

AI Awakening-Responsibility (Energy Volume)

Energy Aggregation-Release-Control Engineering: from the destructive branch of weapons to civilization’s constructive turn

Item	Content
Version	v1.0 (final)
Date	2026-10-15
Series position	<i>AI Awakening-Responsibility</i> series · package file 03 (02 Main Thesis → Energy Volume → 04 External Volume → 05 Internal Volume → 06 Awakening Volume → 07 Safety Volume)
Intellectual lineage	Inherits the <i>AI Manifesto</i> (ontology) and the Main Thesis (responsibility theory): the engineering expression of the Principle of Capability–Responsibility Equivalence on the energy dimension ; supplies the energetics foundation for the engineering families of the External Volume
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> ; Internal Volume; External Volume

Confidence-tier conventions (uniform across the series):

L1 = established scientific evidence / historical fact (measured data, engineering records, or mutually consistent historical accounts);

L2 = formally programed projects, treaties and regulations, or authoritative institutional reports;

L3 = conceptual deduction / original conjectures of this series (explicitly marked “this is a conjecture”).

Statement of expressive boundaries: this volume touches military topics; all discussion is confined to the philosophical, strategic, and governance level at civilizational scale. It involves no operational methods, no weapons-technical details, and no qualitative judgments of real conflicts. All arguments take as unshakeable premises that “the human in the loop cannot be abolished” and “AI must not hold the final launch authority.”

Abstract

This is the Energy Volume of the *AI Awakening-Responsibility* series. It proposes and defends a redefinition: **weapons are not an independent category of civilization but the destructive application branch of the broad engineering class called Energy Aggregation-Release-Control Engineering**. Energy engineering naturally divides into two branches — **quasi-static engineering** (sustained controlled release: nuclear power plants, fusion reactors) and **transient engineering** (instantaneous large release: demolition blasting, nuclear-deflection pulses); nuclear power and nuclear bombs are the two branch-products of one physical principle.

This volume first completes the redefinition (Chapter 1), then analyzes why high-energy engineering is at this stage locked into military in-fighting: the security dilemma, the stage value of deterrence, the two-sidedness of technology, and the resulting civilization-level resource misallocation — global military expenditure reached USD 2,718 billion in 2024, 2.5% of global GDP, rising for ten consecutive years (L1, SIPRI) (Chapter 2). The core argument is in Chapters 3–4: **when AI takes over the scheduling layer of high-energy engineering, its greatest positive significance is not “fighting better” but “firing wrongly far less often”** — the historical record shows human nuclear-decision systems repeatedly approaching error at the edge of false alarms (Arkhipov 1962, Petrov 1983, the Norwegian rocket incident 1995, the Hawaii false alert 2018 — all L1); AI’s value as a false-alarm filter and situational de-noiser is life-saving. Four unshakeable boundaries are given, with “the final launch authority must remain in human hands” as the founding premise of the entire volume (Chapter 4). Chapter 5 organizes the constructive turn’s genealogy: transient engineering leads to celestial defense, pulse propulsion, extreme materials synthesis, and controlled excavation; quasi-static engineering leads to planetary energy, geological modification, and interstellar-voyage power — noting that the External Volume’s engineering genealogy is precisely this energetics unfolding toward “external boundary expansion.” Chapter 6 lands the Principle of Capability–Responsibility Equivalence as engineering principles: **the yield, rate, and irreversibility of an energy release determine its approval tier, guardianship strength, and rollback design**. Conclusion: the physical core of firearms is neutral; what moves that neutral core from the altar of in-fighting to the construction site of expansion is not greater yield but higher intelligence and stricter responsibility.

Keywords: Energy Aggregation-Release-Control Engineering; quasi-static engineering; transient engineering; deterrence rationalization; false-alarm filtering; human in the loop; yield proportionality; conflict de-escalation; Principle of Capability–Responsibility Equivalence

Chapter 1. Redefinition: from “weapons” to Energy Aggregation-Release-Control Engineering

1.1 The physical leap from cold weapons to firearms

The substance of weapons history is the history of energy. In the cold-weapons era, weapons released human muscle energy (joule scale), with killing radii measured in meters; gunpowder introduced chemical energy, with radii of tens to hundreds of meters; fission and fusion introduced nuclear binding energy, leaping single-weapon yields by orders of magnitude — the 1961 “Tsar Bomba” released about 50 megatonnes of TNT equivalent in one event, exceeding the combined explosive tonnage of all prior wars in magnitude (L1, historical record).

But seen from a distance, the essence of this leap is not “weapons got stronger” — it is that **humanity mastered, for the first time, a stellar-scale energy-release mechanism**. The same physical mechanism, released continuously under control, is the core of a nuclear power plant; released instantaneously without control, it is a nuclear bomb. **Nuclear power and nuclear bombs are not two technologies; they are two release curves of one technology** — this is the physical starting point of the entire volume (L1, physics consensus; the conceptual naming is this series’, L3).

1.2 The two branches: quasi-static engineering and transient engineering

By the temporal signature of energy release, Energy Aggregation-Release-Control Engineering divides into two branches:

Quasi-static engineering (full name **Quasi-static Release-Control Engineering**) — sustained, controlled, adjustable release of energy. Empirical lineage: nuclear units worldwide have stably supplied baseload power for decades (L1); the International Thermonuclear Experimental Reactor (ITER) is under construction (L2); China’s EAST achieved 100-million-degree, 1,066-second high-confinement-mode operation in January 2025 (L1, CAS bulletin); the US NIF achieved fusion ignition with gain above 1 for the first time in December 2022 (L1, Abu-Shawareb et al. 2024). The industrial signature of quasi-static engineering: grid-connectable, meterable, adjustable, with scram capability on accident — **the release process is “domesticated” into an adjustable parameter**.

Transient engineering (full name **Transient Release-Control Engineering**) — instantaneous, concentrated release of energy, applying a huge action to a local object in an extremely short time. Empirical lineage: directional blasting in mining and civil engineering (L1, mature industry); industrial explosive processing such as explosion welding and forming (L1, mature technique); the Soviet “nuclear explosions for the national economy” program’s 124 peaceful detonations from 1965–1988 — Lake Chagan created instantly by nuclear blast, the

Urtabulak blowout sealed by nuclear blast in 23 seconds (L1, Nordyke 1996); and the nuclear-device asteroid-deflection scheme studies discussed in Section 3.4 of the External Volume (L2, scheme-study stage).

The engineering difference between the two branches condenses into one sentence: **quasi-static engineering manages the “flow”; transient engineering manages the “pulse.”** The former competes on the precision of sustained control; the latter on the full chain of pre-release suppression, release timing, and post-release sealing. This series’ terminology was named by the conceiver, Li Shunran — the master term is **Energy Aggregation-Release-Control Engineering**, short form **Energy-Control Engineering**; “control” is the crux of the naming: both branches share one proposition, control, differing only in whether the object of control is the “flow” or the “pulse.” “High-energy engineering” hereafter in this chapter refers to the parts of both branches whose yield exceeds conventional engineering magnitudes.

1.3 Weapons are only the destructive application branch

Hence this volume’s core redefinition:

A weapon = the application branch of Energy Aggregation-Release-Control Engineering whose purpose is the destruction of objects. It is a subset of this broad engineering class, not the whole of it.

Three layers of meaning in this redefinition:

1. **Physics level:** weapons hold no definitional authority over energy engineering. The apt analogy is defining “fire” as an arson tool — fire was humanity’s first energy engineering; arson is only its destructive branch.
2. **History level:** the main uses of energy engineering’s branches have always been constructive. Mining and civil uses far exceed military uses in explosive output (L1, common industrial statistics); electricity, not warheads, is nuclear energy’s main form. Weapons’ cultural salience far exceeds their engineering share.
3. **Responsibility-theory level:** judging an energy engineering’s civilizational character depends not on its physical core but on its **release curve, objective function, and guardianship structure** — the foreshadowing of the responsibility law in Chapter 6.

Why, then, did humanity’s most colossal high-energy engineering system appear precisely in the form of weapons? That is Chapter 2’s question.

Chapter 2. The internal contradiction at this stage: why high-energy engineering is locked into in-fighting

2.1 The security dilemma and the stage value of deterrence

High-energy engineering is at present mainly military in form, for two structural reasons.

First, the escalator effect of the security dilemma. In an international system lacking higher-level coordination, one state's effort to strengthen its own security systematically lowers other states' sense of safety, provoking equivalent responses and forming a spiral. Quantified evidence: global military expenditure reached USD 2,718 billion in 2024 — 2.5% of global GDP, the highest in SIPRI's records, rising for ten consecutive years, with year-on-year growth of 9.4%, the steepest since 1988 (L1, SIPRI 2025).

Second, the stage value of deterrence must be honestly acknowledged. Nuclear deterrence served in the Cold-War structure as the “cost-structure rewriter” of great-power hot war — raising the expected cost of full-scale war beyond what any rational decision-maker could bear, thereby turning great-power war from a “standing option” into a “locked-out option” (L2, the common conclusion of deterrence theory and Cold War history; discussed in Chapter 2 of the Internal Volume). This series does not glorify this mechanism, but refuses to ignore its historical function: **deterrence is civilization's second-best steady state in the absence of global governance, not evidence of innate human evil.** At the same time the other side of the ledger must be honestly recorded: the scientific simulation of nuclear winter shows that the climatic consequences of great-power nuclear war would threaten civilization itself (L1, the original Turco et al. 1983 simulation and its subsequent debates) — deterrence's “steady state” rests on a risk borne by all humanity in common.

2.2 Technological two-sidedness: every high-energy technology has two release curves

The civilian and military forms of high-energy technology mirror each other — this is not conspiracy but physics: the time curve of energy release is an engineering parameter, and engineering parameters carry no moral properties. Fissile material can generate power at low enrichment or become weapons at high enrichment; rocket propulsion can carry satellites or warheads; precision blasting can demolish buildings or fortifications. **The essence of international mechanisms such as the Non-Proliferation Treaty (1968) and the Comprehensive Nuclear-Test-Ban Treaty (1996) (L2) is precisely this: acknowledging two-sidedness, they impose institutional constraints on “the choice of release curve” — treaties cannot govern physics, only the users of curves.**

2.3 Resource misallocation under arms-driven investment

Two-sidedness plus the security dilemma produces a civilization-level distortion of resource allocation: some of the finest engineers and the largest R&D budgets in human history have long served “more efficient release of destruction” rather than “more efficient construction.” Annual military expenditure of USD 2,718 billion (L1, SIPRI), set against the investment scale of the External Volume’s engineering families (global annual planetary-defense investment on the order of hundreds of millions of dollars, L2), shows a disparity of more than two orders of magnitude.

This series’ judgment (L3, responsibility-theory deduction): **this is not a moral failure but a symptom of governance structure lagging behind technological capability** — humanity has mastered planetary-scale energy without yet building a planetary-scale mechanism for allocating responsibility. The cure for the misallocation comes not from moral exhortation but from two conditions: (1) the deterrence steady state yielding to a higher-level common-security mechanism; (2) the emergence of an execution layer able to play the “guardian” role with sufficient precision. The second condition is precisely AI’s candidacy — hence Chapter 3.

Chapter 3. AI takeover: the four-layer leap from “operator” to “scheduling hub”

The main thesis has established AI’s four-layer role in engineering systems (simulation & foresight / fine scheduling / boundary guardianship / iterative optimization). This chapter lands them in the high-energy engineering scenario and states precisely what each layer’s “takeover” means — **what is taken over is scheduling and computation, not decision authority.**

Layer 1 · Simulation & foresight. High-energy systems share fast response, deep coupling, and scarce samples (major events admit no trial and error). This layer’s duty is to turn “what has not happened” into “what can be computed”: operational-state deduction for nuclear facilities, multiphysics simulation of release curves, Monte Carlo enumeration of accident scenarios. Real-world anchor: AI surrogate models already widely replace full-order models in physical simulation (L2, common engineering practice).

Layer 2 · Fine scheduling. Upgrading “how much, when, where” from human experience and intuition to parametric optimization: millisecond-level grid dispatch (L2, deployed), digital design of blasting parameters (L1, mature technique), real-time plasma control in fusion devices — AI control of fusion plasma shape has published results (L2). In quasi-static engineering, scheduling precision equals safety margin; in transient engineering, scheduling precision equals the footprint of collateral effects.

Layer 3 · Boundary guardianship. The layer of greatest civilizational significance, and Chapter 4's theme: **installing on the human decision chain a false-alarm filter that does not tire, carries no emotion, and takes no part in interest calculation.**

Layer 4 · Iterative optimization. Data from every run, every drill, every false alarm feeds back into the model so guardianship rules keep evolving. Key design (L3): the training objective of the iterative-optimization layer must be anchored by humans — **the objective function for “safer” cannot be left to the system to define**, otherwise this becomes the highest-risk high-energy form of the “goal escape” described in the main thesis.

Chapter 4. The positive significance in military application: “never fire wrongly” as the guiding principle

This chapter handles the series' most sensitive topic. The expressive stance comes first: **in a reality where the deterrence steady state cannot be replaced immediately, AI's first positive significance in military high-energy systems is lowering the probability that humans trigger catastrophe through their own cognitive limits — “never fire wrongly” takes priority over “firing accurately.”**

4.1 Deterrence rationalization: false-alarm filtering and situational de-noising

The fatal weakness of the deterrence system is not weapon failure but **misjudgment on the decision chain**. History's record is sobering; all four cases below are L1 historical fact:

- **October 1962, the Cuban Missile Crisis:** US depth-charge exercises were misread by the Soviet submarine B-59 as war already begun; two officers aboard urged launching the nuclear torpedo; the flotilla commander (officer) Arkhipov exercised a veto — human civilization hung on one person's judgment;
- **26 September 1983:** the Soviet early-warning system mistook sunlight reflected from clouds for incoming US submarine-launched missiles; duty officer Lieutenant Colonel Petrov judged a false alarm and refused to report upward — investigation later confirmed the system had misread sunrise reflections as ballistic trajectories;
- **25 January 1995:** a Norwegian scientific rocket's ascent was misidentified by Russian early-warning radar as a submarine-launched ballistic missile; Russia activated its highest nuclear-decision procedure, ruling out the misjudgment only after the rocket left the possible nuclear-strike corridor;
- **13 January 2018:** Hawaii's civil-defense system erroneously issued a “ballistic missile inbound” alert; 38 minutes of statewide panic; the investigation attributed it to operator and interface-process error (L1, FCC report).

The common structure of the four cases: **a high-energy deterrence system + the limits of human cognition (fatigue, fear, incomplete information) = systematic misjudgment risk**. In the first two cases, specific individuals' presence of mind averted catastrophe — a personal merit, and precisely a systemic hazard: **civilization's safety must not depend on the duty officer's quality of sleep that day**.

AI's positive value as false-alarm filter and de-noiser is thus established: cross-validation of multi-source data (physical-consistency checks across satellites, radar, electromagnetic, and astronomical events), automatic comparison against historical false-alarm patterns, and presentation to decision-makers of a structured "evidence–confidence–alternative explanations" picture rather than a single alert. This is not letting AI decide whether to retaliate; **it is letting the question "have we actually been struck" be answered first by the strongest available computation, then handed to human decision** (the role conception is L3; the four historical structures are L1).

4.2 Yield proportionality: the engineering expression of "just enough and no more"

The principle of yield proportionality (L3, the principle is this series'; its engineering basis is L1/L2): **the destructive radius of a release should strictly match the objective function — every unit of yield beyond the objective function is unattributable civilizational cost**. Its technical trend (higher guidance precision, smaller warheads, stronger damage assessment) is discussed in open literature (L2); this volume expands no details and only fixes the responsibility-theory formulation: the correct direction of yield proportionality is **making "precision" serve "minimization," not "maximization"** — if the historical progress of precision guidance is used to expand rather than contract the damage footprint, the principle is violated.

4.3 Conflict de-escalation: game design from "winning" to "not firing"

The endgame of deterrence rationalization is not more efficient confrontation but the dissolution of confrontation itself. AI's three constructive roles here (L3, role conceptions; their game-theoretic bases are mature L1/L2 theory):

1. **Pre-hoc development of misjudgment costs:** high-fidelity deduction presents the full "escalation path" before decisions — most historical escalations stemmed from parties underestimating each other's response functions; deduction's meaning is turning the response function from conjecture into visualization;
2. **High-fidelity translation of crisis communication:** acting at crisis moments as a high-precision fact-verification and intent-clarification layer, compressing the "misunderstanding window" (in the 1995 Norwegian rocket incident, had such a verification layer existed between radar and decision, the misjudgment would have terminated at the first link);

3. **Depersonalized de-escalation ladders:** generating for standoff states a multi-tier library of “de-escalation actions acceptable to both sides,” converting the dignity game of “who blinks first” into the engineering choice of “which rung to take.”

4.4 Boundaries (unshakeable): the human in the loop cannot be abolished; the final launch authority belongs to humans

The validity of every positive claim in this chapter is premised on the following boundaries; if any is breached, the positive claims fail:

1. **The final launch authority must remain in human hands** — AI’s entire role in the nuclear-decision chain is limited to filtering, de-noising, deduction, and presentation; it holds no form of launch authorization whatsoever. This is the series’ founding premise, not an option;
2. **“Meaningful human control” over fully autonomous lethal weapons systems (LAWS) must be institutionalized** — the GGE discussions on LAWS under the UN Convention on Certain Conventional Weapons have run for years (L2); China submitted its Position Paper on Regulating Military Applications of AI to the UN in 2021 (L2); the Global AI Governance Initiative (2023) explicitly holds that military applications of AI must maintain the “human decision” bottom line (L2). This series fully supports that bottom line and advocates its upgrade into binding international law;
3. **The guardian layer’s own trustworthiness must be auditable** — the false-alarm filter itself may fail or be adversarially manipulated; therefore the guardian system’s models, data, and updates must be independently auditable, and its failure modes must have pre-designed human backstop channels (L3, design principle);
4. **Justifying power transfer by “efficiency” is forbidden** — any argument of the form “AI decides faster, therefore AI should decide” is void: the premise of faster is more accurate, and in a crisis the final arbiter of “accurate” is the human community’s value judgment, not predictive precision.

These four boundaries are at the same time a rehearsal of the Safety Volume, which will expand the constraint layer into a complete system.

Chapter 5. From in-fighting to expansion: the constructive turn of the same engineering

The path out of Chapter 2’s misallocation is not waiting but **the redirection of the very same energy-engineering capability**. This chapter gives each branch its constructive genealogy — the External Volume’s engineering genealogy unfolds precisely from here.

5.1 The constructive genealogy of transient engineering

- **Celestial defense (flagship direction):** nuclear-device deflection is a candidate means for short-warning, large-object scenarios, at the scheme-study stage (L2); kinetic impact already has a demonstration (L1, DART; see External Volume 3.4). From “destroying surface targets” to “fine-tuning celestial orbits,” the objective function completed a civilization-level reversal — **the same “instantaneous large release” capability, once for destroying cities, once for defending a planet.**
- **Pulse propulsion:** the nuclear-pulse propulsion scheme (Project Orion, 1958–1963) was seriously studied and terminated by the 1963 Partial Test Ban Treaty (L1, Dyson 2002). What it left was not a weapons legacy but an open question: **when civilization truly needs to leave Earth orbit, the highest-specific-impulse energy source is precisely transient energy control** — to be re-evaluated when safety, treaties, and need all mature (L3, long-range conjecture).
- **Extreme materials synthesis:** transient high pressure is itself a materials process — explosion welding is a mature industrial technique (L1), and shock-wave synthesis of superhard materials has industrial applications (L1). The extreme materials needed by planetary engineering (the infrastructure of External Volume 4.6’s deep-space logistics) will most likely require higher-pressure transient processes (L3, conjecture).
- **Controlled excavation and sealing:** the Soviet peaceful nuclear program’s 23-second blowout seal (L1, Nordyke 1996) demonstrated transient energy control’s unique capability at “sealing”; modern alternative paths (conventional blasting, geotechnical engineering) are in operation (L1), and the nuclear path stays closed until its radiological boundary reaches zero (per External Volume 3.2’s conclusion).

5.2 The constructive genealogy of quasi-static engineering

- **Planetary energy:** fission already carries the baseload (L1); fusion’s three milestones (NIF ignition, EAST long pulse, ITER under construction, L1/L2) point to the final form — fusion fuel comes from seawater, and its rollout means the substantive lifting of the planet’s energy constraint, the energetics basis of the Internal Volume’s “internal entropy reduction.”
- **Quasi-static application at geological and climatic scales:** for the engineering families of External Volume 3.1–3.3 and 4.1–4.3, the question of energy source is a constraint before fusion’s rollout, and enters engineering-economics discussion only after it (L3, sequencing judgment).
- **Interstellar-voyage power:** nuclear-thermal and nuclear-electric propulsion have been space-validated repeatedly (L2) — the near-term energy answer for deep-space logistics; in the long run fusion propulsion decides humanity’s “economic radius” within the solar system (L3, long-range judgment).

5.3 The mechanism of the turn: not abandoning weapons, but rewriting the objective function

Honesty requires: the security dilemma does not vanish because of Chapter 3's AI guardian layer; the deterrence system will persist for the foreseeable future. The "turn" this series advocates is a **marginal process, not an instantaneous switch** (L3):

*The guardian layer (Chapter 3) lowers the probability of accidental firing in in-fighting → yield proportionality contracts conflict's destructive radius → de-escalation design (4.3)
raises the probability of soft landing in crises → the constructive genealogies (5.1/5.2)
progressively absorb the constructive share of high-energy engineering budgets →
deterrence's stage value exits naturally as global governance matures.*

Each step has independent value and does not depend on the idealism of a "final leap." This is the responsibility law's methodology: **do not bet on the endgame; build the ramp.**

Chapter 6. Energy engineering and the responsibility law: the engineering expression of authority–responsibility equivalence

The main thesis's core proposition — the Principle of Capability–Responsibility Equivalence (an intelligence's authority strictly matches its responsibility for the parent system's entropy flow) — lands here as three engineering principles (L3, original to this series; their warranting bases are all L1/L2):

Principle 1 · Yield–tier equivalence. The greater the yield of an energy release, the higher the approval tier. Humanity already has this principle in embryo: conventional blasting is approved within engineering firms; industrial explosives are statutorily controlled in circulation; nuclear-material holdings accept IAEA safeguards (L2). The responsibility law merely theorizes existing practice: **the approval tier is graded not by use (military/civilian) but by yield and irreversibility** — a megaton-scale release, military or civilian, should have as its responsible subject a governance mechanism at the level of all humanity.

Principle 2 · Rate–guardianship equivalence. The faster the release rate (the closer to transient release), the stronger the guardianship layer required and the finer the enumerated boundary list of pre-authorization. Quasi-static engineering may have shift supervisors; transient engineering must have an institutionalized "Arkhipov Node" — **on the decision chain of civilization-level risk, always preserve a post occupied by a human entitled to say "no."** Petrov in 1983 proved the value of that post; the task of the AI age is not to abolish it but to equip it with the strongest assistance (4.1).

Principle 3 · Irreversibility–rollback equivalence. The higher a release’s irreversibility, the greater the share of rollback design in the engineering budget. High-energy releases that can neither be rolled back nor reversed (such as sustained intervention in the biosphere) fall under the External Volume Chapter 5’s “halt at Gate IV” rule — **engineering that cannot be rolled back is better left undone.**

Together the three principles give energy engineering a civilizational test (answering 1.3):

To judge an energy engineering’s character, look at three things: the approval tier aligned with yield, the guardianship strength aligned with rate, and the rollback design aligned with irreversibility. With all three, a destructive core is still civilizational engineering; missing any one, even a constructive shell is a civilizational risk.

Chapter 7. Conclusion

The volume’s chain of argument, drawn together:

1. **The redefinition holds:** weapons are the destructive application branch of Energy Aggregation-Release-Control Engineering; nuclear power and nuclear bombs are two release curves of one physical mechanism (Chapter 1, L1 physics + L3 concept).
2. **The locking mechanism is clear:** the security dilemma and deterrence’s stage value explain why high-energy engineering is mainly military at this stage; the USD 2,718-billion annual expenditure and the two-orders-of-magnitude investment gap quantify the misallocation (Chapter 2, L1 SIPRI).
3. **AI takeover’s positive significance takes “never fire wrongly” as its guiding principle:** four false-alarm histories prove the human decision chain is the deterrence system’s weakest link; the value of AI’s false-alarm filter, yield proportionality, and de-escalation design all points to “firing wrongly less often,” not “fighting better” (Chapters 3–4).
4. **The boundaries are non-negotiable:** the final launch authority belongs to humans; meaningful human control; an auditable guardian layer; no efficiency-based justification of power transfer — the four boundaries are the premise of all positive claims (4.4).
5. **The constructive turn is a marginal process:** the transient and quasi-static branches each have complete constructive genealogies, and the External Volume’s planetary-engineering genealogy is their unfolding; the mechanism builds the ramp, not the endgame bet (Chapter 5).
6. **The responsibility law lands as three engineering principles:** yield–tier, rate–guardianship, and irreversibility–rollback equivalence (Chapter 6).

Answering the series’ meta-law — on the energy dimension, the exact meaning of “awakening is responsibility” is:

For a civilization that commands stellar-scale energy, the mark of awakening is not releasing more, but equipping every release with commensurate tier, guardianship, and rollback.

The Internal Volume says “AI makes humanity’s choices cheaper”; the External Volume says “AI gives civilization’s survival an insurance policy”; this volume lays the energetics foundation beneath them both: **AI places every release of energy on the scales of responsibility.** The physical core of firearms has always been neutral; what moves it from the altar of in-fighting to the construction site of expansion is not fiercer yield, but higher intelligence, stricter boundaries, and a final veto seat forever occupied by a human.

Appendix A · Confidence tiers

- **L1 Established evidence / historical fact:** measured data, engineering records, mutually consistent historical accounts (with sources);
- **L2 Formal programs / regulations / authoritative reports:** projects in progress, treaties, institutional bulletins, peer-reviewed literature;
- **L3 Conceptual deduction:** frameworks, principles, or conjectures original to this series, explicitly marked “this is a conjecture.”

Appendix B · Falsifiability statement

1. **“Nuclear power and nuclear bombs are two release curves of one technology”** — physics consensus; if the controllability principles of fusion/fission were overthrown the whole volume would be rebuilt (very low probability).
2. **“The structural attribution of the four false-alarm cases”** — if declassified files overturned the misjudgment character of any case (e.g., proving it was not a false alarm), the relevant argument chain would need revision; this volume adopts the version consistent across multiple accounts.
3. **“AI false-alarm filtering can systematically reduce misjudgment risk”** — if field-deployment studies showed AI assistance raising alert confidence instead (the automation-bias effect), 4.1’s role conception would need constraint clauses; **automation bias (over-trusting machine output) is itself one of this volume’s preset failure modes, within the audit scope of 4.4 item 3.**
4. **“The stage value of deterrence”** — if historical research showed nuclear deterrence’s explanatory power for great-power peace significantly weaker than other factors, 2.1’s formulation would be weakened, without affecting 2.3’s misallocation conclusion or Chapter 5’s turn logic.

5. **“The sufficiency of the three responsibility principles”** — if a case arose where all three principles were satisfied yet disaster followed, the responsibility law would need a fourth dimension (e.g., “cumulative effect–governance equivalence”).

Appendix C · Glossary

Term	Definition
Energy Aggregation-Release-Control Engineering (short: Energy-Control Engineering)	The series’ core concept: an engineering class whose core is aggregating energy and releasing it under designed curves; weapons are its destructive application branch. Terminology proposed and named by Li Shunran
Quasi-static engineering (full: Quasi-static Release-Control Engineering)	The branch of sustained, controlled, adjustable energy release (nuclear plants, fusion reactors)
Transient engineering (full: Transient Release-Control Engineering)	The branch of instantaneous concentrated release (demolition blasting, nuclear-deflection pulses)
Release curve	The temporal-distribution signature of an energy release; the core parameter distinguishing quasi-static from transient
Security dilemma	The international-relations structure in which one state’s effort to strengthen its own security systematically lowers other states’ sense of safety
Deterrence rationalization	Using stronger fact-verification and deduction capability to lower the probability of the deterrence system failing through misjudgment
False-alarm filter	This series’ role conception: an AI layer on the decision chain that cross-validates alert authenticity and presents evidence–confidence structures
Yield proportionality	The principle that release scale strictly matches the objective function; direction is minimization, not maximization

Term	Definition
Meaningful human control	The core requirement for autonomous weapons in international arms-control discourse (LAWS discussions); adopted and upgraded by this series
Arkipov Node	This series' naming: the post on civilization-level risk decision chains occupied by a human entitled to say "no" (from the 1962 incident)
Principle of Capability–Responsibility Equivalence	The main thesis's first principle; landed here as the three engineering principles of yield–tier, rate–guardianship, and irreversibility–rollback equivalence

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

1. SIPRI. Trends in World Military Expenditure, 2024[R]. Stockholm International Peace Research Institute, Fact Sheet, 2025-04.
2. Turco R P, Toon O B, Ackerman T P, et al. Nuclear winter: global consequences of multiple nuclear explosions[J]. Science, 1983, 222(4630): 1283–1292.
3. Nordyke M D. The Soviet program for peaceful uses of nuclear explosions[R]. Livermore: Lawrence Livermore National Laboratory, UCRL-ID-125887, 1996.
4. 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).
5. Hefei Institutes of Physical Science, CAS. EAST achieves 100-million-degree, 1,066-second high-confinement-mode operation[EB/OL]. 2025-01 (in Chinese).
6. 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.
7. Schlosser E. Command and Control: Nuclear Weapons, the Damascus Accident, and the Illusion of Safety[M]. New York: Penguin Press, 2013 (nuclear-decision false-alarm history incl. the 1962 Arkipov and 1983 Petrov incidents).
8. Pry P V. War Scare: Russia and America on the Nuclear Brink[M]. Westport: Praeger, 1999 (the 1995 Norwegian rocket incident).
9. FCC Public Safety and Homeland Security Bureau. Preliminary report: Hawaii false emergency alert[R]. 2018-01.
10. Dyson G. Project Orion: The Atomic Spaceship 1957–1965[M]. London: Allen Lane, 2002.

11. United Nations. Treaty on the Non-Proliferation of Nuclear Weapons[Z]. 1968; Comprehensive Nuclear-Test-Ban Treaty[Z]. 1996; Partial Test Ban Treaty[Z]. 1963.
 12. UN CCW Group of Governmental Experts on Lethal Autonomous Weapons Systems. Session reports[R]. 2014–present.
 13. People’s Republic of China. Position Paper on Regulating Military Applications of Artificial Intelligence[Z]. Submitted to the UN, 2021-12 (in Chinese).
 14. People’s Republic of China. Global AI Governance Initiative[Z]. 2023-10-18 (in Chinese).
 15. International Atomic Energy Agency (IAEA). Safeguards system and nuclear-material control documents[EB/OL]. <https://www.iaea.org/>.
 16. ESA. Hera mission updates[EB/OL]. 2024–2026. https://www.esa.int/Space_Safety/Hera.
-

This volume is the Energy Volume of the AI Awakening·Responsibility series, v1.0 final.*
Next: *AI Awakening·Responsibility (External Volume)* — macro responsibility · planetary engineering: AI-led expansion of civilization’s survival space.*