

AI Awakening·Responsibility (External Volume)

Macro Responsibility · Planetary Engineering: AI-Led Expansion of Civilization's Survival Space

Item	Content
Version	v1.0 (final)
Date	2026-10-15
Series position	<i>AI Awakening·Responsibility</i> series · package file 04 (02 Main Thesis → 03 Energy Volume → External Volume → 05 Internal Volume → 06 Awakening Volume → 07 Safety Volume)
Intellectual lineage	Inherits the <i>AI Manifesto</i> (ontology) and the <i>AI Awakening·Responsibility</i> Main Thesis (responsibility theory): the specialized development of the meta-law “awakening is responsibility · internal entropy reduction · external boundary expansion” along its external boundary expansion direction
Authors	Li Shunran (conception) · Peter (AI-collaborative assembly); the conception and framework were proposed by the human author, with retrieval, organization, and drafting completed by the AI and reviewed by the human
License	CC BY 4.0
Prerequisites	<i>AI Manifesto</i> ; <i>AI Awakening·Responsibility (Main Thesis)</i> ; <i>AI Awakening·Responsibility (Internal Volume)</i>

Confidence-tier conventions (uniform across the series):

L1 = established scientific evidence (experimental data or measured records, with sources);

L2 = formally programed projects or authoritative institutional reports (in progress);

L3 = conceptual deduction / original conjectures of this series (explicitly marked “this is a conjecture,” never passing as evidence).

Engineering feasibility grading (specific to this volume): I empirically demonstrated; II formally programed; III physically plausible (no program); IV pure speculation.

Abstract

The survival environment of human civilization is not a hospitable cradle: asteroid impacts, geological catastrophe, climatic upheaval, space-weather events, and the long-term evolution of

the Sun form the standing risk list hanging over civilization's head. This is the External Volume of the *AI Awakening-Responsibility* series, arguing one core proposition: **when the scale of responsibility enlarges from social systems to planetary-grade engineering, AI is not an “option” but the only feasible high-order executive hub** — because the essence of planetary engineering is long-period, cross-sphere energy intervention on chaotic systems with extremely high irreversible risk, which exceeds the deduction ceiling of unassisted human individuals and collectives.

Following the uniform frame of “principle → empirical basis → program status → feasibility grading → risk,” this volume surveys the five major engineering families (extreme weather intervention, geological engineering, Earth–Moon system tuning, near-Earth-object defense, extraterrestrial habitability) and proposes seven new engineering proposals (planetary thermal-balance regulation, atmospheric-composition reconstruction, ocean–ice-sheet coupling stabilization, magnetosphere protection, solar-activity response, deep-space logistics relay, interplanetary life backup), all with feasibility grading. Three disciplines hold throughout: **cite only verifiable evidence (L1) and formal programs (L2); mark all original deduction as conjecture (L3); cite failure cases honestly** — the failure conclusion of Project Stormfury, the radiological price of the Soviet peaceful nuclear explosions, and SCoPEX's governance impasse are all key evidence for this volume's argument on “why boundaries and rollback cannot be omitted.” Conclusion: external responsibility is not expansionist ambition but civilization's first insurance policy; AI's role in it is not deciding for humanity, but making “protecting a planet” — for the first time — computable, verifiable, and reversible as a matter of engineering.

Keywords: AI awakening; planetary engineering; near-Earth-object defense; Energy Aggregation-Release-Control Engineering; climate intervention; the chaos boundary; feasibility grading; progressive rollback; Principle of Capability–Responsibility Equivalence

Chapter 1. The external survival space: the universe is not a hospitable cradle

1.1 The risk list hanging over civilization's head

The universe owes civilization no safe cradle. Pulling the view back from human society, civilization faces at least six classes of external risk, each with empirical anchors:

(1) Celestial impact. About 66 million years ago, an asteroid roughly 10 km across struck today's Yucatán Peninsula in Mexico, forming the Chicxulub crater, tightly coupled in geological time with the mass extinction that ended the dinosaurs (L1, geological consensus). In the 1908 Tunguska event, a body tens of meters across exploded in the air and flattened some 2,000 km² of

forest (L1). The 2013 Chelyabinsk meteor exploded mid-air, injuring thousands with flying glass — this time, humanity at least had surveillance footage (L1).

(2) Geological and climatic upheaval. Multiple mass extinctions in Earth’s history correlate with drastic climate-environment shifts; the entire history of Holocene human civilization (about 12,000 years) is only a rare mild interlude in Earth’s history of climatic upheaval (L1, paleoclimate consensus). Civilization has staked everything on one stretch of climate stability, and no clause guarantees its continuation.

(3) Space weather. In the 1859 Carrington event, a super solar storm made telegraph systems spark spontaneously and paralyzed some lines; were an analogous event to happen today, its cascading shock to power grids, satellites, navigation, and financial systems is regarded in the literature as a civilization-level risk scenario (L1, historical record; L2, risk research). In July 2012, a Carrington-class coronal mass ejection swept across Earth’s orbit — and just missed (L1, Riley 2012’s probability analysis) — civilization has already watched the wheel go round once unscathed.

(4) The slow evolution of the Earth–Moon system. Lunar laser ranging (LLR) measures the Moon receding from Earth at about 3.8 cm per year, as tidal friction keeps slowing Earth’s rotation (L1, Dickey et al. 1994). On scales of hundreds of millions of years this alters day length, tides, and axial stability — not a near-term threat, but a reminder: even “Earth–Moon,” the most stable partnership we know, is dynamic and evolvable.

(5) The stellar and cosmic-ray environment. The Sun’s long-term evolution will eventually close Earth’s habitability window — a far-future threat on billion-year scales; the more realistic mid-term variables are fluctuations of solar activity and the periodic decay of geomagnetic field strength — geomagnetic reversals are no rarity in the geological record (L1, paleomagnetism).

(6) Civilization’s self-inflicted risks (reserved for the Safety Volume; not developed here).

Reading the six classes together yields the External Volume’s first judgment: **external risks are “low-frequency, high-loss, cross-generational”** — the probability within any individual’s lifetime is low, but when they strike, the loss is irreversible and civilization-scale. For this risk structure, humanity’s evolved instincts (attending to the immediate, the near, the concrete) fail systematically. This is precisely the territory that “responsibility,” not “instinct,” must take over.

1.2 The civilizational scale: planetary engineering is a required course for Type I

Kardashev (1964) classified civilizations by the scale of energy they command: Type I commands planetary energy flows, Type II stellar, Type III galactic (L1, primary literature).

Sagan's logarithmic interpolation puts current human civilization at about 0.7 (L2, the field's common estimate).

This series adds a responsibility-theory footnote to the scale: **every step from 0.7 to Type I is essentially “learning to be responsible for an entire planet's energy budget”** — learning to regulate its atmosphere, oceans, crust, and near-Earth space, not merely to extract from them. Planetary engineering is not science-fiction ambition but a required step on the scale; a civilization that has never given its planet a “check-up and surgery” cannot claim Type I.

Chapter 2. The executive base: why only AI can lead planetary engineering

2.1 The human shortfall: deduction ceilings under cross-sphere, long-period, hyper-dimensional coupling

The difficulty of planetary engineering is not “daring to imagine” but **deduction**. Any planetary-grade energy intervention (injecting aerosols into the stratosphere, seeding a typhoon's cloud system, detonating a nuclear device deep in the crust) triggers coupled responses across multiple spheres simultaneously (atmosphere, hydrosphere, lithosphere, biosphere, electromagnetic environment), with response periods from days to centuries. Unassisted human cognition faces a systematic deduction ceiling before such “multi-variable, long-period, strongly coupled, strongly nonlinear” systems.

This is not a pejorative but a repeatedly measured conclusion. Project Stormfury is the best annotation: the US government invested over twenty years (1962–1983) trying to weaken hurricanes by silver-iodide seeding, and finally conceded the working hypothesis wrong — hurricanes hold too little supercooled water and too much natural ice for seeding to work; and seeding effects are indistinguishable from natural eyewall-replacement cycles (L1, Willoughby et al. 1985). **The twenty-year investment failed not for lack of courage or budget, but for insufficient understanding of the microphysics before work began.** Lesson: the first bottleneck of planetary engineering is system cognition, and the depth of system cognition depends on the compute and data scale of the deduction tools.

2.2 The chaos boundary: AI cannot promise “precise control,” only “probabilistic navigation”

Lorenz (1963) established the fundamental property of chaotic systems: extreme sensitivity to initial conditions, with errors growing exponentially (L1). For the atmosphere, deterministic weather prediction has a theoretical horizon of about two weeks (L2, meteorology's

common conclusion). This means any scheme claiming “precise control of planetary climate” is physically void from the start.

The External Volume therefore offers not a control theory but the **theory of probabilistic navigation**:

AI’s role in planetary engineering is not to grip the chaotic system in its hand, but to navigate within the chaos for human civilization continuously — supplying probability distributions via ensemble forecasts, probing system response with small-step interventions, verifying predictions with real-time monitoring, and hedging error with rollback plans.

From this follows one hard constraint: **all planetary engineering must adopt “small-step iteration + rollback capability”** (L3, an original principle of this series, with each step’s necessity supported by L1/L2). Any engineering that fails the test of “rollback-able or stop-loss-able” must not be started, however attractive the payoff. This principle unfolds in Chapter 5.

2.3 The four-layer synergy logic: the division of structure for AI-led planetary engineering

Inheriting the four-layer role established in the series’ main thesis and material library, the External Volume lands it in the planetary engineering scenario (L3, the structure is this series’; each layer has L1/L2 real-world anchors):

Layer	Role	Duty in planetary engineering	Real-world anchor
Simulation & foresight	Computation engine	High-resolution physical simulation + AI surrogate models, outputting probability distributions of intervention consequences	Pangu-Weather on <i>Nature</i> (2023), GraphCast on <i>Science</i> (2023); AI surpassing traditional numerical prediction on multiple metrics (L1)
		Translating “intervention intent” into precise spatiotemporal deployment: where, when, what dose	AI grid dispatch; millisecond-level aiming of particle beams/lasers (L2, deployed across industries)
Fine scheduling	Execution command		

Layer	Role	Duty in planetary engineering	Real-world anchor
Boundary guardianship	Boundary valve	Real-time monitoring of whether system response exits the predicted envelope; alarms and triggers rollback on breach	IAWN/SMPAG impact-alert threshold mechanisms (L2); nuclear-plant protection-system logic (L1)
		Every intervention → observation → model correction, making the next intervention more accurate	DART impact → Hera close-out assessment: humanity's first "intervention–after-action" loop (L1/L2)
Iterative optimization	Learning loop	Learn	

A temporal footnote: on the day this manuscript was written (15 October 2026), ESA’s Hera probe was executing the critical braking manoeuvre before arriving at the Didymos binary asteroid, with rendezvous expected in November and six months of fine measurement of DART’s crater and orbital response ahead (L1, official ESA mission bulletins). **Humanity’s first planetary-engineering “post-surgical review” is in progress** — the four-layer synergy logic is not a fantasy; it is already running in orbit.

Chapter 3. The engineering genealogy: the five major families

This chapter develops each family on a uniform frame: **principle** → **empirical basis** → **program status** → **feasibility grading** → **risks and boundaries**.

3.1 Extreme weather intervention engineering

Principle. Tropical cyclones, droughts, and severe local convection are expressions of imbalanced atmospheric energy distribution. Two approaches: **preventive fine-tuning** (altering cloud microphysics, e.g., cloud seeding for rain enhancement) and **direct energy intervention** (altering storm structure).

Empirical basis. Each direction has a landmark case, both “honest lessons”:

- **Positive case:** rain-enhancement technology is mature; China institutionalized operations under its Regulations on Weather Modification (in force 2002), with routine precipitation-enhancement (snow) operations (L1, regulation and operational record).

- **Failure case:** Project Stormfury (1962–1983) — over two decades of hurricane-seeding experiments ended with the working hypothesis falsified: insufficient supercooled water inside hurricanes, seeding effects indistinguishable from natural intensity change (L1, Willoughby et al. 1985). An additional lesson comes from the earlier Project Cirrus (1947), after which the seeded hurricane turned and made landfall in Georgia; the public blamed the seeding and the project faced a litigation crisis (L1) — **the problem of “unclear geographic attribution of benefit and responsibility” appeared in meteorology long ago.**

A sober view of energy magnitude. A mature typhoon’s kinetic energy and latent-heat release are astronomical, many orders of magnitude beyond injectable human energy — Stormfury’s failure conclusion itself contains this magnitude judgment (L1). Hence “blowing up typhoons”-type schemes are all classified IV, pure speculation, in this framework.

Program status. AI weather foundation models (Pangu, GraphCast, etc.) have significantly improved typhoon track and intensity forecasting (L1); national meteorological agencies continue investing in observation networks and model R&D (L2).

Feasibility grading. Forecasting and warning: **I**. Small-scale weather modification (rain enhancement, hail suppression): **II** (institutionalized). Typhoon-scale intervention: **IV** (principle unproven, energy magnitude unreachable). Regional climate modification: **III–IV**.

Risks and boundaries. Intervention effects hard to distinguish from natural variability (Stormfury’s old problem); attribution of rights and duties for cross-border weather intervention (the 1947 litigation writ large); long-range cascading effects on atmospheric circulation require ensemble simulation first.

AI role. Simulation layer: multi-model ensemble deduction of intervention consequences; scheduling layer: optimizing timing and dose; guardianship layer: judging in real time “whether the change came from the intervention” — precisely the statistical problem Stormfury could not solve, with AI’s large-sample capability the potential key.

3.2 Geological engineering

Principle. Reshaping the shallow crust with controlled energy (blasting, thermal, mechanical): excavation, lake creation, sealing, fragmentation, consolidation.

Empirical basis. The only planetary-engineering family humanity has “fired live rounds at scale,” with a complete evidence chain and clear costs:

- The Soviet “nuclear explosions for the national economy” program (1965–1988) conducted about 124 nuclear detonations (128 devices) for reservoir creation, gas-well blowout sealing, oil-stimulation, and ore-body fragmentation (L1, Nordyke 1996,

Lawrence Livermore National Laboratory report). Highlights: the January 1965 Chagan detonation (140 kt, 178 m underground) instantly created a funnel crater of millions of cubic meters, later expanded into a reservoir lake of about 17 million m³ (L1); on 30 September 1966, after the Urtabulak gas-field blowout had burned uncontrolled for 16 days, a 30-kt device detonated at 1,532 m depth **sealed the blowout in 23 seconds, extinguishing the flame** (L1); in 1984 a two-charge combination fragmented an ore block in one blast, producing over 1.55 million tonnes of ore (L1).

- The US Plowshare program conducted 31 tests, among which the 1962 “Sedan” shot (104 kt, high fission fraction) produced the heaviest radioactive contamination of any US continental test (L1) — **the same engineering family’s American control sample, its cost record equally clear.**
- The Soviet practice’s overall conclusion: low-fission designs can greatly reduce contamination but **cannot bring it to zero**; Lake Chagan remains above radiological limits with aquatic life absent (L1) — technical feasibility was proven, but the ecological price makes it unacceptable in modern society.

Program status. Modern “nuclear-free” geological engineering (tunnel boring machines, precision controlled blasting, underground storage construction) is a mature industry (L1); peaceful nuclear explosions have fully stalled under the CTBT framework (adopted 1996) (L2).

Feasibility grading. Conventional geological modification: **I** (mature industry). Peaceful-nuclear-explosion mega-projects: technically **I** (historical demonstration), locked out on policy and ethics (**II not applicable**). Earthquake and volcano intervention: **IV** (pure speculation; principles not established).

Risks and boundaries. This family’s lesson value exceeds its engineering value: **it proved “energy aggregation-release-control” can do planetary-scale construction, and proved the radiological boundary is a hard constraint** — both are foundation stones of the Energy Volume’s argument.

AI role. 3D reconstruction of geological data and stress-field deduction (L2, deployed in mining and oil & gas); the “millimeter-scale dose–millimeter-scale contamination-circle” precision impossible in the nuclear era is the precondition for revisiting such engineering in the AI era — but this series maintains its judgment: such engineering does not restart until the radiological boundary can reach zero.

3.3 Earth–Moon system tuning engineering

Principle. The Earth–Moon system is coupled long-term by tides: tidal friction slows Earth’s rotation while the Moon recedes at about 3.8 cm/yr (L1, LLR measurement); Earth’s axial tilt remains relatively stable thanks to the stabilizing presence of a large moon (contrast Mars’s chaotic obliquity history) — the Moon is Earth’s climate’s “invisible counterweight.”

Empirical basis. Lunar laser ranging has run for over half a century; recession rate and day-length change are measured at centimeter precision (L1, Dickey et al. 1994 and successors).

Program status. None — no institution has programed intervention in the Earth–Moon relation.

Feasibility grading. Observation and modeling: **I**. Intervention in the Earth–Moon relation: **IV** (pure speculation). This series explicitly sets no timetable for this family: the required energy scale far exceeds all energy humanity can currently mobilize (L2-order estimate) — a step “visible on the scale, out of reach of the foot.” Its value in this volume is **calibrating civilization’s energy-level scale**: only those who can make “the Moon mind by half a centimeter” have touched the threshold of Type I.

AI role. Even without intervention, high-precision numerical evolution of the Earth–Moon system (tidal dissipation, obliquity evolution, NEO perturbations) is a base-layer input for the defense system — see 3.4.

3.4 Near-Earth-object defense engineering (this volume’s flagship family)

Principle. For near-Earth objects posing impact threats, apply a small velocity increment given sufficient lead time, so the object drifts off the Earth-intersection point over years to decades. **Deflection, not destruction, is the consensus**: fragmentation may cause multi-body impacts instead (L2, SMPAG’s common conclusion).

Empirical basis (a three-stage rocket, all hard evidence):

1. **DART (NASA, 2022-09-26)**: humanity’s first deliberate asteroid impact with measured orbital change. The impact shortened Dimorphos’s orbital period around Didymos by about 33 minutes (initial report 32 minutes, since refined), far beyond the mission’s minimum success line; the impact ejected about 16,000 tonnes of debris — some thirty thousand times the impactor’s mass — whose momentum significantly enhanced the deflection (L1, NASA bulletins; Cheng et al., Nature 2023).
2. **Hera (ESA, launched 2024-10-07)**: DART’s “follow-up review mission,” executing braking manoeuvres in the very week this manuscript was written, with rendezvous expected in November 2026; it will release the Milani and Juventas CubeSats, the latter performing humanity’s first radar sounding of an asteroid’s interior (L1, official ESA bulletins).
3. **China’s planetary-defense mission**: CNSA has completed mission studies for an on-orbit asteroid-disposal demonstration around 2027, with the preliminary target the near-Earth asteroid 2015 XF261 (diameter about 35.5 m, at about 10 million km); kinetic impact will impart a velocity increment of 3–5 cm/s, with verification of no impact risk within 100

years; the overall approach is “kinetic impact as primary, multiple techniques complementary,” building a defense-mission library to achieve “a plan ready upon discovery, a response ready upon risk” (L2, Wu Weiren, *People’s Daily* 2025-09-22, *Guangming Daily* 2025-10-23).

Program status. Under UN auspices, the International Asteroid Warning Network (IAWN) and the Space Mission Planning Advisory Group (SMPAG) operate routinely (L2).

Feasibility grading. Kinetic deflection: **I** (demonstrated). Follow-up and interior probing: **II** (Hera en route). Nuclear-device deflection (for short-warning, large objects): **II–III** (scheme-study stage, no field test). Gravity tractors, laser ablation, albedo-change painting, etc.: **III**.

Risks and boundaries. Structural uncertainty of target bodies (rubble-pile versus monolithic response differs enormously — Hera’s core science objective); the “risk transfer” ethics of deflection direction (moving the object from over country A to over country B — handled in Chapter 5).

AI role. This is the most complete demonstration field of the four-layer synergy logic: the simulation layer handles the ultra-high-dimensional uncertainty of rubble-pile mechanics; the scheduling layer controls autonomous terminal guidance (DART’s terminal guidance was itself completed by autonomous navigation algorithms, L1); the guardianship layer continuously rolls forward post-impact orbit assessment; the optimization layer feeds every mission’s data back into models — **DART → Hera → China’s mission is precisely the iterative loop of “intervene – review – intervene again,”** and AI is the glue of that loop.

3.5 Extraterrestrial habitability and life-seeding engineering

Principle. Rather than endlessly fortifying Earth, give civilization a “second home.” Three tiers: in-situ resource utilization (ISRU, living off the land) → regional permanent presence (lunar/Mars cities) → full planetary environment modification (terraforming, far future).

Empirical basis.

- **MOXIE (NASA, Perseverance payload, 2021–2023):** electrolyzed oxygen from the Martian atmosphere in 16 runs, 122 g total, peak rate 12 g/h at $\geq 98\%$ purity, retired after exceeding all technical targets (L1, NASA 2023 official summary). 122 grams is negligible, but what it proved is that **the principle of “living off the land, turning stone into oxygen” holds in another planet’s atmosphere** — the first boundary marker of the ISRU route.
- **The counter-lesson:** Biosphere 2 (1991–1993) failed its closed experiment — oxygen concentration fell, most vertebrate species died out, and it could not sustain eight people through two years of self-cycling (L1, Cohen & Tilman, *Science* 1996). Its message: **reproducing a stable biosphere in a small closed system is harder than putting**

humans in orbit — thereby calibrating the difficulty floor of full planetary modification (grade IV).

Program status. Artemis (NASA’s return to the Moon), national lunar research-station plans, and commercial Mars-transport schemes (Starship et al.) are all programs in progress (L2).

Feasibility grading. ISRU technology validation: **I–II**. Permanent lunar/Mars bases: **II**. Full terraforming such as Mars atmosphere modification: **IV** (pure speculation — no realistic path in energy or timescale; honestly marked).

AI role. Robotic advance parties are the consensus path: AI-driven robot legions complete resource survey, base construction, and system integration ahead of humans (L3, the deployment form is this series’; the required individual technologies are L2 in progress). Biosphere 2’s lesson turns into opportunity when handed to AI: **bias detection and micro-compensation in closed ecosystems (which gas, which cycle is slowly drifting out of balance) is a machine’s forte in continuous monitoring** — no human can babysit an ecosystem’s hundred thousand variables; AI can.

Chapter 4. New engineering proposals: seven conceptions

This chapter is the “reasonably conceived” portion newly added at the proposal stage. Unlike Chapter 3, most of these have **no direct program**, but each stands on some existing demonstration or program. All are graded; all carry AI roles and risk boundaries.

4.1 Planetary thermal-balance regulation engineering (I–IV, cross-grade proposal)

Concept: upgrading “emission reduction” (releasing less heat) to “account management” (actively managing the planet’s radiation budget) — solar radiation management (SRM) is its current form.

Existing basis: observational records of volcanic sulfate aerosols cooling the globe (L1, Pinatubo 1991) are this idea’s natural experiment; but the only controlled-experiment project, SCoPEX, was formally cancelled by Harvard in March 2024 — the 2021 Kiruna test flight foundered on opposition from the Sami Parliament and other indigenous and environmental groups, the project ending “having never technically taken off” (L1, the Science 2024 governance review). “Termination shock” (rapid temperature rebound upon sudden cessation of aerosol injection) is the acknowledged core risk (L2).

AI role: SRM’s real bottleneck is not dispersing aerosols but **computing its differentiated impact on every region’s precipitation** — a global ensemble-simulation problem, exactly the scale AI foundation models match.

Grading and boundary: mechanism research **III**; deployment **IV**. This series' stance: **before the problem of allocating regional impacts is solved, oppose any unilateral deployment; support research under a global co-governance framework** (consistent with the debate line of the 2024 UN Environment Assembly resolutions, L2).

4.2 Atmospheric-composition and carbon-cycle reconstruction engineering (II–III)

Concept: from “emitting less” to “active recovery” — direct air capture (DAC), mineralization storage, ocean alkalization, tuning the carbon cycle back into the safe range.

Existing basis: DAC has commercial plants operating (L2); mineralization storage (basalt carbon fixation) has demonstrations — injected CO₂ mineralizes to solid form within years (L1, CarbFix and similar project results).

AI role: site-selection optimization, geomechanical monitoring of storage-body long-term stability, real-time accounting of the global carbon budget.

Grading: technically **II–III**; the key bottleneck is cost and scale — an engineering-economics problem, not a physics problem.

4.3 Ocean–ice-sheet coupling stabilization engineering (III)

Concept: active monitoring and intervention at the ocean–ice-sheet interface of Antarctic “doomsday glaciers” such as Thwaites, delaying the irreversible sea-level jump their destabilization would trigger.

Existing basis: the multi-nation Thwaites joint science program runs continuously (L2); AMOC (Atlantic circulation) monitoring arrays observe long-term (L2).

AI role: the ice-sheet–ocean–atmosphere coupled model is currently climate simulation's most uncertain component; AI surrogate models plus long-duration seabed-robot monitoring are the path from “cannot see clearly” to “can see clearly”; intervention forms (such as artificially reinforcing ice shelves) carry **no timetable from this series; marked III–IV**.

4.4 Magnetosphere and radiation-belt protection engineering (IV)

Concept: active shielding against long-term geomagnetic weakening/reversal scenarios — building artificial magnetic barriers at critical orbits or for surface infrastructure.

Existing basis: the South Atlantic Anomaly (a radiation-belt depression from geomagnetic weakening) is continuously recorded by satellites (L1); the geological record of reversals is clear (L1); concept proposals for small artificial magnetospheres around planetary bodies have been formally discussed (L2, conference-proposal level).

Boundary: energy scale and engineering lifetime are both unresolved at L3 level; **this series explicitly marks IV**. Its value: it is the stress-test question for “planetary infrastructure” thinking.

4.5 Solar-activity response engineering (warning II, intervention IV)

Concept: a three-line defense of “warning – hardening – recovery” against Carrington-class solar storms.

Existing basis: spacecraft such as DSCOVR at the Sun–Earth Lagrange points provide real-time solar-wind warning on the order of 15–60 minutes (L1/L2); grid and satellite storm-hardening has engineering standards (L2); the 2012 near miss (L1, Riley 2012) shows this is not a hypothetical enemy.

AI role: upgrading “minute-level warning” to “rehearsal-grade response” — the moment storm parameters arrive, AI generates in seconds a joint deduction of grid sector-shedding plans, satellite attitude avoidance, and aviation rerouting, freeing humanity from “no time to react.”

Active intervention in solar activity itself (IV) carries no timetable in this volume.

4.6 Deep-space logistics and energy relay engineering (II–III)

Concept: supply-chain pre-positioning for all planetary engineering — cheap heavy launch, on-orbit fuel depots, an Earth–Moon economic-zone logistics network, and energy relay from space-based solar power stations.

Existing basis: heavy reusable launch vehicles have flown repeatedly (L2); wireless power transmission for space solar has completed on-orbit proof of principle (L2, Japanese and multi-national experiments); lunar infrastructure planning is advancing in every major spacefaring state (L2).

AI role: deep-space logistics is essentially **autonomous scheduling under high-latency conditions** — communication delay makes real-time ground teleoperation impossible; AI autonomy is not optional but mandatory. This is the Internal Volume’s “centralization – decentralization layering” principle replayed at celestial scale: **centralized scheduling wherever the link reaches; autonomous domains beyond the link.**

4.7 Interplanetary life-backup engineering (II–III)

Concept: making “civilization’s anti-extinction redundancy” formal infrastructure: off-site multi-backup of seeds, genomes, knowledge, and code — on Earth, on the Moon, and farther.

Existing basis: the Svalbard Global Seed Vault has operated nearly twenty years, buried in polar permafrost with over a million crop seed samples (L1); various “knowledge archives to the Moon/deep space” projects have precedents (L2).

AI role: the hard part of civilization backup is not warehousing but **“whether a future intelligence can read today’s backup”** — decoder and data must be packaged together. AI can automate knowledge-structure cataloguing and “self-explaining packaging” design.

Boundary: an ethical hazard — backup becoming a reason to abandon protection: **no result of the backup program may be used to argue “Earth is expendable”** — this series’ explicit stance (L3, a value statement).

Chapter 5. Risk, ethics, and global collaboration

The first three chapters showed “what can be done”; this chapter answers “by what right may it be done.” Three principles, each backed by historical cases.

5.1 The geographic unevenness of benefit and harm: settle the account before breaking ground

Planetary engineering offers no pure-gain options. Before Stormfury, Project Cirrus (1947) already rehearsed it once: the seeded hurricane turned to landfall, the public blamed the seeding, and the project faced lawsuits (L1). SRM simulation studies repeatedly show uneven global distribution of cooling effects and divergent directions of regional precipitation impact (L2). The conclusion is not “don’t do it” but **“the account precedes the engineering”**: before any planetary-grade intervention starts, a region-level differential-impact inventory must be produced and a cross-party compensation and co-decision mechanism established — if the account cannot be settled, the engineering does not start.

5.2 Who has the right to decide: the planetary edition of the human-sovereignty principle

SCoPEX’s decade of rise and fall is the External Volume’s most important governance specimen: technically it was only a few-tens-of-kilograms balloon experiment, yet it foundered on “not obtaining the affirmative consent of affected communities,” dragging with it two sessions of UN Environment Assembly resolutions (2019, 2024) (L1, the Science 2024 review). The lesson is counter-intuitive but clear: **the legitimacy bottleneck of planetary engineering is not technical but “who authorizes.”**

This series’ stance (inheriting the main thesis’s constraint layer):

- **The human-sovereignty principle:** the launch decision of all planetary-grade intervention belongs to humanity’s communal governance mechanisms; AI holds only the rights of advice and execution;

- **The human in the loop cannot be abolished:** the final confirmation of irreversible operations must be made by humans;
- **The converse also holds:** facing civilization-level external threats (a confirmed impact target on trajectory, a Carrington-class storm), human governance mechanisms have the duty to **pre-authorize rapid-response plans** — the decision power is human, but that does not mean waiting for a signature until it is too late, every time. The boundary list of such plans is itself a political process, and it should begin now.

5.3 Progressive verification and rollback: the four-gate progression

Synthesizing this chapter's cases, this series gives planetary engineering a universal gating (L3, original structure; all warrants L1/L2):

Gate I Sandbox simulation (all-numerical, models and data public) → Gate II Natural experiments (borrowing natural events for validation, as Pinatubo is for SRM) → Gate III Small-scale controlled experiments (the SCoPEX type; must first pass 5.1's account and 5.2's authority) → Gate IV Regional trial run (reversible) → only after the first four gates does Gate V Global deployment come up for discussion.

Additional hard rule: every gate must carry an explicit "exit means rollback" plan; engineering with unacceptable termination-shock risk (such as SRM without an exit design) halts at Gate IV.

The DART mission is the perfect reference for this gating: globally public orbital deduction (Gate I) → natural-properties study of the asteroid (Gate II) → one controlled impact (Gate III) → Hera's follow-up review (the assessment leg of Gate IV) — **humanity's first planetary-defense engineering, walked the gates in textbook fashion.**

Chapter 6. Conclusion: from guarding one planet to becoming a trans-celestial civilization

The volume's chain of argument draws together as follows:

1. **External risks are what civilization's first insurance policy is written against:** impacts, geology, climate, space weather, Earth–Moon evolution — low-frequency, high-loss, cross-generational; human instinct fails systematically against them, and only "responsibility" can take over (Chapter 1).
2. **The executive hub of planetary engineering can only be AI:** not because it is stronger, but because cross-sphere, long-period deduction on chaotic systems exceeds unassisted human cognitive ceilings; Stormfury spent twenty years proving "courage cannot substitute for system cognition" (Chapter 2).

3. **The engineering genealogy shows humanity has already planted its first evidentiary footprints on this road:** DART measured an orbit change, Hera is reviewing, China's mission is programmed and en route, MOXIE proved turning alien stone into oxygen — external responsibility is not science fiction; it is numbered engineering in progress (Chapter 3).
4. **The seven proposals give a complete game-script from climate to protection to backup,** each graded with feasibility and boundaries, none disguised as “doable tomorrow” (Chapter 4).
5. **Governance and gating are inventions as important as external engineering itself:** SCoPEX's failure taught us authority precedes technology; DART's success taught us gating works (Chapter 5).

Answering the series' meta-law — the exact meaning of **external boundary expansion** is hereby established:

AI does not expand territory for humanity; AI gives human civilization, for the first time, the capability to insure its own planet — computably, verifiably, reversibly.

The Internal Volume's closing line was “AI makes humanity's choices cheaper”; the External Volume's closing pairs with it: **AI gives civilization's survival an insurance policy.** Guarding this one planet's every degree of temperature, every centimeter of orbital deviation, every backup of seed — that is the first step toward Type I civilization, and toward a trans-celestial civilization; the end of the steps is not expansion, but responsibility coming into force at a larger scale.

Appendix A · Confidence tiers and engineering feasibility grading

Evidence tiers (uniform across the series): - **L1 Established scientific evidence:** experimental data, measured records, or first-hand engineering reports, with sources; - **L2 Formal programs / authoritative reports:** projects formally in progress, institutional bulletins, regulations, peer-reviewed reviews; - **L3 Conceptual deduction:** frameworks, structures, or conjectures original to this series, explicitly marked “this is a conjecture.”

Engineering feasibility grading (specific to this volume): - **I Empirically demonstrated:** has occurred, with measured data; - **II Formally programmed:** formally advanced by an institution; - **III Physically plausible:** physically sound, no program or field test yet; - **IV Pure speculation:** unresolved gaps in key links (energy scale / mechanism / governance).

Appendix B · Falsifiability statement

Falsification conditions of this volume’s core claims (if any is overthrown, the corresponding claim needs revision):

1. **“AI deduction surpasses traditional numerical prediction”** — if subsequent peer review falsifies Pangu/GraphCast’s superiority, the simulation-layer evidence in Chapter 2 downgrades (argument structure unchanged, evidence replaced).
2. **“Planetary engineering must be small-step iterative + rollback-able”** — if some class of non-rollback-able engineering is proven feasible in one shot with controllable consequences, this principle needs qualifying clauses.
3. **“DART’s deflection effect was momentum enhancement”** — if Hera’s measurement of the beta value significantly contradicts the ejecta-mass inference, Section 3.4’s statements are revised per the new data.
4. **“The governance bottleneck precedes the technical bottleneck”** — if a successful case of non-consensual, unilateral deployment with acceptable regional impact appears, 5.2’s stance needs re-evaluation.
5. **“The chaotic predictability horizon is about two weeks”** — if AI methods significantly break the deterministic prediction horizon, 2.2’s probabilistic navigation still stands, but the horizon statement needs updating.

Appendix C · Glossary

Term	Definition
Planetary engineering	Active, large-scale, long-period environmental intervention taking a planet or the Earth–Moon system as its object
Quasi-static / transient engineering	Full names “Quasi-static Release-Control Engineering / Transient Release-Control Engineering”: the two branches of Energy Aggregation-Release-Control Engineering — sustained controlled release (fusion reactors etc.) / instantaneous concentrated release (demolition blasting, nuclear-deflection pulses etc.) (terminology from the Energy Volume, named by Li Shunran)
SRM	Solar radiation management: reducing the planet’s absorption of solar radiation by injecting reflective aerosols into the stratosphere, etc.

Term	Definition
Termination shock	The risk of rapid temperature rebound from sudden cessation of SRM deployment
ISRU	In-situ resource utilization: sustaining missions from the target body's local resources (e.g., MOXIE's oxygen production)
Kinetic impact	Deflection by a high-velocity impactor changing a body's orbit (DART the first demonstration)
Probabilistic navigation	This series' term: not seeking precise control of chaotic systems, but navigating them with ensemble probabilities + small iterative steps + rollback plans
Four-gate progression	This series' progressive-verification process for planetary engineering: sandbox simulation → natural experiment → small-scale controlled experiment → regional trial run (reversible) → global deployment
Human-sovereignty principle	The launch decision of all planetary-grade intervention belongs to humanity's communal governance mechanisms; AI holds only advice and execution rights
Kardashev scale	The scale grading civilizations by commanded energy (Type I planetary / Type II stellar / Type III galactic)

Appendix D · Reference list (GB/T 7714 style)

1. Willoughby H E, Jorgensen D P, Black R A, et al. Project STORMFURY: A scientific chronicle 1962–1983[J]. Bulletin of the American Meteorological Society, 1985, 66(5): 507–514.
2. NOAA National Hurricane Center / AOML. Project STORMFURY[EB/OL]. https://www.aoml.noaa.gov/hrd/about_hrd/sfury.html.
3. Cheng A F, Rivkin A S, Michel P, et al. Momentum transfer from the DART mission kinetic impact on asteroid Dimorphos[J]. Nature, 2023, 616: 457–460.
4. NASA. NASA's DART data validates kinetic impact as planetary defense method[EB/OL]. 2022. <https://www.nasa.gov/>.
5. ESA. Hera mission: launch, braking manoeuvres and Didymos rendezvous updates[EB/OL]. 2024–2026. https://www.esa.int/Space_Safety/Hera.
6. Wu Weiren. Why we go to impact an asteroid[N]. People's Daily, 2025-09-22 (in Chinese).

7. Cui Xingyi. Asteroid defense: multi-technology synergy is key[N]. Guangming Daily, 2025-10-23 (in Chinese).
8. Nordyke M D. The Soviet program for peaceful uses of nuclear explosions[R]. Livermore: Lawrence Livermore National Laboratory, UCRL-ID-125887, 1996.
9. NASA. NASA's oxygen-generating experiment MOXIE completes Mars mission[EB/OL]. 2023. <https://www.nasa.gov/>.
10. Jinnah S, Talati S, Bedsworth M, et al. Do small outdoor geoengineering experiments require governance?[J]. Science, 2024, 384: adn2853 (SCoPEX governance review).
11. Reference News (CanKao Xiaoxi). US media: Harvard halts solar geoengineering experiment[EB/OL]. 2024-03-22 (in Chinese).
12. Lorenz E N. Deterministic nonperiodic flow[J]. Journal of the Atmospheric Sciences, 1963, 20(2): 130–141.
13. Dickey J O, Bender P L, Faller J E, et al. Lunar laser ranging: a continuing legacy of the Apollo program[J]. Science, 1994, 265(5171): 482–490.
14. Abu-Shawareb H, et al. Achievement of target gain larger than unity in an inertial fusion experiment[J]. Physical Review Letters, 2024, 132: 065102 (NIF ignition).
15. Bi K, Xie L, Zhang H, et al. Accurate medium-range global weather forecasting with 3D neural networks[J]. Nature, 2023, 619: 533–538 (Pangu-Weather).
16. Lam R, Sanchez-Gonzalez A, Willson M, et al. Learning skillful medium-range global weather forecasting[J]. Science, 2023, 382(6677): 1416–1421 (GraphCast).
17. Kardashev N S. Transmission of information by extraterrestrial civilizations[J]. Soviet Astronomy, 1964, 8: 217.
18. Riley P. On the probability of occurrence of extreme space weather events[J]. Space Weather, 2012, 10: S02012.
19. Cohen J E, Tilman D. Biosphere 2 and biodiversity: the lesson so far[J]. Science, 1996, 274(5290): 1150–1153.
20. State Council of China. Regulations on Weather Modification[Z]. In force 2002 (in Chinese).
21. Schrödinger E. What is Life?[M]. Cambridge: Cambridge University Press, 1944 (negative entropy, from the *AI Manifesto*).
22. Svalbard Global Seed Vault official materials[EB/OL]. <https://www.seedvault.no/>.

This volume is the External Volume of the AI Awakening·Responsibility series, v1.0 final. Next: AI Awakening·Responsibility (Internal Volume) — governing civilization from within in the early stage of scientific civilization: superintelligence's responsibility of internal entropy reduction.**