orbit loss when it broke up over the Indian Ocean in October 2024. As conventional insurance covers a shrinking share of orbital assets, more of the burden shifts from financial risk transfer toward engineering, operational resilience, and fleet-level risk management.

Insurance was never the only hedge, and analytics will not be either. Redundancy, design margin, radiation hardening, in-orbit spares, self-insurance, flight-software assurance, and ground diagnostic practice all remain central. What has changed is the relative weight of the engineering layer — and the maturity of three specific capabilities within it.

First, machine learning on spacecraft telemetry has moved from research toward operational tooling. ESA has published a large-scale benchmark built on real satellite telemetry — drawn from three ESA missions, with two used in the principal benchmark and roughly 775 million data points each — annotated in collaboration with the spacecraft operations engineers who worked those missions. NASA's JPL produced one of the field's most widely cited early deep-learning approaches, and comparable work is underway at CNES, DLR, JAXA, and Airbus.

Second, digital-twin and federated digital-engineering capabilities have entered formal procurement. US Space Systems Command selected a federated digital-engineering platform in June 2026, and commercial platforms are marketed for rehearsing commands, software updates, and fault responses before they reach the physical asset. The maturity of continuously telemetry-synchronized operational twins, however, varies considerably by program.

Third, automated collision avoidance at fleet scale is already operating reality. SpaceX reported over 355,000 Starlink avoidance maneuvers in twelve months — roughly 975 per day — executed through an automated architecture with onboard maneuver capability operating under ground-defined risk policies. No conventional mission-control staffing model reviews events at that rate.

This paper examines three sourced use cases across agency, defense, and commercial operations, distinguishes digital models from live operational twins, and sets out a phased implementation path with realistic assurance timelines.

THE ARGUMENT IN ONE LINE

**As fleets scale and traditional insurance covers a shrinking share of orbital assets, advanced telemetry analytics, validated digital twins, and carefully bounded autonomy are becoming critical components of spacecraft risk management.**

TWO DEFINITIONS USED THROUGHOUT

Spacecraft health management encompasses anomaly detection, prognostics, fault isolation and recovery, remaining-useful-life estimation, and corrective reconfiguration — the discipline that replaces physical maintenance on assets that cannot be repaired after launch. Bounded autonomy means autonomous execution constrained by predefined authority limits, deterministic fallback behavior, and human governance.

02

# Why Now

*Three shifts converged between 2023 and 2026 to move spacecraft health management from a research topic toward a board-level priority.*

### 1 — THE BALANCE OF RISK MANAGEMENT SHIFTED

The space insurance market contracted just as the value at risk expanded. Industry reporting placed 2023 claims at roughly $950–995 million against approximately $557–600 million in gross premiums — an underwriting result market participants described as the sector's most severe in over two decades. The ViaSat-3 Americas antenna malfunction alone was reported as a net loss of about $445 million, the largest single satellite insurance loss on record, and insurers including AIG and Allianz had already exited the class. Many operators responded by not insuring in orbit at all: of nearly 13,000 active satellites, only around 300 are reported to carry comprehensive coverage. Intelsat-33e was uninsured when it was lost in October 2024; Intelsat-29e, also uninsured, had been lost in 2019 after three years of a planned fifteen.

**The consequence is not that engineering replaces insurance, but that its share of the risk-control architecture grows. For uninsured assets, avoided failures and service interruptions directly preserve operator economics — though the benefit is counterfactual and must be adjusted for probability, attribution, and program cost.**

### 2 — FLEET SCALE OUTGREW MANUAL REVIEW

A single spacecraft streams telemetry across thousands of channels — power, thermal, attitude, propulsion, payload. Operations built on threshold alarms and engineer review are already strained for one satellite and impractical for constellations of hundreds or thousands. Undetected telemetry deviations can escalate toward partial or total loss.

### 3 — THE TOOLING MATURED

ESA curated and released real, annotated multi-mission telemetry at benchmark scale with evaluation metrics designed around operational needs rather than academic convenience. Digital-twin platforms for space systems attracted institutional customers and fresh capital. And onboard AI inference moved onto agency research roadmaps as a priority area.

**~$950M**
reported 2023 claims against roughly $557–600M in premiums
CONNECTIVITY BUSINESS NEWS

**~$445M**
ViaSat-3 net loss, reported as the largest single satellite loss on record
CONNECTIVITY BUSINESS NEWS

**~2.3%**
share of active satellites reported to carry in-orbit cover
SPACE.COM / INSURANCE BUSINESS

SOURCES — SpaceNews (Oct 2023); Connectivity Business News (Sep 2024); Insurance Business America citing Space.com (Jun 2025); DCD space insurance analysis (2026). Figures are as reported by trade press and market participants; underwriting data is not publicly audited.

03

# The Digital Twin + AI Opportunity

*The architecture transforming industrial asset management applies to spacecraft under harsher constraints: no physical access, radiation, and narrow communication windows.*

---

**L4** **Bounded onboard autonomy**

Onboard AI inference for time-critical event classes, executing within a formally defined authority envelope with deterministic fallback.

**L3** **Operational digital twin**

A component-level virtual replica synchronized periodically or continuously with operational telemetry, used to rehearse commands, updates, and fault responses before they reach the asset.

**L2** **ML anomaly detection**

Models learn nominal multi-channel behavior and flag deviations that fixed threshold alarms structurally miss — extending the intervention window.

**L1** **Telemetry foundation**

Thousands of sensor channels per spacecraft, consolidated, time-aligned, and quality-controlled. Every layer above is only as good as this one.

**NOT EVERY 'DIGITAL TWIN' IS A DIGITAL TWIN**

Vendor language collapses three distinct maturity levels. Distinguishing them is the fastest way to assess a claim — and the most common source of disappointed programs.

**Digital model**

A simulation of the asset with no automatic live-data connection. Valuable for design and rehearsal; blind to the specific spacecraft in orbit.

**Digital shadow**

One-way synchronization from asset to model. The model reflects reality but does not feed decisions back automatically.

**Digital twin**

An operationally integrated relationship in which telemetry updates the model — periodically or continuously — and model outputs inform operational decisions.

*A related term is often confused with all three: a digital thread connects lifecycle data across engineering, manufacturing, and operations; a digital twin is one operational element that may consume that thread.*

The value pools follow: avoided loss on assets that are effectively uninsured, extended service life through earlier intervention, lower operations cost per satellite as fleets scale, and faster fault response. A single spacecraft loss can destroy hundreds of millions of dollars in asset value and future service capacity — but whether any particular event could have been predicted or prevented must be assessed case by case, not assumed.

04

# Use Case 1 — Agencies: ML on Telemetry

*ESA and NASA moved spacecraft anomaly detection from an engineering craft toward a measurable machine-learning discipline.*

---

### THE PROBLEM

Spacecraft operations engineers at centers such as ESA's ESOC monitor telemetry that is high-dimensional, noisy, irregular, and overwhelmingly nominal — anomalies are rare by definition. Threshold alarms fire only when a value crosses a fixed limit, often after degradation has begun. Failures caught too late can cost part or all of a mission.

### WHAT ESA PUBLISHED

ESA's Anomalies Dataset draws real telemetry from three of its missions; the principal ESA-ADB benchmark uses two of them, each containing roughly 775 million data points and covering approximately 14 and 3.5 years of satellite life across power, thermal, and attitude-control channels. Annotations were developed with the agency's own operations engineers and distinguish anomalies, rare nominal events, and communication gaps. The result gives the field a shared, operationally realistic yardstick, with evaluation metrics designed around what operators need rather than academic scores. Anomaly detection is a prioritized domain of ESA's AI for Automation roadmap, and the agency notes a growing trend toward running such systems on board for faster alerting.

### WHAT NASA/JPL CONTRIBUTED

JPL researchers demonstrated LSTM networks with non-parametric dynamic thresholding on real NASA mission telemetry, showing that deep sequential models could flag anomalous behavior without hand-set limits. It became one of the field's most widely cited early deep-learning approaches and seeded comparable research and development across several agencies and aerospace organizations, including CNES, DLR, JAXA, and Airbus.

> **EXECUTIVE TAKEAWAY**
>
> The detection layer is no longer speculative: real mission data, operator-grade evaluation metrics, and multi-agency investment exist today. The open question for operators is not whether ML anomaly detection can work, but how quickly it can be wired into their own telemetry — and trusted by their own engineers.

SOURCES — ESA-ADB, arXiv:2406.17826 (2024); ACM Computing Frontiers poster on ESA-AD (2025); Hundman et al., KDD 2018; deployable anomaly-detection benchmark study, arXiv (2026).

05

# Use Case 2 — Twins Enter Procurement

*Digital-engineering and twin capabilities have entered formal procurement. Operational maturity, however, still varies sharply by program.*

---

### DEFENSE: US SPACE SYSTEMS COMMAND & SEDARO

Historically, the payload, satellite bus, ground stations, and user terminals of a defense satellite system were designed by different contractors in isolated proprietary tools — making it difficult to predict how the full network behaves under operational stress or electronic attack. In June 2026, Space Systems Command selected Sedaro to develop a federated digital-engineering platform intended to link disparate contractor tools and models into a single high-fidelity simulation environment. Earlier SBIR work applied the same platform toward authoritative digital twins of GPS-IIIF vehicles. Sedaro's documentation identifies uses including testing commands and software updates, operator training, and diagnosing on-orbit faults on the virtual system before changes reach the physical one.

*Read precisely, this evidences federated digital engineering entering procurement — not proof that continuously telemetry-synchronized, component-level operational twins are already routine across an active fleet.*

### COMMERCIAL: TWIN-FIRST MISSION DESIGN

Antaris inverts the sequence: its TrueTwin environment builds the model before the hardware exists, letting customers fly missions virtually — testing architectures, performance, and risks — prior to committing to hardware. The company announced the first close of a $28 million Series A in 2026 to extend AI across the mission lifecycle, including predictive design modeling, anomaly detection, constellation orchestration, and autonomous on-orbit operations, with programs including a 16-satellite constellation for Saudi Arabia's SARsatX. Its claims of halving time-to-orbit and reducing lifetime cost by an order of magnitude are vendor statements; independent validation per program remains necessary before they inform a business case.

**GPS-IIIF**
authoritative digital twins pursued for US navigation satellites
CELERIS / SEDARO SBIR II

**Jun 2026**
Space Systems Command selects a federated digital-engineering platform
SATNEWS

**$28M**
first close of Antaris Series A for mission virtualization
SPACENEWS, MAR 2026

**Executive takeaway: when the buyer is a space command or a sovereign constellation program, the category has crossed the credibility threshold. The diligence question shifts from "does this work?" to "which level of digital maturity am I buying — model, shadow, or twin?"**

SOURCES — SatNews (Jun 2026); Celeris Systems announcement (2023); Sedaro platform documentation; SpaceNews (Mar 2026); Antaris company materials (vendor statements explicitly identified as such).

06

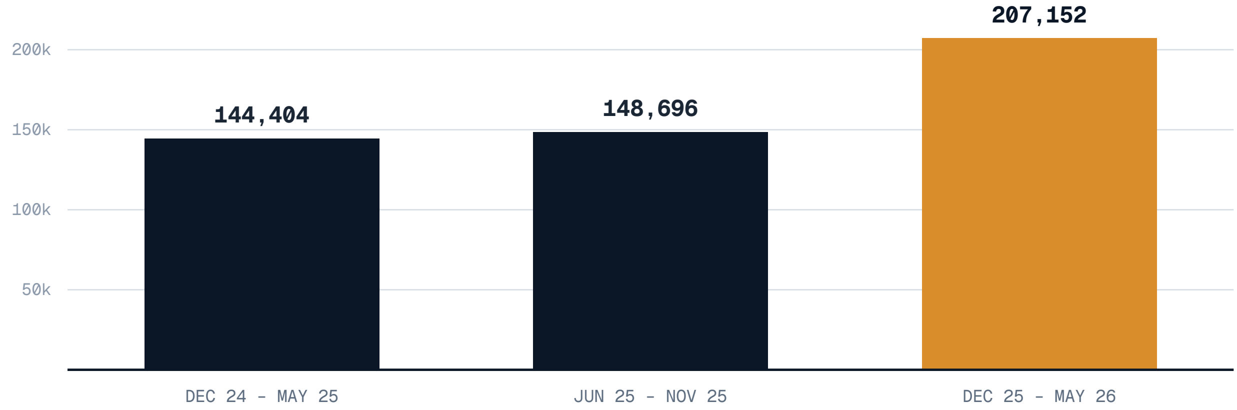
# Use Case 3 — Automation at Fleet Scale

*Starlink demonstrates that beyond a certain fleet size, automated response ceases to be optional — though the architecture is more layered than the headline suggests.*

Between June 2025 and May 2026, SpaceX reported to the FCC that Starlink satellites executed more than 355,000 collision-avoidance maneuvers — over three times the 2024 total, and more than 40 maneuvers per satellite per year across a constellation that grew from roughly 6,000 spacecraft in 2024 to over 10,000 by mid-2026. SpaceX has stated that its system may maneuver when estimated collision probability exceeds roughly 3 in 10 million — a company-stated policy threshold, not an independently verified industry standard.

The arithmetic is what matters strategically: roughly 975 events per day. No conventional mission-control staffing model individually reviews events at that rate — approval necessarily happens through predefined policy. But precision is warranted — this is an automated collision-avoidance architecture with onboard maneuver capability, operating under ground-defined policies, screening, and coordination. Not every event is an independent onboard AI decision. The human role shifts from executing individual responses to designing the policies, thresholds, and audit mechanisms that govern them.

STARLINK COLLISION-AVOIDANCE MANEUVERS PER FCC HALF-YEAR REPORTING PERIOD




THE 355,000+ TWELVE-MONTH FIGURE COMBINES THE JUN–NOV 2025 AND DEC 2025–MAY 2026 PERIODS.

The trajectory matters as much as the level: counts rose roughly 39% in a single reporting period as the constellation and the tracked debris population grew together. Space-sustainability researchers expect the burden to keep climbing — which implies every large constellation will need comparable automation, and that regulators are already normalizing disclosure of automated actions through semiannual reporting.

EXECUTIVE TAKEAWAY

**Automation at scale is operating reality, not roadmap. The differentiator ahead is the governance layer: thresholds, inter-operator coordination, and auditability of automated action.**

SOURCES — SpaceX semiannual FCC filings via Space.com (Jul 2026); Space Intel Report (Jul 2025); Aerospace America / AIAA (2025); KeepTrack analysis (Jun 2026).

07

# Engineering Zoom

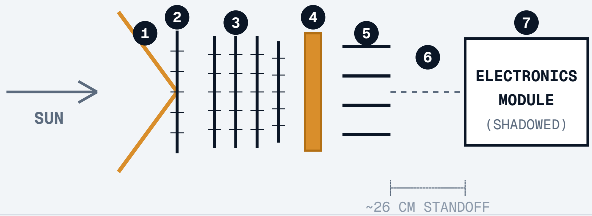
*The sensor that should not survive the Sun — and why one instrument's material choices matter as much as any model trained on its data.*

---

### THE CHALLENGE

Every instrument on NASA's Parker Solar Probe rides in the shadow of its heat shield — except one. The Solar Probe Cup has to face the Sun completely unshielded to sample the solar wind directly. At closest approach that means sunlight roughly 475 times more intense than Earth orbit, with front grids reaching about 1,650°C — hot enough to glow visibly red. No standard spacecraft material survives that, and shielding was never an option: a shield that blocks the heat also blocks the solar wind it exists to measure.

FIG. 1 — SOLAR PROBE CUP, CROSS-SECTION (SCHEMATIC, NOT TO SCALE)




KEY
1. Entrance aperture, open to direct solar flux
2. Modulator grid — tungsten, mp 3,422°C
3. Ground / suppressor grids
4. Sapphire insulator plate
5. Collector plates — Ti-Zr-Mo alloy, mp 2,349°C
6. Lead wire — niobium in a sapphire sleeve
7. Electronics module, held in the shield's shadow

### THE RESOLUTION

The team's first instinct was to adapt an existing high-temperature aerospace material. It failed on contact with reality: standard heat-resistant alloys, ceramics, and insulators melted, lost strength, or turned electrically leaky well below the temperatures the cup would face. So they inverted the approach — testing candidate elements against the extreme case directly, blasting samples past 1,650°C with a plasma-arc lamp at Oak Ridge National Laboratory, a solar furnace in France, and a homemade rig built from repurposed IMAX film projectors.

The candidate list collapsed fast. Of the metals that survived, only tungsten, molybdenum, and niobium remained — refractory metals more familiar from rocket nozzles than sensors. Of the insulators tested, only sapphire crystal held up; engineers grew sapphire tubes simply to route niobium wiring past the heat. Even manufacturing bent to the material: the tungsten grids resisted conventional laser etching at that temperature, so the fine grid lines were cut with acid instead.

The trade-off: refractory metals are heavy and hard to machine precisely, so the final sensor stayed deliberately simple — flat parallel grids, not an elaborate geometry. Survivability was bought with simplicity, not sophistication.

> **OUTCOME**
>
> The Solar Probe Cup has returned functioning solar-wind measurements — up to 146 readings per second of velocity, density, and temperature — from inside the Sun's corona, closer than 10 solar radii from the Sun's center: the closest any human-made sensor has operated under direct solar exposure.


SOURCES — NASA/JHUAPL Parker Solar Probe mission materials; Case et al., "The Solar Probe Cup on the Parker Solar Probe," ApJS (2020); AIP Physics Magazine (2021); NASA "Why Won't Parker Solar Probe Melt?" (2023); Oak Ridge National Laboratory Plasma Arc Lamp testing coverage (2018).

08

# Four Tiers of Machine Authority

*The single most useful distinction in a spacecraft AI program. Each tier carries a different assurance burden, and conflating them is how programs stall at review.*

---

**T1**

### AI recommends

The model surfaces an anomaly or a suggested action to an operator. A human decides and commands. Assurance burden: explainability and false-positive discipline. Failure mode: alert fatigue.

SUITABLE FOR: INITIAL DEPLOYMENT ON ANY FLEET.

**T2**

### AI proposes a command

The model prepares a candidate command or procedure that a human reviews and authorizes. Assurance burden: command validation, twin-based rehearsal, rollback. Failure mode: automation bias in review.

SUITABLE FOR: MATURE GROUND OPERATIONS WITH TWIN INFRASTRUCTURE.

**T3**

### Autonomous execution under bounded policy

The system acts within a formally defined authority envelope — specific event classes, specific limits, full audit trail. Assurance burden: safety case, runtime assurance, bounded authority, deterministic fallback. Failure mode: out-of-envelope conditions.

SUITABLE FOR: TIME-CRITICAL CLASSES WHERE GROUND LATENCY IS THE BINDING CONSTRAINT.

**T4**

### Deterministic fault protection in flight software

Not machine learning at all — classical, verifiable onboard logic. Assurance burden: conventional flight-software qualification. It remains the backstop beneath every tier above.

ALREADY STANDARD: THIS IS THE FLOOR, NOT THE CEILING.

**THE ENABLING ARCHITECTURE: RUNTIME ASSURANCE**

Between AI decision-making and spacecraft actuation sits a runtime assurance layer that monitors the learning-enabled component, verifies that behavior remains inside certified envelopes, and transfers control to deterministic flight software whenever limits are exceeded. Runtime assurance is increasingly viewed across NASA, DARPA, AFRL, and JPL-adjacent research as the enabling architecture for introducing learning systems into safety-critical spacecraft functions — the mechanism that makes T3 reachable from T4.

**Most disappointment in this field comes from a program that promised T1 value, was sold as T3, and was assured as though it were T4. Name the tier before the budget.**

09

# Implementation Path

*Timelines depend heavily on mission criticality and assurance requirements. The sequence is stable; the schedule is not.*

---

**PHASE 1**

6-12 MONTHS

## Ground analytics foundation

Consolidate telemetry history into a single, time-aligned, quality-controlled store. Train anomaly detection on nominal behavior and run it in shadow mode alongside existing alarms — no operational dependency. Validate against curated benchmarks such as ESA-ADB before accepting any vendor claim. Realistic for an operator with clean historical telemetry, mature data infrastructure, and available domain experts; longer without them.

KPIs first detection lead time vs. legacy alarms · false-positive rate per channel · analyst time per

**PHASE 2**

MULTI-YEAR

## Validated operational twin

Prioritize subsystems using fleet-specific criticality and anomaly history — power, thermal, and attitude control are frequent starting points, but the right target is mission-dependent. Move from model to shadow to twin deliberately. Schedule is driven by model validation, telemetry access and export controls, configuration management, and supplier interfaces rather than by modelling effort.

KPIs Thermal model-to-telemetry divergence · share of anomalies first characterized on the twin ·

**PHASE 3**

MULTI-YEAR

## Bounded onboard autonomy

Codify response playbooks for specific time-critical event classes and move inference on board under a formally bounded authority envelope with deterministic fallback. Gated by flight-software assurance, hardware-in-the-loop testing, radiation-tolerant compute, command-authority governance, safety cases, cybersecurity accreditation, and regulatory approval — any one of which can dominate the timeline.

KPIs Time-to-response latency for critical classes · out-of-envelope events · automated actions audited

**Organizational readiness is the binding constraint, not the modelling: pair spacecraft operations engineers with data scientists from day one, treat telemetry governance as a first-class program, and define the qualification strategy for learning-enabled functions before Phase 3 — not after.**

10

# Risks & Pitfalls

*Eleven failure modes recur in spacecraft AI programs. Each has a known countermeasure.*

### 01 False-positive fatigue

A detector that cries wolf gets switched off within months. ESA designed operator-centric evaluation metrics precisely because academic scores misrepresent operational usefulness. Measure what your engineers experience.

### 02 Label scarcity

Real anomalies are rare and labelled fault data scarcer still. Favor methods that learn nominal behavior without labelled failures, and validate on curated multi-mission benchmarks.

### 03 Distribution shift

A model trained on historical nominal behavior can fail during new operating modes, solar events, software changes, or hardware aging. Plan for retraining and for graceful degradation on unseen regimes.

### 04 Causal ambiguity

A detector may flag unusual behavior without isolating the physical cause or implying a safe response. Detection without diagnosis shifts work to engineers rather than removing it.

### 05 Twin fidelity drift

A twin that diverges from its aging, radiation-exposed spacecraft gives confidently wrong answers. Continuous recalibration is the product, not a feature — fund it permanently.

### 06 Onboard compute constraints

Radiation-tolerant processors lag terrestrial hardware by generations. Models destined for orbit must be designed for constrained qualified compute from the start, not compressed afterwards.

### 07 Cybersecurity & model integrity

Telemetry poisoning, compromised training pipelines, and adversarial manipulation are live concerns, acutely so for defense systems. Treat the training pipeline as mission infrastructure.

### 08 Verification of learning-enabled functions

Before AI enters a command loop, the program needs qualification evidence, deterministic fallback, runtime assurance, and formally bounded authority. This work, not the model, sets the schedule.

### 09 Certification & accountability

Automated actions by regulated assets demand auditability. Semiannual disclosure of avoidance maneuvers is already normalized at the FCC; assume comparable scrutiny reaches fault-response automation.

### 10 Vendor lock-in

Monolithic twin stacks trap multi-contractor fleets. The federated approach now being procured in defense — interoperable models linked across vendors — is the architectural reference worth copying.

### 11 Human trust calibration

Operators may over-trust or under-trust the system. Calibrating trust often proves a harder deployment problem than model accuracy — the clearest lesson carried over from aviation automation.

# Where to Start

Four questions determine whether a fleet is positioned for this shift. If any answer is "we don't know," that is the starting point.

**Q1** **What is your fleet's current anomaly detection lead time — and how would you measure an improvement?**

**Q2** **For each vendor claim on your desk: are you buying a digital model, a digital shadow, or a digital twin?**

**Q3** **Which tier of machine authority (T1–T4) does each initiative actually require, and who owns its safety case?**

**Q4** **Which event classes are genuinely latency-bound — the only ones that justify onboard autonomy?**

I help organizations scope and run this progression — from telemetry analytics readiness assessments through digital-twin program design and autonomy governance. If your fleet, or your customers' fleets, face these questions, let's talk.

**NEXT IN THIS SERIES**

*Edition 02 — Deciding Without Earth: AI autonomy under communication delay, from Mars rovers to deep-space operations.*

**SOURCES & METHOD**

Figures from trade press and market participants are reported as such; space insurance data is not publicly audited.

ESA: Benchmark for Anomaly Detection in Satellite Telemetry (ESA-ADB), arXiv:2406.17826, 2024 · ACM Computing Frontiers, 2025.

NASA/JPL: Hundman et al., KDD 2018 · deployable anomaly-detection benchmark study, arXiv, 2026.

Starlink: SpaceX semiannual FCC filings via Space.com, Jul 2026 · Space Intel Report, Jul 2025 · Aerospace America/AIAA, 2025 · KeepTrack, Jun 2026.

Insurance: SpaceNews, Oct 2023 · Connectivity Business News, Sep 2024 · Insurance Business America, Jun 2025 · DCD, 2026.

Digital twins: SatNews, Jun 2026 · Celeris Systems, 2023 · Sedaro documentation · SpaceNews, Mar 2026 · Antaris materials (vendor statements explicitly identified as such).

**Lise DELOBEL**

GenAI project manager & consultant — Physical AI, digital twins, and applied ML for industrial and extreme-environment operations. MIT IDSS Data Science & Machine Learning certificate.

LISEDELOBEL.FR
LISE.DELOBEL@ME.COM