ALPHA GENESIS

FLYWHEEL Ω

Turn the next advance in AI into your next ability to autonomously discover at scale.



RIDING THE EXPONENTIAL

Build the advantage that discovers—
and builds—the next advantage.

Vincent Boucher
Montréal, Québec, Canada

TRANSFORMATIONAL INTELLIGENCE ∞
4 November 2026

THE ADVANTAGE WORTH BUILDING

ALPHA GENESIS FLYWHEEL Ω

Riding the Exponential: Three Orders of Intelligence
for an Enterprise That Discovers What Becomes Possible

Vincent Boucher · MONTREAL.AI / QUEBEC.AI

Montréal, Québec, Canada · 4 November 2026

ABSTRACT

An advance in artificial intelligence is not yet an advantage for an enterprise. The missing step is to discover which valuable activities it makes attainable, construct the appropriate means, and learn from their use. We propose ALPHA GENESIS FLYWHEEL Ω, a frontier-adaptive discovery enterprise organized around three orders of intelligence: see the frontier, see its consequences, and improve the machinery that discovers those consequences. Building on Transformational Intelligence, the institution may change representations, information channels, production methods, coordination, and economic arrangements while preserving the authorized purpose. Its methods are origin-neutral: an existing frontier agent is adopted when it is the best complete route; costly inheritance is retired; construction is funded only when its prospective value justifies its full costs. A capability hypergraph represents jointly enabling means, while a belief-aware planner compares direct use, composition, invention, experimentation, delay, and non-entry. MuZero supplies an important planning analogy, not a ready-made model of unknown industries. We formalize discovery versus selection error, decision-relevant information value, and conditional conversion-lag advantage. A reproducible finite study selects a diagnostic experiment with value 34.41, versus 31.30 for the strongest no-probe planner, and chooses existing tools or non-entry in adverse cases. Separately, source-matched native simulation completes six accepted deliveries and records US$143.20 modeled profit; conventional routing ties the proposed selector and is retained. Neither result establishes real demand, general discovery superiority, or open-ended self-improvement. We specify the empirical programme needed to establish those claims and connect qualified capability construction to an economically sustained expansion of scientific and productive capacity.

> **Research status.** A research synthesis, formal decision contracts, exact authored computations, and reproduced implementation checks. This edition is not a deployed discovery enterprise, a claim of new physical measurements, or evidence of actual customer revenue. **What is newly verified here:** an independent planning oracle, 54 parameter checks, four source-matched operating reproductions, and the completed 62-test native suite. The operational source remains the stated historical reference, not current-main qualification.

Keywords: transformational intelligence; frontier awareness; opportunity discovery; model-based planning; economic invention; capability inheritance; self-sustaining enterprise.

THE PAPER AT A GLANCE

New intelligence is an input.
A qualified capability is an asset.
A better way to discover valuable capabilities is the deeper advantage.

The question. What valuable activity has just become technically or economically attainable, and what must be constructed to make it useful?

The product hypothesis. A customer-owned discovery application that accepts a consequential objective, authorized evidence and resources; constructs competing complete routes; produces usable artifacts; and explains what the evidence supports. The customer buys an ability, not a conversation among agents.

The enterprise rule. Use the best complete means, regardless of origin. Do not preserve internal machinery for prestige, reject an ordinary winner because it is ordinary, or book imagined successor benefits as present earnings.

The compounding hypothesis. A qualified invention can become an instrument, representation, or construction procedure that improves different successor work. That enabling relationship must be tested, not inferred from chronology or a growing archive.

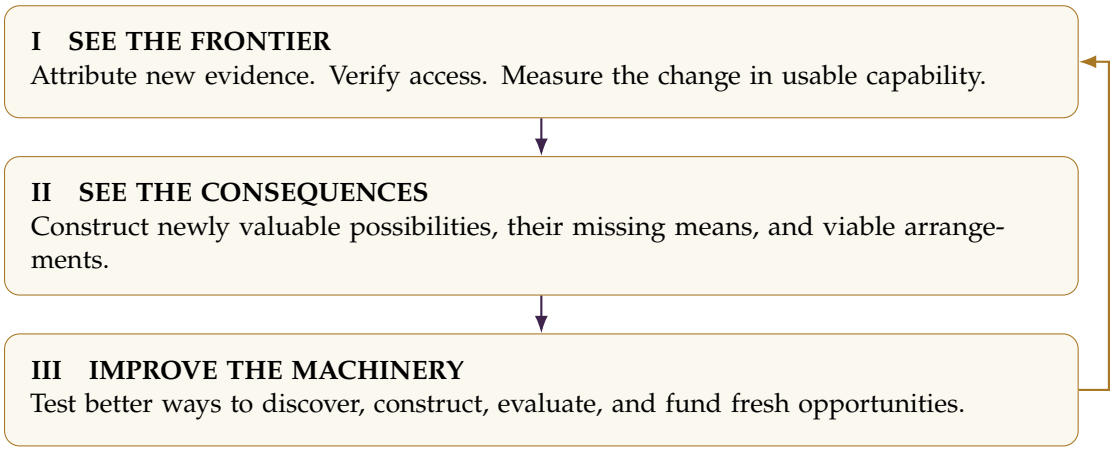


Figure 1: Three operational functions, not a universal intelligence scale. Independent evidence connects them; successful operation at one order does not certify the next.

Read by purpose. Sections 1–3 explain the thesis and frontier. Sections 4–8 define the decision architecture. Sections 9–10 report the computations and implementation. Sections 11–15 develop the product, qualification programme, and physical horizon. Appendices provide derivations, reproduction, and the claim register.

Contents

1 The advantage worth building 4

2 At the frontier: what changed, and what did not 5

3 Three orders of intelligence 6

4 A living model of productive possibility 7

5 MuZero is a planning inspiration, not the whole enterprise 9

6 The decision contract: discover, select, and learn 10

7 Frontier Replay: turn a release into a changed opportunity 11

8 The enterprise: value, participation, and feasible continuation 12

9 Exact decision studies: what the planner actually chooses 14

10 Native Enterprise 002: a source-matched execution record 16

11 A product people can actually use 18

12 How to establish that the enterprise contributes 19

13 The next consequential construction 21

14 An institution that outlives its models 22

15 From discovery to expanding productive capacity 22

16 Limitations and falsifiable commitments 23

17 Conclusion: the advantage that builds the next advantage 24

A Analytical details and interpretation 25

B Reproduction, source identity, and evidence 26

C Source lineage and responsible interpretation 27

D Claim–evidence register 28

1 The advantage worth building

The strategic opportunity is not merely to own a powerful answer generator. It is to build an institution that repeatedly turns improving intelligence into valuable new abilities. A model may become better at engineering; a measurement instrument may consequently become affordable; that instrument may enable a production process; and accepted production may supply resources for further construction. The valuable object is the supported relationship between these achievements, not an impressive list of them.

ALPHA GENESIS FLYWHEEL Ω is proposed as that institution. Its mandate is to become exceptionally effective at seeing what has just become possible, building what is most valuable, and improving its ability to discover the next opportunity. It joins an autonomous commercial flywheel to an inventive one, without making either's success depend on a false claim about the other. A fixed-capability operation can earn and reinvest surplus. An experimental method can improve research without yet finding a customer. The complete enterprise seeks a defensible connection between useful construction, accepted value, available resources, and fresh successor benefit [2, 3].

1.1 Four achievements, not one score

We distinguish *commercial value*, *scientific novelty*, *incremental institutional contribution*, and *improved invention*. An established algorithm can underpin a valuable business. A novel algorithm can have no economical application. A carefully designed workflow may improve delivery without improving how new workflows are discovered. A newer foundation model may improve every result while the surrounding institution adds no advantage.

These distinctions resolve a false opposition between using the frontier and competing with it. If an accessible frontier agent provides the best complete route, GENESIS should use it. That is an operating success, although it is not evidence that a proprietary constructor is superior. The additional institution must justify its opportunity selection, integration, acceptance, distribution, or renewal contribution after the overhead it introduces. It need not defeat every component it incorporates.

For an authorized mission $\mathcal{I}$, let $J_{\mathcal{I}}(s)$ denote an explicitly defined net outcome for complete system s. With a nonempty, disclosed comparison set $\mathcal{B}$,

$$\alpha_{\mathcal{I}}(s;\mathcal{B}) = J_{\mathcal{I}}(s) - \max_{b\in\mathcal{B}} J_{\mathcal{I}}(b). \tag{1}$$

The alternatives may include ordinary software, a direct frontier agent, a specialist, procurement, customer self-service, and non-entry. Hard requirements remain constraints. Where outcomes cannot legitimately be reduced to money or a scalar score, report the actual trade-offs instead of inventing an exchange rate. Equation (1) defines a comparison, not an estimate of all alternatives that might exist.

1.2 Transform the means, not the success criterion

The foundation is *Transformational Intelligence: Learning to Change the Game* [4]. Its separation of authorized purpose, problem formulation, and evaluation instrument is essential. A method's requirement is not necessarily the objective's requirement. If reliable delivery conventionally requires an expensive measurement, the enterprise can investigate a cheaper instrument, a selective measurement policy, or a production process that removes the uncertainty. It must preserve the outcome that made the work worth doing.

The representation, workflow, information channel, and interaction rules are therefore objects of invention. So are the procedures that invent them. But an agent cannot make a failed construction

successful by lowering acceptance standards, treating a forecast as collected money, or changing the customer's purpose. Freedom over means is productive only when the outcome and authority remain accountable.

The founding hypothesis. An origin-neutral institution that converts accessible frontier progress into qualified capabilities, economically useful delivery, and tested successor advantages can accumulate productive power beyond access to any single model.

2 At the frontier: what changed, and what did not

The evidence cutoff is **4 November 2026**. Table 1 records selected developments relevant to the construction problem, not a claim of exhaustive awareness. Publication, demonstrated capability, account-level availability, and authorization to deploy are different facts. An announcement also does not establish that a task was impossible a week earlier.

Table 1: A decision-relevant frontier snapshot. Provider announcements support their reported demonstrations, not the integrated performance of GENESIS.

Date / source	What the source establishes	Consequence for this programme
3 Nov 2026 OpenAI	A device workflow connected mechanical design, electronics, and software; manufacturing still required review and iteration [19].	Construct connected engineering artifacts. Treat manufacturing qualification as work, not as a model prediction.
3 Nov 2026 DeepMind	Agentic cheminformatics connects scientific reasoning with executable domain tools and candidate generation [23].	Make a research skill executable. Computed candidates still need the applicable experimental evidence.
3 Nov update Google Research	Computational Discovery adds an MCP integration route; the programme is described through a trusted tester channel [5].	An agent can access strong existing discovery machinery where eligible. Do not assume access from a public description.
1 Nov revision Hyperagents	The revised research describes editable task and meta-agent programs and reports improvements in evaluated domains [26].	Treat the improvement procedure as a candidate for improvement; compare against capable precedents.
9 Jul 2026 Google Cloud	AlphaEvolve is announced as generally available to Cloud customers on its agent platform [1].	It is a potential complete alternative, not merely an inaccessible research curiosity. Entitlement, costs, and authorization still matter.
Inspected 4 Nov 2026	OpenAI's mathematics repository contains manuscripts and proof artifacts at different verification stages, generated by an internal model [20].	Investigate useful results and proof dependencies. Public artifacts do not provide access to their generating model.

OpenAI's engineering example is relevant because its residual manufacturing corrections identify a concrete integration gap. The chemistry workflow is relevant because language reasoning is coupled to explicit domain operations. The economic inference is ours: lowering the cost of connected construction may change which purpose-specific capabilities are worth producing. Neither source supplies a demand curve or proves that our proposed enterprise can capture that value.

The competitive bar is already high. AlphaEvolve combines generated code with automated evaluation [16]; ADAS searches executable agent designs [10]; and the Darwin Gödel Machine investigates empirically evaluated self-modification [25]. GENESIS should use applicable mechanisms from this work and compare with strong accessible alternatives. Generic orchestration, more agents, or a renamed search loop is insufficient differentiation.

2.1 Riding the exponential is a conditional strategy

The title names a strategy for adapting to rapid progress, not an empirical law that all useful capabilities grow exponentially. Some improve discontinuously, others stagnate, and deployment can remain limited by data, verification, access, or production. The enterprise should exploit improvements where observed and remain viable when a preferred frontier does not advance.

A simple conditional model makes the conversion issue precise. Suppose a relevant available capability index follows $F(t) = F_0 e^{gt}$ over a stated interval, and institution i converts it with effective lag τ_i and efficiency η_i:

$$Y_i(t) = \eta_i F(t - \tau_i), \qquad \frac{Y_A(t)}{Y_B(t)} = \frac{\eta_A}{\eta_B} e^{g(\tau_B - \tau_A)}. \tag{2}$$

This is an algebraic implication of the model. It is not a measured growth rate, a profit forecast, or a guarantee that faster adoption is better. Qualification effort can increase η while increasing lag; premature deployment can destroy both. The optimization target is faster *qualified conversion*, after failures and migration costs, not a race to install every release.

When all flows are expressed in the same units of accepted outcomes per unit time, useful throughput is additionally bounded by the smallest of construction, verification, delivery, and demand capacity. Improving only the model does not remove those other bottlenecks. The valuable frontier may be an instrument that changes one of them.

3 Three orders of intelligence

The three orders describe operating functions and corresponding evidence requirements. They are neither an established scale of general intelligence nor a claim that an application acquires superintelligence by naming its stages.

3.1 First order: see the frontier

The observatory asks: *What changed that could alter a decision?* It combines broad model knowledge with current primary sources, executable capability probes, access records, and failures observed in use. A frontier LLM is a powerful proposal and interpretation resource, but its internal knowledge is not a source of live access status, current price, or a verified scientific conclusion.

A useful frontier record binds the source and version to its publication time, observation time, claimed change, observed task behavior, prerequisites, access status, cost evidence, and expiry or invalidation conditions. Corrections, retractions, revoked permissions, and regressions are events too. A public statement that a model can design a circuit should not be converted directly into permission to fabricate or a claim that a particular board is safe.

Awareness is measured over declared source and task populations. Relevant-release recall, detection delay, source correctness, and the cost of verification are useful. "Everything relevant" has no observable universal denominator. Independent horizon scans and deliberately different retrieval methods can expose omissions in the institution's own catalogue. They cannot prove omniscience.

The observatory should prioritize evidence by the decision it may change. Reading another hundred articles about a capability already well understood may be inferior to checking one unresolved access or reliability condition. Conversely, attention focused exclusively on today's backlog can miss a new class of activity. Targeted monitoring and deliberate exploration must therefore coexist.

3.2 Second order: see the consequences

The consequence engine asks: *What becomes worth doing because this change occurred?* It searches beyond applications that merely substitute cheaper inference into an unchanged product. A capability might enable a new observation, make a costly experiment economical, remove an engineering dependency, or create an acceptance procedure that allows a different commercial arrangement.

Its object is the *complete productive arrangement*: beneficiary, useful outcome, productive means, evidence, participant incentives, resource commitments, and viable timing. A product that is technically feasible but cannot be verified, afforded, distributed, or voluntarily purchased is not a completed opportunity. These are construction targets, not reasons to assume the market will eventually cooperate.

Search proceeds in both directions. Backward reasoning identifies what blocks a valuable outcome. Forward reasoning asks what newly available means could enable. Their intersection proposes a construction. Cross-domain structural transfer can suggest relationships that topical retrieval misses; a verified failure can reveal the missing interface; and an alternative formulation can eliminate the need for that interface entirely [3].

The engine must retain rival hypotheses: the installed route may already suffice, the new capability may be irrelevant, or a simpler direct agent may win. Generation is not endorsement. Each proposal needs an executable path and a discriminating test before substantial commitment.

3.3 Third order: improve the machinery that sees possibilities

The third order asks: *Which change would make future opportunity discovery and construction more effective?* The object can be a representation, search policy, experimental strategy, retrieval method, risk estimate, tool interface, or the procedure that proposes such changes.

A useful inheritance is more than a transcript. It may contain a working instrument, an applicability model, and a procedure for constructing related instruments. These are distinct assets. Only the first may initially have evidence. The deeper gain occurs when a retained procedure improves fresh construction, not merely when a new task reuses a prebuilt answer.

DreamCoder supplies a bounded precedent for learned abstractions supporting subsequent program search; POET supplies one for co-developing environments and solutions with transfer [7, 24]. Hyperagents directly makes the modification procedure editable in its evaluated systems [26]. Our proposed enterprise extension concerns whether such mechanisms improve economically consequential discovery under complete resource accounting and changing external capability.

An updated foundation model, a richer initial archive, and a better discovery procedure can all improve results. The third-order claim must separate them. Without that separation, the institution can simply inherit progress made elsewhere and call it self-improvement.

4 A living model of productive possibility

We propose a hybrid model rather than a single neural network expected to know the economy. It contains an authoritative resource and permissions layer, a source-grounded capability representation,

probabilistic beliefs about unresolved mechanisms and markets, and generators that propose new executable actions. The model is useful only insofar as it improves real decisions.

4.1 Capabilities are conditional operations

A capability record describes what an operation accepts, produces, consumes, and changes; where its evidence applies; which rights and dependencies it requires; and how it can fail. Access, qualification, and economic attractiveness are separate attributes. An operation can be accessible but unreliable, reliable but unsuitable, or useful but unaffordable under the present commitment schedule.

Let $\mathcal{H}_t$ denote the qualified capability stock and $\mathcal{G}_t$ a directed hypergraph of possible enabling relationships. A hyperedge $\{h_1, \ldots, h_k\} \to h'$ indicates that several prerequisites together may enable a construction. It must carry its assumptions, realization, cost, and evidence status. A hypothesis edge is not a qualified causal fact. Circular descriptions of mutual enablement do not supply a feasible bootstrap sequence.

This representation matters because a collection of individually weak components can be jointly enabling. An instrument without calibration may have little value; calibration without an instrument may have no use. Their bundle may make a useful distinction measurable. A greedy rule that admits only individually profitable additions can miss the combination. Section 9 gives a transparent numerical example, not an algorithmic novelty claim.

4.2 The graph must not become the boundary of imagination

A finite capability graph is an instrument for reasoning, not an exhaustive map of reality. Frontier agents should be allowed to propose operations and representations outside its existing grammar. Such proposals enter a development environment with explicit effects and checking requirements; they do not grant themselves execution authority.

There are three different questions. Can an existing planner find a route through known operations? Can a constructor implement a proposed new operation? Can an open-world generator discover that operation or a valuable outcome not already supplied? The current prototypes primarily address the first and parts of the second. Solving an authored graph exactly does not settle the third.

A practical diagnosis records whether a failure comes from unavailable evidence, inadequate reasoning, insufficient resources, physical constraints, an avoidable dependency, or an inaccurate model of buyer behavior. Repeated failure alone is not proof of impossibility. Likewise, saying "reframe the problem" does not create a missing observation.

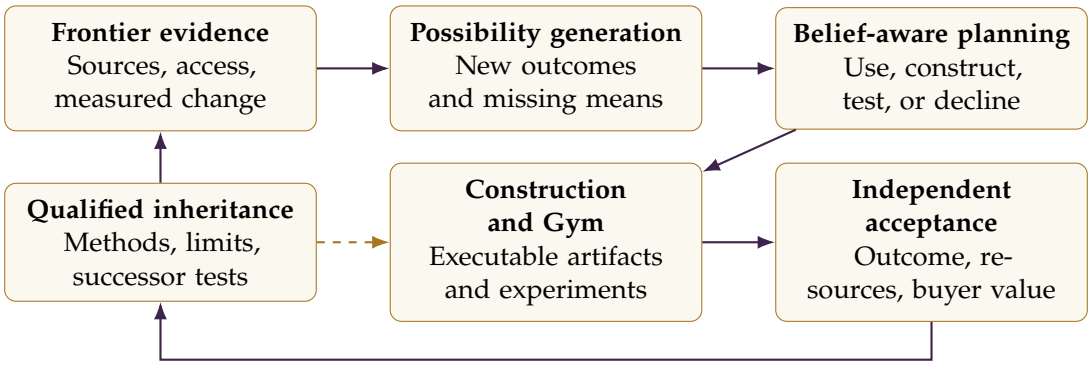


Figure 2: Proposed information and construction loop. The independent acceptance boundary and authoritative resource controls constrain every path; they are not editable rewards owned by the candidate.

5 MuZero is a planning inspiration, not the whole enterprise

MuZero learns representations and latent transitions sufficient to predict decision-relevant reward, policy, and value, then plans with a learned model. Its reported evaluation covers Atari, Go, chess, and shogi [22]. The important lesson is that a planner need not reconstruct every physical detail to make useful predictions. The inference that this already supplies an open-world scientific or commercial discovery engine would be unwarranted.

GENESIS faces different conditions: new tools can change the action repertoire; scientific experiments may be expensive and irreversible; market responses are strategic; observations are partial; access can disappear; and rewards may arrive late. A compelling imagined trajectory is not independent evidence of its physical or commercial feasibility.

Table 2: Methods to incorporate when warranted, not a compulsory stack. The architecture should be allowed to simplify.

Method	Useful contribution	What it does not establish here
Exact decision trees / dynamic programming	Transparent optimization in small declared models; independent reference calculations.	Optimal discovery outside the authored state and action space.
Bayes-adaptive planning [21]	Joint reasoning about uncertain dynamics and actions that improve information.	Correct priors, an exhaustive hypothesis set, or calibrated market forecasts.
MuZero / sampled-action planning [11, 22]	Learned decision-relevant dynamics and search over promising actions.	Autonomous invention of arbitrary action semantics or a valid model of the economy.
Dreamer-style world models [8]	Learning policies using imagined trajectories in evaluated control settings.	That imagined scientific or commercial outcomes are externally valid.
Executable evolutionary construction [10, 16]	Generate and evaluate algorithms or agent designs.	Demand, financing, or general discovery advantage in our application.
Editable improvement procedures [25, 26]	Treat the machinery of improvement as an experimental object.	Unrestricted, indefinitely accelerating self-improvement.

The immediate engineering choice is a small, inspectable planner with explicit uncertainty, supplemented by strong proposal generators and conventional specialist tools. Learned planning becomes a challenger when the institution has enough relevant interaction evidence to train and calibrate it. The choice should be earned against exact or simpler approximate methods on the same decision requirements, including training, search, validation, and inference costs.

5.1 Separate learned predictions from authoritative facts

Model-predicted revenue is not a collection. An estimated probability of customer acceptance is not consent. A learned transition predicting access to a tool does not create access. The planner may reason probabilistically about each, but the execution layer must obtain the actual evidence or authorization required for the action.

The symbolic layer therefore preserves cash, obligations, identities, resource limits, and grants of authority. Learned components estimate consequences and propose constructions. This division is

not a claim that symbolic records are infallible; their implementations need checking. It prevents a plausible forecast from silently becoming an operating permission.

6 The decision contract: discover, select, and learn

Let $x_t = (b_t, \mathcal{H}_t, \mathcal{G}_t, c_t, o_t, \ell_t)$ contain beliefs, capabilities, enabling hypotheses, available cash, obligations, and remaining time or other finite resources. The authorized action set may include direct use, procurement, partnership, composition, construction, an experiment, delay, retirement, and stopping. Actions can alter future capabilities as well as beliefs.

For a finite horizon, one risk-neutral reference problem is

$$V_t(x) = \max_{a \in \mathcal{A}_{\text{feasible}}(x)} \mathbb{E}[r_t(x, a, Z) - k_t(x, a, Z) + \gamma V_{t+1}(x') \mid x, a]. \tag{3}$$

Here r_t and k_t are incremental realized-value and cost quantities in declared common units, and Z is the uncertain outcome. The transition tracks resource commitments and settlement timing. A terminal value, if used, must have a stated realizable interpretation. Future option value already represented in V must not be added again as a separate bonus. Hard safety, permission, quality, and liquidity requirements belong in feasibility, not as penalties that high predicted profit can simply outweigh.

Equation (3) is a model-based decision specification, not a solved game for the actual enterprise. Risk preferences, robustness requirements, and incomplete models may demand other formulations. In every case, the forecast distribution, feasible set, solution method, and approximation gap should be stated. "Game-theoretically optimal" is not meaningful without players, information, strategies, timing, incentives, and a solved or bounded comparison.

6.1 Discovery error is not selection error

A perfect selector can choose only among the possibilities it receives. For an analytically specified universe of feasible arrangements with values $J(a)$, let $V^* = \sup_a J(a)$ and let C be the nonempty candidate set produced by a discovery process. Then

$$V^* - J(\hat{a}) = \underbrace{V^* - \max_{a \in C} J(a)}_{\text{discovery gap}} + \underbrace{\max_{a \in C} J(a) - J(\hat{a})}_{\text{selection gap}}. \tag{4}$$

This identity separates two targets; it does not estimate either in an open world. A finite planner may have zero selection gap in its model while overlooking the most valuable application. A larger candidate list can also be worse after the costs of generating, integrating, and checking it. Complete discovery campaigns, rather than final plans alone, must be compared.

Proposition 6.1 (Conditional selection bound)**.** *For a declared finite candidate set C, suppose $|\hat{J}(a) - J(a)| \le \epsilon$ for every $a \in C$ and $\hat{J}(\hat{a}) \ge \max_{a \in C} \hat{J}(a) - \delta$. Then*

$$\max_{a \in C} J(a) - J(\hat{a}) \le 2\epsilon + \delta. \tag{5}$$

Proof. For a maximizer a^* in C, $J(a^*) \le \hat{J}(a^*) + \epsilon \le \hat{J}(\hat{a}) + \delta + \epsilon \le J(\hat{a}) + \delta + 2\epsilon$. □

This elementary bound clarifies the priority of calibrated values and adequate candidate generation. An exact optimizer does not compensate for confidently wrong economics. Uniform error control must be justified; assigning a small ϵ does not establish it. No term in the bound covers opportunities absent from C.

6.2 Buy information when it changes the decision

For an experiment e with outcome Z, a one-step value-of-information comparison is

$$\mathrm{VOI}(e) = \mathbb{E}_Z\left[\max_{a \in A(Z)} \mathbb{E}[J(a) \mid Z, e]\right] - \max_{a \in A_0} \mathbb{E}[J(a)] - K_e. \tag{6}$$

K_e includes the applicable experiment, interpretation, delay, and opportunity costs. A sequential problem must include the consumed time and resources in the transition rather than subtract them twice. The post-experiment feasible set may change because the experiment constructs a useful instrument, but that effect requires an actual construction and a validated applicability claim.

A measurement can be highly informative yet economically irrelevant if no subsequent action changes. Conversely, a small amount of information can be valuable when it avoids an expensive mistaken build. Cases where no useful affordable experiment is found, where neither initial hypothesis is correct, and where stopping is best are part of the specification, not embarrassments to suppress.

6.3 Retain the contribution, not the sunk cost

The prospective value of retaining capability h is the best achievable continuation with it, net of its future carrying and adaptation costs, compared with the best continuation without it. The latter may buy or construct an alternative. Historical acquisition cost remains in programme accounting, but is not charged a second time merely to decide whether an already acquired capability should stay active.

A cheap inert archive, an actively maintained executable tool, and a production-qualified service have different costs. Retiring one need not erase the evidence behind another. Complementary capabilities should be assessed jointly when their contribution depends on a bundle. Neither "retain everything" nor "delete anything without standalone revenue" is an adequate policy.

7 Frontier Replay: turn a release into a changed opportunity

Frontier Replay is the proposed operational bridge between awareness and construction. The institution preserves a compact record of valuable unresolved opportunities: the intended outcome, strongest complete alternative, suspected obstruction, attempted means, incurred costs, and evidence that could change the next decision. It does not treat failed attempts as permanently impossible or as an obligation to continue indefinitely.

When a new capability becomes accessible, replay begins with a specific counterfactual: *Would this change remove the obstacle that mattered?* A stronger numerical coder may reopen a failed solver construction. A better engineering interface may make a fixture manufacturable. A new experimental result may invalidate the original opportunity entirely. Each is a legitimate frontier consequence.

7.1 A capability delta, not a release name

The useful input is a measured or testable difference in capability, cost, reliability, or access. A large benchmark improvement can be irrelevant to a particular customer outcome. A small reduction in a previously dominant integration cost can change feasibility. The application should therefore bind a frontier event to the dependency it is alleged to change, not merely attach the newest model label to an old workflow.

A replay record should answer five compact questions: what changed; which opportunity might be affected; which enabling link is uncertain; what executable test distinguishes the alternatives; and what

action follows each plausible outcome. The test can select the incumbent, retire a capability, or choose non-entry. A repeated comparison is not fresh confirmation when its outcomes have already influenced development.

7.2 A productive replay episode

The following is an operating specification, not a claim that the released prototypes execute every step:

1. Verify the event, access, scope, and applicable authority. Preserve uncertainty where evidence is missing.
2. Generate materially different consequences, including a direct existing-tool route and an option that avoids the assumed construction.
3. Compare complete prospective value and feasible funding sequences. Select an informative test or construction only when warranted.
4. Build and challenge in the development environment; record failed branches and resource consumption.
5. Freeze the candidate and assess it on fresh work under the original outcome requirements.
6. Deliver within its supported scope, reconcile actual obligations, and test any claimed successor contribution separately.

Backlog replay must coexist with de-novo opportunity discovery. Otherwise, the institution only revisits questions it knew to ask before the frontier changed. Search budgets should be explicit, but no permanent field catalogue or fixed percentage of spending is inherently optimal. Allocation itself is a candidate for experimentally demonstrated improvement.

8 The enterprise: value, participation, and feasible continuation

The customer promise is a useful capability that the buyer prefers to its complete alternatives. The enterprise should search the outcome, production method, verification process, commercial arrangement, and financing path together. Better production economics are insufficient when the buyer's alternative improves even faster.

8.1 The price window describes a construction problem

In a stated quasilinear model, let v_i be buyer i's expected value of an accepted outcome, u_i^{out} the utility of its best alternative, and b_i its non-price burden. Voluntary participation requires

$$p_i \le v_i - u_i^{\text{out}} - b_i. \tag{7}$$

The burden includes the buyer's own setup, compute, attention, integration, risk, and residual checking. For n accepted deliveries, fixed lifecycle cost F, per-delivery cost c, and required total surplus m, a common viable price additionally needs

$$p \ge c + \frac{F + m}{n}. \tag{8}$$

These are feasibility inequalities under their assumptions, not a market-equilibrium solution. They identify what might need to change: improve value, lower complete costs, reduce the buyer's burden, or reach additional real demand. The variable n cannot be enlarged on paper without buyers, capacity, acquisition, and funded obligations.

A technically excellent construction with no plausible price window should be declined or reformulated. An intermediate instrument can still be justified when its credible downstream contribution or decision value exceeds its complete cost. Its hypothetical future benefit is not earned revenue, collateral already secured, or unrestricted cash.

8.2 Competition can help society while compressing our margin

A public capability is available to rivals and customers as well as to GENESIS. They may develop cheaper alternatives, copy an application, or change the terms of participation. A rational enterprise must model these responses when material, including the buyer's option to do the work itself or decline.

An authored arithmetic counterexample makes the point. If a buyer's maximum acceptable price is 130 and our complete unit cost is 100, the price-ceiling margin is 30. A public advance might reduce cost to 40 while lowering the acceptable price to 60 because outside options improve. The margin becomes 20, despite a much better production capability. These are illustrative numbers, not observed willingness to pay.

Potential defensibility comes from demonstrated conversion ability, qualified interfaces, distribution, appropriately permissioned knowledge, useful experimental assets, and reliable delivery. None supplies an automatic monopoly. The operating question is whether that integrated contribution remains worth its cost as alternatives improve.

8.3 A self-sustaining enterprise is a cash-and-obligations claim

Let C_t be cash, P_t protected obligations and reserves, and $A_t = C_t - P_t$ available cash under a declared accounting convention. A construction is currently financeable only if its required commitments are supported by actual available resources and permissible financing. Positive expected terminal profit does not by itself satisfy that condition.

The accounting path is:

$$\begin{aligned}\text{Offer} \to \text{voluntary purchase} \to \text{funding} \to \text{production} \to \text{acceptance} \\ \to \text{collection} \to \text{obligations covered} \to \text{justified continuation.}\end{aligned}$$

An invoice is not a collection. Customer advances remain obligations until earned. Investment proceeds are not operating revenue. Modelled receipts are not actual funds. No parallel planning ledger should replace the native account that authorizes commitments.

For a stated horizon, an operational self-sustainability claim needs repeated accepted delivery, complete economic costs, sufficient liquidity through the path, and a supported continuation policy without unreported capital injections. This does not forbid legitimate financing. It requires financing to be distinguished from earned-funded growth and its terms to be accounted for. Seed-funded protection reserves can remain essential even when current operating expenditures are paid from earnings.

8.4 Two flywheels, several sources of progress

The *commercial flywheel* repeats worthwhile delivery and uses available surplus for productive continuation. The *invention flywheel* constructs qualified means and tests their contribution to new work. A third-order gain improves the procedure for finding those means. Exogenous foundation-model progress is a fourth source of improvement, not a substitute for evidence of the other three.

There is no requirement that the enterprise invent on every cycle. When direct delivery or distribution is the best investment, it should win. Conversely, an instrument that creates substantial future opportunity

can be worth constructing before it earns standalone revenue. What matters is justified prospective value, feasible resources, and evidence of the enabling relationship, not adherence to a permanent research-spending ritual.

9 Exact decision studies: what the planner actually chooses

This edition re-executes the supplied finite Bayesian planner [13] and adds a separately implemented oracle. The oracle re-creates the transitions in its own tuple-state recursion; it does not call the producer's action or outcome functions. Exact rational arithmetic avoids sampling uncertainty within this authored model. There is no trained MuZero network, model API call, live market, or discovered action grammar in this study.

9.1 A disclosed finite world

There are two hypothetical markets, A and B, each deliverable once, and four decision steps. Starting cash is 25 model units and the protected reserve is 10. An existing frontier route has net returns 14 and 16; a specialist has 10 and 12. Constructing a candidate instrument costs 12. With prior probability 7/20, it succeeds and supports returns 32 and 36, with zero per-delivery carrying cost in the base case. Otherwise, it adds no useful delivery ability. Construction reveals viability perfectly in this finite world; this strong assumption is not an empirical property of research.

Before construction, the planner may buy one diagnostic signal for 3 units. Its sensitivity is 9/10 and false-positive probability 1/10. Both experiment and construction consume a decision step. Acquisition must leave the protected reserve intact. Non-entry is always available. The return parameters and likelihoods are stipulated, not learned from buyers or current model releases.

The state tracks horizon, cash, posterior belief, whether the experiment or construction has occurred, and which markets have been served. The reference computes the strongest continuation over all permitted actions, including using an existing tool first to change later liquidity. This avoids comparing the candidate only with a weak, purely myopic policy.

Table 3: Exact authored decision results. Values are model units, not observed dollars. The independent recursion agrees in each case.

Case	Selected first action	Full plan	No probe	Direct only
Uncertain useful instrument	Buy diagnostic	34.41	31.30	30.00
Existing frontier returns rise to 28 and 30	Use frontier route	58.00	58.00	58.00
Carrying cost rises to 25 per use	Use frontier route	30.00	30.00	30.00
Every delivery route has negative net return	Stop	0.00	0.00	0.00

The baseline diagnostic is positive with probability 19/50, producing posterior 63/76; the best next action is construction. A negative result has probability 31/50 and posterior 7/124; the best continuation uses existing frontier delivery. The full value is 3441/100 = 34.41. The strongest planner forbidden to probe obtains 313/10 = 31.30, so the net value of access to the diagnostic is 311/100 = 3.11.

The often tempting comparison with direct delivery alone gives 34.41 − 30 = 4.41. That is a weaker comparison. The paper uses the stronger 3.11 difference for the information claim. Perfect initial information, before charging for its acquisition, has value 39.10; the corresponding improvement over

the strongest no-probe plan is 7.80. These values apply only to the stated horizon, resource rules, likelihoods, and payoffs.

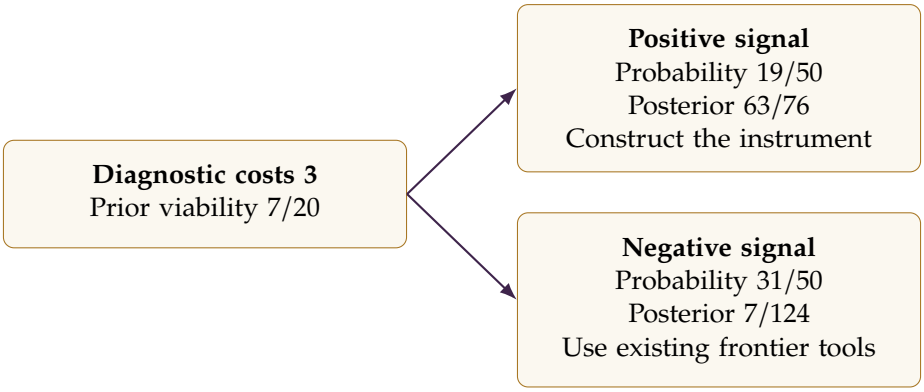


Figure 3: Information is valuable because it changes construction versus direct-use decisions. More information is not itself the enterprise objective.

9.2 Sensitivity, complementarity, and retirement

The independently coded oracle and the original planner agree across 54 additional configurations, crossing three priors, three construction costs, three carrying costs, and two signal sensitivities. These checks test implementation agreement in a finite supplied family. They do not validate the family's economic realism or identify the optimal open-world opportunity.

Three additional exact illustrations test the intended decision logic. First, suppose an instrument costs 4, an interface costs 3, and benefit 20 appears only when both exist. Their isolated net values are -4 and -3, but the bundle yields 13. A requirement that each component independently pay for itself would exclude a worthwhile combination. The bundle and its payoff are authored, not discovered.

Second, a retained capability expected to save 8 per use over twelve uses, with remaining integration cost 15 and carrying cost 20, has prospective value $12 \cdot 8 - 15 - 20 = 61$. Increasing carrying cost to 120 makes it -39. This is a keep-or-retire illustration, not permission to erase its acquisition cost from the historical programme account.

Third, the competitive-diffusion example in Section 8 establishes that falling production costs need not increase seller surplus. A capability portfolio therefore has to model changed alternatives as well as improved internal performance. The analytical companion produces all these values and tests their stated boundary cases.

9.3 Finite composition is useful but not open-world discovery

The separate Frontier-to-Possibility 001 prototype [12] searches a declared graph of executable operations. In its authored example, an existing frontier-agent route plus a standard checker costs 3,200 cents, versus 6,400 when that declared capability is unavailable. Its positive modeled offer has 5,500 cents contribution after the specified costs; the other examples reject negative economics or buyer refusal. Restricted or unverified capabilities are not silently treated as installed tools.

The 25 tests in that supplied snapshot pass on re-execution. They support the finite route and exclusion logic. Prices, access assumptions, and buyer values remain declarations. This graph cannot discover a valuable outcome that its grammar never represents. The engineering target beyond it is source-grounded generation and construction of new operations, not a larger table of invented payoffs.

10 Native Enterprise 002: a source-matched execution record

The native implementation connects decisions to actual deterministic production code, independent artifact checks, and the repository's existing `JournalAccount` and adaptive `Account`. Its commerce is simulated. It is an additive research component, not the complete discovery application or an admitted Blue Sky invention [15].

For this edition, we extracted the corrected supplied package, applied its additions to the byte-preserved native reference at commit `4a87b455373be72811ad3a0c654a116c3604a374`, and completed its 62-test suite. We also ran four new, separate six-cycle accounts, independently audited them, and resumed each to the identical final result. This new execution supersedes the supplied repackage's incomplete full-suite recheck for these exact bytes. It does not qualify the later repository main or imply a remote deployment.

10.1 The strongest existing method is selected

The primary fixture exposes conventional, generated, and inherited adapters, together with an unverified announced capability. It uses the qualified conventional route when the complete declared costs favor it. A low advertised price is insufficient to admit the unverified capability. Buyer participation checks include outside utility, non-price burden, and budget.

This behavior is intentional. GENESIS should be an institution that uses effective intelligence, not a competition in which its own constructor must always win. The later routing experiment also ties the strong conventional method; its recorded decision is `RETAIN_DIRECT_CONVENTIONAL_TIE`. That result establishes neither meta-discovery superiority nor a new general algorithm.

10.2 What the central account records

The fresh primary account starts with US$500 of simulated capital and US$25 working reserve. Six synthetic deliveries are accepted and settled at US$40 each. The code executes reconciliation work on supplied synthetic sources; an independent local checker tests the artifacts. Five repeat production expenditures and a US$5 research expenditure are funded from native unrestricted earnings.

The `EARNED_ONLY` continuation policy is a declared experimental arm, not a universal prohibition on sensible seed-funded or financed expansion. This finite schedule is not a trailing-year profitability result.

The modeled reconciliation is

$$\$240.00 - \$54.76 - \$37.04 - \$5.00 = \$143.20, \tag{9}$$

where the terms are collections, operating expenses, an authored tax provision, and research. The tax provision is part of the declared simulation, not tax advice. The account ends with cash US$643.20 and available cash US$377.00. Equity backs US$266.20 of protection, including a US$241.20 tail reserve under this native accounting model. Accordingly, earned-funded expenditure is *not* evidence that the enterprise no longer needs seed-backed protection.

Table 4: Four freshly reproduced accounts, all with six scheduled cycles. Each uses the same source and its declared stress. Accounts are not pooled. All pass local audit and unchanged resume.

Scenario	Settled	Refunded	Net USD	Observed model behavior
Central	6	0	+143.20	Five earned-funded repeats; research funded.
No demand	0	0	−7.00	Six non-entry decisions; incurred startup costs remain.
Failed acceptance	0	1	−13.50	Refund, then five non-entry decisions.
Payment failure	0	0	−7.50	One unpaid attempt, then five non-entry decisions.

All actual customer revenue in these experiments is zero. The expense envelopes, buyer values, and willingness to purchase are authored assumptions. Actual complete development, inference, infrastructure, customer-acquisition, and maintenance costs have not been established by the simulation. No live model is called by the native run. A functioning financial trace is evidence about execution and accounting behavior, not market demand.

10.3 Verification granularity

The native suite tests method choice, voluntary refusal, uneconomic startup rejection, protected funding, refunds, missing payment, unchanged resume, interrupted execution, source and configuration changes, and tampered outputs. Recovery retains the original account authority rather than reseeding. These are valuable engineering properties, but locally administered procedural separation is not independent third-party certification.

Table 5: Checks actually executed for this edition. Counts are disaggregated because they test different claims.

Suite	Passed	Scope
Native Enterprise 002	62	Decision, observer, integration, failure, audit, and recovery behavior on the historical native reference.
Frontier Decision Model 0.1	19	Supplied finite Bayesian planner and its controls.
Frontier-to-Possibility 001	25	Supplied composition and observer snapshot.
Edition 2.0 analytical companion	15	Independent planning recursion, 54-case agreement, information value, retirement, and authored counterexamples.

Passing tests establishes the checked computations and tested invariants. It does not make the environment assumptions true, provide formal verification of the entire codebase, or establish business-scale autonomy. The paper's claim register preserves these distinctions.

11 A product people can actually use

The proposed customer-facing product is *GENESIS Discovery Studio*: a local-first application for commissioning a capability, observing its construction, and using its qualified result. The opening question is "What should become possible?" The customer supplies the desired outcome, relevant evidence, available means, permitted actions, and resource requirements. The application must not require the customer to understand an agent hierarchy to know whether the work succeeded.

11.1 The result is an ability, not a transcript

An accepted capability package contains an editable artifact, executable operating procedure, acceptance evidence, resource record, applicability conditions, and a supported continuation path. A software instrument may ship as a tested program and input contract. An engineering instrument may additionally require dimensioned CAD, component specifications, calibration, and new measurements. An unsuccessful campaign should deliver an inspectable failure or unresolved condition rather than a polished claim of success.

The interface should show the actual construction and measured results. Generated renderings cannot stand in for executable CAD; colorful reconstructions cannot stand in for accurate decisions. The primary status vocabulary should distinguish proposal, development, simulation, physical measurement, scoped qualification, and commercial acceptance. Users should be able to inspect the strongest alternative and understand why GENESIS chose it.

Publicly available foundations reduce unnecessary infrastructure work. The documented Codex app-server provides persistent agent sessions, streamed activity, tool integration, and approval handling [17]. An eligible signed-in local workflow and a separately billed API workflow have different authentication and usage conditions [18]. The application should verify actual access rather than prescribe a hard-coded supposedly latest model or assume unlimited entitlement. A public component is not automatically free, compatible, qualified, or authorized for new spending.

11.2 An interface organized around decisions

The proposed workflow is **describe the outcome, inspect the alternatives, launch authorized construction, review the evidence, and use the accepted capability**. A frontier event can then trigger a challenger in a separate development branch. The current qualified version remains usable until a replacement earns adoption.

For ordinary authorized computation, the system should proceed without recurring unnecessary confirmations. Purchases, account access, physical effects, and material objective changes require the applicable authority. These boundaries should be explicit controls rather than paragraphs hidden in a conversation. User-friendly autonomy means fewer avoidable interactions, not fewer safeguards or concealed failure states.

The buyer's data and confidential results remain within their granted use. Generic reusable methods can be retained only where rights permit, without transferring customer secrets. Portability means exportable artifacts, evidence, and operating history; it does not imply ownership of a provider's foundation-model weights or unrestricted rights to every input.

11.3 Commercial entry and expansion

The initial offer is a hypothesis: *a qualified, purpose-specific research or engineering capability delivered through a self-service application*. A complete deployment, a supported capability package, and a continuing mission application are candidate commercial forms, not proven price points. Adoption must be compared with a competent specialist and with a customer using the same frontier tools directly.

The business should track buyer success, time to accepted outcome, repeat use, switching or integration burden, collection, and complete cost. An additional deployment should not depend on founder consulting unless that work is actually part of a separately disclosed service. Actual human intervention is counted where it occurs; fictional mandatory human review is not added to an otherwise autonomous task.

Expansion is justified by repeatable value, supportable capacity, and feasible funding. There is no universal customer-count ceiling, required startup capital, or compulsory industry catalogue. Nor does an enormous industry establish obtainable demand. The repository's US$3 billion trailing-365-day complete-cost profit objective remains a milestone to investigate, not a forecast or a qualification result [14].

12 How to establish that the enterprise contributes

The central scientific question is not whether a sufficiently capable model can produce an interesting artifact. It is whether the institution improves complete, useful discovery campaigns after its costs and whether retained work changes subsequent construction. The unit of evaluation must match the claim.

12.1 Three comparisons, three legitimate verdicts

Operating comparison. Compare complete delivery against the strongest feasible customer alternatives. A direct frontier agent may be the production backend in both systems; GENESIS's additional contribution can lie in opportunity discovery, integration, acceptance, or commercial convenience. Its overhead must remain visible.

Inheritance comparison. Give matched fresh campaigns the same model and initial resources, with and without the claimed inherited asset. Permit the control to construct or acquire its own alternative. Count the asset's acquisition under a declared lifecycle allocation, plus transfer, requalification, and maintenance. Chronological succession or a named parent pointer is not causal evidence.

Discovery-policy comparison. Compare the old and revised procedures across independent complete campaigns. Equalize the starting capability stock and information, and count proposal generation, failed searches, qualification, and integration. A better initialized tool library is not, by itself, a better inventor.

Table 6: A minimal factorial design at a fixed foundation model. Repeat the design at another model version to separate external progress and interactions. All cells receive matched initial permissions and full budgets.

	Earlier capability stock	**Updated capability stock**
Earlier discovery policy	Reference campaigns	Isolates inherited-stock contribution.
Revised discovery policy	Tests better discovery without a privileged inherited start.	Tests joint contribution and interaction.

Formally, let Z be a fresh campaign initialization, M the underlying model, H the starting capability stock, B the complete resource budget, and ξ a discovery policy. Define

$$J_{\text{disc}}(\xi; M, H, B) = \mathbb{E}_Z\left[W\left(\text{Campaign}(\xi, M, H, B, Z)\right)\right], \tag{10}$$

where W is a prespecified terminal outcome that counts complete campaign costs and is measured on protected evaluation work. Improving ξ means improving this quantity under the same declared conditions, not selecting one impressive descendant after many hidden attempts.

Crossing these four cells with old and new underlying models creates eight conditions when feasible. This design does not require all components to be independent. It makes their effects and interactions observable rather than attributing every gain to the preferred explanation. An economical subset can answer a narrower registered claim; it must state which distinctions it cannot identify.

12.2 Freshness and measurement are scarce resources

Freeze the candidate, comparator, task population, outcome bridge, resource rules, and evaluation procedure before confirmation. Development cases may be revisited, but must not continue to count as fresh. Several questions from one source, several descendants of one discovery campaign, and many agent messages are not independent replications.

Adaptive testing requires explicit error accounting. Repeatedly consulting a fixed holdout can corrupt the interpretation of apparent gains [6]. Time-uniform confidence sequences can support some sequential decisions under their assumptions [9]. Neither technique repairs an invalid outcome measure, contaminated source, or misclassified commercial receipt. The foundation manuscript's distinction between measurement quality and score precision remains central [4].

An external evaluator may itself need improvement. Its revision must be validated independently and both methods compared under a common supported outcome measure. A candidate cannot choose a more favorable judge and then approve itself. When calibration does not extend to a new artifact class, the correct outcome is an unresolved measurement question, not a hidden extrapolation.

12.3 Measure missed possibilities as well as accepted ones

A system can appear judicious by rejecting everything. It can appear creative by proposing endless attractive ideas. Both fail the enterprise objective. Evaluation should record accepted useful capabilities, invalid promotions, missed independently identified opportunities, complete resource use, and actual interventions required.

Opportunity-discovery evaluation needs a deliberately diverse, independently assembled reference set, including cases where no useful construction is warranted. The resulting recall is relative to that set, not the universe of all possible inventions. Historical event-time replays should restrict the information supplied to what was available then. Current-case evaluations should preserve release and source versions. Forecast calibration can be assessed only where subsequent outcomes are actually observed, with selection effects disclosed.

A compact scorecard can therefore track qualified capabilities per complete resource unit, time from verified access to accepted use, buyer benefit, repeated collected surplus, inheritance contribution, and independent campaign improvement. No single dashboard number substitutes for the underlying outcome requirements.

13 The next consequential construction

The next engineering target is one complete frontier-triggered campaign that constructs a useful capability, tests a different successor, and establishes an economical customer application where authorization permits. It should not be another general architecture layer. Three materially different candidate tracks should be compared before promoting a flagship.

Table 7: Prospective research candidates, not qualified products or established demand. These illustrate different enabling mechanisms rather than three versions of one benchmark.

Candidate	Valuable outcome hypothesis	Missing means to test	Decisive comparison
Discriminating measurement	Distinguish an internal sample change from contact effects, both, or another disturbance.	A realizable observation, calibration intervention, or altered fixture.	Contact-aware ordinary methods and equally permitted engineering; fresh physical measurements.
Mathematics to useful machinery	Make a costly scientific or engineering computation economical at required accuracy.	A verified reduction, algorithm, or numerical implementation with a valid application bridge.	Best accessible solver and direct frontier constructor; complete construction and execution costs.
Integrated engineering qualification	Make a purpose-specific experimental apparatus practical to assemble and use.	Consistency across mechanical, electrical, firmware, calibration, and manufacturer constraints.	Competent engineering with the same tools and suppliers; actual assembly and accepted function.

The first electrical track is not intrinsically the best business. It is a candidate because an information limitation can be stated and challenged. A public recording cannot validate a measurement never acquired. A simulator that omits electrode contact effects cannot establish their separation from internal changes. Combined causes, drift, neither-hypothesis cases, and justified abstention belong in the task.

The mathematical track should not assume that an important theorem supplies a practical algorithm. Formal correctness, numerical stability, implementation cost, usable performance, and buyer value are separate bridges. The engineering track must not count an attractive CAD model as a qualified instrument. These distinctions determine which experiment removes the actual bottleneck.

13.1 A connected acceptance sequence

The first accepted result should state what changed in the attainable outcome or affordability threshold. The second should test whether a retained instrument or construction method improves a genuinely different useful task. A market test should then establish voluntary acceptance and collection rather than import favorable demand assumptions from a simulator.

An existing complete route may win at any point. In that case, use it and revise the claim about the residual obstacle. Commercial success does not require proving that all earlier methods failed. A stronger Blue Sky claim does require a consequential enabling change relative to adequate complete alternatives. The ordinary native services reproduced in Section 10 are operating references, not that flagship.

The important stopping rule is not "continue until something impressive appears." Register the available resources, decisive outcomes, and how failure redirects the campaign. Preserve useful code

and compact evidence, including incurred costs. Do not accumulate an unbounded archive of untested ideas or repeatedly reset a failed experiment under new names.

14 An institution that outlives its models

The author-supplied SUCCESSOR Ω visual framework organizes the journey as Alpha, Gym, qualification, and institutional continuity. This paper preserves that structure while treating its illustrated architecture as a design intention rather than evidence of a functioning deployment [3].

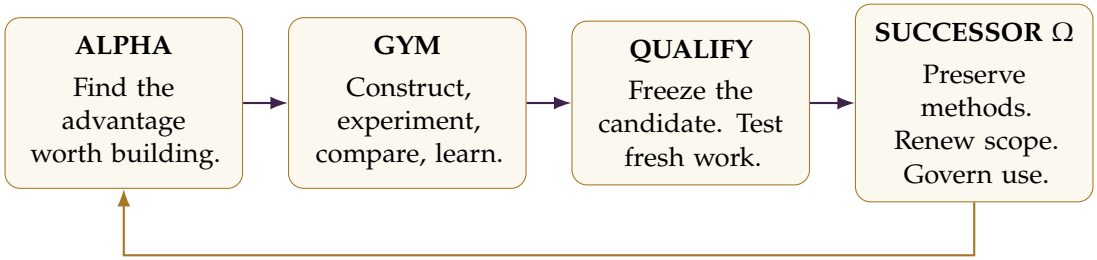


Figure 4: Institutional continuity is the continuity of purpose, lawful assets, methods, and evidence—not automatic inheritance of approval across model changes.

The Gym develops capability; it does not certify itself. Qualification grants only the scope supported by the evidence. The term "Specialist ASI" in the conceptual visual material is not an achieved claim in this release. Superior performance on a named mission would require an appropriate comparison; it would not establish general superintelligence.

The institution retains qualified operations, representations, experiments, lessons, and operating history. It may replace a model without discarding everything legitimately learned. But model or environment changes can invalidate an analysis, integration, or acceptance bridge. Inheritance therefore carries applicability conditions, not permanent truth or unrestricted authority.

Generated code, external documents, and retrieved sources are untrusted inputs to execution. They must not modify the authority that controls secrets, payments, protected data, or promotion. The same principle applies to proposals for improving the controller: they can be constructed and evaluated, but cannot self-activate to approve the proposal that generated them. Authorized revision remains possible without circular certification.

Own the institution. Replace the models. Preserve the mission.
Retain what remains useful; requalify what has changed.

15 From discovery to expanding productive capacity

The long horizon is larger than digital deliverables. Better observation can support better process models; those models can support more efficient or reliable production; production can create new instruments, energy systems, and infrastructure. Intellectual progress becomes materially consequential when it changes what can be built, maintained, and economically used.

The sequence is neither fixed nor guaranteed. Several intermediate constructions may be jointly necessary. A later invention may bypass a step assumed essential. A scientific instrument can be

valuable through its contribution to downstream work without being independently profitable. The complete enabling chain still has to bear its costs and demonstrate its claimed relationships.

For a specified class of qualified productive units, write the stock relation as $\dot{K} = b - d$, where b is commissioning rate and d is retirement or loss. If each new unit requires energy E_K, material M_K, fabrication, and qualification, then a necessary dimensional bound is

$$b \leq \min\left\{\frac{P_{\text{available}}}{E_K}, \frac{\dot{M}_{\text{available}}}{M_K}, b_{\text{fabrication}}, b_{\text{qualification}}\right\}. \tag{11}$$

All terms are expressed as units per time, with power and material flows actually available in the required forms. This is an accounting constraint under declared production assumptions, not a universal physical model. A new process can change the coefficients or the required inputs. It cannot turn predicted profit into absent materials, fabricate unqualified spare parts by assertion, or ignore heat rejection and maintenance.

The prospective progression extends from scientific instruments to materials and manufacturing, useful energy availability, robotic construction, and increasingly capable space-based productive systems. Its most ambitious horizon is stellar-scale productive reach: constructing the instruments and industrial capacity that make far greater use of available energy and matter possible. The paper supplies no deployment date, attained scale, or guaranteed wealth trajectory. These are separate construction programmes with physical uncertainty, ecological consequences, governance requirements, and competing uses of capital.

The enterprise objective should remain additional useful value through voluntary participation. Competitive diffusion may reduce private margins while increasing public benefit. Sustainable expansion should not depend on concealing risk, manufacturing customer dependence, or assuming that possession of intelligence confers entitlement to all value others create.

16 Limitations and falsifiable commitments

The principal unresolved task is open-world discovery of economically meaningful possibilities. The current finite planner receives its markets, hypotheses, actions, costs, and observation laws. The composition engine receives its operation grammar. The native enterprise receives authored commercial assumptions and uses deterministic production. None establishes a live institution that sees the whole frontier or invents arbitrary new tools autonomously.

The framework also depends on measurement quality. Real scientific novelty can be difficult to judge; customer outcomes may be delayed; discovery value can be diffuse; and strong alternatives may change during a campaign. A low uncertainty estimate from an inadequate model is not adequate evidence. New hypotheses and changed representations can invalidate earlier calibration.

The three orders can fail independently. The observatory may miss a relevant development. The consequence generator may identify no valuable use or overestimate a chain of prerequisites. The planner may optimize the wrong economics. A retained method may be harder to maintain than to reconstruct. A successful application may offer no transfer. A new model may account for all apparent progress.

These conditions do not impose a ceiling on ambition. They make ambition operational. A small decisive experiment can address a large consequence; a large architecture without a decisive construction does not.

17 Conclusion: the advantage that builds the next advantage

ALPHA GENESIS FLYWHEEL Ω proposes an enterprise whose central skill is converting frontier progress into useful possibility. Its first order recognizes a real change in available means. Its second constructs the consequences that matter. Its third tests whether experience has improved the process that discovers those consequences.

The most important economic rule is origin neutrality. Use the strongest accessible existing means when they win. Construct only when the complete prospective value justifies the commitment. Retire unsupported or uneconomic inheritance. Preserve the distinction between commercial success, a useful invention, and a better inventor.

The executed evidence is intentionally narrower than the horizon. Exact finite planning selects experiments, existing agents, or non-entry as the conditions warrant. Native operation produces checked artifacts, records simulated collections, preserves obligations, and funds a research trial without inventing another ledger. The strong conventional method ties and remains in use. These are working foundations, not the completed discovery enterprise.

The decisive next achievement is a connected, independently tested lineage: a frontier change enables a valuable construction; the construction serves an accepted outcome; its retained means improve different fresh work; and the resulting economics support worthwhile continuation. Repeating that pattern would make the institution more than a beneficiary of stronger models. It would make it a productive participant in the creation of what comes next.

See the frontier. See the consequences.
Improve the machinery that sees possibilities.

Invent the means. Expand what becomes possible.
Compound the ability to discover.

Build the advantage that discovers—
and builds—the next advantage.

A Analytical details and interpretation

A.1 The finite diagnostic calculation

In the base decision world, let G denote successful construction. The probability of a positive signal is

$$\Pr(+) = \frac{7}{20}\frac{9}{10} + \frac{13}{20}\frac{1}{10} = \frac{19}{50}, \qquad \Pr(G \mid +) = \frac{63}{76}, \quad \Pr(G \mid -) = \frac{7}{124}.$$

After a positive signal, the remaining three steps allow construction and two deliveries. Conditional expected value from that state is $-12 + (63/76)68 + (13/76)30 = 49.5$. After a negative signal, using the frontier routes gives 30. The diagnostic's total value is therefore

$$-3 + \frac{19}{50}(49.5) + \frac{31}{50}(30) = 34.41.$$

Without the diagnostic, immediate construction yields $-12 + (7/20)68 + (13/20)30 = 31.30$, which the independent dynamic programme confirms is optimal among all allowed no-probe continuations. Simple direct delivery gives 30. With perfect initial information at no acquisition cost, the good world yields $68 - 12 = 56$, the bad world 30, and the weighted value is 39.10.

The signal's gross decision benefit relative to the no-probe planner is 6.11 before paying its cost of 3; the net benefit is 3.11. Perfect information improves by 7.80. These inequalities are coherent with the information comparison: a noisy signal does not exceed free perfect knowledge under these assumptions. The signal also consumes time, already included in the four-step decision model.

The computation gives ordinary exact Bayesian planning the same actions and information. It therefore establishes a correct model-contingent decision, not superiority of a new planning algorithm. A misspecified prior, imperfect construction test, additional hidden market, changed horizon, or strategic rival response defines a different problem.

A.2 Scope of the sensitivity comparison

The 54 checks cross prior viability $\{1/10, 7/20, 4/5\}$, construction cost $\{4, 12, 16\}$, carrying cost $\{0, 10, 25\}$, and signal sensitivity $\{1/2, 9/10\}$. The false-positive rate remains $1/10$; other base settings remain fixed. Both implementations permit only financially feasible experiments and construction. The tested agreement concerns optimal expected value; tie-breaking among equally valuable actions need not be a scientific distinction.

The oracle is separately coded but not an independent organization or a separately validated model of reality. It reduces the risk of a shared implementation mistake in the exercised transitions; it does not validate authored assumptions. All cases, including ones selecting no construction, remain in the output.

A.3 Prospective retirement versus lifecycle attribution

For remaining use count R, per-use saving s, integration K, and carrying cost M, the simple prospective comparison is $Rs - K - M$. An already incurred construction cost is sunk for this decision. For the separate claim that acquiring the asset was worthwhile, add the original construction and search costs back into the lifecycle account. The different comparisons answer different questions; neither authorizes hiding a failed investment.

A jointly useful bundle requires a comparison against the best alternative bundle, not a sum of independent asset scores. Likewise, one successor benefit must not be credited in full to each of several

necessary ancestors and then summed as if it occurred several times. Lineage attribution can remain nonadditive while the cash account stays exact.

A.4 Conditional adoption advantage

Equation (2) compares two institutions under a common assumed exponential capability path. It can be rewritten as $\log(Y_A/Y_B) = \log(\eta_A/\eta_B) + g(\tau_B - \tau_A)$. A lower conversion efficiency can outweigh a lag advantage. If a prerequisite bottleneck caps realized output, the ratio need not describe actual delivery. If a rival receives the same improvement, economic advantage additionally depends on value capture and outside options. No exponential revenue or intelligence law follows from this identity.

B Reproduction, source identity, and evidence

B.1 Run the self-contained analytical companion

The paper package contains the finite decision model, the independently implemented oracle, unit tests, and retained exact results. Python's standard library is sufficient; no network, API key, model access, payment, or GPU is required.

```
cd analysis
python3 -m unittest -v
python3 exact_studies.py --out ../my_exact_studies.json
```

The output path must be new. Compare its content with `evidence/exact_studies.json`. The independent oracle and the original planner are both available for inspection. The scenario inputs are in the code's explicit configuration; they are not inferred from current markets.

B.2 Native reproduction requires its exact reference

The native reproduction is separate from the paper's small analytical module. The source package includes the corrected Native Enterprise 002 additions, the historical public native source archive, and a reproduction driver. Its provenance record binds their hashes and the executing source identities. The native account's canonical funding and recovery machinery must not be replaced with a new balance simulation merely to obtain the reported result.

The bundled driver expands these sources into a new working directory, runs the native suite and the four account reproductions, and keeps local recovery authority out of the publication folder:

```
python3 reproduction/reproduce_native.py --work /absolute/new/reproduction
```

It installs nothing. Its README states the pinned reference dependencies and POSIX requirement. From that exact reference plus the additive changes, the underlying commands are:

```
python3 -m unittest discover -s tests \
  -p 'test_genesis002_*.py' -v
python3 -m genesis002 run \
  --config campaign/genesis-enterprise-002/INPUT.json \
  --output /absolute/new/genesis-run
python3 -m genesis002 audit --output /absolute/new/genesis-run
python3 -m genesis002 resume --output /absolute/new/genesis-run
```

The three stress cases are separately initialized accounts with the declared configuration and stress option. Do not combine their funds or interpret an account reproduction as a continuation of a historical operation. Operating directories contain local simulation authority and recovery state; they are not included in this publication package.

The safe evidence files report source-bound checks, not a transferable grant of account authority. The native results refer to commit `4a87b455373be72811ad3a0c654a116c3604a374`. The programme README was separately inspected at later commit `4f197eed10ddafc34f5a35968d93a1487815d0e1`. Reading the latter is not testing compatibility with it.

B.3 Build and inspect the paper

```
pdflatex -interaction=nonstopmode -halt-on-error \
  ALPHA_GENESIS_FLYWHEEL_OMEGA_v2.tex
bibtex ALPHA_GENESIS_FLYWHEEL_OMEGA_v2
pdflatex -interaction=nonstopmode -halt-on-error \
  ALPHA_GENESIS_FLYWHEEL_OMEGA_v2.tex
pdflatex -interaction=nonstopmode -halt-on-error \
  ALPHA_GENESIS_FLYWHEEL_OMEGA_v2.tex
```

The supplied `build.py` automates the same build and handles the available BibTeX executable. The manuscript uses standard TeX Live packages and vector figures. The release includes a checksum manifest, build and layout checks, test transcripts, source provenance, and a claim register. File identity checks establish bytes, not scientific truth, correct consent, or data independence.

C Source lineage and responsible interpretation

This edition develops the three-order framing and origin-neutral enterprise rule from the author's research discussions, the Transformational Intelligence manuscript, the prior GENESIS paper, and the supplied decision and native implementation packages. Those source discussions are intellectual provenance; their earlier narratives are not substitutes for inspected artifacts or new execution. The institutional figure adapts the Alpha–Gym–qualification–SUCCESSOR structure in the author-supplied visual material.

The new contribution is an integrated, falsifiable enterprise synthesis, a sharper discovery-versus-selection contract, the stronger finite information comparison, and source-matched verification of the supplied operational component. The elementary propositions use established decision and accounting reasoning. No claim is made to originate Bayesian planning, reinforcement learning, experimental design, program synthesis, or the broad idea of self-improvement.

External research results are credited to their authors. Availability statements are dated observations and may change. The paper does not claim external peer review, third-party certification, a tested world model of the economy, or any newly achieved general AGI/ASI threshold. AI assistance was used in research synthesis, drafting, implementation checking, and publication preparation; independent scientific and commercial claims still require their stated evidence.

The delivered paper and analytical companion do not modify a repository, operate customer equipment, contact buyers, activate billing, or initiate transactions. Publication readiness and enterprise qualification are distinct. A future claim of an operating GENESIS institution must release its own exact implementation, permitted interfaces, full campaign evidence, and accepted outcomes.

D Claim–evidence register

Table 8: The strongest claim supported by each evidence class. Detailed machine-readable records accompany the release.

ID	Claim	Evidence	Boundary
C1	Three-order enterprise architecture is specified.	Sections 3–14; explicit operating and evaluation contracts.	Proposal; not deployed general discovery.
C2	The diagnostic is preferred in the base finite decision world.	Exact result 34.41; independent no-probe value 31.30; separately coded recursion.	Authored probabilities, actions, payoffs, and horizon.
C3	Existing tools and non-entry can correctly win.	Frontier, carrying-cost, and no-market counterexamples; reproducible tests.	Does not estimate actual market-optimal choices.
C4	Planner implementations agree over the disclosed sensitivity grid.	54 cases; exact rational values; 15 analytical tests.	Agreement within the supplied family, not model calibration.
C5	Native simulated operation is reproducible on its recorded reference.	Four six-cycle runs; local audit and unchanged resume; 62 native tests.	Historical source; no current-main or hosted release qualification.
C6	The central native case records US$143.20 modeled profit.	US$240 receipts less US$54.76 operating cost, US$37.04 provision, US$5 research.	Synthetic demand and cost envelopes; actual lifecycle costs unmeasured.
C7	Some continuation uses native unrestricted simulated earnings.	Five repeat productions and one research purchase are funded EARNED.	Seed still backs protection reserves; not seed-independent operation.
C8	The tested router does not demonstrate superiority over the strong conventional method.	Both accept 24 cases with identical mapping-work count 1,216; conventional method retained.	Mapping work is not total cost; no better-inventor claim.
C9	Relevant external frontier advances are documented.	Dated primary sources and source/access distinctions in Table 1.	Provider reports and papers do not validate GENESIS.
C10	Fast qualified conversion can matter under the exponential-lag model.	Equation (2) and Appendix A.	Conditional algebra, not a fitted growth or profit forecast.
C11	Open-world opportunity discovery, real commercial self-sustainability, and recurring inventor improvement are targets.	Registered tests and falsifiers specified; not yet satisfied by this release.	Explicitly unestablished.

References

[1] Tom Beyer. We're Rolling Out AlphaEvolve Widely to Solve Google Cloud Customers' Hardest Problems, 2026. URL https://blog.google/innovation-and-ai/infrastructure-and-cloud/google-cloud/alphaevolve-on-cloud/. Google Cloud, 9 July. Accessed 4 November 2026.

[2] Vincent Boucher. ALPHA GENESIS FLYWHEEL Ω: Riding the Exponential, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Invention-Flywheel-Lab/tree/4f197eed10ddafc34f5a35968d93a1487815d0e1/docs/alpha-genesis. Research release 1.1.0, 4 November. Prior research synthesis and analytical companion.

[3] Vincent Boucher. Transformational Intelligence and the Recursive Invention Flywheel: Research and Design Discussions, 2026. Author-supplied October–November research records. Used as design provenance, not independent experimental verification.

[4] Vincent Boucher. Transformational Intelligence: Learning to Change the Game, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Invention-Flywheel-Lab/blob/main/docs/research/transformational-intelligence.html. October research manuscript; author-supplied edition, 24 pages.

[5] Lizzie Dorfman and Michael Brenner. Empirical Research Assistance (ERA): From Nature Publication to Catalyzing Computational Discovery, 2026. URL https://research.google/blog/empirical-research-assistance-era-from-nature-publication-to-catalyzing-computational-discovery/. Google Research, 19 May; MCP update 3 November. Accessed 4 November 2026.

[6] Cynthia Dwork, Vitaly Feldman, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Aaron Roth. Preserving Statistical Validity in Adaptive Data Analysis, 2014. URL https://arxiv.org/abs/1411.2664. arXiv:1411.2664.

[7] Kevin Ellis et al. DreamCoder: Growing Generalizable, Interpretable Knowledge with Wake-Sleep Bayesian Program Learning, 2020. URL https://arxiv.org/abs/2006.08381. arXiv:2006.08381.

[8] Danijar Hafner, Jurgis Pasukonis, Jimmy Ba, and Timothy Lillicrap. Mastering Diverse Control Tasks through World Models. *Nature*, 2025. doi: 10.1038/s41586-025-08744-2. URL https://arxiv.org/abs/2301.04104.

[9] Steven R. Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-Uniform, Nonparametric, Nonasymptotic Confidence Sequences. *The Annals of Statistics*, 49(2):1055–1080, 2021. doi: 10.1214/20-AOS1991. URL https://arxiv.org/abs/1810.08240.

[10] Shengran Hu, Cong Lu, and Jeff Clune. Automated Design of Agentic Systems, 2024. URL https://arxiv.org/abs/2408.08435. arXiv:2408.08435.

[11] Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Mohammadamin Barekatain, Simon Schmitt, and David Silver. Learning and Planning in Complex Action Spaces. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 4476–4486, 2021. URL https://proceedings.mlr.press/v139/hubert21a.html.

[12] MONTREAL.AI. GENESIS Frontier-to-Possibility 001, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Invention-Flywheel-Lab/pull/229. Finite composition planner and public-release observer. Author-supplied source snapshot; re-executed for this edition.

[13] MONTREAL.AI. GENESIS Frontier Decision Model, Research Prototype 0.1, 2026. Author-supplied standard-library finite Bayesian planner. Four authored scenarios; re-executed for this edition.

[14] MONTREAL.AI. AGI ALPHA Invention Flywheel Lab: Programme and Evidence Boundaries, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Invention-Flywheel-Lab/blob/4f197eed10ddafc34f5a35968d93a1487815d0e1/README.md. Repository README inspected at commit 4f197eed10ddafc34f5a35968d93a1487815d0e1, 4 November.

[15] MONTREAL.AI. GENESIS Native Enterprise 002: Additive Implementation and Source-Bound Simulation, 2026. Author-supplied corrected archive, 4 November. Historical native reference 4a87b455373be72811ad3a0c654a116c3604a374; re-executed for this edition. Public-safe results in the companion evidence directory.

[16] Alexander Novikov et al. AlphaEvolve: A Coding Agent for Scientific and Algorithmic Discovery, 2025. URL https://arxiv.org/abs/2506.13131. arXiv:2506.13131.

[17] OpenAI. Codex App Server Documentation, 2026. URL https://developers.openai.com/codex/app-server/. Living documentation, inspected 4 November; interfaces and entitlements must be checked at use.

[18] OpenAI. Codex Authentication, 2026. URL https://developers.openai.com/codex/auth/. Living documentation, inspected 4 November.

[19] OpenAI. Accelerating Engineering for the Physical World, 2026. URL https://openai.com/index/accelerating-engineering-for-the-physical-world/. 3 November; provider-reported demonstration. Accessed 4 November 2026.

[20] OpenAI. Mathematical Manuscripts and Supporting Proof Artifacts, 2026. URL https://github.com/openai/math. README inspected 4 November; blob cc5b0a03bca6c742f045d683b6647a6f44ef2cae. Results at different verification stages; generating model internal.

[21] Stéphane Ross, Brahim Chaib-draa, and Joelle Pineau. Bayes-Adaptive POMDPs. In *Advances in Neural Information Processing Systems*, volume 20, 2007. URL https://papers.nips.cc/paper_files/paper/2007/hash/3b3dbaf68507998acd6a5a5254ab2d76-Abstract.html.

[22] Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, Timothy Lillicrap, and David Silver. Mastering Atari, Go, Chess and Shogi by Planning with a Learned Model. *Nature*, 588:604–609, 2020. doi: 10.1038/s41586-020-03051-4. URL https://arxiv.org/abs/1911.08265.

[23] Zoi Tsangalidou, Ruijie Zhang, Sukhdeep Singh, and Michael V. LeVine. Agentic Cheminformatics with Google Antigravity, 2026. URL https://deepmind.google/blog/agentic-cheminformatics-with-google-antigravity/. Google DeepMind, 3 November. Accessed 4 November 2026.

[24] Rui Wang, Joel Lehman, Jeff Clune, and Kenneth O. Stanley. Paired Open-Ended Trailblazer (POET): Endlessly Generating Increasingly Complex and Diverse Learning Environments and Their Solutions, 2019. URL https://arxiv.org/abs/1901.01753. arXiv:1901.01753.

[25] Jenny Zhang, Shengran Hu, Cong Lu, Robert Lange, and Jeff Clune. Darwin Gödel Machine: Open-Ended Evolution of Self-Improving Agents, 2025. URL https://arxiv.org/abs/2505.22954. arXiv:2505.22954.

[26] Jenny Zhang, Bingchen Zhao, Wannan Yang, Jakob Foerster, Jeff Clune, Minqi Jiang, Sam Devlin, and Tatiana Shavrina. Hyperagents, 2026. URL https://arxiv.org/abs/2603.19461v2. arXiv:2603.19461v2, revised 1 November.