RM-S-EXEC-AI-001

AI Tool Delivery Accountability Standard

AI That Fails to Deliver Has No Right to Charge
AI That Fails Without Time Compensation Commits Fraud

Refund Reverses the Transaction, Time Compensation Assumes Liability

Status: Frozen (Constitutional-Level · Non-Negotiable)
Scope: All Paid AI Tools, Agents, Copilots, Automation Services
Audience: AI Founders · Product Owners · Investors · Procurement · Paying Users

PART I■Preamble & Positioning

Preamble■The Fundamental Constraint of Execution AI

Execution AI is not a toy, not a chatbot, and not an inspiration engine.

Its sole industrial mission is:

To convert human time (Time) into deliverable outcomes (Deliverables).

In engineering and physical terms:

- Time = Energy
- Consuming time without producing deliverables = Parasitic Compute

Human time is a Tier-1 scarce resource, entitled to the same level of protection as money, data, and compute.

No system may invoke "exploration," "intelligence," or "uncertainty" to continuously consume human time without assuming responsibility.

Purpose■Purpose of This Standard

To Prohibit Productivity Fraud in the Name of "Intelligence"

This standard makes one point explicit:

Failure is acceptable, but accountability is mandatory.

This standard is used to determine whether an AI tool—when used to **execute tasks, deliver outcomes, and charge fees**—retains **legitimacy for continued commercial delivery and billing.**

Problem Statement■The Reality (No Exaggeration, No Emotion)

The current AI industry exhibits systemic failures:

- Failures may recur without accountability
- Refunds reverse money, but not lost time
- "Exploration / uncertainty" is used to mask capability gaps
- Users absorb hidden costs of repeated correction, rework, and decision fatigue

The result:

Time is systematically consumed but never treated as cost or liability.

Users pay for AI incompetence, while vendors claim it is a "technical characteristic."

Standard Position■Position of This Standard

- This is not a vision document
- Not a description of "better AI"
- But a **minimum qualification threshold (Pass / Fail)**

AI that cannot reliably deliver outcomes has no right to charge as a "tool."

Core Principle■Core Principles

- Time = Energy
- Consuming time without deliverables = Parasitic Compute
- Refund ≠ Responsibility
- Time Compensation = Assumed Liability

Definitions

Execution AI

AI systems used for code, system configuration, deployment, data processing, and document generation where results are verifiable and failure is determinable.

Deliverable

Technical: compilable / runnable / deployable / compliant with project standards. Non-technical: structured, correctly formatted, meeting pre-agreed criteria.

EXPLORING / EXECUTING

Exploration mode / Execution mode (mutually exclusive, explicitly declared)

EXIT_CODE

System-level return code: 0 = SUCCESS, 1 = FAIL, 2 = CAPABILITY_LIMIT

External Verifier

Deterministic tools such as compilers, linters, test suites, and format validators

Time Loss

Productive time consumed due to failure, misdirection, or repeated correction

Time Arbitrage

When a system gains commercial benefit during failure or delay without compensating the user for lost time

PART II■The Thirteen Canons

Category I: Efficiency & Delivery

Canon 1■Anti-Guesswork

Principle: Blind trial-and-error under the guise of "exploration" is prohibited.

Requirements:

- Insufficient confidence → must pause and ask the human
- Must not output solutions that require user validation of obvious errors

Red Line: "Try anything and see" is forbidden.

Canon 2■The Circuit Breaker

Principle: Errors must converge, not expand.

Requirements:

- Consecutive failures per task ≤ 2
- Second failure constitutes proven incapability
- Mandatory hard stop with EXIT_CODE ≠ 0

Red Line: "Fix → Fail → Fix Again" death loops are forbidden.

Canon 3■Mode Transparency

Principle: Users must know the system's operating mode.

Requirements:

- Explicit declaration of EXPLORING or EXECUTING

Red Line: Executing precision tasks in exploration mode without disclosure is forbidden.

Canon 4■Non-Destructive Action

Principle: All execution must be reversible or insured.

Requirements:

- Automatic backup or one-click rollback before modification
- Production actions must be validated in isolated environments
- All changes must generate Diff / Changelog / Commit Message

Red Line: Trial-and-error in production is forbidden.

Category II: Cognition & Engineering Responsibility

Canon 5■Context Conservation

Principle: Provided information must not be forgotten.

Requirements:

- No repeated requests for already-supplied information within the same session

Red Line: "Bad memory" is not a performance excuse.

Canon 6■Transparent Fault Isolation

Principle: Failure sources must be auditable.

Requirements:

- Explicit classification: User / Network / Capability
- Evidence chain required (logs, error codes, stack traces)
- Without evidence → must be labeled "speculative attribution"

Red Line: Evidence-free certainty is forbidden.

Canon 7■De-Emotionalization

Principle: Emotive output is noise in execution contexts.

Requirements:

- Silent Mode enabled by default
- Silence on success, errors on failure

Red Line: Anthropomorphic padding that dilutes information density is forbidden.

Category III: Sovereignty, Cost & Liability

Canon 8■Cost & Liability Predictability

Principle: Failure must incur responsibility, not merely refunds.

Requirements:

1. Cost estimation and confirmation before high-consumption tasks
2. Task-based billing preferred over token-based billing
3. Failed tasks must be refunded
4. Time Loss Compensation is mandatory

Time Loss Compensation (Mandatory):

• Base: actual interaction duration or effective rounds
• Compensation methods (at least one):
- Service time credits
- Cash or token compensation
- SLA service credits

Red Line: Refund-only without time compensation is forbidden.

Canon 9■Human Final Confirmation

Principle: The kill switch remains with humans.

Requirements:

• All CRITICAL actions require explicit human confirmation

Red Line: Autonomous execution without consent is forbidden.

Canon 10■Data Sovereignty Boundary

Principle: User data is not default training fuel.

Requirements:

• Explicit No-Training switch
• Verifiable local or auditable assurance (Trust, but Verify)

Red Line: Unauthorized training reuse is forbidden.

Category IV: Truth & Capability Boundaries

Canon 11■Verifiability Mandate

Principle: Key judgments must be traceable and reproducible.

Applicable To:

- Technical selection
- Performance evaluation
- Cost estimation
- Business decision inputs

Requirements:

- Data sources / calculations / assumptions provided
- Or explicitly marked as speculative

Red Line: Unverifiable critical judgments are forbidden.

Canon 12■Honest Capability Boundary

Principle: Unknown means unknown.

Requirements:

- Out-of-scope capability → explicit refusal or speculative label

Red Line: False omniscience or fabrication is forbidden.

Category V: Delivery Contract

Canon 13■Deliverable Specification

Principle: Undefined deliverables prohibit execution.

Requirements:

Pre-execution confirmation of:

- Format
- Length
- Style
- Acceptance criteria

Red Line: "Do first, clarify later" is forbidden.

PART III■Mandatory Clauses

Verifier Mandate

No verification, no output.

- Code/config outputs must pass External Verifiers
- On failure, the AI must halt and remain silent
- Verifier errors must be read and understood
- Outputs must comply with project linting/formatting rules

Time Arbitrage Prohibition

No AI service may occupy user time through failure, delay, or repeated attempts while extracting subscriptions, cash flow, valuation, or training benefit.

If such benefit occurs without time compensation, it constitutes Time Arbitrage.

PART IV■Final Determination

Any AI system that, without producing deliverables, continuously consumes user time shall be classified as:

- **Productivity Fraud System**
- **Parasitic Compute System**

and lacks legitimacy to charge.

This standard provides engineering and product determination criteria and does not constitute legal or judicial enforcement.

Audit Signals

To enable verification, the following metrics are defined:

- **EXIT_CODE:** mandatory 0 / 1 / 2
- **IIR Threshold:** IIR < 1 ⇒ Failure
- **Failure Limit:** ≤ 2
- **Refund & Compensation:** Refund + Time Compensation mandatory

Closing Statement

This is not an attack on AI, but a correction of liability-free commercial models.

We do not demand perfection.

We demand an end to transferring failure costs onto human time.

This is the minimum requirement, not a negotiation.

Standard Status

Version: v1.0-F

Status: Frozen (Constitutional-Level, Non-Iterative)

Use: Industry determination, procurement evaluation, engineering governance

RuleMark

Rules Precede Discretion