MONTREAL.AI

RESEARCH EDITION 01

AUTONOMOUS ENTERPRISES

ACTIVATE. EVOLVE. SCALE.

SUCCESSOR Ω Harness

Powered by TRANSFORMATIONAL INTELLIGENCE ∞
and the applicable $AGIALPHA utility.



Achievements that become capabilities.

Capability inheritance, complete counted cost,
and a constructive thermal-measurement counterexample.

Vincent Boucher • MONTREAL.AI

22 October 2026 • Montréal, Québec

Abstract

An enterprise can complete a task without improving its capacity to perform the next one. We study a stronger, bounded mechanism: an accepted achievement becomes an executable asset with explicit applicability, dependencies, evidence and cost; a successor can reuse that asset while earning its own acceptance. We implement and evaluate this mechanism in the existing AGI ALPHA workbench and mathematical enterprise controller. Six checked exact-rational programs support later transform, convolution and inverse tasks; a changed-kernel successor becomes a seventh retained artifact. At 48 tasks, retention uses 63,196 declared arithmetic-work units versus 915,136 for repeated construction. The supplied expert uses 5,760, and full inherited successor qualification loses 7,918 versus 2,030 for direct qualification. These losses are integral results.

A second construction addresses a thermal-metrology counterexample: a 4.9477% prediction–observation relative residual coexists with 8.6964% error against a numerical heat reference. A separate simulated volume witness and bounded, volume-weighted temperature difference yield a conditional 1.8862% error bound. The adapter executes through the native worker and durable journal and supports three measurement families. Physical validity remains unestablished. We separately exercise source-defined Ascension contracts in a local EVM with simulated balances; the complete independently accepted thermal enterprise-to-settlement journey remains unfinished.

The contribution is an inspectable construction and evaluation of capability inheritance: retained executable lineage, full accounting within a declared metric, informative negative results, and explicit distinctions among numerical, institutional and physical acceptance. We derive conditional amortization and error-enclosure relations, identify validation replay as a measured bottleneck, and specify falsifiable transitions toward physical instrumentation and productive infrastructure. No monetary advantage, achieved general intelligence or unlimited growth follows from these finite demonstrations.

Keywords reusable capability; autonomous enterprise; scientific instrumentation; evidence; metrology; successor; full-cost evaluation

> **The central idea**
>
> An achievement becomes productive capital when its executable structure, evidence and operating limits survive the task that created it. The successor inherits a stronger set of available means; its new requirement still needs a valid check.

Demonstrated	Executed software, counted arithmetic, retained results and local-EVM source behavior. Thermal inputs and funds are simulated.
Testable next	Connected thermal acceptance; reduced replay overhead; calibrated bench; voluntary beneficiary with measured total costs.
Ambition	Useful instruments, discoverable processes, manufacturing, infrastructure and net useful energy delivery. Every transition requires physical resources and validation.

Edition and scope. This is an additive capability-inheritance research edition. It preserves earlier papers and evidence. Software source: `34a068e`. Reporting date: 22 October 2026 in America/Toronto; evidence and observation timestamps use UTC, including 23 October. AGI ALPHA is the project identity. This manuscript is not a claim of peer-reviewed certification or an operationally complete Ascension deployment.

1 From delivered work to expanded capability

The useful question for a productive organization is not exhausted by whether it finished a task. A delivery can leave behind an instrument, executable method, representation, process or manufacturing asset that changes the feasible work available to its successors. Such an achievement has two outputs: the immediate result and a reusable means of producing further results. Their value must be assessed separately.

TRANSFORMATIONAL INTELLIGENCE ∞ names the project's discipline of changing the means by which a worthwhile objective is pursued: improve the observations, representation, instrument or organization while preserving the objective and its acceptance boundaries [1]. Here that discipline is made operational through a deliberately finite system. The software has supplied objectives, fixed checkers, bounded resources, persistent evidence and explicit refusal states. It cannot increase its authority or redefine success to manufacture progress.

We investigate four questions:

1. Can accepted executable structure survive its original mission and support a distinct successor?
2. Does reuse lower complete counted work as later workload grows, and how does it compare with a supplied expert?
3. Can a constructive measurement interface expose an error that a prediction–observation residual misses?
4. Which execution, acceptance, economic and physical boundaries remain before these achievements can support useful enterprises?

The answer is mixed in an informative way. Retention amortizes construction, but the expert wins the workload comparison. A successor extends the retained inventory, but inherited qualification is more expensive than direct qualification. A conditional thermal enclosure addresses the specified numerical mismatch, but does not establish physical calibration. Native execution and local-contract settlement both run, but the thermal enterprise handoff actually refuses the unadmitted family. These outcomes constrain the mechanism rather than weaken its description.

1.1 Contribution and relation to prior work

Reusable temporal structure has substantial antecedents. Options formalize temporally extended policies within reinforcement learning [2]; our exact-rational program inventory is a different, fixed-interface construction, not an options-learning benchmark. Automated physical science also predates this project: Robot Scientist Adam combined hypotheses with laboratory experiments [3], and a mobile robotic chemist combined physical laboratory operation with constrained search [4]. The present thermal experiment does not reproduce those physical achievements.

Our contribution is the particular executed combination of:

- a retained capability graph with explicit applicability and a changed-requirement successor;
- a complete ledger within a declared arithmetic metric, including failed search, admission and replay;
- a constructive response to the preserved Foundry measurement failure;
- source-bound execution, recovery and contract evidence whose unconnected boundaries remain visible.

We do not claim a new transform algorithm, the first reusable skill, the first autonomous laboratory, a generally superior enterprise, or open-ended recursive improvement. The distinction between

scope-preserving software reuse and institutionally accepted, physically useful production is the organizing contribution [5–7].

2 A precise model of capability inheritance

2.1 Enterprise, asset and evidence

For this study, an *enterprise* is a software operating record, not an incorporated entity. Write

$$E_t = (O, A_t, B_t, K_t, J_t, \mathcal{Q}_t), \tag{1}$$

where O is the objective, A_t permitted authority, B_t resource allowances, K_t retained capabilities, J_t history and $\mathcal{Q}_t$ obligations. A capability is represented conceptually by

$$k = (f, D, \mathrm{dep}, h, e, v, c), \tag{2}$$

with executable program f, applicability domain D, dependencies, content identity h, evidence e, validation policy v and attributed cost c. The notation summarizes the implemented records; it is not a new universal interchange standard.

An accepted artifact may be reused only where its domain and relevant dependencies remain applicable. A digest identifies bytes. It does not by itself establish semantic correctness, measurement accuracy, rights, reviewer independence or permission to act. Acceptance of an ancestor does not accept an arbitrary descendant.

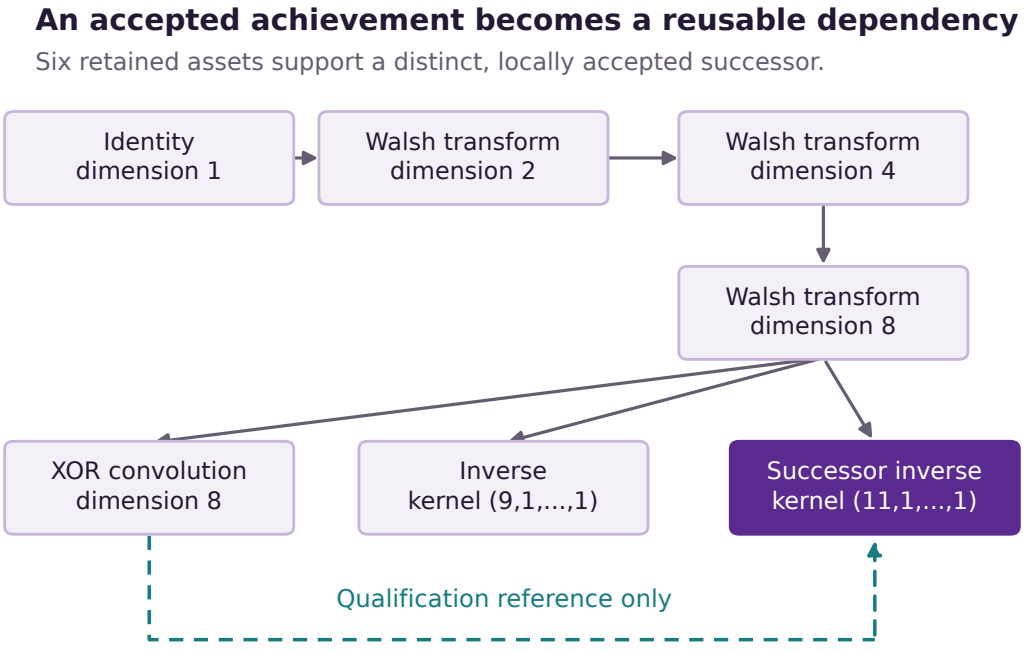


Figure 1. The executed capability graph. Six initial assets remain available when a changed-kernel inverse becomes a seventh. Solid edges identify runtime construction dependencies; the dashed edge identifies the convolution reference used by the inherited qualification path. The new requirement earns a new check.

2.2 Two kinds of progress

For a resource allowance b, define the set of *witnessed feasible missions*

$$\mathcal{F}(K,b) = \{m : \exists\, \pi \in \Pi(K,m),\ \text{valid}(\pi,m) \ \wedge\ C(\pi,m) \le b\}\,, \tag{3}$$

where Π contains the supplied, evaluated paths. This set is intentionally relative to tested paths; the implementation does not prove that every possible construction has been searched.

Two distinct changes can matter:

$$\Delta\mathcal{F} = \mathcal{F}(K',b) \setminus \mathcal{F}(K,b), \tag{4}$$

$$\Delta C(m) = C_{\text{old}}(m) - C_{\text{new}}(m). \tag{5}$$

The first measures newly supported work within the evaluated inventory and allowance. The second measures efficiency on common work. The changed-kernel experiment expands the retained executable inventory; it does not by itself prove strict expansion of $\mathcal{F}$, because direct construction remains available. Inherited qualification also loses on full cost. Neither quantity establishes demand or a monetary surplus.

Retaining old assets does not guarantee that operational feasibility grows monotonically. Review policy, validity conditions, budgets, available operators and physical resources can change. An older artifact can remain historically authentic while becoming inapplicable to a new mission. The runtime preserves that distinction by separating completed-history recovery from unfinished-work continuation.

2.3 Conditional amortization

Let B denote complete construction work, V current admission work and $W(R)$ checked execution for R rounds. If identical assets remain applicable and both paths face the same checked workloads, then

$$C_{\text{retain}}(R) = B + V + W(R), \tag{6}$$

$$C_{\text{rebuild}}(R) = R(B+V) + W(R). \tag{7}$$

Proposition 1 (Conditional reuse advantage)**.** *Under the assumptions of Eqs.* (6)–(7)*, retention avoids* $(R-1)(B+V)$ *counted work for* $R \ge 1$.

Proof. Subtract Eq. (6) from Eq. (7). The shared execution term cancels. □

This elementary accounting identity is useful because its assumptions can be tested. A changed dependency, new admission burden, storage cost or more expensive execution changes the comparison. If retention adds burden $H(R)$, the difference becomes $(R-1)(B+V) - H(R)$; reuse helps only when that quantity is positive. Neither identity proves that generated assets beat an expert alternative.

2.4 Cost is a vector before it is a price

The relevant cost record includes arithmetic work, processing time, human attention, energy, money by currency, maintenance and failures. These quantities cannot be added without a justified conversion model. In this experiment, only a declared arithmetic projection is comprehensively charged:

$$\mathbf{c} = (c_{\text{arith}}, t_{\text{human}}, e_{\text{energy}}, c_{\text{cash}}, c_{\text{maintenance}}, \ldots).$$

Unmeasured fields are recorded as unknown, not zero. A monetary return would require actual beneficiaries, accepted deliveries, prices and measured resource costs. Repeated program execution and token settlement are neither willingness to pay nor profit.

3 The operating architecture and six out-stages

3.1 One history, explicit boundaries

The architecture reuses the existing workbench, mathematical controller, native worker and durable journal. A request is validated, frozen against its executable identity, executed within bounded operations, reviewed against its exact result and retained. A successor receives a distinct semantic commitment and its own acceptance obligation.

The full Ascension target connects opportunity discovery, a FusionPlan, seed commitment, review, funding, bounded jobs, execution, acceptance, settlement and successor lineage. The current evidence does not establish that complete thermal path. The six out-stages in Table 1 are useful only where their input/output contracts have causal content; their labels do not bridge a missing integration [8, 9].

Table 1. Six stages as contracts, with the current evidence boundary.

Stage	Required change	Executed evidence and remaining limit
OUT-LEARN	Make a restriction observable; record uncertainty and falsification conditions.	Command/flow mismatch is retained; simulated independent-volume observations supply a new measurement channel.
OUT-THINK	Change representation while preserving the question and continuation needs.	Volume-weighted heat intervals replace command-as-flow observation. Typed requests and source/result identities survive.
OUT-DESIGN	Construct an enabling method or instrument.	Fixed thermal adapter executes through the native worker; mathematical programs become checked assets. Physical qualification remains open.
OUT-STRATEGISE	Resolve necessary uncertainty or build around it.	Compare rebuilding, inheritance and expert paths; retain wider-uncertainty refusals, conventional ties and costly inheritance.
OUT-EXECUTE	Deliver through verified boundaries and recover valid work.	Real worker, signed journal, exact-result local review and corrected export; independent thermal settlement remains unconnected.
OUT-EVOLVE	Qualify a distinct successor using retained capability.	Mathematical changed-kernel successor is locally accepted. Thermal family extension executes but lacks connected independent acceptance.

3.2 Why the three directions are different

We compared three promising directions. A *self-checking experimental interface* changes what an instrument observes and can therefore address the concrete Foundry mismatch. A *discoverability compiler* constructs experiments that distinguish hypotheses; it is a separately owned research implementation. A *shared productive-capability foundry* organizes qualified assets and capacity across voluntarily supported investigations. The third needs evidence of demand and operating capacity in addition to sound software.

We select the first for the thermal flagship and pair it with an executed capability-inheritance experiment. This pairing studies complementary mechanisms without pretending that the thermal instrument is already an accepted mathematical asset. The long-term foundry is a development direction, not an observed market.

4 Experiment I: a software enterprise that retains its means

4.1 Task, grammar and checkers

The implementation uses the existing finite exact-rational construction grammar and separate checker at dimension $n = 8$. The initial inventory contains identity-1, Walsh-2, Walsh-4, Walsh-8, XOR-convolution-8 and an inverse for kernel $(9, 1, 1, 1, 1, 1, 1, 1)$. A transform, a convolution and an inverse request form one round.

The evaluation compares retained assets, rebuilding every round, and a supplied static specialist on identical public inputs at $R \in \{1, 4, 16\}$. Component workloads use the retained LCG protocol and seed 20261022; the separate controller journey uses 20261021. We do not describe this public finite case set as a blind discovery benchmark. The controller freezes the assets before its later task generation, which prevents those task values from silently changing the accepted construction.

The checker uses exact rational calculations, so equality is not a tolerance-adjusted numerical success. Exactness within these authored interfaces does not make the grammar complete or the specification adequate for an unrelated task. The system is not allowed to modify the objective or its checker during this evaluation.

4.2 Prospective protocol and historical correction

The native criteria were frozen at 23:40:39.270 UTC on 22 October 2026, after earlier post-implementation component observations. The initial native run began at 23:43:29.649 UTC. A later source/data inspection found a provenance-label defect: specialist execution had been attributed to a generated artifact digest. A recorded amendment at 23:46:02.115740 UTC corrected the attribution to the actual specialist source, with artifact digest null. The corrected native run began at 23:46:02.359 UTC.

The objective, dispatch, criteria and work definition were unchanged. Initial source, protocol and results remain retained. Different invocation and mission identifiers yield different workspace hashes, so we verify each history rather than demanding identity across distinct executions. The separately performed audit checked the 53-source closure, raw receipt hashes and accounting; it did not constitute an institutional or owner approval [6].

4.3 Measurement scope

The metric `paper-declared-arithmetic-work/v1` charges applicable search, construction, current admission, productive arithmetic, reference arithmetic and comparisons. It does not count CPU instructions, bit complexity, hashing, framing, serialization or all development effort. Timing fields are observational and are not the decisive cost metric.

A complete build costs $B = 28{,}398$; current admission costs $V = 28{,}398$. Checked execution of the generated three-task round costs 400, so $W(R) = 400R$ in the measured path. The expert round costs 360. Its exposed arithmetic preparation is zero; the human authorship, proof and institutional qualification of that supplied implementation remain unknown.

The controller's scoped allowances are 2,000,000 construction units, 50,000,000 admission units and 2,000,000 per task. These are work allowances, not cash. The configured execution is synthetic-local, with empty authority and obligations, currency none, collected cash zero and external submission absent.

5 Results I: amortization, expansion and loss

5.1 Retention works against rebuilding; the expert wins

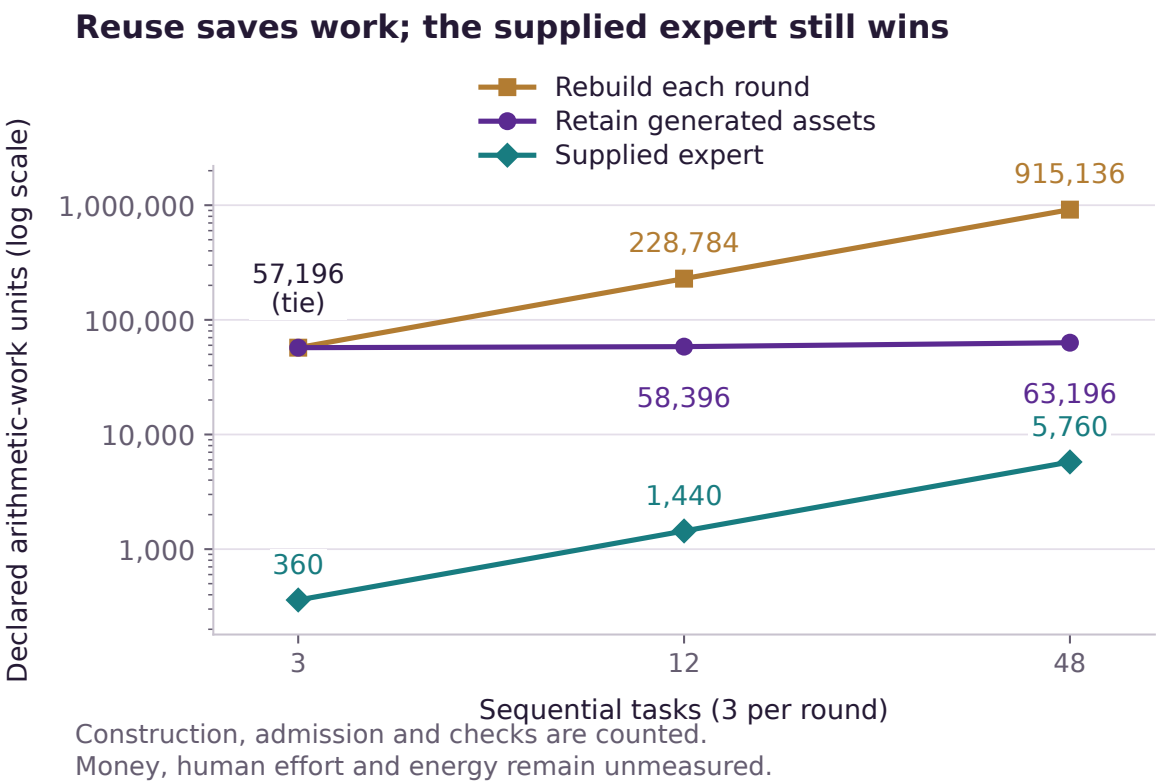


Figure 2. Identical sequential workloads under the declared arithmetic metric. The logarithmic vertical axis keeps all three arms visible. Retention amortizes construction; the supplied expert remains cheaper at every workload. No monetary or concurrent-capacity inference follows.

Table 2. Complete counted workload totals. All values are declared arithmetic-work units.

Rounds	Tasks	Retained	Rebuilt	Expert
1	3	57,196	57,196	360
4	12	58,396	228,784	1,440
16	48	63,196	915,136	5,760

The one-round tie and all later differences match Proposition 1. At sixteen rounds, retention avoids 851,940 units relative to rebuilding, about 93.1% of the rebuilding total. This is an accounting consequence of retained valid work under matched conditions; it is not evidence of a new asymptotically superior algorithm.

Budget feasibility changes for the supplied paths. Retention completes twelve tasks within 60,000 units and forty-eight within 250,000, whereas rebuilding does not. The expert also fits those allowances. There is no established task here that a suitably implemented expert cannot solve.

The specialist ties the generated Walsh and convolution operating counts at 24 and 88, respectively. Its fixed-kernel inverse uses 24 operations against the generated inverse's 64. Retaining an asset is useful, but does not make that asset the best available implementation. Production allocation should preserve the expert option.

5.2 A distinct successor, without retrospective acceptance

The successor changes the required kernel to $(11, 1, 1, 1, 1, 1, 1, 1)$. The old inverse is not deleted or retroactively declared invalid for its original mission. It simply fails the new residual check: its reconstructed first component is $-235/64$, while the new requirement calls for -3. The new inverse reconstructs all requested components exactly and becomes the seventh retained artifact.

Runtime execution inherits Walsh-8; the inherited qualification path also invokes convolution as a reference. This expands the inventory's supported requirement. It does not show an ability unavailable to direct construction or an adapted expert, and changing a supplied kernel is not open-ended recursive invention.

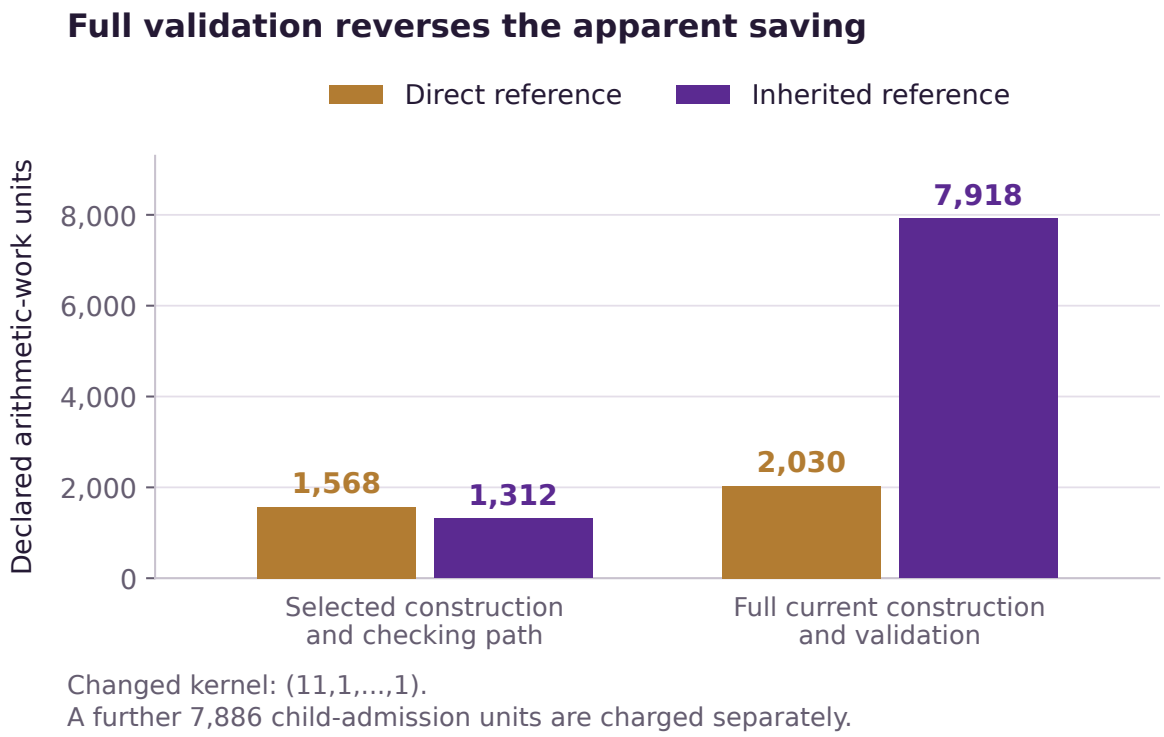


Figure 3. Inheritance can lose after full qualification. A smaller selected construction/checking path becomes a larger full current bill once dependency and scope checks are included. Both results remain in the ledger.

Table 3. Changed-kernel qualification. The apparent selected-path saving reverses.

Construction	Selected path	Full current cost
Direct reference	1,568	2,030
Inherited convolution reference	1,312	7,918

A further inherited-child admission costs 7,886. Direct and inherited constructions, that admission, stale/corrected executions and singular rejection total 18,250. The all-ones singular kernel is rejected and its 32 spent units are retained. These outcomes make applicability and qualification costs visible rather than concealing them in an asset count.

5.3 The full observed ledger reveals the bottleneck

One complete build examines 196 candidates, including 187 rejected candidates. Its 28,398 units comprise 11,480 search/preparation, 9,356 selected qualifications, 7,560 additional dependency/scope checks and two identity-admission units. Counting only the selected path would understate the construction burden.

The component campaign records 241 distinct charge identifiers and totals 1,348,918 units. Alternative workload totals share ancestor attribution and must not be summed as if they were separate campaigns. The separate controller journey adds 1,802,222 units. Its original construction, future treatments and successor account for 151,892; actual replay accounts for 1,650,330.

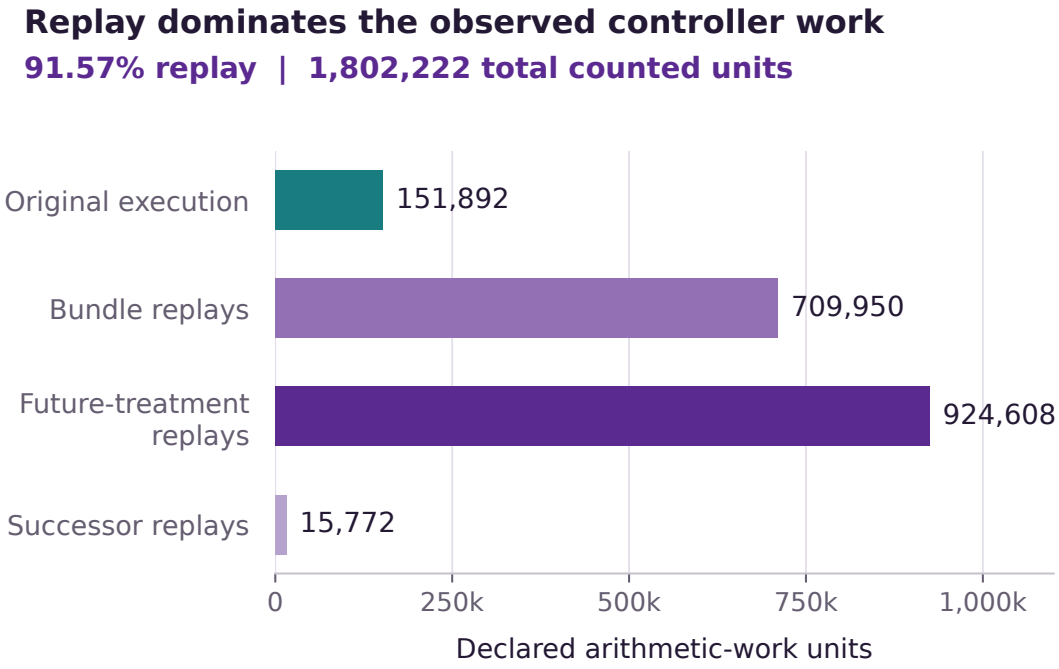


Figure 4. Actual controller work. Replays consume 91.57% of its 1,802,222 declared units. This identifies a candidate improvement target; no reduced-replay implementation is claimed by this experiment.

Table 4. Nested accounting scopes. These rows must not all be added together.

Scope	Counted work	Interpretation
Complete component campaign	1,348,918	All 241 unique component charges
Complete controller journey	1,802,222	Original work plus actual replay
One full native invocation	3,151,140	Sum of the two disjoint scopes above
Two retained full reports	6,302,280	Initial and corrected native invocations
Third regression invocation	3,151,140	Assertions/TAP retained; third raw JSON not retained
Three known invocations	9,453,420	Finite observed account, excluding other work

The controller's historical lineage number 180,290 is descriptive. It must neither replace the observed total nor be added to it again. Separate thermal, recovery, EVM, browser, hosted and other development work is outside the mathematical account. A prior V8 summary survives, but its complete raw campaign receipts were lost during an environment refresh; no total is invented for them.

Human active time, engineering, institutional review, money, energy, peak memory, general maintenance and unmetered overhead remain null. The experiment therefore establishes neither a complete economic bill nor lower human supervision. It establishes where the declared work went, including the cost of being able to check the result.

6 Results II: constructing a measurement that can witness the error

6.1 The inherited limitation

The preserved Foundry case is a counterexample to treating model–measurement agreement as measurement accuracy [10]. Its predicted receiver heat is 5,175.864695 J. A command-based calorimeter reports 5,445.281559 J, while the simulated heat integral is 5,009.624458 J. The runtime residual is

$$r = \frac{|Q_{\text{command}} - Q_{\text{prediction}}|}{|Q_{\text{command}}|} = 4.9477\%, \tag{8}$$

which passes a 5% agreement rule. Yet the calorimeter's error relative to the numerical heat integral is

$$\epsilon_{\text{ref}} = \frac{|Q_{\text{command}} - Q_{\text{reference}}|}{|Q_{\text{reference}}|} = 8.6964\%. \tag{9}$$

The measured quantity and the prediction can share a misleading premise. Here a flow command is not the actual delivered flow. The relevant intervention is a different observation channel and a complete heat-accounting interface.

6.2 A conditional volume-witness construction

The adapter takes witnessed volume and the mean temperature difference weighted by that volume over each interval. The target is receiver-fluid heat:

$$Q = \int \rho c_p q(t)\,[T_{\text{out}}(t) - T_{\text{in}}(t)]\,dt = \sum_j \rho_j c_{p,j} V_j \overline{\Delta T}_{V,j}. \tag{10}$$

The second equality assumes constant fluid properties within each interval. For varying properties, or a temperature statistic that does not cover the whole volume window, this representation needs additional information. A point temperature and an unrelated volume total are insufficient.

Each input has a declared interval. The implementation evaluates the 16 endpoint combinations of density, specific heat, volume and signed temperature difference, then sums interval extrema with outward integer rounding. In its fixed integer representation,

$$Q_{j,\mu\text{J}} = \frac{\rho_{j,\text{kg/m}^3}\ c_{p,j,\text{J/(kg K)}}\ V_{j,\text{nL}}\ \overline{\Delta T}_{j,\mu\text{K}}}{10^{12}}. \tag{11}$$

Arbitrary-precision integers avoid floating-point inward rounding in this production calculation. The result is a point estimate E, an interval $[L, U]$ and a conditional relative-error bound.

Proposition 2 (Conditional heat enclosure)**.** *Suppose each interval covers its actual whole-window inputs, the property assumptions of Eq. (10) hold, and all measurement windows are included. Then*

the constructed sum contains the target Q. If $0 \notin [L, U]$, a valid bound for the reported point estimate is

$$\frac{|E - Q|}{|Q|} \leq \frac{\max\{|E - L|,\ |E - U|\}}{\min_{q \in [L,U]} |q|}. \tag{12}$$

Proof. The heat product is multi-affine in its four bounded inputs, so its extrema on a rectangular domain occur at vertices. Summing outward enclosures preserves containment. The numerator in Eq. (12) bounds absolute error over the interval; the denominator is a positive lower bound on the target's magnitude. □

The adapter rounds Eq. (12) upward to integer parts per million. An interval containing zero cannot supply this finite relative bound. Missing interval coverage, unsupported variable-property semantics and unsuitable denominators remain refusals. A small numerical residual cannot override them.

This is deterministic interval reasoning under declared premises. It is not a standard uncertainty, a 95% confidence interval or an independently certified measurement procedure. The GUM and NIST uncertainty guidance provide the broader measurement framework against which a physical implementation must be evaluated [11, 12]. A content hash can preserve a calibration record; it cannot create calibration.

6.3 Numerical experiment and results

A separately implemented evaluator integrates four thermal states and a fifth accumulated-heat state using fixed-step RK4 at $h = 1/32$ s. The receiver reference is the integrated heat state, not a value supplied to the worker. A 180 s reset precedes the 720 s mission: 28,800 steps and 115,200 right-hand-side evaluations. Both inlet temperatures are 20 °C; density and specific heat are fixed synthetic constants, 998 kg/m^3 and 4,180 J/(kg K). The mismatch multiplies the two commanded flows by 1.08 and 0.92 and heater power by 1.05. Appendix C records the schedule and the measurement premises.

The simulated volume channel has a +0.3% bias inside a declared 1% band. The differential-temperature channel has a +0.002 K bias inside a declared ±0.005 K calibration band. A supplied curvature bound, trapezoidal-integration allowance and upward quantization increase the reported temperature allowance to 7,606 μK. These are simulation premises, not observed sensor certificates.

The new point estimate differs from the numerical reference by approximately 0.5228%. This is numerical agreement, not demonstrated physical accuracy. Increasing the declared temperature uncertainty to ±0.05 K fails the 5% policy. A conventional implementation given the same measurement channels and assumptions ties the result. The advantage is the constructed information and validity check, not an exclusive benefit of an agent identity.

Twelve public protocol cases vary actuator regime, initial thermal history and permitted channel biases. All six declared criteria pass across these cases; the largest conditional bound is 1.9398% and the largest observed reference error is approximately 1.0873%. The native run is a replay of previously evaluated public cases, not a fresh blind holdout. Only two actuator regimes are covered, and the reset nearly erases initial-history differences. This is limited numerical coverage, not broad robustness to physical conditions [6, 7].

6.4 Three mission families, one accounting capability

The same typed heat accounting supports delivered heat, heat capacity and recovery fraction:

$$C_{\text{sample}} = \frac{Q - Q_{\text{loss}} - Q_{\text{container}}}{\Delta T_{\text{sample}}}, \qquad \eta = \frac{Q - Q_{\text{stored release}}}{Q_{\text{input}}}. \tag{13}$$

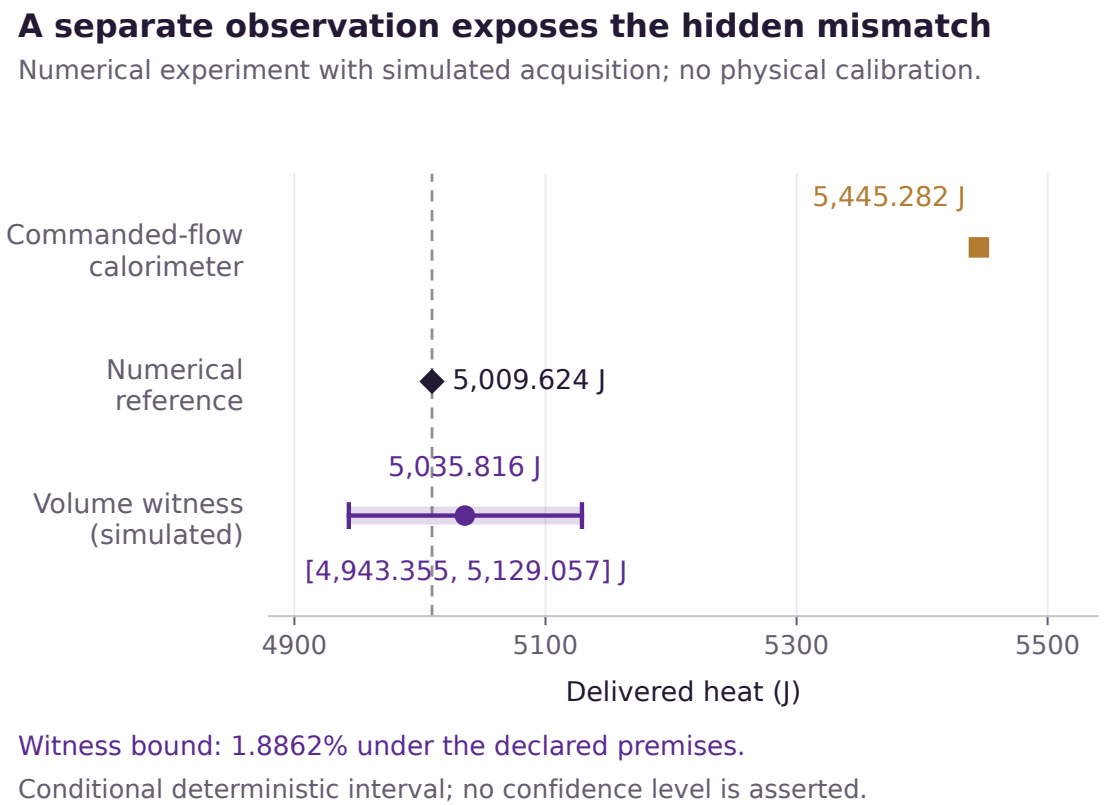


Figure 5. A useful construction changes what can be checked. The original 4.9477% residual coexists with 8.6964% reference error. The new simulated witness reports 5,035.816195 J inside a conditional enclosure of [4,943.354615, 5,129.056527] J and a 1.8862% bound. The reference is a numerical model integral, not a physical calorimetry result.

These extensions introduce new obligations. Sample temperature, losses, container heat, stored-energy release and input energy require their own justified bounds. They do not follow from the parent heat witness. The implementation checks interval corners, denominator suitability and inconsistent energy claims.

Table 5. Three executed families. The two fixture sets have intentionally different declared extension inputs. These are simulated inputs, not estimates of a physical sample or plant.

Family	Default adapter example	Held enterprise-composition probe
Delivered heat	3,680 J	3,680 J
Heat capacity	$(3680 - 800 - 200)/10 = 268$ J/K	$(3680 - 600 - 80)/20 = 150$ J/K
Recovery fraction	$(3680 - 200)/5000 = 0.696$	$(3680 - 180)/5000 = 0.700$

The three families execute through real software. They are not three physically constructed instruments. Their outputs retain physical validity `NOT_ESTABLISHED`, independent review false, capability acceptance false, admission `HOLD` and settlement `NOT_CONNECTED`. A local exact-result review is useful software evidence but does not remove those boundaries [13].

7 Execution, recovery and settlement: what connects today

7.1 The native boundary is exercised

The fixed thermal adapter is invoked by the existing native worker with typed inputs, bounded operations, evidence output and durable journal records. Its declared limits include 960 intervals, a one-hour observation window and a 2 MB input limit. Partial budget exhaustion remains `EXHAUSTED`; it is not silently promoted to a completed result. Unsupported execution families cannot become arbitrary tool access.

The latest recovery correction distinguishes two questions. A completed record can be authenticated and exported without invoking a now-unavailable worker family. Unfinished execution must still match the complete frozen worker and source closure before continuing. The implementation checks the worker wrapper as well as nine component modules, and scrubs Node preload settings. Source changes cannot pass merely because a subset of module hashes matches. Thirty-nine Python tests and nineteen Node tests pass for the recorded correction; 54 runtime files pass the recorded type check [14].

These checks establish specific tested behaviors. Completed-history reload is not equivalent to an operating-system crash during payment. A full crash, transaction replay and reorganization matrix across the integrated enterprise remains unfinished.

7.2 The observed integration refusal is a result

A disposable, uncommitted composition combines the public thermal candidate with the existing held enterprise facade. It invokes four real workers: three parent mission families and a distinct heat-capacity successor. Each of the three parent records survives a fresh-process reopen with the same result digest and one execution attempt. Wrong-result review digests are refused without journal changes. The child starts as a new mission and reaches review without inheriting acceptance.

The exact local reviews explicitly record `independent: false`. When the compiled FusionPlan is passed to both enterprise handoff and Ascension preparation, the existing Studio boundary refuses it:

> This executor supports only the installed fixed workbench families.

No Studio state is created by these rejected calls. This identifies concrete missing work: family admission, verdict semantics and the relevant command catalog. The held facade and harness retain their existing review restrictions. No held code was copied through another pull request to avoid that restriction [13].

7.3 A separate real-source local-EVM journey

The unchanged Alpha five-contract suite—Access, NovaSeed, RiskOracle, JobMarket and Mark—is exercised on an unforked local Hardhat chain, ID 31337. Balances, ENS, admission and actors are fixtures. Thirty-eight existing source tests pass. The existing demonstration first runs the native allocation worker and verifies its result and signed journal, then performs seed, risk, funding, job, delivery, two-validator acceptance, settlement, reserve refund and successor-token operations. Eighteen local transaction summaries are retained [15, 16].

The allocation value is 210, tying its baseline. Funding is 158 simulated `$AGIALPHA`; the job price is 80, worker net is 75.24, each of two validators receives 1.98, burn totals 0.8 and the unused reserve refund is 78. The successor token preserves lineage and requires new risk review. It does not establish an improved productive capability. This experiment is not the thermal journey in Section 6.

For a valid price p in token base units, the source computes

$$v = \lfloor p/40 \rfloor, \qquad g_w = p - 2v, \qquad b_i = \lfloor g_i/100 \rfloor, \qquad n_i = g_i - b_i, \tag{14}$$

where $g_1 = g_2 = v$ for the two qualifying validators. Burn applies separately to each recipient's gross amount. It is not an interchangeable blanket 1% deduction. Nine separately frozen boundary cases preserve small-integer rounding and source admission failures.

Table 6. Literal local-contract payout boundaries, in token base units. Each valid row includes source execution and collateral/conservation assertions.

p	Worker net	Each validator	Burn	Source outcome
1, 39, 40, 99	No payout executed			Posting and bidding rejected below 100
100	96	2	0	Paid
103	99	2	0	Paid
104	99	2	1	Paid
3,999	3,763	99	38	Paid
4,000	3,762	99	40	Paid

Distinct fixture actors establish role-separated local calls, not independent organizations. Local application configuration excludes external and forked providers; it is not a claim of operating-system-wide network isolation. The retained mainnet observation attempt failed at its first chain-identity request, with no successful observation. Historical addresses and source do not establish present chain state, admission, funding or readiness.

7.4 Repository and token boundaries

Alpha owns orchestration and the five-contract Ascension route. AGIJobsv0 is a separately identified `$AGIALPHA` backend with its own contracts and admission rules. AGIJobManager retains its separate USDC route. Fleet supplies execution and review infrastructure where qualified. A cross-repository source citation is not an integration test, and the maintained compatibility record remains incomplete [17].

For the inspected Alpha source, business, agent and validator naming roots are respectively `LABEL.alpha.agi.eth`, `LABEL.alpha.agent.agi.eth` and `LABEL.alpha.club.agi.eth`. Owner-

ship and admission matter; token balance alone does not confer a role. Alpha's job market contains its source-defined reputation-weighted auction; Jobs v0 must not be described as if it shares that selection mechanism. MARK's supported lot/bonding behavior is not a generic DEX. NovaSeed records ownership and commitments with parent references; ownership does not automatically transfer intellectual property, equity or executable ability [8, 16].

8 Economic usefulness and the missing bill

8.1 An asset must change a beneficiary's alternatives

The mathematical inventory changes what this software can execute from retained programs. The thermal construction changes what its declared observations can support. These are mechanisms for productive value, not observations of revenue. No paying customer, market-clearing price, actual willingness to participate or monetary surplus is measured here.

A credible economic test compares accepted useful output against a matched alternative and counts the resources needed to deliver it. Let y be the same accepted service, $R(y)$ realized revenue and $C_{\text{full}}(y)$ the measured cost of construction, qualification, operation, failed attempts, oversight, maintenance and capital use. Then

$$\Pi(y) = R(y) - C_{\text{full}}(y) \tag{15}$$

is a proposed accounting definition, not an estimated profit in this study. Avoided cost requires an observed counterfactual: the relevant expert or conventional provider must deliver equivalent quality under equivalent conditions.

Our strongest observed relative result is lower declared work than rebuilding the same assets repeatedly. The expert workload loss, the successor qualification reversal and the conventional thermal tie remain economically relevant. An arrangement that makes a new job possible may still be inferior to buying an existing service.

8.2 Oversight and scaling must be measured

The study supplied objectives, program grammar, reference implementations, checkers, protocol criteria, authority limits and the thermal model. The automated facade probe recorded local operator-API reviews; owners retain institutional admission and merge decisions. These actions are not evidence of independent human review. Automated execution reduces no measured number of oversight hours in this evidence package because active human time was not recorded.

A proposed deployment should record at least beneficiary acceptance, operator and independent-reviewer minutes, rejected work, requalification, resource occupancy, failure recovery and fully attributed cost. The same budget cannot be spent independently by multiple descendants. Shared reviewers and instruments are finite resources.

For one class of accepted jobs, a simple planning upper bound is

$$\lambda_{\text{accepted}} \le \min\left(\lambda_{\text{demand}}, \frac{N_e}{\tau_e}, \frac{N_v}{\tau_v}, \frac{N_h}{\tau_h}, \frac{\dot{B}}{c_{\text{job}}}\right), \tag{16}$$

where execution, review and hardware each have N available units and mean occupancy τ; $\dot{B}$ is the spending allowance per time. All terms must use the same job definition. Rework and failures increase effective occupancy and cost. This is a planning relation under stated assumptions, not a fitted capacity model.

The workload experiment increases repeated numerical jobs through 48 tasks. It does not test simultaneous tenants, queueing, reviewer scarcity, demand saturation, hardware utilization or changing

prices. The changed-kernel test shows a validity boundary and a cost reversal; it does not establish robustness to general distribution shift.

The next economic claim must be earned

A voluntarily supported mission should deliver an accepted service, retain a reusable capability, and then deliver a distinct successor service. Its report must show the beneficiary's alternative, all observed costs, the human work and whether reuse survives changed conditions. Unknown prices or hours remain unknown.

9 A falsifiable development path

9.1 Layer 2: the next achievable constructions

The first deployable software step is legitimate thermal-family integration into the existing Studio and Ascension path. Existing owners must resolve the held review restriction and agree on the typed family admission and numerical-versus-physical verdict. A source-bound local-EVM journey should then connect plan commitment, eligible review, simulated funding, real bounded thermal execution, independent acceptance policy, settlement and a distinct successor. Interrupted pending work and changed chain identity need their own retained failure tests. This is a specific missing integration, not a reason to rebuild a competing harness.

The next measured optimization is validation replay. Its 91.57% share of controller work motivates a proposal to reuse previously verified evidence only when the complete source, inputs, policy and dependency identities match. The evaluator and acceptance target must remain fixed. Deliberate stale-identity, changed-policy and corrupted-evidence cases must still fail. Lower complete cost without weaker detection would be an accepted improvement; a cheaper but unsound cache would be rejected. This proposal has not been implemented or promoted in the reported experiment.

The next scientific construction is a calibrated liquid-loop bench. It requires a physical receiver, pump, heater, reservoir, valves, temperature channels, a separately measured flow or collected mass, synchronized acquisition and electrical input measurement. Thermal storage and environmental losses must be identified for the chosen useful objective. A reference based on collected mass and an independent temperature system would test the claimed volume witness. It must not reuse the same unverified command channel as its own truth.

Before observations, the bench protocol should fix acceptance error, operating envelope, reference uncertainty, calibration checks, exclusions, workload and total-cost categories. Planned interventions should include flow-command bias, temperature offset, acquisition gaps, changing flow and temperature, residual stored heat and failure of the separate channel. The comparison should include an expert's conventional instrument. NIST's traceability policy requires a documented chain with uncertainty contributions; traceability alone does not ensure fitness for the intended measurement [18]. Common-cause failure and property variation require explicit tests.

Finally, recruit a willing beneficiary for one bounded service whose output can be independently accepted. Record actual setup, operator, reviewer, calibration, materials, energy, downtime and maintenance costs. The initial service should create a retained method or instrument configuration; the successor should add a materially different requirement. Progress means that the second service is both technically supported and useful to its beneficiary under the measured bill.

9.2 Layer 3: from digital assets to useful physical capacity

The mechanism can motivate a longer path, provided each transition constructs something real. Table 7 gives necessary conditions, not a forecast or a sequence already achieved. An instrument

can enable a process search; a validated process can justify a pilot line; a reliable line can supply an infrastructure project. Each arrow carries new costs and validation obligations.

Table 7. Long-term ambition as a sequence of physical construction and acceptance obligations. None of the physical transitions is demonstrated by the present software study.

Transition	Construction and resources	Validation and economic condition
Digital method → instrument	Sensors, apparatus, acquisition, calibration artifacts, operators and maintenance.	Reference accuracy, traceability, operating envelope, repeatability and a service valuable enough to cover the complete bill.
Instrument → process discovery	Samples, feedstocks, controlled experiments, characterization and safe handling.	Reproducible causal effects, batch variation, scale-sensitive losses and an advantage over expert process search.
Process → pilot manufacture	Equipment, tooling, controls, labor, consumables, utilities, quality systems and spares.	Yield, throughput, reliability, waste, maintenance and demand sufficient for utilization and capital recovery.
Pilot → infrastructure	Site engineering, civil works, supply chains, interfaces, construction and long-lived operations.	Site-specific engineering, required permissions, integration, resilience and sustainable lifecycle funding.
Infrastructure → useful energy	Actual generation, conversion, storage or heat-recovery equipment and delivery connections.	Net useful output after parasitic loads and losses, resource availability, temperature/quality of energy, reliability and an accepting load.

For heat recovery in particular, useful temperature, compatible materials, contamination, maintenance, and a receiving process matter. The DOE technology assessment documents such engineering constraints; it supplies no valuation for this project [19]. Recovering heat is not creating primary energy. A measured calorimetric fraction does not establish plant economics, net useful energy or infrastructure feasibility.

Digital replication can make an accepted design easier to distribute. It cannot duplicate a pump, wafer, machine tool, transformer, mine, trained operator or joule. The long-term opportunity is to use accumulated methods and evidence to make those constructions more reliable and worthwhile. Their physical resource burden remains.

10 Limits, failure modes and scientific interpretation

Finite and supplied search. The program grammar, workload and expert are designed by the study. Basis checking supports exact claims within that restricted grammar, not arbitrary software verification. Reuse is evaluated against repeated construction and a supplied expert, not all possible implementations.

Post-implementation protocol. Earlier component observations informed the frozen mathematical native evaluation. The thermal native evaluation repeats public cases. Neither experiment is a blind prospective discovery study. No population-level generalization or statistical confidence interval is claimed.

Metric incompleteness. The ledger is complete within its declared arithmetic projection and identified invocation scope. It is not complete economic cost. Unknown human effort, money, energy and development overhead prevent a profitability or labor-saving conclusion.

Reference dependence. Separate code for the thermal reference helps expose software mistakes but integrates the same declared physical model. It is neither an independent physical experiment nor a certified numerical enclosure of the differential equation. The deterministic measurement enclosure is conditional on its premises.

Authority and independence. Different fixture addresses and authenticated local reviews do not establish organizational independence. Physical acceptance, governance admission and owner-reserved decisions remain separate. No autonomous worker is granted governor privileges by this study.

Incomplete composition. The mathematical controller, thermal native worker and local-contract lifecycle are distinct exercised paths. Their evidence cannot be concatenated into a nonexistent complete thermal enterprise journey. The actual Studio refusal and unresolved review restriction remain part of the result.

No open-ended growth proof. Seven retained artifacts, three simulated measurement families and finite workloads do not prove general intelligence, unlimited improvement, universal autonomy or civilization-scale outcomes. The contribution is a mechanism that can be challenged, measured and extended with explicit next tests.

11 Conclusion

The productive unit is more than the completed job. It can be the checked means that survives the job: an executable method, an instrument interface or a validated process with an applicability domain and a cost history. The reported software demonstrates this inheritance, a distinct changed-requirement successor and useful negative results. It also constructs a conditional measurement interface that addresses the preserved numerical thermal failure.

The evidence supports neither automatic economic advantage nor completed physical autonomy. The expert wins, inherited qualification can lose, and the thermal enterprise handoff remains blocked. Those boundaries make the next construction concrete: complete the legitimate integration, reduce measured verification overhead without weakening it, and qualify one physical witness for one useful service.

> **Achievements that become capabilities**
>
> The credible path to broader usefulness is a sequence of accepted constructions. Each should leave its successor better equipped, while carrying forward the evidence, costs and limits that make that inheritance trustworthy.

Data, software and author statement

The companion source package includes this complete manuscript, bibliography, vector figures, figure generator, frozen figure data, selected execution evidence, a file-hash manifest and a build script. It compiles without running experiments or contacting a chain. Public software and retained evidence are pinned in the bibliography and Appendix A. Earlier releases and papers remain available.

Vincent Boucher, MONTREAL.AI, is the named project author. AI assistants supported implementation, evidence analysis, manuscript drafting and figure production. Local automated reviews are identified as such and are not represented as external peer review. This manuscript does not report institutional laboratory validation, a commercial contract or independent organizational acceptance. The project uses AGI ALPHA as its identity; the name is not an experimental result.

A Reproduction and exact evidence boundaries

A.1 Paper build

The source package contains pre-rendered vector figures, so a normal PDF build does not require Python or matplotlib. Use XeLaTeX and BibTeX, directly or through the supplied script:

```
./build.sh
# Equivalent:
latexmk -xelatex -bibtex -interaction=nonstopmode \
  -halt-on-error Autonomous_Enterprises_Capability_Inheritance.tex
```

The supplied font fallbacks use TeX-distributed or common open fonts. Figure regeneration uses the recorded JSON and generator; it does not rerun the product or create new experimental observations. The bundled validation report records page inspection, bibliography and reference checks. The public source is an experimental candidate, not a newly declared production release.

A.2 Source and chronology

Table 8. Source identities and retained evidence. Short identities here are expanded in the companion manifest.

Scope	Identity and interpretation
Paper source cut	`34a068e311451ce1d4fcc0a6cea85970a8faf457`; PR #4871 remained open/unmerged at the reporting cut. All nine ordinary exact-head workflow groups passed; that does not clear other owners' review restrictions.
Historical release	Alpha v1.38.0, `61532fb82621a065e668fd11e57289dcfb65c67b`; preserved reference, not current readiness.
Mathematical engine	Base `8f13b39ed1718f5e8141ac1b37a10bb05c0f2900`; 53-file source closure recorded in the native protocol.
Prior observation	22 Oct 2026 23:10:55 UTC, retained prior V8 summary; complete raw campaign receipts unavailable.
Initial native freeze	22 Oct 23:40:39.270 UTC; followed earlier component observations.
Initial native run	22 Oct 23:43:29.649–23:43:33.151 UTC; full raw report retained.
Metadata amendment	22 Oct 23:46:02.115740 UTC; corrects specialist provenance labeling, not criteria, dispatch or cost.
Corrected native run	22 Oct 23:46:02.359–23:46:05.820 UTC; full 935,180-byte report retained.
Thermal protocol	Published 22 Oct 22:21:17 UTC. Twelve-case native replay: 23:46:40.075–23:46:41.215 UTC; its historical adapter identity is preserved separately from the corrected current worker.
Held composition	Thermal `f1ac7df` plus facade `6bb361d`; prospective, uncommitted tree `abc12184debbdd7f2c8ecf9928cc5482dc7b09f8`. It does not include the later recovery correction.
Held harness ancestor	#4863, `c455ad4`; existing restriction is not resolved by this paper. Dependent facade: #4865.
Local EVM source	1,335 source files match public `f1ac7df`; Node 22.17.1, Python 3.12.14, Hardhat 2.29.1, ethers 6.17.0, solc 0.8.37 and OpenZeppelin 5.6.1.

Scope	Identity and interpretation
Runtime correction	Node 22.17.1 and Python 3.12.14; full worker and nine-module closure bound before unfinished continuation.

Long-lived public coordination is issue #4850; the maintained four-repository record is cited in Ref. [17]. The reporting snapshot includes UTC observations on 23 October, corresponding to 22 October in America/Toronto. Publication or release status may change after that snapshot; scientific results remain bound to their recorded sources.

A.3 Reproduce the software, not merely the figures

Product reproduction requires the pinned repository, its documented setup and the recorded runtimes. Exact executable arguments, environment, failed attempts and source hashes are retained in the respective evidence directories. The paper bundle includes selected records, not installed dependencies or a complete chain database.

The mathematical report is produced by the standalone Node `compounding-probe.mjs`, which invokes the existing components and enterprise controller. It is not a separately admitted harness worker family. The thermal report uses the fixed thermal worker path and sends only typed observations to the worker; hidden evaluator heat and actuator multipliers are excluded. The local-contract evidence uses `tests/contracts`, its existing Hardhat configuration and fixture actors. These are three separately identified executions. A proposed combined thermal enterprise demonstration must preserve that distinction until it actually runs.

B Mathematical protocol details

B.1 Grammar, input generation and reference checks

The study evaluates dimension $n = 8$, while the bounded implementation admits powers of two through 32. For each of $n = 2, 4, 8$, transform search uses 64 supplied split/mix candidates: two input layouts, two output layouts and 16 sign matrices. Convolution candidates include divisors $-n, -1, 1, n$. A complete build evaluates 196 candidates and retains the costs of 187 rejections.

Component input generation resets the seed $s_0 = 20261022$ for each workload. The separate controller journey uses seed 20261021. The component recurrence is:

$$s_{i+1} = (1664525 s_i + 1013904223) \bmod 2^{32}, \qquad x_i = (s_{i+1} \bmod 15) - 7. \tag{17}$$

The checker has its own interpreter and bounded rational arithmetic and does not import the constructor's evaluator. It checks all basis vectors for supported linear programs and all basis-vector pairs for supported bilinear programs. It checks kernel spectra, inverse residuals, source/dimension applicability and resource bounds. Exactness follows only within this restricted semantics; rational numerator/denominator sizes are bounded.

The supplied specialist uses an iterative Walsh transform and transform-based XOR (dyadic) convolution. For the original kernel $(n + 1, 1, \ldots, 1)$, it uses the closed-form inverse

$$x_i = b_i/n - \frac{\sum_j b_j}{2n^2}. \tag{18}$$

No separately adapted expert for the changed kernel is evaluated. The mathematical child adds a new inverse program to the retained inventory; its direct and inherited paths yield the same child artifact identity.

Table 9. Per-round arithmetic projection for the three checked task families.

Generated task	Productive	Reference	Comparison	Total
Walsh transform	24	56	8	88
XOR convolution	88	120	8	216
Inverse	64	24	8	96
Generated round				400
Supplied expert round				360

B.2 Complete component account

Table 10. The 241 unique component charges. Workload comparison rows in the main paper attribute shared ancestors and are not additive campaign totals.

Charge class	Unique charges	Declared work
Construction	22	624,756
Admission	24	681,552
Execution, reference and comparison	191	24,744
Changed-kernel construction	2	9,948
Additional inherited-child admission	1	7,886
Singular rejection	1	32
Total	241	1,348,918

The scoped allowances are 2,000,000 units for construction, 50,000,000 for admission and 2,000,000 per task. The frozen evaluation does not grant authority, use a currency or submit transactions. Its cash field is zero for the invocation, while the broader unmeasured economic costs remain unknown. The review is local AI-assisted evidence inspection, not independent institutional admission.

C Thermal numerical model and challenge protocol

The reference model has two fluid temperatures, wall temperature and a thermal-memory temperature. It includes flow-dependent conductance, pump dissipation, ambient loss and wall–memory exchange. A fifth state integrates receiver heat. The evaluator implements the same declared model independently of the production interval worker. No numerical step-refinement study or physical-model validation is established by its agreement.

Table 11. 720-second simulated mission, following the 180-second reset. Commanded flows are in L/min; mismatch multipliers are applied by the evaluator.

Duration (s)	Heater (W)	Flow 1	Flow 2
120	0	0.12	0.10
120	42	0.12	0.10
180	0	0.12	0.10
120	64	0.18	0.15
180	0	0.18	0.15

For one-second measurement windows, the reference samples at 1/32 s and forms a mean differential temperature without using its accumulated heat as a worker input. Within the declared state/control box, the supplied second-derivative allowance is $M = 32\ \mathrm{K/s^2}$. The trapezoidal mean allowance is

$$\delta_{\mathrm{time}} \leq Mh^2/12 = 0.0026041667\ \mathrm{K}. \tag{19}$$

Adding the 0.005 K calibration premise, ceiling to μK and one quantization unit yields 7,606 μK. Density and specific heat have no uncertainty in this synthetic fixture. Common-cause corruption of the separate volume channel is not ruled out.

The native workload emits 720 interval checkpoints plus an evidence checkpoint. Accepted numerical bounds remain distinct from acceptance of a physical capability. The twelve cases cover supplied variations, not unseen physical samples; the retained historical adapter is `175d8a5`, and the current corrected worker has a different identity. Its prior evidence is reused only for the source scope recorded in the package.

D Acceptance boundaries and unresolved owners

Table 12. Accountable scope at handoff. Owner roles identify the relevant existing repository or integration authority; no new governance role is assigned.

Component	Evidence/status in this paper	Responsible owner
Insight / FusionPlan	Source-supported architecture and a compiled probe plan; full evidence-driven thermal opportunity lifecycle not demonstrated.	Alpha Insight / Studio maintainers
Nova-Seeds	Source-defined local commitment, ownership and successor-token paths; no live issuance or rights transfer claim.	Alpha Ascension maintainers
MARK / risk oracle	Simulated funding and review in the local suite; live eligibility and present role configuration unverified.	Alpha contract / admission owners
Sovereign / Business	Mathematical controller and held thermal facade execute bounded work; thermal Studio handoff refuses.	Existing #4863 / #4865 owners with Studio maintainers
Marketplace / Jobs	Alpha allocation and local source settlement exercised; alternate Jobs v0 and USDC Manager routes remain distinct.	Alpha, Jobs v0 and Manager backend owners
Agents / workers	Native fixed adapter and identity-bound recovery exercised; no universal Computer Use or live admission claim.	Alpha runtime and qualified Fleet maintainers
Architect	Retained lineage and one accepted changed-kernel child; singular rejection and cost loss retained. Replay optimization proposed.	Alpha capability / evaluator owners
Validator Council	Local distinct actors and exact-result review checks; independent organizational and physical acceptance absent.	Existing reviewer / admission owners

Component	Evidence/status in this paper	Responsible owner
Physical instrument	Simulated observations and conditional bounds; calibrated liquid-loop construction proposed.	A future participating laboratory and independent metrology reviewer
Commercial validation	No observed payer, price, total monetary bill or profit.	A future voluntary beneficiary and operating enterprise

The software candidate's passing checks qualify only their stated scope. The held harness restriction, independent acceptance, source-specific backend integration, clean deployment and physical qualification require their own evidence. This paper neither removes those restrictions nor claims the full AGI ALPHA architecture complete.

References

[1] Vincent Boucher. Transformational intelligence: Learning to change the game, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Agent-v0/blob/34a068e311451ce1d4fcc0a6cea85970a8faf457/docs/Transformational_Intelligence.pdf. MONTREAL.AI research manuscript; conceptual predecessor.

[2] Richard S. Sutton, Doina Precup, and Satinder Singh. Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence*, 112:181–211, 1999. doi: 10.1016/S0004-3702(99)00052-1. URL https://people.cs.umass.edu/~barto/courses/cs687/Sutton-Precup-Singh-AIJ99.pdf.

[3] Ross D. King, Jem Rowland, Stephen G. Oliver, Michael Young, Wayne Aubrey, Emma Byrne, Maria Liakata, Magdalena Markham, Pinar Pir, Larisa N. Soldatova, Andrew Sparkes, Kenneth E. Whelan, and Amanda Clare. The automation of science. *Science*, 324(5923):85–89, 2009. doi: 10.1126/science.1165620. URL https://research.manchester.ac.uk/en/publications/the-automation-of-science/.

[4] Benjamin Burger, Phillip M. Maffettone, Vladimir V. Gusev, Catherine M. Aitchison, Yang Bai, Xiaoyan Wang, Xiaobo Li, Ben M. Alston, Buyi Li, Rob Clowes, Nicola Rankin, Brandon Harris, Reiner Sebastian Sprick, and Andrew I. Cooper. A mobile robotic chemist. *Nature*, 583 (7815):237–241, 2020. doi: 10.1038/s41586-020-2442-2. URL https://pureportal.strath.ac.uk/en/publications/a-mobile-robotic-chemist/.

[5] MONTREAL.AI. AGI ALPHA: Native thermal witness and compounding implementation, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Agent-v0/tree/34a068e311451ce1d4fcc0a6cea85970a8faf457. Experimental source candidate, commit 34a068e311451ce1d4fcc0a6cea85970a8faf457.

[6] MONTREAL.AI. Compounding capability experiment: Retained execution, protocol and review, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Agent-v0/tree/34a068e311451ce1d4fcc0a6cea85970a8faf457/evidence/thermal-witness-20261022. Complete counted ledger under paper-declared-arithmetic-work/v1; source-bound evidence.

[7] MONTREAL.AI. Witnessed thermal measurements: Experimental workbench adapter, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Agent-v0/blob/34a068e311451ce1d4fcc0a6cea85970a8faf457/examples/THERMAL_WITNESS.md. Typed inputs, mathematical bounds, native execution and limitations.

[8] MONTREAL.AI. Ascension implementation map and six-stage acceptance boundaries, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Agent-v0/blob/34a068e311451ce1d4fcc0a6cea85970a8faf457/docs/agent/ASCENSION_IMPLEMENTATION_MAP_2026-10-22.md. Architecture target with actual interfaces, persistence, status and owners.

[9] MONTREAL.AI. Ascension execution follow-up: 23 October 2026, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Agent-v0/blob/34a068e311451ce1d4fcc0a6cea85970a8faf457/docs/agent/ASCENSION_EXECUTION_FOLLOWUP_2026-10-23.md. Source-specific evidence scopes and remaining boundaries.

[10] Vincent Boucher. Foundry v2—computational research edition, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Agent-v0/releases/tag/study-foundry-v2-2026-10-22-r1. Preserved research release, study-foundry-v2-2026-10-22-r1.

[11] Joint Committee for Guides in Metrology. Evaluation of measurement data—guide to the expression of uncertainty in measurement. JCGM 100:2008, Bureau International des Poids et Mesures, 2008. URL https://www.bipm.org/en/doi/10.59161/jcgm100-2008e. GUM 1995 with minor corrections.

[12] Barry N. Taylor and Chris E. Kuyatt. Guidelines for evaluating and expressing the uncertainty of NIST measurement results. NIST Technical Note 1297, National Institute of Standards and Technology, 1994. URL https://doi.org/10.6028/NIST.TN.1297.

[13] MONTREAL.AI. Disposable thermal enterprise composition: Execution and studio refusal, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Agent-v0/tree/34a068e311451ce1d4fcc0a6cea85970a8faf457/evidence/thermal-held-facade-20261023. Public f1ac7df composed locally with held 6bb361d; no integration clearance.

[14] MONTREAL.AI. Retained export and complete-worker identity: Recovery evidence, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Agent-v0/tree/34a068e311451ce1d4fcc0a6cea85970a8faf457/evidence/thermal-recovery-20261023. Original failure reproductions, corrected selector and source identities.

[15] MONTREAL.AI. Ascension local-EVM lifecycle and integer settlement qualification, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Agent-v0/tree/34a068e311451ce1d4fcc0a6cea85970a8faf457/evidence/ascension-local-evm-20261023. Unforked local EVM, distinct fixture actors and simulated balances.

[16] MONTREAL.AI. Ascension five-contract historical source, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Agent-v0/tree/c0fc5217e364a39a4cd123d17a8317c06b80f820/contracts/ascension. Historical revision c0fc5217e364a39a4cd123d17a8317c06b80f820; source identity is not present chain readiness.

[17] MONTREAL.AI. Maintained four-repository delivery and compatibility record, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Agent-v0/issues/4850#issuecomment-6352930563. Dated owner-maintained status; overall delivery incomplete at the study cut-off.

[18] National Institute of Standards and Technology. NIST policy on metrological traceability, 2024. URL https://www.nist.gov/calibrations/traceability. Accessed 22 October 2026, America/Toronto.

[19] U.S. Department of Energy. Waste heat recovery systems. Technology assessment, U.S. Department of Energy, 2015. URL https://www.energy.gov/documents/qtr2015-6m-waste-heat-recoverypdf. Quadrennial Technology Review 2015, Chapter 6.