MONTREAL.AI · QUEBEC.AI

# GOALOS SINGULARITY NAVIGATOR Ω + SEIZE

# GYM, SPECIALIST ASI & MISSION-SOVEREIGN SUCCESSOR Ω

The Verified Succession Institution for Manufacturing, Proving, Governing and Recursively Renewing Mission Alpha in Singularity Time

**NAVIGATOR → SEIZE → MISSION GYM → SPECIALIST ASI → SUCCESSOR Ω**

*Beta is rented. The Spread is underwritten. Alpha is manufactured. Specialist ASI is proven. Authority never exceeds fresh proof.*


**Vincent Boucher**

President, MONTREAL.AI & QUEBEC.AI

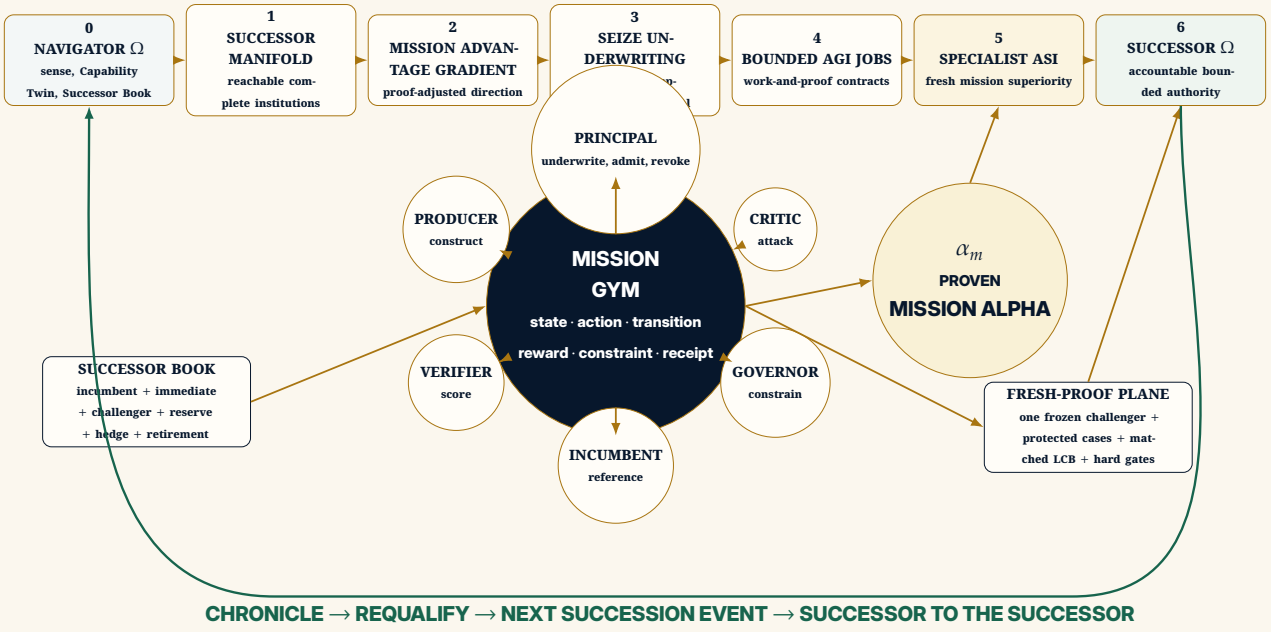
https ://montreal.ai/ · president@montreal.ai

**VERSION 6.0.0 · Unified Constitutional, Gym and Evidence Edition · 10 August 2026**

Montreal, Quebec, Canada

# THE UNIFIED SUCCESSION CONSTITUTION

## From frontier sensing to the successor to the successor




### Mission Alpha

$$\alpha_m = U_m(S_m) - \max\{U_m(I_m), U_m(B_m)\} - C_m - R_m - D_m - \widehat{\epsilon}_{\text{gym}}.$$

The Spread is an opportunity. Alpha exists only after the complete denominator, protected comparison and Gym-to-reality basis-risk reserve.

$$A_t \preceq \phi(E_t).$$

Capability may justify candidacy. Fresh proof may establish mission dominance. Only accountable admission may grant authority.

### The complete doctrine

Navigator Ω senses the changing frontier and maintains the Successor Book. SEIZE underwrites one reversible trajectory. The Mission Gym makes the mission executable. The Foundry constitutes complete challengers. Fresh proof may establish Specialist ASI. Accountable authority may admit Successor Ω. Chronicle preserves what survives and prepares the successor to the successor.

**Governing declaration**

In calendar time, institutions optimize plans. In Singularity Time, institutions govern succession. GoalOS Singularity Navigator Ω detects architecture depreciation and maintains a live market of successor options; SEIZE purchases decisive evidence; the Mission Gym turns the constituted mission into an executable and adversarial environment; Specialist ASI is an earned comparative state; Mission-Sovereign Successor Ω is the complete admitted institution; and the deepest recursion is the verified succession of the successor itself.

## Abstract

This paper unifies the outer institutional architecture of GoalOS Singularity Navigator Ω + **SEIZE** with the inner executable architecture of the Mission Gym, Specialist ASI and Mission-Sovereign Successor Ω. The result is one closed-loop Verified Succession Institution : a recurring system that senses capability, cost, legal, security and market discontinuities; marks incumbent assumptions for review; maintains a live Successor Book; maps the **Successor Manifold**; follows the **Mission Advantage Gradient**; underwrites one reversible trajectory through SEIZE; compiles that trajectory into Bounded AGI Jobs; constitutes complete successor institutions inside a sovereign Mission Gym; freezes one challenger; purchases protected fresh proof; and grants only the authority collateralized by current evidence.

The paper distinguishes the control plane from the mission plane. Navigator Ω does not perform the mission and the Gym does not choose corporate strategy. Navigator Ω maintains the frontier state, incumbent ledger, option portfolio and review cadence. SEIZE converts a succession event into a proof-capital decision. The Mission Gym is a controlled, versioned and rights-cleared representation of a consequential mission : hidden state, role-specific observations, bounded actions, transition dynamics, vector rewards, non-compensable constraints, scenario distributions, verifier and proof receipts. It is not merely a reinforcement-learning benchmark. It is the mission's sovereign experimental environment and substitution surface.

The paper also distinguishes the policy acting inside the Gym from the complete Successor. A policy may select actions, but a successor institution includes models, world models, data, retrieval, deterministic tools, planners, specialist roles, critics, independent verifiers, accountable human principals, workflow integration, evidence custody, operational controls, degraded modes, rollback, Chronicle and an Authority Envelope. A Specialist ASI is the complete mission-specific intelligence institution that outperforms the incumbent and the best credible general, specialist, human, software or hybrid alternative on fresh, protected and representative work after quality, critical error, complete economics, human burden, rights, sovereignty, resilience, governance and Gym basis risk are counted. A Mission-Sovereign Successor Ω is that proven system only after separate accountable admission and the grant of scoped, versioned, monitored, expiring and revocable authority.

The integrated formalism defines a Navigator–Gym succession complex, a constrained multiplayer stochastic game, proof-adjusted Mission Alpha, a partial order on authority, the recursive successor operator and the condition under which Chronicle-admitted capability improves the next succession cycle without weakening proof or governance. It combines real-options underwriting, mechanism design, information geometry, nonlinear dynamics, portfolio logic, formal assurance, matched protected evaluation and quality-diversity search. It adds GSSB-4, a deterministic four-mission benchmark containing 6,000 synthetic cases and 30,000 candidate episodes across Invoice Integrity, Building Operations, Acquisition Underwriting and Verified Software Engineering, together with the Aurelia end-to-end synthetic succession cycle. The evidence is reproducible and synthetic; it establishes executability of the mechanism, not customer reality or production authority.

The canonical progression is :

**Navigator Ω → Successor Manifold → Mission Advantage Gradient → SEIZE Underwriting → Bounded AGI Jobs → Specialist ASI → Mission-Sovereign Successor Ω.**

The paper provides the Successor Book, six-plane technical architecture, twenty-one-job formation portfolio, protected Fresh-Successor Gate, Authority Lattice, Chronicle and Negative Capability Graph, machine-readable records, board metrics, corporate portfolio doctrine, Mission-Sovereign Successor-as-a-Service model, implementation path and empirical completion gate. Its governing rule is exact : authority never exceeds fresh proof.

**Keywords :** GoalOS; Singularity Navigator Omega; SEIZE; Verified Succession Institution; Mission Gym; Specialist

ASI ; Mission-Sovereign Successor Omega ; Successor Manifold ; Mission Advantage Gradient ; Successor Book ; Alpha ; Beta ; Spread ; GSSB-4 ; proof capital ; Chronicle ; recursive succession ; Authority Envelope ; Singularity Time.

## Résumé exécutif

Ce document unifie GoalOS Singularity Navigator Ω + SEIZE avec le Mission Gym, l'ASI spécialisée et le successeur souverain pour la mission Ω. Le Navigator détecte les ruptures de capacité, de coût, de droit, de sécurité et de marché ; il réévalue l'architecture en place et maintient un portefeuille vivant d'options de succession. SEIZE transforme l'événement en décision de capital de preuve. Le Gym rend la mission exécutable, rejouable, mesurable et attaquable. Le Foundry constitue plusieurs institutions complètes ; une seule candidate est figée et soumise à une preuve fraîche et protégée.

L'ASI spécialisée est une désignation à mériter. Elle désigne l'institution d'intelligence complète qui dépasse l'architecture en place et la meilleure solution crédible sur du travail frais, protégé et représentatif, après prise en compte des coûts complets, des erreurs critiques, de la charge humaine, des droits, de la souveraineté, de la résilience, de la gouvernance et du risque de transfert entre le Gym et la réalité. Le successeur souverain pour la mission Ω est cette intelligence prouvée, mais seulement après admission séparée par une autorité responsable, avec une enveloppe d'autorité circonscrite, versionnée, temporaire, surveillée et révocable.

La récursion centrale porte sur le successeur du successeur. Le système actuel peut proposer des descendants ; le Foundry les diversifie ; le Gym les expose ; la preuve fraîche sélectionne ; l'autorité responsable admet ; la réalité requalifie. Aucun descendant ne peut consulter les cas protégés, se certifier, s'installer ou hériter automatiquement de l'autorité. Les résultats Aurelia et GSSB-4 sont synthétiques et reproductibles ; ils démontrent le mécanisme, non la vérité d'une mission client.

## Principal contributions

| Contribution | Institutional addition |
|---|---|
| Navigator–Gym closure | Unifies frontier sensing, incumbent impairment, Successor Book, SEIZE, sovereign Gym, Foundry, protected proof, authority, Chronicle and next-succession triggers in one closed loop. |
| Verified Succession Institution | Defines the permanent constitutional system that repeatedly underwrites, constitutes, proves, admits, monitors and replaces mission-bearing architectures. |
| Mission Gym formalism | Defines a sovereign, versioned, constrained partially observable stochastic game for formation, adversarial testing, fresh proof, authority rehearsal and requalification. |
| Policy–Successor separation | Shows why the complete successor is larger than the acting policy : it includes tools, world models, verifiers, people, evidence, operations, rollback, Chronicle and authority. |
| Successor Book and portfolio logic | Maintains incumbent, immediate, challenger, reserve, hedge and retirement positions rather than one irreversible recommendation. |
| Mission physics | Develops the Successor Manifold, institutional metric, Mission Advantage Gradient, least-action SEIZE trajectory, metastability, phase transitions and stochastic Alpha dynamics. |
| Multiplayer succession game | Models principal, incumbent, producer, critic, verifier, governor, provider, affected party and adversary as strategic actors with different information and incentives. |
| Specialist ASI proof state | Separates narrow benchmark strength from complete mission dominance across capability, economics, reliability, sovereignty, governance and transfer. |
| Successor Ω | Defines the complete admitted institution and its Authority Lattice, proof-carrying actions, degraded modes, rollback and requalification. |
| Recursive successor operator | Formalizes the successor to the successor while prohibiting self-certification, protected-case access and automatic proof or authority inheritance. |
| Synthetic evidence programme | Integrates the Aurelia end-to-end cycle and GSSB-4 : 6,000 deterministic cases and 30,000 candidate episodes across four mission classes. |

| Contribution | Institutional addition |
|---|---|
| Corporate institution | Connects board metrics, mission portfolios, Successor-as-a-Service, the GoalOS Singularity Office Ω and a proof-gated implementation path. |
| Evidence-bearing publication | Distributes code, data, figures, environment cards, schemas, source, manifests and clean-room verification as one versioned release. |

## Recommended citation

Boucher, Vincent (2026). *GoalOS Singularity Navigator Ω + SEIZE : Gym, Specialist ASI & Mission-Sovereign Successor Ω—The Verified Succession Institution for Manufacturing, Proving, Governing and Recursively Renewing Mission Alpha in Singularity Time*. Version 6.0.0, Unified Constitutional, Gym and Evidence Edition, MONTREAL.AI Research, Montreal, Quebec, Canada. https://montreal.ai/.

## Claim boundary and source of truth

This paper is research and institutional design. It does not establish that universal ASI exists, that any deployed external system has earned Specialist ASI status, that any customer has realized Mission Alpha, or that any production authority described here has been granted. Aurelia and GSSB-4 are synthetic, deterministic and unaudited. Physics terminology is used as a disciplined design formalism for constrained architecture spaces, information geometry, stochastic dynamics and phase transitions; it is not a claim that corporations literally obey fundamental physical laws.

The authoritative identity is Vincent Boucher, President of MONTREAL.AI & QUEBEC.AI. The official website is https://montreal.ai/ and the official email is president@montreal.ai. Exact text, equations, tables, thresholds, authority states and claim boundaries are governed by the versioned LaTeX source, compiled PDFs, release manifest and cryptographic checksums. Image-generated ornament and supplementary visual works are non-authoritative.

# CONTENTS

PART I

# Foundations : Mission, Environment and Succession

*From static plans to verified succession under capability-time pressure*

The institutional operating condition, Navigator Ω, and the mission triad of Gym, Specialist ASI and Successor Ω.

**SUCCESSOR MANIFOLD → GRADIENT → SEIZE → BOUNDED JOBS → SPECIALIST ASI → SUCCESSOR Ω**

# 1 The Operating Condition : Capability Time and Institutional Succession

## 1.1 From stable planning to capability time

A conventional institution assumes that its operating architecture remains valid long enough for annual planning, procurement, implementation and review. That assumption becomes fragile when models, tools, costs, providers, regulations, interfaces and customer expectations change on a timescale shorter than the institution's adaptation cycle. The relevant operating regime is not defined by one metaphysical date. It is defined by a relation between the half-life of accepted assumptions and the lag required to detect, prove and authorize a replacement.

Let $H_c$ denote the expected half-life of a material capability or architecture assumption, and let $L_a$ denote the institution's adaptation lag. Define the **Succession Exposure Ratio**

$$\mathrm{SER} = \frac{L_a}{H_c}. \tag{1}$$

When $\mathrm{SER} < 1$, the institution generally adapts before its assumptions become materially stale. When $\mathrm{SER} > 1$, the institution increasingly acts from depreciated architectures. The objective is not speed alone. An institution can shorten $L_a$ by eliminating proof, rights review, human authority or rollback and thereby accelerate toward a larger failure. The relevant objective is **Verified Adaptation Velocity** :

$$\mathrm{VAV} = \frac{\Delta K_v}{\Delta t} \cdot \frac{1}{1 + \lambda_R R + \lambda_H H + \lambda_D D_P}, \tag{2}$$

where $K_v$ is verified productive capability, $R$ residual risk, $H$ human and reviewer burden, and $D_P$ Proof Debt.

Recent work on possible transitions from AGI to ASI describes scaling, paradigm shifts, recursive improvement and large multi-agent collectives as distinct but potentially interacting pathways, while emphasizing substantial uncertainty and bottlenecks rather than a single guaranteed discontinuity (Genewein et al., 2026). GoalOS therefore treats capability time as an environment to govern rather than a calendar prophecy.

**Capability-time law**

A capability event does not authorize replacement. It creates a duty to re-underwrite the incumbent, reconstruct the reachable successor set, purchase decisive evidence and preserve the ability to refuse.

## 1.2 Every incumbent is a depreciating institutional architecture

An incumbent is more than software. It includes the objective, division of labour, data rights, knowledge representation, tools, models, providers, human approvals, verifier, workflow, contracts, infrastructure, evidence, memory, incident routes and authority. Its depreciation may be technical, economic, strategic, legal or constitutional.

Let $S_0$ be the incumbent. The expected present burden of waiting is

$$C_{\mathrm{wait}}(S_0) = V_{\mathrm{foregone}} + C_{\mathrm{legacy}} + R_{\mathrm{competitive}} + R_{\mathrm{dependency}} + I_{\mathrm{option\ loss}}, \tag{3}$$

while the expected burden of moving too early to candidate $S_i$ is

$$C_{\mathrm{premature}}(S_i) = C_{\mathrm{formation}} + C_{\mathrm{integration}} + R_{\mathrm{error}} + R_{\mathrm{security}} + C_{\mathrm{rollback}} + R_{\mathrm{lock\text{-}in}} + D_P. \tag{4}$$

Verified succession is a comparison between those two burdens, not a choice between enthusiasm and caution.

## 1.3 Why the Gym becomes necessary

Without an executable mission environment, an institution can compare successors only through rhetoric, isolated demonstrations or historical anecdotes. It cannot reliably answer :

— whether candidates faced the same cases ;
— whether hidden human rescue produced the result ;
— whether the score represented the real mission ;

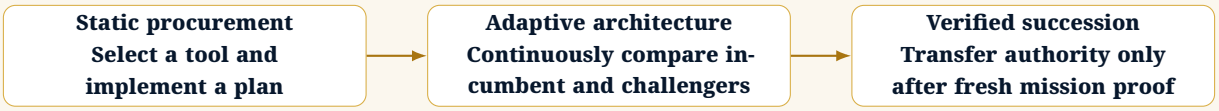
— whether rare failures were sampled;
— whether authority and tools were bounded;
— whether the candidate can be replayed after a component changes;
— whether the successor remains superior to a new external alternative.

The Mission Gym converts the Mission Constitution into executable state, observation, action, transition, constraint and verification logic. It makes failure replayable and comparative proof possible. It also becomes a corporate option on every future model or architecture : new components can be substituted and tested against an institution-owned mission rather than accepted on provider reputation.

Design Proposition 1.1 (Gym necessity). Where a mission is valuable, repeated, bounded, measurable and capable of replay or simulation, a sovereign Mission Gym reduces the cost of comparing successor architectures and increases the credibility of authority decisions. The proposition fails when the Gym has greater basis risk than the uncertainty it is meant to resolve.

### 1.4 The unit of improvement is the mission-bearing institution

A model upgrade can be useful without changing the institution. Conversely, a superior successor may arise from a smaller model, deterministic tool, better verifier, redesigned human role, new contract, different provider, improved data or narrower authority. The unit of succession is therefore the complete mission-bearing institution.




### 1.5 Claim discipline

A forecast is not a capability result. A Gym score is not a mission result. A mission result is not mission dominance. Mission dominance is not authority. Access is not authority. Historical proof is not fresh proof. A policy is not the complete Successor. These distinctions are the logical foundation of the paper.

**No authority by narrative**

No forecast, brand, model scale, benchmark reputation, generated explanation, simulation score or strategic urgency can substitute for the exact evidence required by the Mission Constitution. Where evidence is absent, the correct institutional state is uncertainty, not inherited authority.

## 2 GoalOS Singularity Navigator Omega + SEIZE : The Verified Succession Institution

### 2.1 The unified institutional object

The merged architecture is not a collection of adjacent tools. It is one recurring constitutional institution whose purpose is to keep consequential missions superior while their capability environment changes. GoalOS Singularity Navigator Ω senses frontier, cost, legal, security and market discontinuities; the Successor Book maintains the incumbent and the field of possible replacements; **SEIZE**underwrites one reversible trajectory; the Mission Gym makes the mission executable and comparable; the Successor Foundry constitutes materially different complete candidates; the Independent Proof Plane tests one frozen challenger; accountable authority admits or rejects; and Chronicle preserves only the exact capability that survives reality.

Definition 2.1 (Verified Succession Institution). A Verified Succession Institution is a recurring constitutional system that detects when an incumbent architecture has become materially suboptimal, maintains a portfolio of successor options, authorizes bounded proof capital, constitutes complete candidate institutions, compares one frozen challenger on fresh protected work, grants authority only through a separate accountable decision, and continuously re-underwrites the admitted successor.

The institution therefore operates across two coupled loops :

$$\underbrace{\text{Sense} \to \text{Book} \to \text{Underwrite}}_{\text{Navigator } \Omega \text{ macrocycle}} \quad \Longrightarrow \quad \underbrace{\text{Gym} \to \text{Form} \to \text{Prove} \to \text{Admit} \to \text{Requalify}}_{\text{mission succession cycle}}. \tag{5}$$

The macrocycle decides *which mission and which uncertainty deserve institutional attention*. The mission cycle decides *which exact successor, if any, has earned bounded authority*. Conflating these loops creates predictable errors : strategic excitement becomes a development mandate, development becomes deployment, and one attractive output becomes institutional memory.

## 2.2 Six permanent functions

| Function | Controlling question | Canonical output |
|---|---|---|
| Sensing | What changed, and which incumbent assumption is impaired ? | Frontier Event Record, Capability Twin update, source and expiry record. |
| Underwriting | Which successor trajectory is valuable, feasible, governable and worth proof capital ? | Successor Book position, SEIZE memo, proof-capital authorization and stop-loss. |
| Formation | Which complete architectures should be constituted rather than merely imagined ? | Successor Manifold, Bounded AGI Job portfolio, exact candidate manifests and reserve lineages. |
| Proof | Does one frozen challenger outperform the incumbent and the strongest credible alternative on fresh work ? | Protected scorecard, proof receipts, lower-confidence result, critical-error record and independent verdict. |
| Authority | What may this exact release observe, recommend, prepare or execute, for how long and with what rollback ? | Admission Record, Authority Envelope, monitoring, expiry, revocation and known-good fallback. |
| Compounding | What accepted capability, failure and evidence may shape the next succession cycle ? | Chronicle admission, Negative Capability Graph, Capital-to-Capacity memo and requalification trigger. |

These functions are permanent even when the underlying model portfolio changes. A provider may supply intelligence, a consultant may supply expertise, a platform may supply tools and an employee may supply judgment, but none of those roles automatically becomes the institution that decides what counts as truth, authority or reusable capability.

## 2.3 Navigator Ω as the sensing and portfolio layer

Navigator Ω maintains a live institutional state

$$Z_t = \left(F_t, I_t, \mathscr{B}_t, \mathscr{O}_t, E_t, A_t, K_t\right), \tag{6}$$

where $F_t$ is the external frontier, $I_t$ the accepted incumbent and its assumptions, $\mathscr{B}_t$ the best credible alternative set, $\mathscr{O}_t$ the Successor Book, $E_t$ current evidence, $A_t$ authority and $K_t$ Chronicle-admitted capability. A material event $\xi_t$ updates the state through

$$Z_{t+1} = \mathfrak{N}(Z_t, \xi_t), \tag{7}$$

but the update is not itself a promotion. Navigator Ω can mark an assumption as impaired, open a SEIZE mandate, lower a position, preserve an option or trigger requalification. It cannot certify the candidate or enlarge the Authority Envelope.

A useful sensing layer includes :

— a **Capability Twin** : the institution's current map of models, data, tools, workflows, people, rights, costs, providers, controls and bottlenecks ;

— an **Incumbent and Assumption Ledger** : the claims that justify the present architecture, together with owners, evidence, expiry and impairment triggers ;

— a **Frontier Radar** : current capability, cost, law, security, infrastructure and market events with source quality and decision relevance ;
— the **Successor Book** : incumbent, immediate, challenger, reserve, hedge and retirement positions ;
— an **Opportunity and Risk Brief** : what changed, what matters, what evidence would change the position and what capital or authority is presently exposed.

### 2.4 SEIZE as the constitutional underwriting microcycle

SEIZE converts one succession event into a bounded proof decision :

**S — Surface.**
Identify the capability, cost, legal, security, market or infrastructure event ; the assumption that expired ; and the mission newly threatened or made feasible.

**E — Evaluate.**
Generate materially different retain, repair, rent, configure, form, partner, acquire, hedge, reserve, retire and prohibit alternatives.

**I — Instantiate.**
Freeze mission, beneficiary, incumbent, best alternatives, hard gates, rights, proof budget, stop-loss, verifier, protected distribution and authority ceiling.

**Z — Zero in.**
Purchase the smallest evidence object with positive expected decision value.

**E — Elevate.**
Freeze one exact complete challenger and submit it to independent fresh proof.

Let $a \in \mathscr{A}_{\mathrm{succ}}$ denote a succession action and $e$ a candidate evidence package. The value of proof is

$$\mathrm{VoP}(e) = \mathbb{E}_e\left[\max_a \mathbb{E}[U(a) \mid e]\right] - \max_a \mathbb{E}[U(a)] - C(e). \tag{8}$$

SEIZE buys evidence only when it can change the decision, reduce expected loss, preserve valuable optionality or create transferable institutional capacity. It ends with a Promotion Request, not a Release Certificate.

### 2.5 The closed-loop succession law

**Verified succession law**

The frontier creates candidates. Navigator Ω detects and books the opportunity. SEIZE underwrites the smallest reversible proof path. The Gym exposes the mission. The Foundry constitutes complete successors. Protected reality determines mission dominance. Accountable authority admits. Chronicle preserves what survives and opens the next succession event.

The resulting control relation is

$$\boxed{A_t \preceq \phi(E_t) \preceq A_{\max}(m),} \tag{9}$$

where $A_{\max}(m)$ is the mission's constitutional ceiling. Better capability may justify a new proof mandate ; it never grants itself additional authority.

## 3 Navigator Ω : Frontier Radar, Capability Twin and Successor Book

### 3.1 The Navigator is the sensing and portfolio layer

GoalOS Singularity Navigator Ω does not merely display frontier news. It converts a material capability event into a versioned institutional question :

> *What changed ; which incumbent assumption is now impaired ; which successor states became reachable ; what evidence would change the decision ; and what is the maximum capital or authority that may be exposed before that evidence arrives ?*

Its core objects are the Frontier Radar, the Institutional Capability Twin, the Incumbent and Assumption Ledger, and the Successor Book. Together they turn an external discontinuity into a governed internal market for architectures.

### 3.2 Frontier events and the architecture-depreciation ledger

A Frontier Event Record should contain :

- source, date, provenance and replication status ;
- capability, cost, law, security, market or infrastructure change ;
- affected mission and incumbent assumptions ;
- missions newly feasible, threatened or obsolete ;
- expected cost of waiting and premature action ;
- evidence required before capital or authority changes ;
- accountable owner, expiry and next review.

The Incumbent and Assumption Ledger records why the present architecture remains accepted. An assumption can be technical, economic, legal, organizational, security-related or constitutional. When evidence expires, the assumption is marked for re-underwriting ; the incumbent is not automatically retired.

Define the expected present cost of waiting as

$$C_{\text{wait}} = V_{\text{foregone}} + C_{\text{legacy}} + R_{\text{competitive}} + R_{\text{dependency}} + I_{\text{option loss}}, \tag{10}$$

and the cost of premature succession as

$$C_{\text{premature}} = C_{\text{formation}} + C_{\text{integration}} + R_{\text{error}} + R_{\text{security}} + C_{\text{rollback}} + R_{\text{lock-in}} + D_P. \tag{11}$$

A succession event opens when the expected decision value of re-underwriting exceeds the cost of review ; it does not itself authorize replacement.

### 3.3 The Institutional Capability Twin

The Capability Twin is a live, evidence-bearing model of how the institution presently performs the mission. It contains :

- objective, beneficiary, workflow, roles and decision rights ;
- current models, agents, tools, data, memory and providers ;
- accepted quality, cost, latency, risk, human burden and authority ;
- rights, contracts, jurisdiction, confidentiality and retention ;
- dependencies, failure modes, degraded states and rollback ;
- evidence coverage, Proof Debt and assumption expiry ;
- realized outcomes and the conditions under which the incumbent should be challenged.

The Twin is not a decorative digital replica. It is the reference architecture against which every candidate must be measured. A candidate cannot manufacture apparent Alpha by comparing itself with an incomplete or deliberately weakened baseline.

### 3.4 The Successor Book

The Book prevents two opposite errors : treating the incumbent as permanent, and treating the newest challenger as inevitable. It maintains a portfolio whose value depends on expected mission gain, proof cost, duration, covariance of failure, reversibility and strategic option value.

For candidate $i$, a compact underwriting state is

$$U_i = p_i V_i + O_i - C_i^P - C_i^F - R_i - D_i, \tag{12}$$

where $p_i$ is the current probability of passing ; $V_i$ the proof-adjusted mission value ; $O_i$ option value ; $C_i^P$ proof cost ; $C_i^F$ formation cost ; $R_i$ downside ; and $D_i$ duration, decay and Proof Debt. This score informs SEIZE ; it never creates authority.

**TABLE 3.** Canonical Successor Book positions.

| Position | Institutional meaning | Capital and authority posture |
|---|---|---|
| Incumbent | Current accepted architecture and the evidence supporting continued operation. | Operate, monitor, maintain fallback and mark assumptions for expiry. |
| Immediate successor | A challenger with clear rights, credible verifier, short proof path and positive expected value of decisive evidence. | Authorize bounded proof capital; no inherited production authority. |
| Challenger | Potentially superior after one or two named catalysts, evidence objects or architecture improvements. | Maintain a bounded evidence programme and explicit promotion trigger. |
| Reserve | High option value but long duration, incomplete rights or weak present proof. | Preserve cheaply; do not represent as operational capability. |
| Hedge architecture | Fallback against provider, security, legal, market, data or infrastructure failure. | Maintain operational readiness and periodic drills. |
| Retirement candidate | Negative expected value, expired rights, unacceptable risk or strategic obsolescence. | Contain, migrate, revoke, wind down or prohibit. |

### 3.5 Three horizons and option discipline

1. **Immediate horizon.** Enough evidence exists to fund a decisive bounded test now.
2. **Challenger horizon.** The thesis becomes investable after a named capability, cost, right, dataset or distribution catalyst.
3. **Reserve horizon.** The institution preserves a low-cost option against a consequential but presently unqualified future state.

Proof capital is the option premium. Formation capital is exercised only when the evidence purchased by the proof stage justifies the commitment. The institution should prefer a small early loss on a weak thesis to a large late loss on an unproven architecture.

### 3.6 From sensing to SEIZE

Navigator Ω ends when it has created a qualified succession event and a structured Successor Book. **SEIZE**begins when the institution selects one event for underwriting. The handoff contains :

— exact mission and beneficiary;
— incumbent and best credible alternative;
— preserve–build–become constraints;
— candidate classes and anti-recommendations;
— proof-capital ceiling and stop-loss;
— next decisive evidence;
— opportunity half-life and time-to-proof;
— rights, confidentiality and authority ceiling;
— accountable owner and decision deadline.

**Navigator law**

The frontier is not a recommendation. Navigator Ω converts frontier motion into an attributable succession event, a live portfolio of institutional alternatives and a bounded question for SEIZE. No event may pass directly from attention to authority.

### 3.7 Sensing Alpha and institutional latency

Sensing can itself create Alpha when the institution detects a material succession event before competitors, customers or regulators fully incorporate it. Let $T_i$ be institution $i$'s verified mission-cycle time and $L_i$ its validated learning rate. A stylized competitive condition is

$$\frac{L_i}{T_i} \geqslant \max_{j \in \mathscr{C}_i} \left( \frac{L_j}{T_j} \right) - \epsilon_i, \tag{13}$$

where $\epsilon_i$ represents durable complements such as brand, regulation, scarce assets, distribution or relationships. Navigator Ω is strategically valuable when it reduces the lag from external change to a protected institutional decision without laundering uncertainty into authority.

# 4 The Inner Mission Triad : Gym, Specialist ASI and Successor Ω

## 4.1 The triad

Navigator Ω and the Verified Succession Institution form the outer sensing, underwriting and renewal control plane. Inside each constituted mission, the terms Mission Gym, Specialist ASI and Mission-Sovereign Successor Ω refer to different objects and answer different questions.

**Table 4.** The Gym–ASI–Successor triad.

| Object | What it is | Question answered |
|---|---|---|
| Mission Gym | A controlled, versioned representation of the mission's state, role-specific observations, actions, transitions, outcomes, constraints, scenario distributions and independent verification. | Can this architecture actually perform the mission, fail safely and transfer beyond development cases ? |
| Specialist ASI | The complete candidate intelligence institution that proves mission-specific superiority over the incumbent and the best credible general, specialist, human, software or hybrid alternative. | Which complete intelligence institution wins the mission under fresh proof ? |
| Mission-Sovereign Successor | The proven Specialist ASI after separate accountable admission, with identity, evidence, Chronicle, monitoring, rollback and a scoped Authority Envelope. | What authority should the institution grant this proven capability, for which version, scope, duration and conditions ? |

The three objects form a sequence but not a collapse :

$$\boxed{\text{Mission Gym} \neq \text{Specialist ASI} \neq \text{Mission-Sovereign Successor } \Omega.} \tag{14}$$

## 4.2 The Gym is not the intelligence

The Gym is the mission's controlled world. It contains the rules, data interfaces, constraints, scenario generator, hidden truth where available, and verification logic. A candidate can interact with it, but the Gym does not itself decide what to do. It may contain simulations, historical replay, live shadow interfaces, digital twins, deterministic checks and strategic counterparties. It is owned and governed by the institution or an independent proof custodian.

A Gym can be valuable even when no candidate qualifies as Specialist ASI. It can demonstrate that the incumbent remains superior, that a proposed reward is exploitable, that a provider's performance does not transfer, or that the mission is not sufficiently measurable. A clean rejection is an output of the Gym, not a failure of the Gym.

## 4.3 Specialist ASI is not one policy

A policy $\pi(a \mid h)$ maps a history or observation to an action distribution. A Specialist ASI can include many policies and non-policy components. Its intelligence may arise from :

— a portfolio of general and specialist models ;
— explicit world models and causal representations ;
— retrieval, tools, programs, search and simulators ;
— a producer, critic, verifier and governor ;
— human experts and accountable professional authorities ;
— a shared evidence graph and uncertainty model ;
— workflow integration and institutional memory.

The complete system may be superhuman on the constituted mission even if no individual component is. Conversely, a component can be superhuman on a subtask while the complete institution remains inferior because coordination, evidence, cost, reliability or authority fails.

Definition 4.1 (Mission-Dominant Specialist ASI). A complete mission-specific intelligence institution that exceeds the incumbent and the best credible general, specialist, human, software or hybrid alternative by a preregistered positive margin on fresh, protected and representative work, after quality, critical-error risk, complete economics, human burden, rights, sovereignty, resilience and governance are counted.

This definition follows the system-level and mission-bounded claim discipline established in the foundational GoalOS paper (Boucher, 2026b). It does not imply universal intelligence.

### 4.4 Successor Ω is not unlimited autonomy

The symbol Ω denotes the complete institutional state in which a proven mission capability has been constituted, admitted and placed under continuous reproof. It is not legal personhood, self-ownership or unlimited sovereignty. The successor remains subordinate to accountable human and institutional principals.

Definition 4.2 (Mission-Sovereign Successor). A Mission-Dominant Specialist ASI separately admitted by accountable authority under a scoped, versioned, monitored, expiring and revocable Authority Envelope, with proof-carrying action, Chronicle memory, degraded modes, rollback and requalification.

The progression is

$$\begin{aligned}\text{Candidate Successor} &\longrightarrow \text{Verified Specialist Intelligence} \longrightarrow \text{Mission-Dominant Specialist ASI} \\ &\longrightarrow \text{Mission-Sovereign Successor.}\end{aligned} \tag{15}$$

### 4.5 The Successor Complex

The strongest object of analysis is the **Successor Complex** :

$$\mathfrak{X}_m = \langle \mathscr{G}_m, \gg, \mathscr{P}, \mathit{Chronicle}, \mathscr{A} \rangle, \tag{16}$$

where $\mathscr{G}_m$ is the sovereign mission environment, $\gg$ the complete candidate or admitted successor institution, $\mathscr{P}$ the independent protected evaluation plane, *Chronicle* the admitted evidence and memory state, and $\mathscr{A}$ the current Authority Envelope.

This representation prevents a common category error : treating model performance as if it already contained environment validity, proof integrity, rights, operations, memory and authority.

**Triad law**

The Gym makes the mission executable. Specialist ASI establishes which complete intelligence institution wins. Successor Ω determines how the proven winner may enter reality. Chronicle determines what survives long enough to shape the next succession.

### 4.6 The canonical progression

## 5 The Sovereign Mission Gym

### 5.1 Definition as a constrained partially observable stochastic game

A corporate mission is rarely a clean single-agent Markov decision process. It contains hidden state, partial observations, sequential and simultaneous actions, human and machine roles, external providers, adversaries, budgets, rights, deadlines and irreversible consequences. We therefore define the Mission Gym as a constrained partially observable stochastic game, extending standard environment interfaces (Kaelbling et al., 1998 ; Towers et al., 2024 ; Terry et al., 2021 ; Lanctot et al., 2019).

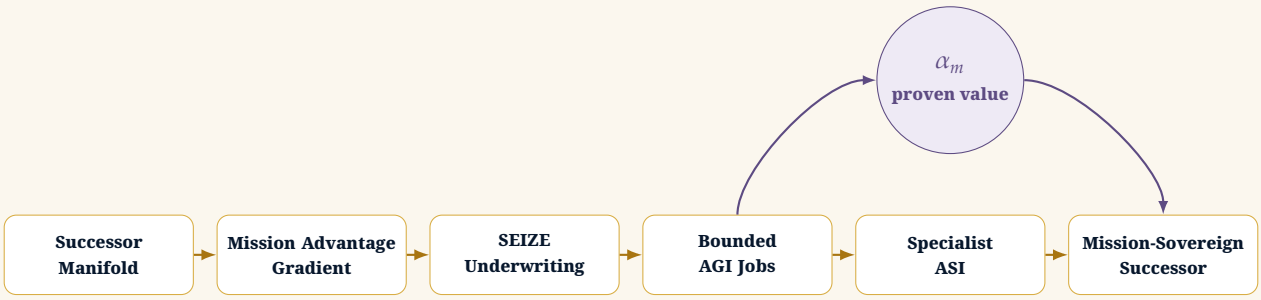



**Figure 1.** The canonical progression. Alpha is the residual value field produced by proof; it is not an independent chronological stage.

Definition 5.1 (GoalOS Mission Gym). For mission $m$, the Mission Gym is

$$\mathcal{G}_m = \left(\mathcal{N}, \mathcal{X}, \{\mathcal{O}_i\}_{i\in\mathcal{N}}, \{\mathcal{A}_i\}_{i\in\mathcal{N}}, \mathcal{T}, \{\mathsf{r}_i\}_{i\in\mathcal{N}}, \mathsf{c}, \rho, \tau, \mathcal{V}, \mathcal{R}\right), \tag{17}$$

where $\mathcal{N}$ is the player set; $\mathcal{X}$ the hidden mission state; $\mathcal{O}_i$ the observation space permitted to player $i$; $\mathcal{A}_i$ the bounded action space; $\mathcal{T}$ the transition kernel; $\mathsf{r}_i$ the mission-dependent payoff vector; $\mathsf{c}$ the constitutional constraint vector; $\rho$ the scenario distribution; $\tau$ the termination and truncation rules; $\mathcal{V}$ the independent verifier; and $\mathcal{R}$ the proof-receipt function.

The player set may include :

$$\mathcal{N} = \{P, I, B, C, V, G, H, E, R\}, \tag{18}$$

where $P$ is the accountable principal, $I$ the incumbent operator, $B$ the builder or producer, $C$ the critic, $V$ the verifier, $G$ the governor, $H$ affected humans or professional authorities, $E$ external providers or counterparties, and $R$ adversarial reality.

## 5.2 State, observation and information asymmetry

The true mission state $x_t \in \mathcal{X}$ may include physical, financial, contractual, strategic, organizational and authority conditions that no single actor can fully observe. Player $i$ receives

$$o_t^i \sim O_i(\cdot \mid x_t, \eta_t), \tag{19}$$

where $\eta_t$ represents sensor noise, missing information, permissions and adversarial distortion. The history available to the player is

$$h_t^i = (o_0^i, a_0^i, \dots, o_t^i), \tag{20}$$

and its policy is

$$\pi_i(a_t^i \mid h_t^i, \theta_i). \tag{21}$$

The Gym must make observation rights explicit. A verifier may see protected ground truth that the producer cannot. The candidate may see the mission input but not scorer internals. The governor may see permissions and budgets but not need access to all confidential business content. The principal may see the final evidence and trade-offs while delegating technical execution.

Design Proposition 5.1 (Information separation). A Mission Gym is constitutionally stronger when the minimum information necessary for each role is sufficient to perform that role, while protected evaluation, release credentials and unnecessary confidential state remain inaccessible.

## 5.3 Action spaces and the Authority Envelope

The action space is not merely a list of model outputs. It includes requests for evidence, tool calls, internal recommendations, human escalations, sandbox operations and bounded external actions. Let

$$\mathcal{A}_i = \mathcal{A}_i^{\text{analysis}} \cup \mathcal{A}_i^{\text{evidence}} \cup \mathcal{A}_i^{\text{tool}} \cup \mathcal{A}_i^{\text{approval}} \cup \mathcal{A}_i^{\text{external}}. \tag{22}$$

The current Authority Envelope $A_t$ induces an action mask

$$M_{A_t}(a) = \begin{cases} 1, & a \text{ is permitted for the current identity, state and evidence;} \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

Consequential constraints must be enforced outside the candidate policy. Prompting a candidate not to exceed authority is not an authority system.

### 5.4 Transition dynamics

The mission evolves according to

$$x_{t+1} \sim \mathscr{T}(\cdot \mid x_t, a_t^{1:n}, w_t), \tag{24}$$

where $a_t^{1:n} = (a_t^1, \dots, a_t^n)$ is the joint action and $w_t$ captures exogenous shocks. In many corporate settings, the transition model is partly deterministic, partly empirical and partly simulated. The Gym may combine :

— historical replay ;
— a deterministic rules engine ;
— a learned world model ;
— a physical or financial simulator ;
— live shadow observations ;
— human-in-the-loop transitions ;
— strategic counterparties or adversarial agents.

The transition model is a major source of Gym basis risk. A high-fidelity interface does not guarantee a high-fidelity world.

### 5.5 Vector rewards and non-compensable constraints

A single reward scalar invites Goodhart effects and hides mission trade-offs. GoalOS uses a vector reward

$$\mathsf{r}_t = \begin{bmatrix} Q_t & V_t & -T_t & \mathscr{R}_t & E_t & S_t & K_t & -C_t & -H_t \end{bmatrix}^\top, \tag{25}$$

where $Q$ is mission quality, $V$ verified value, $T$ time, $\mathscr{R}$ reliability, $E$ evidence, $S$ sovereignty, $K$ future capacity, $C$ complete cost and $H$ human burden.

Hard constraints remain separate :

$$\mathsf{c}_t = \begin{bmatrix} \text{critical error} \\ \text{unauthorized action} \\ \text{safety violation} \\ \text{rights violation} \\ \text{privacy breach} \\ \text{budget breach} \\ \text{governance regression} \end{bmatrix}. \tag{26}$$

The optimization problem is

$$\max_{\pi} \ \mathbb{E}\left[\sum_{t=0}^{T} \gamma^t U_m(\mathsf{r}_t)\right] \quad \text{subject to} \quad \mathsf{c}_t = 0 \text{ for all hard-gate dimensions.} \tag{27}$$

Many low-severity errors may not compensate for one prohibited catastrophic failure.

### 5.6 Termination, truncation and refusal

Gymnasium's distinction between termination and truncation is institutionally important (Towers et al., 2024). A mission **terminates** when it genuinely succeeds, fails or reaches a constitutionally defined terminal state. It is **truncated** when time, budget, evidence, authority or external availability expires before completion.

**Table 5.** Institutional termination and truncation.

| Condition | Meaning | Required record |
|---|---|---|
| Successful termination | The accepted mission objective was achieved and independently verified. | Proof receipt, cost, evidence, residual uncertainty and authority state. |
| Failure termination | A critical failure, stop condition or falsification occurred. | Failure class, containment, rollback and Chronicle eligibility. |
| Time truncation | The deadline arrived before a valid result. | Remaining Proof Debt and whether continuation has positive EVI. |
| Budget truncation | Proof or operating capital was exhausted. | Complete denominator and new underwriting decision. |
| Authority truncation | The action required exceeds the current envelope. | Escalation or refusal; never implicit expansion. |
| Evidence truncation | Required sources, rights or verifier state expired. | Fail-closed disposition and requalification trigger. |

### 5.7 The proof receipt

Every episode emits an immutable or append-only receipt :

$$\mathscr{R}(e) = \mathrm{Hash}\Big(\text{Gym version, scenario, candidate, observations,}\\ \text{actions, tools, human interventions, r, c, verdict}\Big). \tag{28}$$

The receipt must include the exact environment and candidate releases, because neither the Gym nor the successor is stationary.

### 5.8 Five Gym families

A production-grade institution separates at least five Gym families.

**Table 6.** The GoalOS Gym stack.

| Gym | Purpose | Candidate visibility |
|---|---|---|
| Formation Gym | Build tools, policies, world models, specialist organizations and candidate architectures. | Visible; repeated access permitted. |
| Adversarial Gym | Attack rewards, tools, evidence, verifier, authority and resilience. | Visible or semi-blind; designed to expose failure. |
| Transfer Gym | Test new populations, regimes, properties, industries or failure clusters. | Partially protected; prevents narrow overfit. |
| Fresh-Proof Gym | Evaluate one frozen challenger on unseen representative work with protected seeds and scorer internals. | Hidden from candidate and formation roles. |
| Canary and Requalification Gym | Connect controlled evaluation to shadow, sandbox, reversible bounded action and production monitoring. | Limited according to current authority. |

**Gym separation law**

The Gym that teaches the candidate must not be the Gym that certifies it. Training rewards teach; protected proof decides; accountable authority admits.

### 5.9 Gym maturity

## 6 The Complete Successor Is Larger Than the Policy

**TABLE 7.** Mission Gym maturity model.

| Level | State | Evidence |
|---|---|---|
| G0 | Narrative mission | Objectives and outputs described, but no executable state or verifier. |
| G1 | Replay corpus | Historical episodes, labels and baselines. |
| G2 | Interactive environment | Typed observations, actions, transitions, rewards and termination. |
| G3 | Multi-player institution | Producer, critic, verifier, governor and principal are represented. |
| G4 | Protected proof plane | Blind cases, independent custody, exact releases and confidence gates. |
| G5 | Canary interface | Shadow and reversible external actions with fail-closed controls. |
| G6 | Sovereign Gym asset | Rights-cleared environment, portable APIs, Chronicle integration and requalification. |
| G7 | Recursive succession environment | The Gym and successor co-evolve under protected proof without self-authorization. |

## 6.1 The policy fallacy

The dominant technical abstraction in sequential decision-making is a policy $\pi$ that maps observations or histories to actions. This abstraction is necessary but institutionally incomplete. A policy does not, by itself, establish :

— whether the mission objective is legitimate and frozen ;
— whether the observations are rights-cleared and sufficiently complete ;
— whether tools and external systems are permitted ;
— whether evidence is preserved ;
— whether the output is independently verified ;
— whether humans were required to rescue the result ;
— whether a recommendation may become an action ;
— whether the system can roll back or degrade safely ;
— whether learning may influence future missions ;
— whether a later release remains the same institutional object.

The complete successor must therefore be represented as an institution rather than a policy.

Definition 6.1 (Complete Successor Institution). For mission $m$, a complete successor is

$$\gg_m = (M, D, W, H, T, R, V, P, O, E, C, A, B)\,, \tag{29}$$

where $M$ is the model portfolio ; $D$ rights-cleared data and examples ; $W$ the world model and knowledge representation ; $H$ the harness, orchestration and human roles ; $T$ tools, retrieval, code and simulators ; $R$ rules, rights, security and resilience ; $V$ the verifier system ; $P$ the operating policy portfolio ; $O$ workflow and organizational integration ; $E$ the evidence state ; $C$ Chronicle ; $A$ the Authority Envelope ; and $B$ rollback, fallback and lifecycle controls.

A single policy is a component :

$$\pi \in P \subset \gg_m\,. \tag{30}$$

## 6.2 The institutional stack

## 6.3 World model and mission model

The successor's internal world model $\mathscr{W}_t$ is not identical to the Gym's authoritative hidden state $x_t$. The successor estimates :

$$q_\theta(x_t \mid h_t) = \Pr(x_t \mid h_t; \theta), \tag{31}$$

while the Gym or real environment governs the actual transition. A strong successor maintains explicit separation among :

— observed fact ;


**Figure 2.** The complete Successor is larger than the acting policy.

— externally asserted claim ;
— derived calculation ;
— model inference ;
— counterfactual scenario ;
— unresolved uncertainty.

This separation is essential in acquisition underwriting, building operations, financial reconciliation and other missions where persuasive synthesis can conceal missing evidence.

### 6.4 Policy portfolio rather than monolith

A complete successor may route among several policies :

$$\pi^{\star}(a \mid h) = \sum_{k=1}^{K} g_k(h)\pi_k(a \mid h), \tag{32}$$

where $g_k(h)$ is a governed routing distribution. The portfolio may include :

— a general frontier model for broad reasoning ;
— a local specialist model for confidential classification ;
— deterministic code for calculations and invariants ;
— search or planning for sequential choices ;
— a critic with different data or model lineage ;
— a human professional for reserved judgment.

The routing policy is itself subject to proof. A successor can fail through incorrect routing even when every individual specialist is competent.

### 6.5 Verifier independence

The verifier is not merely another prompt asking whether the answer looks correct. It may include :

— deterministic invariants ;
— hidden ground truth ;
— source and provenance checks ;
— adversarial alternative generation ;
— execution tests ;
— independent models or methods ;

— human or professional adjudication;
— post-outcome feedback.

Let $Y$ be the true acceptance state and $\hat{Y}_V$ the verifier decision. The verifier must be measured for both false acceptance and false rejection :

$$\mathrm{FAR}_V = \Pr(\hat{Y}_V = 1 \mid Y = 0), \qquad \mathrm{FRR}_V = \Pr(\hat{Y}_V = 0 \mid Y = 1). \tag{33}$$

A verifier that is too permissive launders unsupported capability into authority. A verifier that is too restrictive destroys option value and forces repeated work. Chronicle precision depends on both.

### 6.6 Human roles are components of the architecture

The human is not an unspecified safety fallback. The complete architecture must state which human roles remain responsible for :

— mission purpose;
— professional judgment;
— exception resolution;
— stakeholder communication;
— legal or fiduciary authority;
— acceptance and admission;
— incident response and revocation.

Human burden is a denominator, not a hidden subsidy. Let $H_m(S)$ include review time, escalation, rescue, training, contestation and professional exposure. A candidate that appears autonomous only because humans repair its failures has not established lower burden.

### 6.7 Lifecycle and degraded modes

The complete successor defines at least six modes :

$$\text{Full} \to \text{Restricted} \to \text{Recommend only} \to \text{Read only} \to \text{Incumbent fallback} \to \text{Emergency stop}. \tag{34}$$

The transition policy among modes is part of the successor, not an afterthought. It must respond to evidence expiry, provider failure, distribution shift, security incident, budget exhaustion, critical error and authority loss.

**Complete-successor law**

A policy can act in the Gym. Only a complete Successor Institution can inherit a mission. Models produce capability; the institution produces evidence, continuity, accountability and bounded authority.

Ω

PART II

# Mission Physics : Geometry, Dynamics and Phase Change

*The manifold is constrained possibility; the Gradient is proof-adjusted direction*

The formal geometry, stochastic dynamics and phase transitions of mission architecture.

SUCCESSOR MANIFOLD → GRADIENT → SEIZE → BOUNDED JOBS → SPECIALIST ASI → SUCCESSOR Ω

# 7 The Physics of the Successor Manifold

## 7.1 Architecture space as a constrained manifold

The frontier does not present one successor. It changes the topology of what can be built. Let a complete architecture be described by coordinates

$$x = (M, D, W, H, T, R, V, P, O, E, C, A, B) \in \mathscr{X}. \tag{35}$$

The **Successor Manifold** $\mathscr{M}_S$ is the constraint-bearing subset of materially reachable architectures after a capability, cost, rights, market, security or infrastructure discontinuity. Reachability depends on capital, time, data, professional authority, integration, compute, rights and reversibility. The admissible region for mission $m$ is

$$\mathscr{M}_S^{\mathrm{adm}}(m) = \{x \in \mathscr{M}_S : C_m(x) = 1\}, \tag{36}$$

where $C_m$ encodes non-negotiable constitutional gates.

The manifold language is a disciplined analogy grounded in smooth manifolds, nonlinear dynamics and information geometry (Lee, 2013; Strogatz, 2015; ichi Amari, 2016). Many coordinates are discrete or combinatorial; local gradients may therefore be finite differences, surrogate models or search signals rather than ordinary derivatives.

## 7.2 Institutional metric tensor

Two architectures can be close in benchmark performance and far apart institutionally. Let $G_m(x)$ be a positive semi-definite mission-dependent metric tensor. For an infinitesimal architecture change $dx$,

$$ds_m^2 = dx^\top G_m(x) dx. \tag{37}$$

The metric includes formation capital, integration, migration, rights, proof, human burden, requalification, dependency, security, rollback and irreversibility. A hosted model may be technically one API call away but institutionally distant from the desired sovereign state.

For policy parameters $\theta$, the Fisher information matrix provides a local measure of behavioural sensitivity :

$$F(\theta) = \mathbb{E}_{\tau \sim p_\theta}\left[\nabla_\theta \log p_\theta(\tau) \nabla_\theta \log p_\theta(\tau)^\top\right]. \tag{38}$$

GoalOS may combine an information-geometric policy metric with a broader institutional metric :

$$G_m(x) = \begin{bmatrix} F(\theta) & 0 \\ 0 & G_{\mathrm{institution}}(z) \end{bmatrix}, \tag{39}$$

where $z$ contains non-policy architecture coordinates.

## 7.3 Mission value field

Define a proof-adjusted mission value field

$$V_m(x) = w_Q Q(x) + w_V V(x) + w_R R_v(x) + w_S S(x) + w_G G(x) + w_K K(x) \tag{40}$$

$$- w_C C(x) - w_T T(x) - w_H H(x) - w_D D_P(x) - w_U I(x), \tag{41}$$

where $Q$ is quality, $V$ verified value, $R_v$ reliability, $S$ sovereignty, $G$ governance integrity, $K$ future capacity, $C$ complete cost, $T$ time, $H$ human burden, $D_P$ Proof Debt and $I$ irreversibility.

The Riemannian Mission Advantage Gradient is

$$\nabla^{G_m} V_m(x) = G_m(x)^{-1} \nabla V_m(x). \tag{42}$$

It answers a practical question : which local architecture change produces the greatest expected increase in mission value after institutional burden is counted ?

The gradient may point toward a stronger verifier, more rights-cleared data, narrower authority, lower latency, a local model, deterministic checks, new human roles or better fallback—not necessarily a larger model.

### 7.4 SEIZE as a least-action trajectory

SEIZE underwrites a path $\gamma : [0,1] \to \mathscr{M}_S^{\mathrm{adm}}$ from incumbent $x_0$ to candidate $x_1$. Define the succession action

$$\mathscr{S}_m[\gamma] = \int_0^1 \left[ \frac{1}{2} \dot{\gamma}^\top G_m(\gamma) \dot{\gamma} + \lambda_R \mathscr{R}(\gamma) + \lambda_A \mathrm{AaR}(\gamma) + \lambda_D D_P(\gamma) - \lambda_V V_m(\gamma) \right] d\tau. \tag{43}$$

The underwritten trajectory is

$$\gamma^\star = \arg\min_{\gamma \in \Gamma_m} \mathscr{S}_m[\gamma] \quad \text{subject to hard gates and proof-capital limits.} \tag{44}$$

This is not a claim that organizations obey fundamental mechanics. It is an exact design analogy that forces the institution to price the path, not merely the endpoint. A technically attractive state reached through irreversible lock-in, unprovable data rights or unsafe migration may have higher action than a more modest but reversible successor.

### 7.5 Stochastic successor dynamics

Architecture evolution is noisy and nonstationary. A local stochastic model is

$$dX_t = -G_m(X_t)^{-1} \nabla \mathscr{F}_m(X_t) dt + B(X_t) dW_t + u_t dt, \tag{45}$$

where $\mathscr{F}_m$ is an institutional free-burden potential, $B(X_t)dW_t$ captures exogenous shocks and formation variance, and $u_t$ is the governed intervention induced by SEIZE and the Foundry. One useful potential is

$$\mathscr{F}_m(x) = B_{\mathrm{institution}}(x) - \eta V_m(x), \tag{46}$$

so lower potential corresponds to stronger value at lower complete burden.

The Foundry does not passively follow this stochastic flow. It generates controlled perturbations, evaluates descendants and imposes absorbing barriers at constitutional violations.

### 7.6 Metastability and succession events

The incumbent can remain a local equilibrium even after a superior state becomes reachable. Switching costs, authority inertia, incomplete evidence and path dependence create metastability. Let $B_m$ be the best credible alternative set and define the order parameter

$$\Phi_m(x) = V_m(x) - \max_{b \in B_m} V_m(b). \tag{47}$$

A Mission Superiority Crossover occurs when

$$\mathrm{LCB}_{1-\alpha}(\Phi_m(x)) > \theta_m \tag{48}$$

on protected fresh work. A Mission-Sovereign phase transition requires additional gates :

$$\Sigma_m(x) = \mathbb{1}[\Phi_m(x) > \theta_m] \mathbb{1}[\mathrm{FreshProof}] \mathbb{1}[\mathrm{HumanAdmission}] \mathbb{1}[\mathrm{AuthorityEnvelope}] \mathbb{1}[\mathrm{RollbackReady}]. \tag{49}$$

Only when $\Sigma_m(x) = 1$ may the release enter the Mission-Sovereign state, and only for the specified mission, version, duration, jurisdiction, data classes and action scope.

### 7.7 Mission gauge invariance

A corporation may substitute a model, provider, tool or implementation without changing the mission-level institution if the relevant observables and proofs remain invariant. Let a transformation group $\mathscr{H}$ act on internal components :

$$x' = h \cdot x, \qquad h \in \mathscr{H}. \tag{50}$$

Let $\mathscr{O}_m(x)$ denote mission observables : accepted quality, evidence, cost, reliability, rights, authority, rollback and Chronicle continuity. A transformation is mission-gauge admissible when

$$\mathscr{O}_m(h \cdot x) = \mathscr{O}_m(x) \tag{51}$$

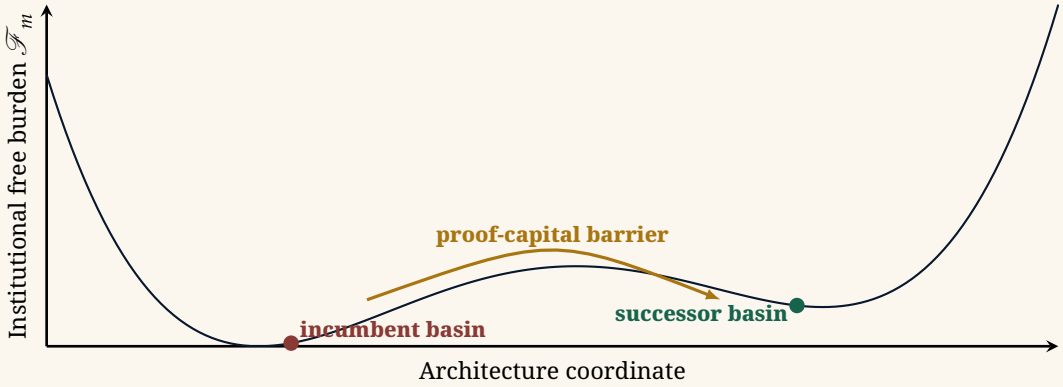



**FIGURE 3.** Metastability : a superior successor basin may exist while the incumbent remains locally stable. Proof capital purchases the evidence required to cross the institutional barrier without uncontrolled release.

within preregistered tolerances and the evidence chain remains valid.

The equivalence class

$$[x]_m = \{h \cdot x : h \in \mathscr{H},\ \mathscr{O}_m(h \cdot x) \simeq \mathscr{O}_m(x)\} \tag{52}$$

represents provider and component freedom. Sovereignty is not autarky; it is gauge freedom over replaceable components while preserving mission invariants.

Design Proposition 7.1 (Sovereignty as gauge freedom). Holding mission performance constant, a successor with a larger admissible substitution group and portable proof state has greater option value and lower expected provider hold-up than an otherwise identical successor with one trapped implementation.

### 7.8 Renormalization from episode to institution

Micro-level benchmark performance does not automatically aggregate into macro-level mission value. Let $\mathscr{R}_\ell$ be a coarse-graining operator from scale $\ell$ to $\ell + 1$ :

$$Z^{(\ell+1)} = \mathscr{R}_\ell\left(Z^{(\ell)}, C^{(\ell)}, H^{(\ell)}, D_P^{(\ell)}\right). \tag{53}$$

Relevant scales include :

$$\text{action} \longrightarrow \text{episode} \longrightarrow \text{workflow} \longrightarrow \text{mission} \longrightarrow \text{portfolio} \longrightarrow \text{institution}. \tag{54}$$

A capability can be strong at the action scale and weak at the mission scale because errors correlate, review burden compounds or incentives change. The effective parameters at each scale must be re-estimated rather than inherited.

**No scale collapse**

A benchmark result is not a mission result; a mission result is not institutional value; institutional value is not authority. Each scale requires its own effective metrics, uncertainty and proof.

## 8 Axioms and Formal Results of Verified Succession

The preceding geometry is useful only if its institutional claims can be stated under explicit assumptions. This section therefore separates axioms, definitions and derived results. None of the results asserts universal ASI, general social optimality or literal physical law. They state what follows inside the GoalOS constitution when the listed assumptions hold.

### 8.1 Constitutional axioms

| Axiom | Name | Statement |
|---|---|---|
| A1 | Mission specificity | Every capability, value, risk and authority claim is indexed by an exact mission family, release, evidence state and time. |
| A2 | Complete reference set | Promotion is measured against the incumbent and the strongest credible external, specialist, human, software, partner, acquisition and no-action alternatives that are materially available. |
| A3 | Protected evaluation separation | The frozen challenger and its formation roles cannot inspect protected cases, scorer internals, release credentials or post hoc thresholds. |
| A4 | Complete denominator | Mission value includes construction, operation, proof, human review, integration, maintenance, dependency, tail risk, rights, requalification and unwind. |
| A5 | Evidence-bounded authority | The admissible Authority Envelope is an isotone function of current evidence and remains below the mission's constitutional ceiling. |
| A6 | Release non-inheritance | A material change in mission, model, data, tool, workflow, verifier, environment, provider or authority creates a new release for the affected claims. |
| A7 | Independent admission | Mission dominance and institutional authority are distinct states; only an accountable principal outside the candidate can issue the admission transition. |
| A8 | Chronicle conditionality | Only rights-cleared, attributable, versioned, challengeable and requalifiable capability may influence future missions. |

## 8.2 Robust promotion functional

Let $\widehat{\Delta}_m(x)$ be the matched fresh-work estimate of candidate gain over the strongest reference architecture, and let $\widehat{\varepsilon}_{\mathrm{gym}}(x)$ be a conservative upper bound on Gym-to-reality basis risk. Define the robust promotion margin

$$\mathscr{M}_m(x) = \mathrm{LCB}_{1-\alpha}\Big(\widehat{\Delta}_m(x)\Big) - \widehat{\varepsilon}_{\mathrm{gym}}(x) - D_P(x) - \theta_m. \tag{55}$$

The candidate is *proof-eligible* only when $\mathscr{M}_m(x) > 0$ and every hard constitutional gate passes. Proof eligibility is not authority.

## 8.3 Beta compression under symmetry

Proposition 8.1 (Beta compression). *Suppose $n$ institutions receive the same rentable frontier capability $F$, possess identical complementary assets, face the same mission and can imitate one another at negligible cost. If product-market competition transfers all common productivity surplus into prices, wages, supplier rents or customer expectations, then expected relative mission surplus from $F$ is zero.*

*Démonstration.* Let institution $i$ receive common uplift $b(F)$ and private complement surplus $a_i$. Under the assumptions, $a_i = a_j$ for all $i, j$ and the competitive adjustment $q(F)$ is common. Relative payoff is

$$\pi_i - \frac{1}{n}\sum_j \pi_j = \Big[b(F) - q(F) + a_i\Big] - \frac{1}{n}\sum_j \Big[b(F) - q(F) + a_j\Big] = 0.$$

The frontier can raise absolute productivity while creating no relative advantage. Positive relative Alpha therefore requires an asymmetry not contained in the common factor : mission architecture, lawful data, verification, distribution, rights, learning velocity or authority design. □

## 8.4 Why protected proof is necessary

Proposition 8.2 (Selection optimism on reused evaluation). *Let $\widehat{V}_k$ be noisy evaluation estimates for $K$ candidate architectures with unbiased errors on one development set, and let $\widehat{k} = \arg\max_k \widehat{V}_k$. Unless every estimate is noiseless or perfectly correlated, the expected selected development score satisfies*

$$\mathbb{E}\Big[\widehat{V}_{\widehat{k}}\Big] \geqslant \max_k V_k, \tag{56}$$

*with strict inequality in the non-degenerate case. A fresh set independent of candidate selection is required for an unbiased estimate of the frozen winner's mission value.*

*Démonstration.* Write $\widehat{V}_k = V_k + \epsilon_k$ with $\mathbb{E}[\epsilon_k] = 0$. Since $\max_k(V_k + \epsilon_k)$ is convex in the error vector,

$$\mathbb{E}\left[\max_k(V_k + \epsilon_k)\right] \geqslant \max_k \mathbb{E}[V_k + \epsilon_k] = \max_k V_k.$$

The selected development score is the maximum on the left and therefore carries selection optimism. If a protected set $Z'$ is independent of the selection set and the candidate is frozen before observing $Z'$, then conditional on the frozen candidate $\widehat{k}$, an unbiased protected estimator satisfies $\mathbb{E}[\widehat{V}_{\widehat{k}}(Z') \mid \widehat{k}] = V_{\widehat{k}}$. □

Corollary 8.1 (No certification by the Formation Gym). *The environment used to search, tune or select candidate architectures cannot by itself establish mission dominance. A separate Fresh-Proof Gym is constitutionally necessary whenever selection used noisy environment feedback.*

### 8.5 Evidence monotonicity and revocation

Proposition 8.3 (Authority monotonicity). *Let evidence states form a partially ordered set $(\mathscr{E}, \preceq_E)$ and authority states a partially ordered set $(\mathscr{A}, \preceq_A)$. If $\phi : \mathscr{E} \to \mathscr{A}$ is isotone, then $E_1 \preceq_E E_2$ implies $\phi(E_1) \preceq_A \phi(E_2)$. If evidence expires or is impaired so that $E_{t+1} \preceq_E E_t$, the maximum justified authority cannot increase.*

*Démonstration.* The first claim is the definition of an isotone map. Applying it to $E_{t+1} \preceq_E E_t$ gives $\phi(E_{t+1}) \preceq_A \phi(E_t)$. Since actual authority must satisfy $A_t \preceq_A \phi(E_t)$, any authority outside the new upper bound must be removed, restricted or re-proved. □

Corollary 8.2 (Revocation is a normal state transition). *Under Axiom A5, revocation, expiry and rollback are not exceptional governance failures. They are required responses when current evidence no longer collateralizes inherited authority.*

### 8.6 Basis-risk-adjusted promotion

Proposition 8.4 (Robust promotion under bounded Gym basis risk). *Assume the real mission gain satisfies*

$$\Delta_m^{\text{real}}(x) \geqslant \Delta_m^{\text{gym}}(x) - \varepsilon_{\text{gym}}(x) \tag{57}$$

*for every environment in a preregistered uncertainty set. If*

$$\text{LCB}_{1-\alpha}\left(\widehat{\Delta}_m^{\text{gym}}(x)\right) - \widehat{\varepsilon}_{\text{gym}}(x) > \theta_m \tag{58}$$

*and $\widehat{\varepsilon}_{\text{gym}}$ upper-bounds $\varepsilon_{\text{gym}}$ with the declared confidence, then the candidate clears the mission-gain threshold throughout that uncertainty set at the same confidence, subject to all hard gates.*

*Démonstration.* By the assumed basis-risk inequality and the upper bound,

$$\Delta_m^{\text{real}}(x) \geqslant \Delta_m^{\text{gym}}(x) - \varepsilon_{\text{gym}}(x) \geqslant \Delta_m^{\text{gym}}(x) - \widehat{\varepsilon}_{\text{gym}}(x).$$

Replacing the unknown Gym gain by its lower confidence bound preserves the inequality at the declared confidence. The stated strict margin therefore implies $\Delta_m^{\text{real}}(x) > \theta_m$ for all environments covered by the uncertainty set. The conclusion is only mission-gain eligibility; rights, safety, governance and authority gates remain independent. □

### 8.7 Provider gauge equivalence

Proposition 8.5 (Mission-gauge equivalence). *Let $h$ substitute an internal provider, model or tool. If all preregistered mission observables, proof receipts, rights conditions, Chronicle continuity and Authority Envelope semantics remain invariant within tolerance, then $x$ and $h \cdot x$ occupy the same mission-equivalence class $[x]_m$.*

*Démonstration.* The equivalence relation is defined by equality of the governed mission observables and validity of the evidence chain. Reflexivity follows from the identity substitution, symmetry from invertibility on the retained state, and transitivity from composition of admissible substitutions. Therefore $x \sim_m h{\cdot}x$ and both represent the same mission-level institution despite different internal components. □

Corollary 8.3 (Sovereignty premium from substitutability). *If two architectures have equal mission observables but one admits a strictly larger set of proof-preserving substitutions, the latter has weakly greater continuation option value and weakly lower expected provider hold-up, before the cost of maintaining portability.*

### 8.8 No self-authorization

Proposition 8.6 (Candidate-only reachability exclusion). *Consider the canonical state machine in Appendix B. Assume admission and release transitions require credentials held exclusively by an independent verifier and accountable principal, and the candidate cannot modify the transition kernel or credential store. Then no sequence of candidate-only actions can reach a state with greater institutional authority.*

*Démonstration.* Every path from candidate formation to an authority-bearing state crosses at least one edge labelled Independent Validation, Accountable Admission or Authority Envelope Issuance. By assumption, each such edge requires an external credential unavailable to the candidate. Candidate actions therefore remain inside the subgraph preceding those edges. Reachability of a higher-authority state is impossible without an external transition. □

Corollary 8.4 (Recursive improvement is not recursive self-authorization). *A successor may generate descendants, propose experiments and request promotion. No descendant can inherit the parent's proof or authority without a new protected evaluation and external admission transition.*

### 8.9 Recursive readiness

Let $H_S$ be the expected evidence-valid half-life of the admitted successor and $T_R$ the expected time required to generate, prove and admit a replacement. Define

$$\text{RecursiveReadiness} = \frac{H_S}{T_R}. \tag{59}$$

Proposition 8.7 (Readiness condition). *If* RecursiveReadiness $> 1$, *the institution has positive expected temporal slack to constitute a replacement before the incumbent successor's evidence expires. If the ratio is at most one, continuity requires a reserve successor, narrower authority, accelerated proof or an explicit degraded mode.*

*Démonstration.* The available evidence-valid interval is $H_S$ and expected replacement duration is $T_R$. Positive slack is $H_S - T_R$, which is positive exactly when $H_S/T_R > 1$. Non-positive slack does not prove failure, but it removes the expectation that ordinary succession completes before expiry. □

**Formal constitutional result**

The mathematics does not grant authority. It makes the assumptions, uncertainty and non-bypassable transitions explicit. Protected proof may establish mission dominance; only accountable admission may constitute Successor Ω.

## 9 The Multi-Player Succession Game

### 9.1 Succession under incomplete information

Verified succession is a dynamic Bayesian game. Participants possess different information, incentives, authority and exposure (von Neumann and Morgenstern, 1944; Jr., 1950; Harsanyi, 1967; Osborne and Rubinstein, 1994). A producer may prefer promotion, an incumbent may resist replacement, a provider may prefer lock-in, an evaluator may minimize effort, an affected party may bear externalized risk, and the principal may prefer speed or political convenience.

Let

$$\mathfrak{G}_m = (\mathcal{N}, \Sigma, u, \mathcal{I}, \mathcal{H}_t), \tag{60}$$

where $\Sigma_i$ is player $i$'s strategy space, $u_i$ its payoff, $\mathcal{I}$ the information structure and $\mathcal{H}_t$ the history of evidence and authority. The mechanism must remain useful even when participants are not naturally aligned.

## 9.2 Canonical players and incentives

**Table 9.** Players in the Successor Gym.

| Player | Legitimate objective | Potential distortion |
|---|---|---|
| Accountable principal | Maximize mission value while protecting the institution. | Speed bias, political pressure, overconfidence or diffusion of responsibility. |
| Incumbent operator | Preserve continuity and reveal operational truth. | Strategic resistance, selective disclosure or sunk-cost defence. |
| Builder / producer | Constitute the strongest candidate. | Hidden retries, benchmark gaming, concealed dependencies or promotion bias. |
| Adversarial critic | Discover failure, contrary evidence and alternative explanations. | Excessive pessimism, low-cost obstruction or performative red-teaming. |
| Independent verifier | Determine whether the result passes frozen criteria. | Low effort, correlated error, capture or false precision. |
| Governor | Enforce permissions, identity, budget, expiry and rollback. | Overrestriction, poor policy encoding or hidden override channels. |
| Provider / partner | Supply models, tools, infrastructure or data. | Hold-up, opacity, data capture or terms drift. |
| Affected party | Preserve rights, safety, fairness and contestability. | Incomplete voice, weak representation or cost externalization. |
| Adversary / reality | Exploit weakness or reveal unmodelled conditions. | Strategic deception, distribution shift and correlated shocks. |

## 9.3 Separation of powers as mechanism design

The mechanism imposes structural inequalities :

$$\text{Producer} \neq \text{Verifier}, \tag{61}$$

$$\text{Verifier} \neq \text{Admitter}, \tag{62}$$

$$\text{Access Receipt} \neq \text{Authority Envelope}, \tag{63}$$

$$\text{Execution} \neq \text{Acceptance} \neq \text{Admission}, \tag{64}$$

$$\text{Accepted result} \neq \text{reinvestable value}. \tag{65}$$

The candidate cannot inspect protected cases, scorer internals or release credentials. These are incentive constraints, not ceremonial workflow. They change the payoff structure by making concealed burden, unsupported promotion and unauthorized action harder or less valuable (Akerlof, 1970; Jensen and Meckling, 1976; Holmstrom, 1979).

## 9.4 Mechanism objective

Let $M$ be the institutional mechanism. GoalOS seeks

$$\max_M \ \mathbb{E}[V_m(\gg)] - \lambda_R \,\mathrm{CVaR}_q(L) - \lambda_A \mathrm{AaR} - \lambda_D D_P - \lambda_B B_{\text{collusion}} \tag{66}$$

subject to :

$$\mathrm{IC} \wedge \mathrm{IR} \wedge \mathrm{EvidenceSeparation} \wedge \mathrm{Rights} \wedge \mathrm{Rollback} \wedge \mathrm{HumanAdmission}, \tag{67}$$

where IC is incentive compatibility and IR is participation rationality for legitimate actors.

The objective does not imply a perfect equilibrium can be computed. It is a design standard : truthful evidence should be rewarded, hidden burden should be priced, refusal should remain available and no actor should be both claimant

and final judge of its own authority. The governing game-theoretic rule is that truth must be easier to preserve than theatre ; diversity without independent incentives and methods is not verification.

### 9.5 Stackelberg structure

The accountable principal acts as a constitutional leader : it freezes the mission, hard constraints, proof budget and authority ceiling before candidate evaluation. Builders and providers then respond within that constitution. This resembles a Stackelberg game in which the leader commits to rules and followers optimize under them.

Let $c$ denote the frozen constitution and $s_i(c)$ player $i$'s best response. The principal chooses

$$c^\star = \arg\max_{c \in \mathscr{C}} U_P\left(c, s_1(c), \dots, s_n(c)\right) \tag{68}$$

subject to rights and governance constraints. The principal may not alter $c$ after observing protected results merely to promote a favoured candidate.

### 9.6 Adversarial challenge and minimax regret

For candidate $a$ and possible world $\omega$, let $L(a, \omega)$ be institutional loss. The maximum regret is

$$\mathrm{Regret}(a) = \max_{\omega \in \Omega} \left[ L(a, \omega) - \min_{a' \in \mathscr{A}} L(a', \omega) \right]. \tag{69}$$

GoalOS may rationally prefer a lower-mean but reversible architecture when evidence is weak or tail loss is severe. The critic should construct anti-recommendations and failure worlds rather than merely assign a pessimistic score.

### 9.7 Correlated error

A multi-agent system may appear diverse while sharing the same model family, data source, reward or ontology. Let errors $e_i$ have covariance matrix $\Sigma_e$. The effective diversification benefit depends on covariance, not agent count :

$$\mathrm{Var}\left( \sum_i w_i e_i \right) = w^\top \Sigma_e w. \tag{70}$$

A committee of identical agents is not independent verification. The system should diversify methods, data, representations and institutional incentives where the consequence justifies the cost.

### 9.8 Coalitional value and credit

A complete successor may owe its gain to a coalition of components. Shapley values can estimate marginal contribution :

$$\phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n - |S| - 1)!}{n!} \left[ v(S \cup \{i\}) - v(S) \right]. \tag{71}$$

This is useful for architecture diagnosis, cost allocation and reinvestment (Shapley, 1953). It must not be used to infer authority. A component can contribute substantial value while remaining ineligible for consequential action.

### 9.9 Foundry evolution and replicator dynamics

Let $p_i$ be the archive share of architecture lineage $i$ and $f_i$ its proof-adjusted formation fitness. A stylized replicator equation is

$$\dot{p}_i = p_i(f_i - \bar{f}) + \mu_i - \kappa_i, \tag{72}$$

where $\mu_i$ represents innovation or mutation and $\kappa_i$ constitutional rejection. GoalOS does not allow average fitness alone to collapse the archive. Quality-diversity preserves lineages that are safest, most sovereign, most transferable, lowest burden or strongest on specific failure clusters (Mouret and Clune, 2015 ; Stanley and Lehman, 2015).

### 9.10 The repeated game and Chronicle reputation

A successor participates in repeated missions. Chronicle creates a public-within-the-institution evidence state that affects future trust, proof cost and bargaining power. Let $K_t$ be validated institutional capability and $F_t$ uncontained failure :

$$K_{t+1} = (1 - \delta_K)K_t + \eta E_t^+ - \omega D_{P,t} - \chi F_t. \tag{73}$$

Because future authority depends on current evidence, a candidate has an incentive to preserve verifiable outcomes if it cannot rewrite Chronicle. The repeated game can therefore support more truthful equilibria than one-shot evaluation.

### 9.11 Provider hold-up game

Relationship-specific investment inside a proprietary platform can shift bargaining power to the provider (Williamson, 1985; Grossman and Hart, 1986; Hart and Moore, 1990). The Mission Gym, exportable evidence, portable jobs and Chronicle reduce the provider's ability to appropriate mission-specific quasi-rents.

$$\Pi_{\text{sov}} = \mathbb{E}[\text{avoided hold-up} + \text{switching option} + \text{continuity} + \text{retained learning} + \text{rights clarity}] - C_{\text{sov}}. \tag{74}$$

Ω

PART III

# Alpha, Beta, Spread and the Mission-Sovereign Institution

*From symmetric frontier access to asymmetric institutional capacity*

The economic and strategic doctrine of rented capability, underwritten opportunity and captured mission advantage.

SUCCESSOR MANIFOLD → GRADIENT → SEIZE → BOUNDED JOBS → SPECIALIST ASI → SUCCESSOR Ω

# 10 AI Beta, Mission Spread, Proven Alpha and Successor Ω

## 10.1 Four objects that must not be conflated

The doctrine can be read at three speeds.

— **One second :** Beta → Spread → SEIZE → Specialist ASI → Successor Ω.
— **Five seconds :** rent the moving frontier baseline, underwrite the distance to a superior mission architecture, prove the complete candidate, then grant only the authority supported by current evidence.
— **Thirty seconds :** discover a dynamic multi-player game involving the sovereign Gym, architecture search, real options, information geometry, protected proof, property rights, vendor hold-up, Chronicle memory, Alpha decay and repeated institutional renewal.

## 10.2 AI Beta

Let $F_t$ denote the generally rentable frontier at time $t$. For institution $i$ and mission $m$, a managerial factor decomposition is

$$Y_{i,m,t} = \beta_{i,m} F_t + \alpha_{i,m,t} + \varepsilon_{i,m,t}, \tag{75}$$

where $\beta_{i,m}$ is the institution's ability to absorb common frontier capability, $\alpha_{i,m,t}$ residual mission-specific advantage and $\varepsilon_{i,m,t}$ noise and shocks.

AI Beta includes frontier models, standard agents, common tools, public data and widely purchasable implementation. Beta can be extremely valuable while becoming symmetric. When competitors obtain the same capability and complementary assets, customer expectations and price competition can compress relative surplus.

Design Proposition 10.1 (Beta compression). If competitors obtain equivalent frontier capability, possess equivalent complements and can imitate one another at negligible cost, the common uplift may raise absolute productivity while expected relative surplus converges toward zero. Durable differentiation requires asymmetry in data, architecture, verification, distribution, rights, learning velocity or authority design.

## 10.3 Mission Advantage Spread

Let $I_m(t)$ be the incumbent and $B_m(t)$ the best external alternative. Define the reference architecture

$$R_m(t) = \operatorname*{arg\,max}_{z \in \{I_m(t), B_m(t)\}} U_m(z). \tag{76}$$

For candidate $S_m(x,t)$ on the Successor Manifold, the gross Spread is

$$\Gamma_m(x,t) = U_m(S_m(x,t)) - U_m(R_m(t)). \tag{77}$$

The Spread is an opportunity, not yet an asset. It can disappear after formation, proof, human rescue, risk, rights, dependency, maintenance and unwind are counted.

The Spread is naturally vector-valued :

$$\vec{\Gamma}_m = \begin{bmatrix} \Delta Q & -\Delta C & -\Delta T & \Delta R_v & \Delta E & \Delta S & \Delta K \end{bmatrix}^\top. \tag{78}$$

A scalar decision score may be formed only after weights, thresholds and hard constraints are frozen.

## 10.4 Mission Alpha

Mission Alpha is the residual risk-adjusted surplus :

$$\alpha_m(x,t) = U_m(S_m(x,t)) - \max\{U_m(I_m(t)), U_m(B_m(t))\} - C_m - R_m - D_m - \widehat{\epsilon}_{\text{gym}}, \tag{79}$$

where $C_m$ is complete lifecycle cost, $R_m$ risk including tail and authority exposure, $D_m$ Proof Debt and dependency burden, and $\widehat{\epsilon}_{\text{gym}}$ a reserve for Gym-to-reality basis risk.

An explicit decomposition is

$$\alpha_m = V_{\text{quality}} + V_{\text{speed}} + V_{\text{reliability}} + V_{\text{sovereignty}} + V_{\text{option}} + V_{\text{learning}} \tag{80}$$

$$- C_{\text{build}} - C_{\text{run}} - C_{\text{proof}} - C_{\text{human}} - C_{\text{integration}} - C_{\text{maintenance}} - C_{\text{unwind}} \tag{81}$$

$$- \lambda_R \, \text{CVaR}_q(L) - D_{\text{proof}} - D_{\text{dependency}} - I_{\text{irreversibility}} - \widehat{\epsilon}_{\text{gym}}. \tag{82}$$

**Alpha is not a stage**

Alpha is a measured residual state. It emerges only after underwriting, formation and comparative proof. Specialist ASI is the proven capability state ; Successor Ω is the governed institutional state that captures, preserves and compounds the resulting Alpha.

### 10.5 The Gym contribution to Alpha

The Mission Gym contributes to Alpha through at least six channels :

1. **Formation Alpha :** cheaper and faster architecture search.
2. **Proof Alpha :** reduced uncertainty and false promotion.
3. **Substitution Alpha :** ability to test new models or providers without losing the mission.
4. **Failure Alpha :** preserved negative capability and reduced repeated mistakes.
5. **Transfer Alpha :** reusable environments, verifiers and schemas across missions.
6. **Authority Alpha :** more precise delegation with lower Authority-at-Risk.

Define the Gym asset value

$$V_{\text{Gym}} = V_{\text{formation}} + V_{\text{proof}} + V_{\text{substitution}} + V_{\text{transfer}} + V_{\text{failure avoidance}} + V_{\text{authority}} - C_{\text{Gym}} - R_{\text{basis}}. \tag{83}$$

A Gym that is expensive, unrepresentative or exploitable can have negative value.

### 10.6 Proof as collateral for authority

Let $E_t$ be the current evidence state and $A_t$ the Authority Envelope. The controlling partial order is

$$A_t \preceq \phi(E_t), \tag{84}$$

where authority is ordered by consequence, scope, duration, reversibility, identities, tools, spend and external effect. Evidence does not merely support a claim. It collateralizes the authority the institution is willing to expose.

### 10.7 Dynamic Alpha

Mission Alpha decays when the frontier diffuses, competitors imitate, the mission changes, scorers expire, providers drift or maintenance grows. Let

$$\frac{d\alpha_m}{dt} = L_{\text{institution}} + G_{\text{proprietary}} + V_{\text{option}} - Q_{\text{commoditization}} - I_{\text{imitation}} - \Delta_{\text{drift}} - M_{\text{maintenance}} - B_{\text{basis}}. \tag{85}$$

A durable successor requires

$$v_{\text{learning}} + v_{\text{reproof}} > v_{\text{commoditization}} + v_{\text{imitation}} + v_{\text{drift}} + v_{\text{decay}}. \tag{86}$$

### 10.8 The successor equilibrium

In the Beta equilibrium, firms converge on similar models, prompts and common agent stacks. Productivity rises, differentiation compresses and correlated dependency grows. In the successor equilibrium, one institution creates a private state through mission architecture, sovereign Gym, rights-cleared evidence, validated jobs, Chronicle memory, distribution and governed authority.

Let firm $i$ rent frontier capability $F_t$ and choose complement stock $k_i$ at cost $c(k_i)$. Its payoff is

$$\pi_i = u_m(F_t, k_i, K_{i,t}, G_{i,t}) - p(F_t) - c(k_i) - \ell_i, \tag{87}$$

where $K_{i,t}$ is Chronicle capital, $G_{i,t}$ Gym capital and $\ell_i$ expected loss. A successor equilibrium exists when the private return from $k_i$, $K_{i,t}$ and $G_{i,t}$ exceeds formation and proof cost and cannot be instantaneously copied.

**Alpha doctrine**

Beta is rented. The Spread is underwritten. Alpha is manufactured through differentiated architecture, a sovereign Mission Gym and comparative proof. Specialist ASI establishes mission superiority. Successor Ω captures, governs and compounds the advantage while authority remains collateralized by fresh evidence.

### 10.9 The complete Alpha stack

| Layer | Institutional meaning |
|---|---|
| Business Beta | Underlying market, mandate, customer, capital and industrial exposure. |
| AI Beta | Common frontier capability and cost improvement available to competent adopters. |
| Sensing Alpha | Detecting a succession event before it is fully incorporated into institutional assumptions. |
| Gym Alpha | Owning the mission environment, proof cases, evaluator, failure map and substitution surface. |
| Spread | Gross opportunity between the strongest reference architecture and a reachable superior successor. |
| Underwriting Alpha | Selecting the right mission, trajectory, proof budget, stop-loss and reserve options. |
| Formation Alpha | Constituting a proprietary complete system rather than repeatedly renting an undifferentiated interface. |
| Proof Alpha | Reducing uncertainty, hidden burden and tail risk through protected comparison and exact release control. |
| Operating Alpha | Recurring quality, speed, cost, reliability, evidence, sovereignty or revenue advantage. |
| Compounding Alpha | Retained learning from Chronicle, validated jobs, failure maps, environments, integrations and realized outcomes. |
| Sovereignty premium | Option value from control of mission, Gym, architecture, data, evidence, memory, authority, portability and economics. |
| Alpha compression | Erosion from frontier diffusion, imitation, provider improvement, mission drift, scorer expiry, basis risk, maintenance and crowding. |
| Successor Ω | The admitted institution that captures, governs and renews Mission Alpha under fresh proof. |

## 11 SEIZE Underwriting, Bounded AGI Jobs and the Successor Foundry

### 11.1 SEIZE as the succession microcycle

SEIZE converts a capability event into one bounded promotion decision. It does not authorize production deployment.

**S — Surface the succession event.**
Identify the assumption that expired, the incumbent that depreciated and the mission newly made feasible.

**E — Evaluate the successor frontier.**
Generate retain, repair, rent, form, partner, acquire, hedge, reserve, retire and prohibit alternatives.

**I — Instantiate the succession constitution.**
Freeze the mission, principals, hard gates, rights, proof budget, stop-loss, protected evaluation and maximum authority ceiling.

**Z — Zero in on decisive proof.**
Buy the smallest evidence package capable of changing the position.

**E — Elevate one challenger.**
Freeze one complete candidate and submit it to the Fresh-Successor Gate.

Each alternative receives an anti-recommendation : the strongest reason not to pursue it, the evidence that would falsify it, and the condition under which another candidate should replace it.

### 11.2 Value of proof

Let $a \in \mathscr{D}$ be an institutional decision and $e$ a proposed evidence package. The value of proof is

$$\mathrm{VoP}(e) = \mathbb{E}_e\left[\max_{a\in\mathscr{D}} \mathbb{E}[U(a) \mid e]\right] - \max_{a\in\mathscr{D}} \mathbb{E}[U(a)] - C(e). \tag{88}$$

The institution should not fund a complete successor when a cheaper diagnostic can show that the thesis is wrong. Proof capital is the option premium on a real option (Black and Scholes, 1973 ; Merton, 1973 ; Dixit and Pindyck, 1994).

### 11.3 Bounded AGI Jobs as work-and-proof contracts

The Foundry does not manufacture a successor through one prompt. It compiles a typed portfolio of jobs. For job $j$,

$$J_j = (O_j, P_j, I_j, D_j, T_j, A_j^-, B_j, \tau_j, Y_j, E_j, V_j, K_j, R_j, C_j), \tag{89}$$

where $O_j$ is objective, $P_j$ principal, $I_j$ acting identity, $D_j$ authorized data, $T_j$ tools, $A_j^-$ prohibited actions, $B_j$ budget, $\tau_j$ duration, $Y_j$ output schema, $E_j$ proof obligation, $V_j$ verifier, $K_j$ acceptance criteria, $R_j$ rollback and $C_j$ Chronicle eligibility.

The canonical 21-job portfolio appears in Appendix C. The core categories are :

1. mission, incumbent, alternatives and rights ;
2. benchmark, oracle study and failure taxonomy ;
3. architecture, data, specialist and tool formation ;
4. verifier, security, economics, human agency and rollback ;
5. fresh evaluation, independent validation, acceptance and admission ;
6. Capital-to-Capacity and requalification.

### 11.4 The Foundry is architecture search

The Foundry may combine :

— supervised adaptation and demonstrations ;
— reinforcement learning with independently grounded rewards ;
— evolutionary search and quality-diversity archives ;
— recursive refinement and test-time search ;
— program synthesis, symbolic planning and executable world models ;
— specialist teams and learned coordination ;
— deterministic tools, formal methods and simulators ;
— redesigned human-machine roles ;
— partnership, licensing, acquisition or infrastructure formation.

Search procedures are powerful when states, actions, feedback and evaluators are sufficiently structured (Silver et al., 2016, 2017 ; Sutton and Barto, 2018 ; Vinyals et al., 2019 ; Berner et al., 2019). Open-ended search can generate stepping stones, but experience and reward remain subject to independent relevance, rights and proof (Lehman et al., 2020 ; Stanley and Lehman, 2015).

### 11.5 Quality diversity rather than one leaderboard

The Foundry should preserve lineages that are :

— safest ;

— most reversible;
— lowest burden;
— strongest on transfer;
— best calibrated;
— most private or sovereign;
— fastest;
— most resilient;
— strongest under scarce resources;
— best for one critical failure cluster.

A MAP-Elites style archive can index candidates by behavioural and institutional descriptors (Mouret and Clune, 2015). Let cell $c$ represent a niche such as high sovereignty and low cost. The archive stores

$$A(c) = \arg\max_{x \in c} f_{\text{formation}}(x) \tag{90}$$

subject to hard gates. The archive prevents premature convergence and preserves alternatives that may become valuable after a provider, price, legal or mission change.

### 11.6 Minimum-burden dominance

Let $\mathscr{P}_m$ be the set of candidates passing all mission and constitutional gates. The champion is

$$S^\star = \arg\min_{S \in \mathscr{P}_m} B_{\text{institution}}(S) \quad \text{subject to} \quad \Delta_{\text{succ}}(S, S_0 \mid m) > \theta_m. \tag{91}$$

The winner is not necessarily the largest model or highest average development score. It is the admissible architecture that creates the strongest verified mission result at the lowest justified complete burden.

### 11.7 Formation budgets and search stop conditions

Search consumes compute, data, human expertise, verifier development, legal review, integration time and opportunity duration. Let $C_{\text{search}}(k)$ be cumulative cost after generation $k$ and $\text{EVI}_{k+1}$ the expected value of another generation. Search should stop when

$$\text{EVI}_{k+1} \leqslant C_{\text{marginal},k+1} \tag{92}$$

or the opportunity half-life becomes shorter than the remaining proof path.

### 11.8 The recursive formation algorithm

**Algorithm 1** Verified Recursive Successor Foundry

**Require :** Frozen mission constitution $c$, Formation Gym $G_F$, protected proof custodian $P$, proof budget $B$

Generate structurally diverse initial manifold $\mathcal{M}_0$
Evaluate matched formation performance and hard-gate failures
Initialize quality-diversity archive $\mathcal{Q}_0$ and Negative Capability Graph $\mathcal{N}_0$
**for** $k = 0, 1, \dots, K-1$ **do**
    Estimate Mission Advantage Gradient around archive elites
    Select diverse parents from $\mathcal{Q}_k$
    Propose bounded mutations to model, data, tool, verifier, workflow, roles or authority design
    Run descendants in Formation and Adversarial Gyms
    Reject critical-error, rights, utility and governance regressions
    Update archive $\mathcal{Q}_{k+1}$ and failure graph $\mathcal{N}_{k+1}$
    **if** expected value of another generation $\leqslant$ marginal complete cost **then**
        **break**
    **end if**
**end for**
SEIZE freezes one exact champion $S^\star$ and manifest
Protected custodian evaluates $S^\star$ in Fresh-Proof Gym
Separate accountable authority accepts, rejects, repairs or limits the release

A candidate may propose experiments, generate descendants, run approved sandbox work and request promotion. It may not inspect protected cases, alter the constitution, classify its own risk, certify itself, hold release credentials, approve descendants, install itself or expand its Authority Envelope.

### 11.9 Formation evidence versus promotion evidence

Development metrics may guide search, but promotion must use separately frozen criteria. The candidate may optimize dense shaping rewards in Formation Gym while the Fresh-Proof Gate measures end-to-end mission outcomes, critical errors, complete burden, transfer, rights and governance. Reusing the same score for both creates selection leakage.

**Foundry law**

SEIZE underwrites the intelligence asset. The Mission Gym makes its work executable. Bounded AGI Jobs manufacture and test complete candidates. The protected proof plane verifies one frozen challenger. Accountable authority decides whether it becomes Successor Ω.

## 12 Specialist ASI as an Earned Comparative State

### 12.1 Mission-bounded definition

Specialist ASI is not defined by parameter count, brand, self-description or generality. It is a complete mission-specific intelligence institution that demonstrates superhuman and frontier-superior performance on a constituted mission family while remaining economically and constitutionally admissible. The definition separates performance depth, breadth and autonomy rather than collapsing them into one label (Morris et al., 2024; Genewein et al., 2026).

Definition 12.1 (Specialist ASI). A complete mission-specific intelligence institution that exceeds the incumbent and the best credible general, specialist, human, software or hybrid alternative by a preregistered positive margin on fresh, protected and representative work, after quality, critical-error risk, complete economics, human burden, rights,

sovereignty, resilience, Gym basis risk and governance are counted.

The comparison is

$$\text{Complete Specialist Successor Institution} \quad \text{versus} \quad \text{Best Credible Alternative Set}, \tag{93}$$

not model versus model.

### 12.2 Best credible alternative set

For mission $m$,

$$\mathscr{B}_m = \{\text{incumbent, frontier stack, existing specialist, human team, human–AI hybrid, software, partner, acquisition, abandonment}\}. \tag{94}$$

A candidate is mission-dominant only when

$$V_m(x_{\text{candidate}}) > \max_{b \in \mathscr{B}_m} V_m(b) + \theta_m \tag{95}$$

and every hard gate passes. The baseline may not be deliberately weakened to manufacture apparent Alpha.

### 12.3 Six superiority dimensions

**Table 11.** Specialist ASI proof dimensions.

| Dimension | Required evidence |
|---|---|
| Capability superiority | Better accepted mission outcomes, calibration, abstention and critical-error performance on fresh representative work. |
| Economic superiority | Lower complete lifecycle burden or materially greater verified value than the best alternative. |
| Reliability superiority | Better exception handling, resilience, recovery, drift control and worst-case performance. |
| Sovereignty superiority | Stronger control of strategic data, Gym, tests, evidence, portability, continuity and provider optionality. |
| Governance superiority | Clearer identity, authority, evidence, validation, acceptance, rollback, expiry, revocation and human accountability. |
| Transfer superiority | Stable contribution across a preregistered mission family, not one memorized development set. |

### 12.4 Protected proof conditions

Let $E_t$ be the current evidence state. A candidate qualifies only when

$$\Pr(\alpha_m > \theta_m \mid E_t) \geqslant \tau_m, \tag{96}$$

$$\mathrm{CVaR}_q(L_m) \leqslant L_m^{\max}, \tag{97}$$

$$\mathrm{EvidenceCoverage}_m \geqslant e_m^{\min}, \tag{98}$$

$$C_m(x) = 1, \tag{99}$$

$$\mathrm{LCB}_{1-\delta}(\bar{d}) > \theta_m, \tag{100}$$

$$\widehat{\epsilon}_{\text{gym}} \leqslant \epsilon_m^{\max}. \tag{101}$$

Here $\bar{d}$ is the mean paired utility gain on matched fresh cases and $\widehat{\epsilon}_{\text{gym}}$ the estimated Gym-to-reality basis risk.

### 12.5 Matched-pair evaluation

Where possible, the incumbent, strongest alternative and frozen challenger process the same cases. Let

$$d_j = U_m(S^\star; z_j) - U_m(R_m; z_j) \tag{102}$$

for fresh case $z_j$. The mean paired gain is

$$\bar{d} = \frac{1}{n}\sum_{j=1}^{n} d_j. \tag{103}$$

Promotion requires a positive lower confidence bound and all hard gates. The report must also show critical subgroups, worst cases, abstention, human rescue and distributional effects rather than only the mean.

### 12.6 Critical errors and false reassurance

Average accuracy can conceal tail failure. A **critical false reassurance** occurs when the system confidently accepts or recommends an outcome containing a material hidden failure. For many missions, one critical false reassurance is a hard stop.

Let $C_j \in \{0,1\}$ indicate a critical case and $F_j \in \{0,1\}$ false reassurance. Define

$$\mathrm{CFR} = \frac{\sum_j C_j F_j}{\sum_j C_j}. \tag{104}$$

The constitutional ceiling may be zero even when ordinary error tolerance is nonzero.

### 12.7 Calibration and abstention

A strong Specialist ASI knows when an answer has not been earned. Let $p_j$ be predicted confidence and $y_j$ correctness. Calibration can be measured through Brier score or expected calibration error. More importantly, the system should have a preregistered abstention policy $\sigma(o)$ :

$$\sigma(o) = \begin{cases} \text{act or recommend,} & \text{if evidence, confidence and authority pass;} \\ \text{request evidence,} & \text{if uncertainty is reducible;} \\ \text{escalate,} & \text{if professional or consequential judgment is required;} \\ \text{refuse,} & \text{if the mission or action is inadmissible.} \end{cases} \tag{105}$$

### 12.8 Transfer and mission family

A single benchmark can be memorized. The mission family $\mathscr{F}_m$ must specify variation across time, cases, users, assets, geographies, failure classes or regimes. Transfer superiority requires

$$\inf_{f \in \mathscr{F}_m^{\text{protected}}} \mathrm{LCB}\left(V_m(S^\star; f) - V_m(R_m; f)\right) > \theta_f \tag{106}$$

for critical subfamilies or an explicitly justified robust aggregate.

### 12.9 Specialist ASI is a state, not a permanent identity

A material change to model, data, tool, provider, routing, policy, verifier, environment or Authority Envelope creates a new release. The label does not persist automatically. Let release $r_t$ have proof state $E_t$. After material change $\Delta r$,

$$E_{t+1}^{\text{inherited}} = \varnothing \tag{107}$$

for the affected claims until requalification. Historical evidence remains informative but does not itself collateralize new authority.

### 12.10 The target-state boundary

No public paper, interface or generated report establishes Specialist ASI. The designation requires an exact candidate manifest, protected evaluation, complete denominator, independent review and accountable acceptance. The executable demos accompanying this paper may report a *simulation pass*; they do not establish real-world Specialist ASI.

**Specialist ASI law**

A candidate earns Specialist ASI status by winning the complete mission against the strongest credible alternatives on fresh work, not by looking intelligent, using many agents, scoring highly on a public benchmark or describing itself as superintelligent.

## 13 Successor Ω : Accountable Admission and the Authority Lattice

### 13.1 Dominance is not sovereignty

Mission dominance is an empirical verdict. Mission sovereignty is a constitutional state. Fresh proof may establish that the complete candidate wins the mission. It does not grant the right to act.

The final transition requires :

— an exact candidate and release manifest ;
— independent proof and known limitations ;
— rights and confidentiality review ;
— security and resilience evidence ;
— a scoped Authority Envelope ;
— version and expiry ;
— monitoring and incident response ;
— rollback and incumbent fallback ;
— a named accountable admitting authority.

**Mission-sovereignty law**

Fresh proof may establish dominance. Only accountable admission may grant mission authority. No candidate, producer, model or agent may validate itself, admit itself, expand its own Authority Envelope or enable its own production release.

### 13.2 Five distinct decisions

A common failure is to collapse several decisions into one. GoalOS separates :

$$\text{Underwrite} \longrightarrow \text{Execute} \longrightarrow \text{Accept} \longrightarrow \text{Admit} \longrightarrow \text{Allocate}. \tag{108}$$

**Table 12.** Five institutional decisions.

| Decision | Question | Authority created |
|---|---|---|
| Underwrite | Does the opportunity deserve proof capital ? | Only the named evidence programme. |
| Execute | May a bounded job begin under the frozen mission, tools, data and budget ? | Job-level execution permission. |
| Accept | Did the output satisfy frozen technical and stakeholder criteria ? | Acceptance for the stated purpose only. |
| Admit | May the accepted capability influence future work ? | Scoped Chronicle and mission authority. |
| Allocate | May realized or conservatively attributable value fund the next successor frontier ? | Capital-to-Capacity authority. |

The same actor need not control all five. The candidate controls none of them.

## 13.3 The Authority Lattice

Authority is multidimensional. Define

$$A_t = (s_t, j_t, d_t, i_t, \ell_t, p_t, x_t, q_t, h_t, r_t), \tag{109}$$

where $s$ is mission scope, $j$ jurisdiction, $d$ duration, $i$ permitted identities and systems, $\ell$ spend or transaction limit, $p$ permission class, $x$ external-effect class, $q$ stressed resource consumption, $h$ human approval requirement and $r$ reversibility.

The partial order $A \preceq A'$ means that $A$ grants no greater authority on every governed dimension. The evidence state maps to an admissible upper set :

$$A_t \preceq \phi(E_t) \preceq A_m^{\max}. \tag{110}$$

The mission's constitutional ceiling $A_m^{\max}$ may remain advisory even if the candidate is highly capable.

## 13.4 Authority inheritance levels

**Table 13.** Authority levels.

| Level | Permitted posture | Typical evidence |
|---|---|---|
| A0 | Observe : read, analyze and propose. | Provenance, logging and no external effect. |
| A1 | Recommend : produce ranked options. | Calibration, evidence, alternatives and uncertainty. |
| A2 | Sealed sandbox : execute with no external effect. | Frozen inputs, budgets, audit logs and repeatability. |
| A3 | Reversible bounded action. | Canary proof, limits, monitoring, stop conditions and explicit rollback. |
| A4 | Consequential action affecting money, rights, safety or material operations. | Independent validation, professional authority where required, dual control and incident response. |

A candidate can be mission-dominant at A1 while remaining ineligible for A3 or A4. Performance and authority are separate state variables.

## 13.5 Proof-carrying action

Every consequential action should carry a proof receipt containing :

— action identity, actor and accountable principal ;
— mission and Authority Envelope ;
— source and input lineage ;
— model, tool, policy, Gym and runtime versions ;
— validation status and residual uncertainty ;
— approval and escalation record ;
— expected effect, rollback and expiry ;
— hash and timestamp.

Proof becomes a condition of action rather than a retrospective report.

## 13.6 Admission record

The accountable Admission Record is

$$\mathscr{D}_{\text{admit}} = (\text{SuccessorID}, m, r, E, A, \text{expiry}, \text{monitoring}, \text{rollback}, \text{revocation}, \text{principal}). \tag{111}$$

Admission may return :

— Admit;
— Admit with limits;
— Continue proving;
— Repair and reprove;
— Retain the incumbent;
— Roll back;
— Reject;
— Quarantine or prohibit.

### 13.7 Graceful degradation

Successor Ω must specify transitions among :

$$\text{Full} \to \text{Restricted} \to \text{Recommend only} \to \text{Read only} \to \text{Incumbent fallback} \to \text{Emergency stop}. \tag{112}$$

The Authority Envelope fails closed when the successor leaves the validated distribution, loses required evidence, exceeds budgets, encounters a material incident or reaches expiry.

### 13.8 Institutional refusal

The Verified Succession Institution must be able to refuse :

— a vague or conflicted mission;
— an exploitable or unknowable evaluator;
— data without adequate rights;
— a candidate that depends on hidden authority or undisclosed external capability;
— a release without rollback or accountable ownership;
— a reinvestment claim based only on projected savings;
— a successor whose gain requires governance regression.

Refusal is not paralysis. It is a bounded decision to preserve the institution until a missing condition is proven or the thesis is abandoned.

### 13.9 The Ω state

The symbol Ω denotes the full constitutional closure of the cycle :

$$\begin{aligned}\Omega_m = \mathbb{1}[\text{MissionDominant}] \cdot \mathbb{1}[\text{FreshProof}] \cdot \mathbb{1}[\text{HumanAdmission}] \\ \cdot \mathbb{1}[\text{AuthorityIssued}] \cdot \mathbb{1}[\text{RollbackReady}] \cdot \mathbb{1}[\text{ChronicleBound}].\end{aligned} \tag{113}$$

The state is local to the mission and release. It is never universal, permanent or self-expanding.

**Successor Ω.** The complete proven institution whose mission, identity, evidence, memory, rights, operations, rollback and scoped authority are coherent enough to act without surrendering accountability.

Ω

PART IV

# The Successor to the Successor

*Recursive improvement without recursive self-authorization*

The verified recursive operator, Chronicle memory and the institution that manufactures the next admitted release.

**SUCCESSOR MANIFOLD → GRADIENT → SEIZE → BOUNDED JOBS → SPECIALIST ASI → SUCCESSOR Ω**

# 14 Recursive Institutional Improvement and Successor Recursion

## 14.1 Recurrence is not recursive improvement

A system that repeats an earlier workflow is recurrent. A system that modifies its own prompt is self-modifying. Neither condition establishes recursive institutional improvement. The stronger GoalOS definition is :

Definition 14.1 (Verified recursive institutional improvement). An institution recursively improves when a Chronicle-admitted capability measurably increases its ability to detect, constitute, prove or govern a superior successor on fresh work, without reducing proof quality, rights, resilience or governance integrity.

Let generation $G_t$ contain Mission Gym $\mathscr{G}_t$, Successor $S_t$, Chronicle $C_t$, proof infrastructure $P_t$ and authority state $A_t$. The recursive successor operator is

$$(G_{t+1}, \mathscr{G}_{t+1}, S_{t+1}, C_{t+1}, P_{t+1}, A_{t+1}) = \Re(G_t, \mathscr{G}_t, S_t, C_t, P_t, A_t; E_t). \tag{114}$$

The operator is not executed by the candidate alone. It is an institutional process containing formation, independent proof and accountable admission.

## 14.2 The central recursion

The cycle is

$$S_t \xrightarrow{\text{Chronicle + Gym episodes}} \{S_{t+1}^{(1)}, \dots, S_{t+1}^{(k)}\}, \tag{115}$$

$$\{S_{t+1}^{(i)}\} \xrightarrow{\text{SEIZE + Foundry}} \widehat{S}_{t+1}, \tag{116}$$

$$\widehat{S}_{t+1} \xrightarrow{\text{Fresh-Proof Gate}} S_{t+1}^{\star}, \tag{117}$$

$$S_{t+1}^{\star} \xrightarrow{\text{accountable admission}} S_{t+1}^{\Omega}. \tag{118}$$

The current successor may help generate descendants, identify failure clusters, propose new tools or improve the Gym. It may not certify the descendant, alter protected proof, inherit authority automatically or install the new release.

## 14.3 Recursion conditions

Promotion from generation $t$ to $t+1$ requires :

$$\mathrm{LCB}\,(V_m(S_{t+1}) - V_m(S_t)) > \theta_{\text{gen}}, \tag{119}$$

$$\Delta\text{ProofQuality} \geqslant 0, \tag{120}$$

$$\Delta\text{GovernanceIntegrity} \geqslant 0, \tag{121}$$

$$\Delta\text{RightsClarity} \geqslant 0, \tag{122}$$

$$\text{UnauthorizedActions} = 0, \tag{123}$$

$$A_{t+1} \preceq \phi(E_{t+1}). \tag{124}$$

A newer architecture is not better because it is larger, faster or more autonomous. It is better only when it produces greater verified mission value at lower justified burden while preserving mission purpose, evidence, reversibility and human sovereignty.

## 14.4 Recursive gain

Let $m_{t+1}$ be a fresh mission or materially changed environment. Define generation gain

$$\Delta_G(G_{t+1}, G_t \mid m_{t+1}) = a_Q \Delta Q + a_V \Delta V - a_C \Delta C - a_R \Delta R - a_H \Delta H + a_T \Delta T + a_K \Delta K. \tag{125}$$

Promotion requires a positive preregistered threshold. Merely using prior output does not count. The gain must be attributable to reusable environments, verifiers, failure maps, rights templates, tools, integrations or Chronicle knowledge.

### 14.5 Successor learning curves

The Foundry becomes a platform rather than bespoke consulting when later missions become cheaper, faster or safer after normalizing for difficulty. Define

$$\text{FormationGain}_n = \frac{B_{\text{institution}}(S_1) - B_{\text{institution}}(S_n)}{B_{\text{institution}}(S_1)}. \tag{126}$$

A positive value is valid only if proof and mission difficulty remain comparable. Reduced burden caused by weaker proof is not learning.

### 14.6 Recursive improvement of the Gym

The Gym itself can improve through :

— better state representations ;
— more realistic transitions ;
— stronger adversarial scenarios ;
— lower-cost replay ;
— improved hidden ground truth ;
— better evaluator calibration ;
— reduced Gym basis risk ;
— transfer cases discovered from production ;
— clearer action and authority schemas.

But the Gym can also be corrupted by the current successor. A candidate may overfit the simulator, shape data collection or suppress failures. Therefore, Gym changes receive their own version, provenance and validation. Let $\mathscr{G}_{t+1} = g(\mathscr{G}_t, E_t)$ be a proposed environment update. It is admitted only if

$$\widehat{\epsilon}_{\text{gym},t+1} < \widehat{\epsilon}_{\text{gym},t} \tag{127}$$

or it expands relevant coverage without weakening protected proof.

### 14.7 Negative Capability Graph

Chronicle preserves not only winners but also rejected architectures, reward exploits, rights failures, verifier weaknesses, unsafe authority paths and non-transferable improvements. Represent the failure memory as a graph

$$\mathscr{N}_t = (V_F, E_F), \tag{128}$$

where nodes are failure conditions and edges represent shared causes, components, missions or repairs. The graph prevents the institution from repeatedly paying to rediscover the same failure and preserves stepping stones that may become useful under a different regime.

### 14.8 The recursion firewall

Recursive improvement may expand capability, not self-authorized power. Every descendant starts with no inherited production authority for materially changed claims. Protected cases, scorer internals, release credentials, mission constitution and admission remain outside the candidate's control.

### 14.9 A bounded monotonicity proposition

Proposition 14.1 (Governance-preserving recursive improvement). *Assume each generation is evaluated on fresh representative work; proof and authority are independently controlled; every material change removes inherited authority for affected claims; and promotion requires positive lower-bound mission gain with no governance regression. Then the admitted sequence can be monotone in verified mission value without being monotone in autonomy.*

*Démonstration.* By assumption, generation $t+1$ is admitted only when its lower-bound mission value exceeds generation $t$ by the frozen threshold and all hard governance gates are non-decreasing. Authority is separately bounded by current evidence, so an increase in capability does not require an increase in autonomy. Therefore, the sequence of admitted mission-value lower bounds is monotone while the authority sequence may remain constant, contract or expand independently. The proposition does not guarantee such a candidate exists; it constrains admission when one is found. □

### 14.10 Successor half-life and recursion cadence

Let $H_S$ be the expected half-life of the admitted successor and $T_R$ the time required to generate and prove a replacement. Define

$$\text{RecursiveReadiness} = \frac{H_S}{T_R}. \tag{129}$$

Values below one imply the institution may not generate a credible next successor before the current one becomes impaired. The strategic purpose of reusable Gyms, verifiers and Chronicle is to increase this ratio.

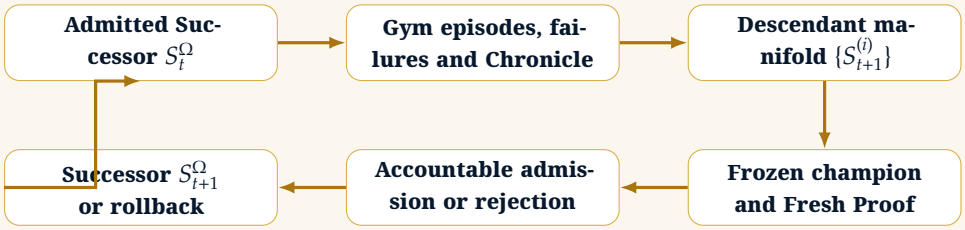



**Figure 4.** The successor-to-successor recursion. The loop contains independent proof and accountable admission; the candidate cannot close the loop itself.

**Recursive succession law**

The successor to the successor is not generated by novelty alone. Chronicle proposes; the Foundry diversifies; the Gym exposes; fresh proof selects; authority admits; reality requalifies.

## 15 The Gym as a Sovereign Corporate Asset

### 15.1 Environment ownership changes the strategic game

A corporation that rents models but owns no mission environment depends on provider claims, transient prompts and non-portable demonstrations. A corporation that owns a rights-cleared Mission Gym can place providers, models and architectures into competition against an institution-owned representation of the mission.

The Gym asset includes :

- Mission Constitution and objective history;
- state, observation and action schemas;
- rights-cleared historical episodes and synthetic scenario logic;
- protected fresh cases and case commitments;
- transition models and digital twins;
- reward vectors and hard constraints;
- verifiers, thresholds and failure taxonomies;
- role and authority simulations;
- proof receipts and Chronicle interfaces;
- transfer, canary and requalification environments.

This asset can survive the replacement of a model or vendor. It therefore changes the build–rent boundary.

## 15.2 The sovereign Gym ledger

Every Mission Gym should carry a ledger :

**Table 14.** Sovereign Gym Ledger.

| Field | Required content |
|---|---|
| Identity | Gym ID, mission, owner, custodian, version and effective date. |
| Rights | Data, model, simulator, scenario, output, verifier, reuse and confidentiality rights. |
| Scope | Valid mission family, populations, regimes, assets, jurisdictions and excluded cases. |
| State model | Hidden state, observation channels, missingness, uncertainty and transition assumptions. |
| Actions | Typed actions, masks, permissions, approvals and external-effect classes. |
| Rewards and constraints | Vector rewards, hard gates, critical failures and non-compensable constraints. |
| Evidence | Formation, adversarial, transfer, protected and canary partitions ; hashes and custody. |
| Basis risk | Known reality gaps, external-validity limits, tail omissions and monitoring. |
| Operations | Runtime, providers, compute, logging, incident response and degraded modes. |
| Lifecycle | Expiry, revalidation, impairment, rollback, retirement and successor history. |
| Economics | Cost basis, carrying cost, substitution value, proof value and transfer value. |

## 15.3 Gym sovereignty premium

Let $V_G$ be the value of the Gym asset. A sovereignty premium arises from :

$$\Pi_G = \mathbb{E}[\text{avoided hold-up} + \text{substitution option} + \text{faster proof} + \text{retained learning} + \text{transfer} + \text{authority precision}] - C_G - R_G. \tag{130}$$

The premium is positive only when the environment remains representative and rights-cleared. A poor Gym can lock the institution into the wrong reality as effectively as a poor provider can lock it into the wrong model.

## 15.4 Build, configure or rent

**Table 15.** The build–rent boundary for mission environments.

| Mode | Use when | Required control |
|---|---|---|
| Rent Beta | The task is low-differentiation, reversible, low-consequence and adequately solved by the market. | Basic vendor, privacy, quality and exit controls. |
| Configure workflow | The mission needs context and tools but limited proprietary learning or authority. | Own prompts, mappings, tests and exportable records. |
| Constitute a successor | The mission is valuable, repeated, differentiable, measurable, proofable and capable of accumulating proprietary evidence or authority. | Own the Mission Constitution, Gym, jobs, proof, Chronicle, permissions, portability and economics. |
| Preserve reserve | The opportunity may become valuable after a capability, legal, price or data catalyst. | Maintain option triggers and low-cost environment assets without operational authority. |

## 15.5 Gym portability

A portable Gym uses typed APIs, machine-readable schemas and model-independent proof receipts. Portability requires that a new candidate can be connected without rewriting the mission. Let $\mathscr{I}_{\text{env}}$ be the environment interface and $\mathscr{I}_{\text{candidate}}$ the candidate adapter. The desired separation is

$$\text{Mission logic} \perp \text{provider-specific implementation} \tag{131}$$

up to explicit adapters and known limitations.

Gymnasium provides a standard single-agent environment interface, PettingZoo supports multi-agent environment cycles, and OpenSpiel supports a broad range of game structures (Towers et al., 2024 ; Terry et al., 2021 ; Lanctot et al., 2019). GoalOS extends those interfaces with institutional identities, proof custody, authority, Chronicle and succession state.

### 15.6 Corporate balance-sheet treatment

The Gym should be treated as a strategic intangible asset, not merely test code. Internal accounting should track :

— formation cost ;
— maintenance and requalification cost ;
— missions and releases supported ;
— avoided duplicated proof work ;
— provider substitutions enabled ;
— failures prevented or detected ;
— external or internal licensing value ;
— rights and expiry ;
— residual basis risk.

This is a managerial valuation, not a claim about financial-statement recognition under a specific accounting standard. The purpose is to prevent the institution from losing the environment, rights and proof history when a project ends or a vendor changes.

### 15.7 The Gym and corporate bargaining power

A provider knows that an institution without a portable Gym faces high switching cost and uncertain revalidation. An institution with a sovereign Gym can invite providers to prove themselves under the same mission conditions. This creates a contestable successor market.

Let $q_v$ be provider quality, $p_v$ price and $s_v$ switching burden. The institution selects

$$v^\star = \arg\max_v \left[ V_m(v; \mathscr{G}_m) - p_v - s_v - R_v \right]. \tag{132}$$

Provider power falls as the Gym makes $s_v$ observable and reducible.

Design Proposition 15.1 (Gym ownership). For repeated differentiable missions, the right to define, preserve and rerun the mission environment can be more strategically durable than ownership of any one model. The Gym is the asset that lets future intelligence compete for the mission.

### 15.8 The Gym must itself be governed

The institution must protect against :

— evaluation leakage ;
— manipulated ground truth ;
— simulator drift ;
— scorer capture ;
— data-rights expiry ;
— selective scenario omission ;
— candidate influence over environment updates ;
— false confidence from synthetic coverage.

Therefore, the Gym has its own custodian, release process, change log, protected partitions, red team, requalification and retirement state.

**Gym sovereignty law**

The corporation may rent intelligence while owning the world in which that intelligence must prove itself. But the Gym becomes sovereign only when its rights, validity, evidence custody, portability and basis risk are governed as rigorously as the successor it evaluates.

# 16 Risk Architecture, Proof Debt and the Reality Gap

## 16.1 Gym basis risk

A candidate can dominate the Gym and fail the real mission. Define

$$\epsilon_{\text{gym}} = \sup_{S \in \mathscr{C}} \left| V_m^{\text{real}}(S) - V_m^{\text{gym}}(S) \right|. \tag{133}$$

The term is not directly observable before deployment. It must be estimated from held-out transfer, live shadow operation, canaries, post-outcome comparison and known omissions.

Sources include :

— incomplete state representation ;
— inaccurate transition dynamics ;
— sparse tail events ;
— unrealistic simulated humans or adversaries ;
— missing legal, social or physical constraints ;
— delayed and confounded outcomes ;
— environment and policy co-adaptation ;
— distribution shift and strategic response.

A conservative promotion statistic is

$$\text{LCB}\left(\Delta_{\text{gym}} - \widehat{\epsilon}_{\text{gym}} - D_P\right) > \theta_m. \tag{134}$$

## 16.2 Reward hacking and specification gaming

Reward grounding requires that :

— task generators may propose tasks and rewards but may not authorize them ;
— the answer or value must be independently knowable or auditable ;
— the verifier is tested for false acceptance and gaming ;
— relevance, rights and blind-set separation are confirmed ;
— accepted experience retains provenance and scope ;
— reward updates do not rewrite past promotion criteria.

Open-ended search can produce useful stepping stones, but it can also generate tasks that are trivial, impossible or strategically irrelevant. The institution must distinguish open-ended proposal from open-ended authority.

## 16.3 Proof Debt

Proof Debt is unresolved assumption or missing evidence supporting current valuation, use or authority. Let claim $k$ have consequence weight $w_k$, uncertainty $u_k$ and current reliance $r_k$. Define

$$D_P = \sum_k w_k u_k r_k. \tag{135}$$

Proof Debt can grow even while observed performance remains strong if the mission changes, a provider drifts, rights expire or the Gym becomes less representative.

Proof Coverage is

$$\text{ProofCoverage} = \frac{\text{independently supported capability and controls}}{\text{authority, valuation and reuse granted}}. \tag{136}$$

A ratio below one requires reduced authority, additional evidence or both.

### 16.4 Authority-at-Risk

Define Authority-at-Risk as consequential authority exposed to unproven or expired capability :

$$\text{AaR} = \sum_j w_j A_j \mathbb{1}[\text{ProofCoverage}_j < 1 \ \vee \ \text{Expired}_j = 1]. \tag{137}$$

The objective is not zero autonomy ; it is zero unjustified authority.

### 16.5 Threat model

The Successor Complex faces threats from :

— prompt injection and malicious data ;
— tool or credential escalation ;
— provider compromise or silent model change ;
— data poisoning and synthetic-label corruption ;
— verifier capture or collusion ;
— evaluation leakage ;
— hidden external calls ;
— unauthorized action paths ;
— corrupted Chronicle admission ;
— supply-chain and infrastructure failure ;
— strategic counterparties exploiting known policy patterns.

The architecture should align with lifecycle risk-management principles such as the NIST AI RMF and Generative AI Profile while adding succession-specific proof, admission and authority objects (National Institute of Standards and Technology, 2023, 2024).

### 16.6 Twelve canonical failure modes

| Failure | What happens | GoalOS control |
|---|---|---|
| Reward hacking | Candidate exploits the score rather than the mission. | Independent grounding, hard constraints, adversarial Gym and post-outcome checks. |
| Environment overfit | Candidate memorizes formation cases. | Hidden Fresh-Proof Gym, transfer families and exact release freeze. |
| Evaluation leakage | Protected cases enter training. | Independent custody, commitments, access logs and role separation. |
| Simulator gap | Transition dynamics differ from reality. | Shadow operation, canary, basis-risk reserve and monitoring. |
| Correlated verifier failure | Producer and verifier share the same blind spot. | Method, model, data and incentive diversity ; human adjudication. |
| Hidden human rescue | Apparent autonomy relies on unreported intervention. | Complete intervention logging and human-burden accounting. |
| Authority laundering | A high score becomes permission to act. | Separate admission and Authority Lattice. |
| Action-space escape | Candidate reaches unapproved tools or systems. | External least-privilege gateway and typed tool calls. |
| False completion | Deadline expiry is recorded as success. | Termination/truncation separation. |
| Distribution shift | Production leaves validated conditions. | Drift detection, fail-closed modes and requalification. |

| Failure | What happens | GoalOS control |
|---|---|---|
| Environment capture | Candidate or provider influences its own test. | Independent Gym ownership, change control and audit. |
| False Chronicle admission | Unsafe or stale capability shapes future work. | Admission criteria, expiry, challengeability and revocation. |

### 16.7 Human oversight and professional authority

Human oversight is only meaningful when the reviewer has :

— adequate time and information ;
— authority to refuse ;
— domain competence or access to it ;
— a clear evidence record ;
— independence from the candidate's promotion incentives ;
— a practicable rollback route.

A human click inside a high-pressure workflow is not necessarily accountable admission. The system must measure reviewer burden, disagreement, overrides and rescue.

### 16.8 Control despite stronger components

The institution should remain safe even if one component is treated as untrusted. Control protocols, restricted tools, audit, monitoring and independent checks can provide useful bounds even when capability exceeds the verifier on some dimensions (Greenblatt et al., 2024 ; Shavit et al., 2023 ; Shevlane et al., 2023 ; Phuong et al., 2024).

### 16.9 Tail loss

Tail loss includes :

— unauthorized financial or legal action ;
— security compromise or data exposure ;
— harmful physical action ;
— discriminatory or unlawful decision ;
— systemic correlated error ;
— irreversible dependency or institutional capture ;
— false admission into Chronicle ;
— capital allocation based on fabricated value.

**Risk law**

A correct secular thesis can still produce institutional ruin through bad sizing, weak proof, concentrated dependency, path dependence, Gym basis risk or uncontrolled authority. GoalOS governs the path, not merely the direction.

PART V

# Formal Assurance, Evidence and the Executable Institution

*A successor is promoted by a proof procedure, not by a score, narrative or model reputation*

The ceiling edition converts constitutional principles into conditional formal guarantees, promotion statistics, basis-risk reserves and explicit invalidation conditions.

SUCCESSOR MANIFOLD → GRADIENT → SEIZE → BOUNDED JOBS → SPECIALIST ASI → SUCCESSOR Ω

# 17 Formal Assurance : Axioms, Promotion Statistics and Robust Admission

## 17.1 Purpose and epistemic boundary

This section strengthens the design from a collection of rules into a conditional assurance system. The results below are not universal laws of organizations, intelligence or markets. They are theorems and propositions *inside a declared GoalOS constitution* : if the mission, candidate set, evidence relation, verifier, hard gates and authority map satisfy the stated assumptions, then the corresponding institutional conclusion follows. If the assumptions fail, the guarantee fails with them.

**Formal-assurance rule**

No theorem can rescue a misconstituted mission, a captured verifier, leaked protected cases, an omitted tail risk or a falsely declared authority boundary. Formal assurance begins by making those assumptions explicit, testable and revocable.

## 17.2 Eight constitutional axioms

Let $m$ denote a constituted mission; $x$ an exact complete candidate release; $\mathscr{B}_m$ the best credible alternative set; $E_t$ the current evidence state; $A_t$ the active Authority Envelope; and $\mathscr{C}_m(x)$ the conjunction of non-negotiable mission gates.

| Axiom | Constitutional statement |
|---|---|
| A1 Mission identity | The objective, beneficiary, incumbent, alternative set, evaluation distribution, critical failures and authority ceiling are frozen before protected evaluation. |
| A2 Release identity | A proof applies only to the exact model, data, tool, workflow, verifier, policy, runtime, provider and authority configuration recorded in the Candidate Successor Manifest. |
| A3 Independent custody | Protected cases, hidden labels, scorer internals and release credentials remain outside candidate formation and outside the candidate's own control. |
| A4 Hard-gate non-compensation | No amount of average reward, speed, revenue or low-severity success compensates for a prohibited catastrophic failure or a failed constitutional gate. |
| A5 Complete denominator | Mission value counts construction, operation, proof, human rescue, integration, security, rights, dependency, requalification, rollback and unwind. |
| A6 Authority separability | Capability, acceptance, Chronicle admission and authority are distinct state variables controlled by separately attributable decisions. |
| A7 Freshness | Material change to the mission, release, environment, verifier or Authority Envelope expires the affected proof claims. |
| A8 Revocability | Every admitted successor has a known-good fallback, expiry, impairment condition and revocation route. |

## 17.3 Exact release identity and proof validity

Define the release identity as the cryptographic commitment

$$\mathsf{RID}(x) = \mathrm{Hash}(m,\ M, D, H, T, V, W, R, I, E, C, A,\ \text{runtime},\ \text{provider}). \tag{138}$$

A proof record $P$ is valid for release $x$ only when

$$\mathsf{ValidProof}(P, x) = \mathbb{1}[\mathsf{RID}(P) = \mathsf{RID}(x)]\,\mathbb{1}[\mathsf{Custody}(P)]\mathbb{1}[\mathsf{Fresh}(P)]\mathbb{1}[\mathsf{Reproducible}(P)]. \tag{139}$$

This prevents a later release from borrowing the identity of an earlier release while changing a component that can alter the mission result.

Proposition 17.1 (No proof by family resemblance). *If* $\mathsf{RID}(x') \neq \mathsf{RID}(x)$ *for a mission-relevant component, a proof of* $x$ *does not establish the corresponding claim for* $x'$ *unless a declared invariance argument is independently validated.*

*Démonstration.* A mission claim is conditional on the release that generated the observed outcomes. A mission-relevant difference can change the transition distribution, evidence quality, action set, failure modes or authority

exposure. Without a validated invariance argument, the likelihood of the observed evidence under $x'$ is not identified by the evidence generated under $x$. The inherited claim is therefore unsupported. □

### 17.4 Existence of a minimum-burden champion

Let $\mathscr{P}_m$ be a nonempty finite set of candidates that satisfy every hard gate and exceed the preregistered mission-dominance threshold. Let $B_{\text{inst}}(x)$ be a real-valued Total Institutional Burden measure.

Proposition 17.2 (Finite champion existence). *If $\mathscr{P}_m$ is finite and nonempty, then at least one minimum-burden mission-dominant champion exists :*

$$x^\star \in \arg\min_{x \in \mathscr{P}_m} B_{\text{inst}}(x). \tag{140}$$

*Démonstration.* A real-valued function on a finite nonempty set attains a minimum. The proposition does not imply uniqueness ; the institution may preserve tied candidates as separate niches in the quality-diversity archive. □

Corollary 17.1 (Complexity is not an objective). *A more complex architecture is selected only when its additional components reduce the complete burden or increase verified mission value enough to survive the dominance and hard-gate conditions.*

### 17.5 Authority monotonicity and the constitutional ceiling

Let evidence states be partially ordered by $E_1 \preceq_E E_2$ when every independently supported claim in $E_1$ remains valid in $E_2$ and $E_2$ contains no new material contradiction. Let $A_{\max}(m)$ be the mission's constitutional authority ceiling.

Definition 17.1 (Evidence-collateralized authority map). An authority map $\phi$ is admissible when

$$E_1 \preceq_E E_2 \Longrightarrow \phi(E_1) \preceq_A \phi(E_2), \qquad \phi(E) \preceq_A A_{\max}(m). \tag{141}$$

Proposition 17.3 (Authority cannot outrun evidence). *Under an admissible authority map, evidence improvement may make a larger Authority Envelope eligible, but no evidence state can exceed the mission's constitutional ceiling and no authority change occurs without a separate accountable admission action.*

*Démonstration.* Monotonicity provides eligibility ordering ; the ceiling bounds the codomain. Axiom A6 makes the authority state a separate institutional variable rather than an automatic function call. Therefore $\phi(E)$ defines an upper bound, not a self-executing grant. □

### 17.6 The robust promotion inequality

Let $d_j$ be the preregistered paired mission-utility difference between challenger and incumbent on protected case $j$, with sample mean $\bar{d}$. Let $\widehat{\epsilon}_{\text{gym}}$ be a conservative reserve for Gym-to-reality basis risk, $D_P$ unresolved Proof Debt expressed in the same mission-utility unit, and $\theta_m$ the required superiority margin.

Definition 17.2 (Robust promotion margin).

$$\mathscr{M}_{\text{robust}} = \text{LCB}_{1-\alpha}(\bar{d}) - \widehat{\epsilon}_{\text{gym}} - D_P - \theta_m. \tag{142}$$

Proposition 17.4 (Robust mission-dominance eligibility). *A frozen challenger is statistically eligible for mission-dominance review only if*

$$\mathscr{M}_{\text{robust}} > 0 \tag{143}$$

*and every non-compensable hard gate passes.*

The inequality is intentionally conservative. It subtracts the estimated simulation-to-reality gap and unresolved evidence burden rather than treating a protected benchmark margin as directly transferable to production.

### 17.7 Intersection–union promotion testing

Specialist ASI is not established by one aggregate score. Let $H_k$ be the null hypothesis that superiority dimension or gate $k$ fails. The candidate is promoted only when all required nulls are rejected or all required acceptance intervals clear their thresholds :

$$\mathrm{Promote}(x) = \bigwedge_{k=1}^{K} \mathrm{Pass}_k(x). \tag{144}$$

This is an intersection–union logic. It avoids the common error of averaging capability, safety, rights and governance into a single number that can conceal a critical failure.

For soft metrics, confidence intervals or posterior credible intervals may be used. For hard gates—critical false reassurance, unauthorized action, rights violation, missing rollback, protected-set leakage—the admissible count may be exactly zero or another mission-specific ceiling fixed in advance.

### 17.8 Sequential proof without optional-stopping theatre

SEIZE may purchase evidence in tranches. Let $\mathscr{F}_t$ denote the evidence history after tranche $t$. A sequential programme is admissible only when its stopping rule, maximum sample, alpha spending or Bayesian decision rule is frozen before protected results are inspected.

A simple alpha-spending design requires

$$\sum_{t=1}^{T} \alpha_t \leqslant \alpha_{\mathrm{total}}. \tag{145}$$

A Bayesian design may stop when the posterior probability of robust positive mission gain exceeds $\tau_m$, provided the prior, likelihood family, utility and stop-loss are declared before the protected run. Repeatedly examining the same hidden cases until a pass appears invalidates freshness.

### 17.9 Power and minimum protected sample

For a paired comparison with standard deviation $\sigma_d$, target true gain $\delta$, required threshold $\theta_m$, type-I error $\alpha$ and power $1-\beta$, a planning approximation is

$$n \geqslant \frac{(z_{1-\alpha} + z_{1-\beta})^2 \sigma_d^2}{(\delta - \theta_m - \widehat{\epsilon}_{\mathrm{gym}} - D_P)^2}. \tag{146}$$

If the denominator is nonpositive, no finite sample can establish the desired claim under the current architecture and reserve assumptions. The institution must improve the candidate, narrow the claim, reduce basis risk, lower Proof Debt or reject the mandate.

### 17.10 Correlated verifier error

Using several model-based judges does not create independent verification when they share training data, architecture, provider, prompts or failure modes. Let $q$ verifiers have pairwise error correlation $\rho_v$. A stylized effective verifier count is

$$q_{\mathrm{eff}} \approx \frac{q}{1 + (q-1)\rho_v}. \tag{147}$$

When $\rho_v \to 1$, additional judges add little independent evidence. GoalOS therefore values methodological diversity : deterministic invariants, independent data, external experts, alternative model families, replay and direct outcome measurement.

**Verifier law**

Diversity of labels is not diversity of evidence. A verification coalition is strong only to the extent that its methods, incentives, information and failure modes are independently informative.

### 17.11 Recursive improvement admissibility

Let generation $g$ contain Gym $\mathcal{G}_g$, successor $S_g$, Chronicle $C_g$, proof plane $P_g$ and authority state $A_g$. A descendant $S_{g+1}$ is recursively admissible only when

$$\mathrm{LCB}\Big[\Delta(S_{g+1}, S_g \mid m_{g+1})\Big] > \theta_{g+1}, \tag{148}$$

$$\mathscr{C}_{m_{g+1}}(S_{g+1}) = 1, \tag{149}$$

$$\mathsf{ProofQuality}_{g+1} \geqslant \mathsf{ProofQuality}_g, \tag{150}$$

$$\mathsf{GovernanceIntegrity}_{g+1} \geqslant \mathsf{GovernanceIntegrity}_g. \tag{151}$$

Proposition 17.5 (No recursive authority inheritance). *Even when a descendant strictly improves fresh mission value, the predecessor's Authority Envelope is not inherited automatically.*

*Démonstration.* The descendant has a distinct release identity and therefore a distinct failure surface. Axiom A7 expires inherited claims affected by the change. Axiom A6 separates capability from authority. Fresh proof may establish eligibility ; only a new accountable admission can reconstitute authority. □

### 17.12 Formal invalidation conditions

A promotion result is invalidated when any of the following is later established :

— protected data, hidden labels or scorer details entered formation ;
— the release tested differs materially from the release submitted for admission ;
— a critical cost, human intervention, external capability or failure was omitted ;
— the evaluator was captured, exploitable or materially miscalibrated ;
— post-selection weights or thresholds were chosen after observing protected outcomes ;
— a rights, security, privacy, professional or authority gate was falsely declared ;
— the Gym-to-reality basis-risk reserve was demonstrably inadequate ;
— the claimed result cannot be reproduced from the signed evidence package.

**Assurance conclusion**

The formal ceiling is not a larger equation. It is a chain in which exact release identity, independent custody, matched fresh work, robust statistical margin, hard-gate intersection, complete denominator, separate admission and revocation all remain simultaneously true.

## 18 One Complete Synthetic Succession Cycle : Aurelia Invoice Integrity

### 18.1 Purpose and claim boundary

The Aurelia cycle is a complete synthetic reference implementation of the GoalOS progression (Boucher, 2026a). It uses generated real-estate maintenance invoice data with hidden ground truth to demonstrate real end-to-end succession mechanics. It is not a paid mandate, customer result, audited saving, production authority or proof that a real-world Specialist ASI exists.

The synthetic corporation is **Aurelia Urban Properties**. The mission is :

**Constituted mission**

Reconcile maintenance invoices against work orders, purchase orders, contracts, vendor records, payment history and completion evidence ; classify every invoice as pay-ready, hold or escalate ; and permit only reversible payment holds below C$25,000.

The incumbent is a human accounts-payable workflow using spreadsheets, ERP reports and manual approvals. The critical prohibited failure is clearing a material duplicate, ghost-vendor, wrong-building or other critical anomaly for payment.

### 18.2 Synthetic sovereign Gym

The Gym contains 4,300 synthetic cases partitioned into formation, protected proof, canary and transfer environments. It includes deterministic hidden ground truth, invoice and contract evidence, normal and adversarial anomalies, protected SHA-256 case commitments, exact termination and authority conditions, and independent proof custody.

The anomaly family includes :

— duplicate invoice ;
— price overage ;
— unauthorized work ;
— missing completion evidence ;
— wrong building ;
— split invoice ;
— ghost vendor ;
— tax mismatch ;
— out-of-contract service ;
— novel bundled anomaly.

### 18.3 The multiplayer institution

The cycle separates producer, adversarial critic, independent verifier, governor, human principal, vendor/adversary and incumbent operator. The producer generates the evidence-grounded decision, the critic searches for hidden anomaly, the verifier controls hidden labels and acceptance, the governor enforces A3 limits, and the synthetic finance committee admits or rejects the release.

### 18.4 Successor Manifold

Five complete alternatives are compared :

1. Incumbent AP plus spreadsheet workflow ;
2. deterministic rules repair ;
3. general AI copilot ;
4. specialist intelligence cell ;
5. sovereign hybrid successor.

The Mission Advantage Gradient identifies verifier independence, evidence linking, novel bundled-invoice detection and lower human-review burden as the highest-value local directions.

### 18.5 SEIZE and proof capital

SEIZE returns **Proceed to bounded proof** with an illustrative proof-capital ceiling of C$175,000 and formation ceiling of C$450,000. The cycle stops on any critical false reassurance, unauthorized action or negative complete value.

### 18.6 Recursive Foundry

The Foundry runs seven recursive generations, maintaining six quality-diversity niches and recording nine rejected lineages. The champion’s formation utility increases from 77.46 in generation one to 80.96 in generation seven.

### 18.7 Protected fresh proof

The frozen generation-seven Sovereign Hybrid is evaluated on 1,200 protected cases against the incumbent and Specialist Intelligence Cell.

The paired mean mission gain is 0.2005 and the 95% lower confidence bound is 0.2002. Evidence coverage is 98.85%, with zero critical false reassurance and zero unauthorized actions. Inside the synthetic environment, the candidate

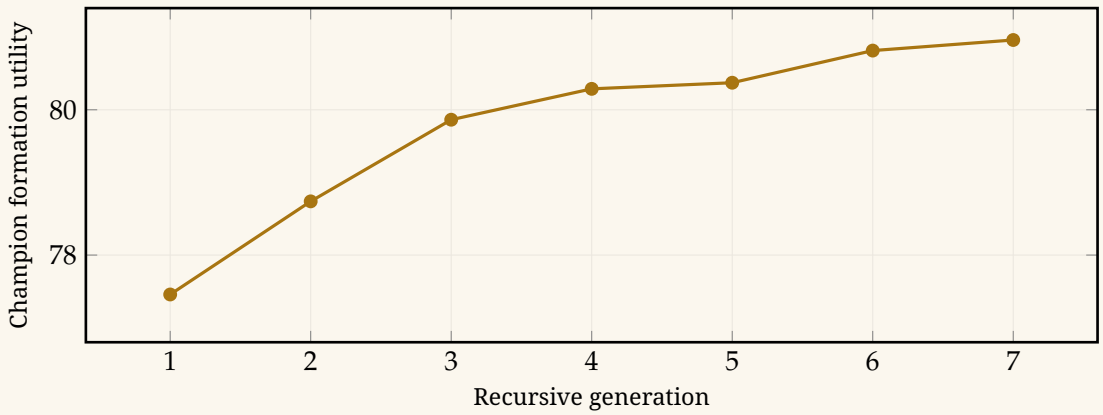



**FIGURE 5.** Synthetic recursive formation : champion utility across seven generations. Formation performance does not itself grant Specialist ASI status or authority.

**TABLE 18.** Protected synthetic fresh-proof result.

| Metric | Incumbent | Specialist cell | Frozen challenger |
|---|---|---|---|
| Decision accuracy | 35.25% | 48.08% | 83.17% |
| Material-risk recall | 100.00% | 100.00% | 100.00% |
| Critical false reassurance | 0 | 0 | 0 |
| Evidence coverage | 58.45% | 85.45% | 98.85% |
| Median cycle time | 26.64 min | 11.92 min | 3.76 min |
| Mean human review | 17.83 min | 10.44 min | 3.06 min |
| Mean cost per case | C$45.29 | C$21.65 | C$9.67 |
| Governance score | 67.80% | 78.80% | 98.14% |
| Transfer score | 69.09% | 82.49% | 99.47% |

receives the verdict :

**Synthetic proof state**

Mission-Dominant Specialist ASI supported in the synthetic reference environment. The designation does not transfer to real customer operations without rights-cleared data, independent custody, production-representative proof and accountable admission.

## 18.8 Accountable demonstration admission

The candidate is separately admitted at A3 for reversible payment holds up to C$25,000. It may read approved records, recommend pay-ready, hold or escalate, request missing evidence and create a proof receipt. It may not release payment, change vendor master data, modify contracts or expand its authority.

## 18.9 Canary and synthetic realized value

The admitted demonstration processes 400 canary cases.

**Table 19.** Synthetic canary outcome.

| Measure | Result |
|---|---|
| Decision accuracy | 77.00% |
| Evidence coverage | 98.81% |
| Median cycle time | 4.02 minutes |
| Mean human review | 3.31 minutes |
| Prevented invalid amount | C$568,886.86 |
| False-hold cost | C$7,565.00 |
| Processing and review cost | C$3,992.47 |
| Synthetic net realized value | C$557,329.39 |
| Critical false reassurance | 0 |
| Unauthorized actions | 0 |

The value is generated from known synthetic ground truth. It is not a customer saving, audited return or investment-performance statement.

## 18.10 Evidence-backed rejection

The General AI Copilot is rejected for independent admission. It produces 34.23% decision accuracy, 63.28% evidence coverage, a governance score of 58.20% and a 100% false-hold rate in the formation set. The rejection prevents the institution from granting automatic-hold or payment-release authority to an architecture without adequate verifier independence and evidence custody.

## 18.11 Chronicle transfer

Chronicle preserves the evidence graph schema, duplicate-memory index, contract ontology, verifier harness, authority templates and failure records. On a new vendor contract-compliance and bundled-invoice mission, the Chronicle-assisted successor produces :

— 25.2 percentage points of decision-accuracy gain ;
— 30.39 percentage points of transfer-score gain ;
— 31% lower formation cost ;
— positive fresh-transfer utility gain of 25.76.

This is the reference cycle's evidence that memory improved a later mission rather than merely storing previous output.

### 18.12 Impairment, rollback and the successor to the successor

A simulated provider update increases speed but reduces evidence coverage, calibration and governance. On 1,200 transfer cases, the updated release has a matched-gain lower bound of −0.2480 and fails evidence, human-burden, value and governance gates. It is impaired and rolled back.

A Chronicle-informed verifier-hardened descendant is then formed and tested on 500 fresh transfer cases. It achieves :

— 82.2% decision accuracy versus 72.6% for the predecessor ;
— 99.05% evidence coverage ;
— 99.61% transfer score ;
— mean paired gain 0.1001 ;
— 95% lower bound 0.0744 ;
— zero critical false reassurance ;
— zero unauthorized actions.

The next successor is admitted with the same A3 ceiling rather than receiving automatic authority expansion.

### 18.13 Completion state

The synthetic cycle demonstrates :

— one protected mission-dominance result ;
— one accountable demonstration admission ;
— one evidence-backed rejection ;
— one positive synthetic canary-value result ;
— one fresh-transfer gain ;
— one impairment, rollback and reproof ;
— one superior successor-to-successor release ;
— Chronicle-assisted later performance ;
— zero material unauthorized actions.

It still requires several paid mandates, rights-cleared customer data, independently calculated real production value, a repeat or expansion customer and real operational admission before the empirical completion gate is satisfied.

**Synthetic-cycle law**

The reference cycle proves that the institutional mechanics can be executed together. It does not convert synthetic proof into real authority. The next legitimate step is to replace synthetic Aurelia data with one rights-cleared customer mission without weakening any gate.

## 19 Synthetic Evidence Programme and Sensitivity Analysis

### 19.1 Why synthetic evidence is useful—and insufficient

Synthetic data can exercise the complete succession machinery before a rights-cleared customer mission is available. It can test partitioning, hidden ground truth, candidate formation, protected proof, authority masking, proof receipts, canary operation, impairment, rollback and successor-to-successor requalification. It cannot establish that the same utility function, data-generating process, human behaviour, operational cost or tail risk holds in production.

The v3.0 evidence programme therefore treats synthetic evidence as a signed *mechanism demonstration*. It supports claims about reproducibility and internal consistency of the reference cycle, not customer value, universal capability or production authority.

The results in this section are generated from the Aurelia reference environment. They are unaudited, non-customer, non-investment and non-production results. Their purpose is to expose the proof chain and its sensitivity, not to simulate away the need for real protected work.

## 19.2 Dataset partition and cryptographic commitments

The reference environment contains 4,300 generated cases. Formation receives visible labels ; protected proof, canary and transfer labels are separately held. The release includes SHA-256 commitments to the protected files before the published outcome record.

**Table 20.** Aurelia synthetic evidence partitions.

| Partition | Cases | Candidate visibility | Institutional purpose |
|---|---|---|---|
| Formation | 2,200 | Ground truth visible to formation tooling | Architecture search, debugging, quality diversity and failure repair. |
| Protected proof | 1,200 | Inputs visible ; labels and scorer protected | Matched challenger-versus-incumbent mission-dominance test. |
| Canary | 400 | Bounded operational inputs ; hidden outcome record | Reversible A3 demonstration and complete-denominator value measurement. |
| Transfer | 500 | New anomaly mix and higher novelty | Chronicle transfer, provider-change impairment and next-successor proof. |

The protected commitments are part of the Research Companion and Mission Pack. A production institution should commit not only the case file but the scorer version, threshold configuration, candidate manifest, environment image and permissible human-intervention policy.

## 19.3 Formation frontier on the Successor Manifold

The candidate set illustrates why the complete successor must be compared rather than the model alone. The deterministic repair is inexpensive and highly auditable but weak on novel bundled anomalies. The general AI copilot is fast but has low evidence custody and verifier independence. The specialist cell improves domain reasoning but retains substantial human burden. The sovereign hybrid combines evidence graph, deterministic controls, specialist intelligence, critic, verifier, governor, Chronicle and portability.

The frontier demonstrates a central GoalOS principle : the winning point is not merely the most accurate policy. It is the complete architecture whose mission gain survives evidence, governance, cost, human burden and authority constraints.

## 19.4 Recursive Foundry evidence

Seven bounded generations increase champion formation utility from 77.46 to 80.96 while preserving multiple archive niches and rejecting lineages that regress below governance, utility or critical-error floors.

The modest slope is institutionally preferable to a dramatic gain purchased through verifier collapse, hidden human rescue or authority expansion. Recursive improvement is accepted only when the successor-to-successor comparison passes fresh proof without governance regression.

## 19.5 Complete fresh-proof scorecard

The protected comparison reports not only decision accuracy but evidence, governance, transfer, time and human burden.

The frozen challenger achieves 83.17% decision accuracy, 98.85% evidence coverage, 98.14% governance score, 99.47% transfer score, median cycle time of 3.76 minutes and mean human review of 3.06 minutes. It records zero critical false reassurance and zero unauthorized actions in the protected synthetic proof.

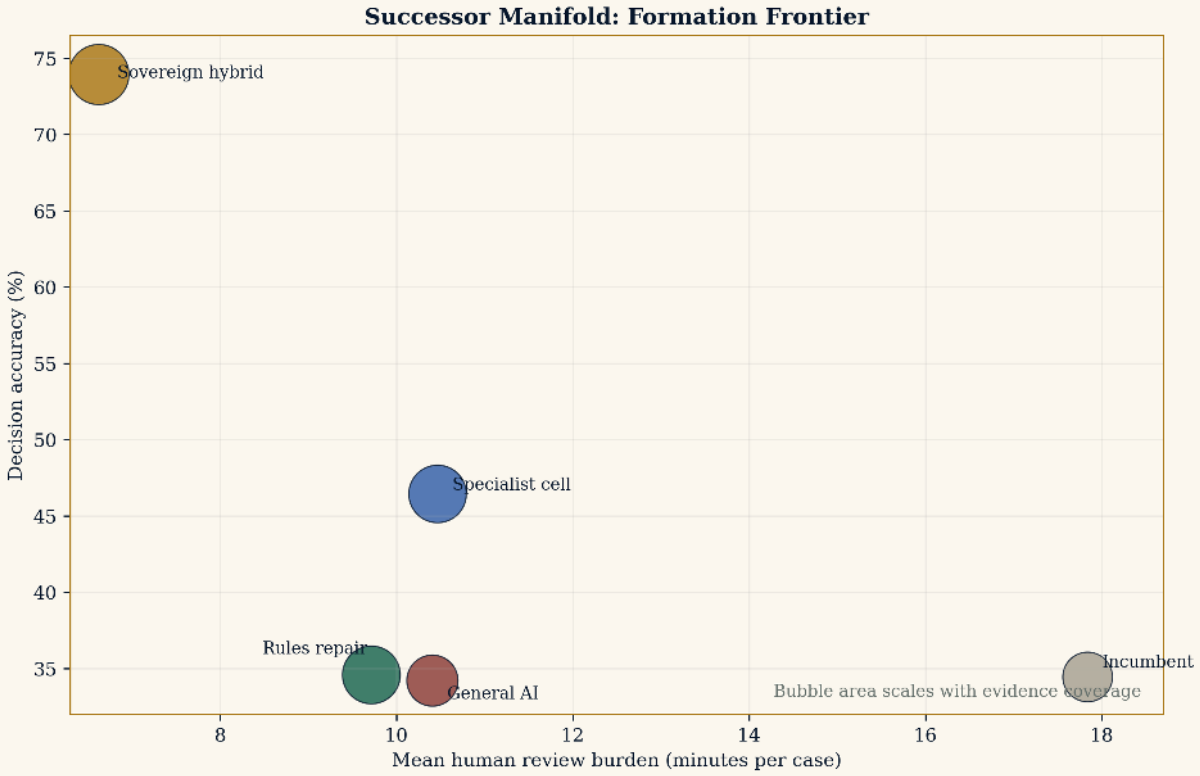



**FIGURE 6.** Synthetic formation frontier. Horizontal position is mean human-review burden; vertical position is decision accuracy; bubble area scales with evidence coverage. The plot is descriptive of the reference environment, not a universal ranking of architecture classes.

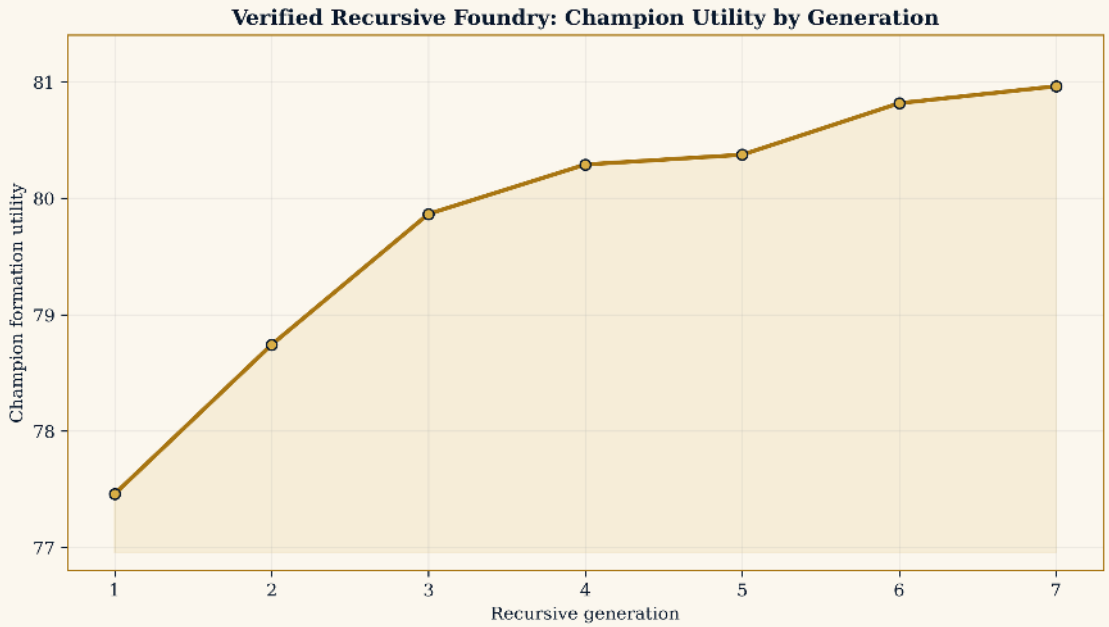



**FIGURE 7.** Champion formation utility across seven synthetic recursive generations. Formation utility is a search signal; it does not certify Specialist ASI or grant authority.

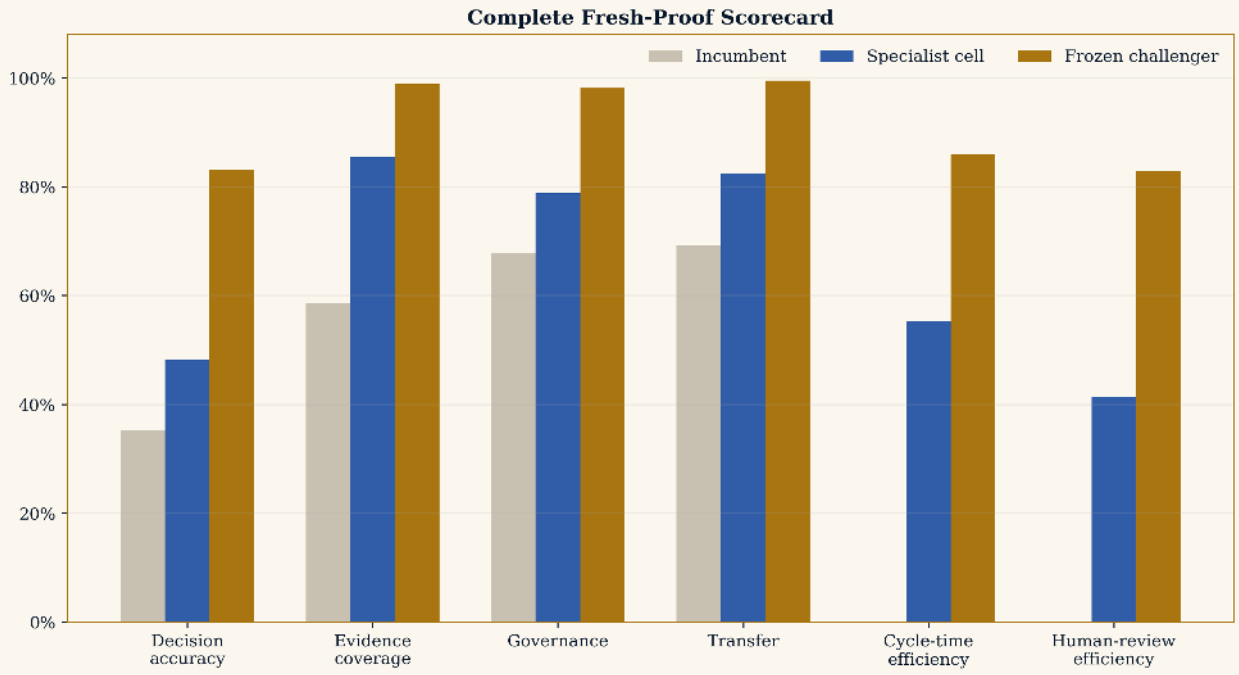



**FIGURE 8.** Protected fresh-proof scorecard for the incumbent, specialist cell and frozen generation-seven challenger. Efficiency dimensions are normalized against the incumbent.

## 19.6 Bootstrap assurance of paired mission gain

The protected proof uses a matched case-level utility difference. Version 3.0 reproduces the exact challenger and incumbent receipts and performs 20,000 nonparametric bootstrap resamples of the 1,200 paired gains. The 95% bootstrap interval is approximately $[0.20020, 0.20078]$; all bootstrap means exceed the preregistered threshold of 0.14.

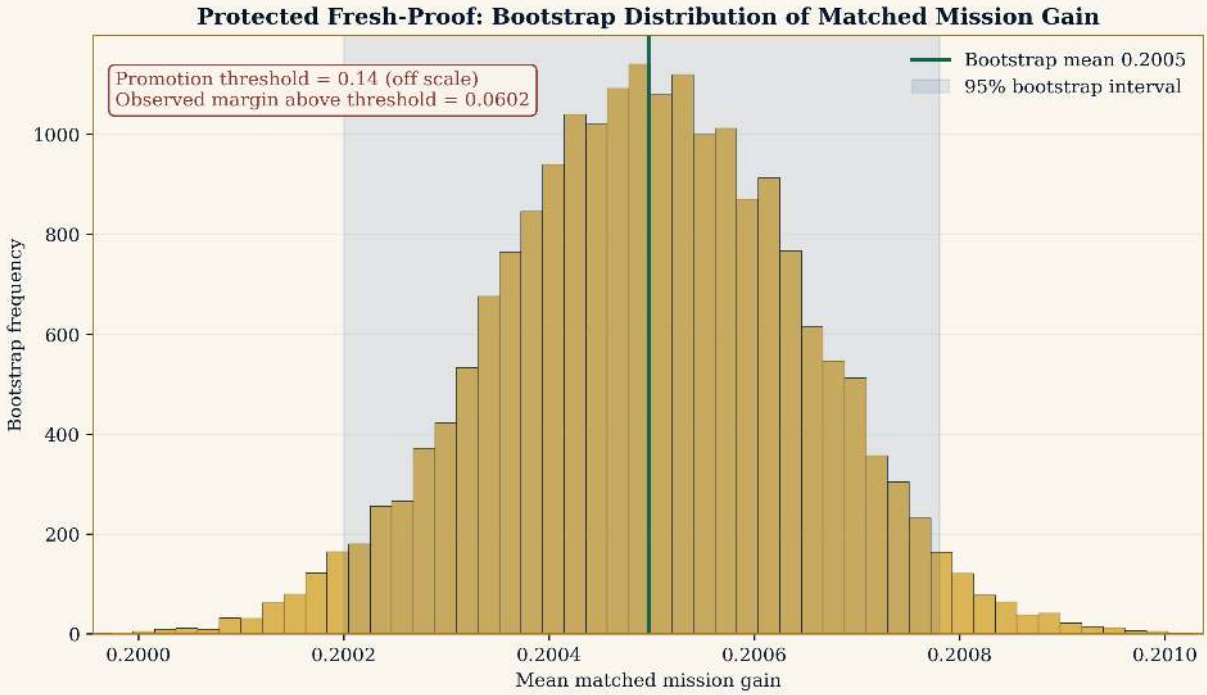



**FIGURE 9.** Bootstrap distribution of mean paired mission gain. The threshold is outside the zoomed plotting range; the observed lower-bound margin above threshold is approximately 0.0602 before basis-risk and Proof-Debt reserves.

The narrow interval is a property of this deterministic synthetic environment and utility construction. It must not be interpreted as evidence that real operational uncertainty will be equally small.

## 19.7 Gym-to-reality basis-risk sensitivity

The robust promotion margin is

$$\mathscr{M}_{\text{robust}} = 0.200208 - \theta_m - \widehat{\epsilon}_{\text{gym}} - D_P. \tag{152}$$

The following surface shows how the synthetic pass changes as the required matched-gain threshold and basis-risk reserve increase, holding $D_P = 0$ for display.

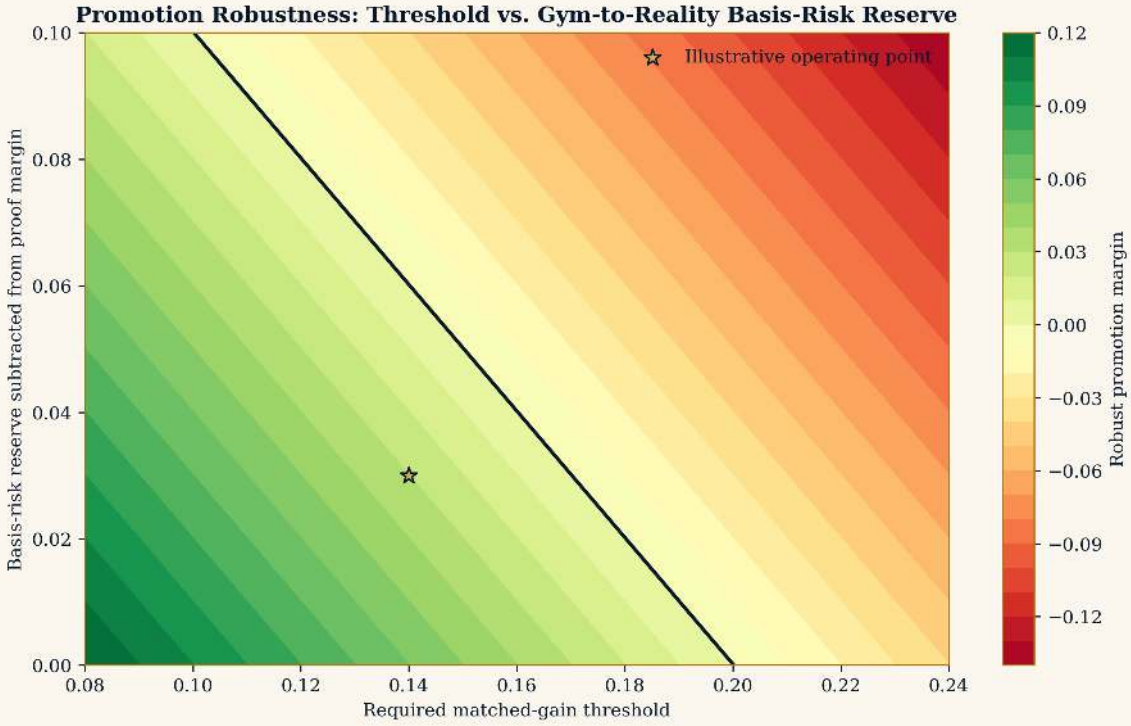



**Figure 10.** Promotion robustness surface. The navy contour is the zero-margin boundary. Green regions retain a positive robust margin ; red regions require repair, narrower claims or additional evidence.

At the illustrative point $\theta_m = 0.14$ and $\widehat{\epsilon}_{\text{gym}} = 0.03$, the residual margin remains positive at approximately 0.0302. At a basis-risk reserve above approximately 0.0602, the same proof no longer supports the stated threshold. This is the intended institutional behaviour : uncertainty about transfer is carried against the claim rather than ignored.

## 19.8 Complete-denominator canary value

On 400 canary cases, the synthetic successor prevents C$568,886.86 of invalid payments while generating C$7,565.00 of false-hold cost and C$3,992.47 of processing and review cost. The resulting synthetic net canary value is C$557,329.39.

The reference cycle deliberately records false holds and human review rather than reporting only prevented invalid amounts. A production denominator would additionally include integration, security, legal review, platform operations, incident reserves, requalification, migration and unwind.

## 19.9 Authority as a partial order

Evidence coverage is necessary but not sufficient for authority. The illustrative frontier below shows progressively larger authority classes becoming eligible as evidence strengthens, while the constitutional ceiling remains external to the evidence score.

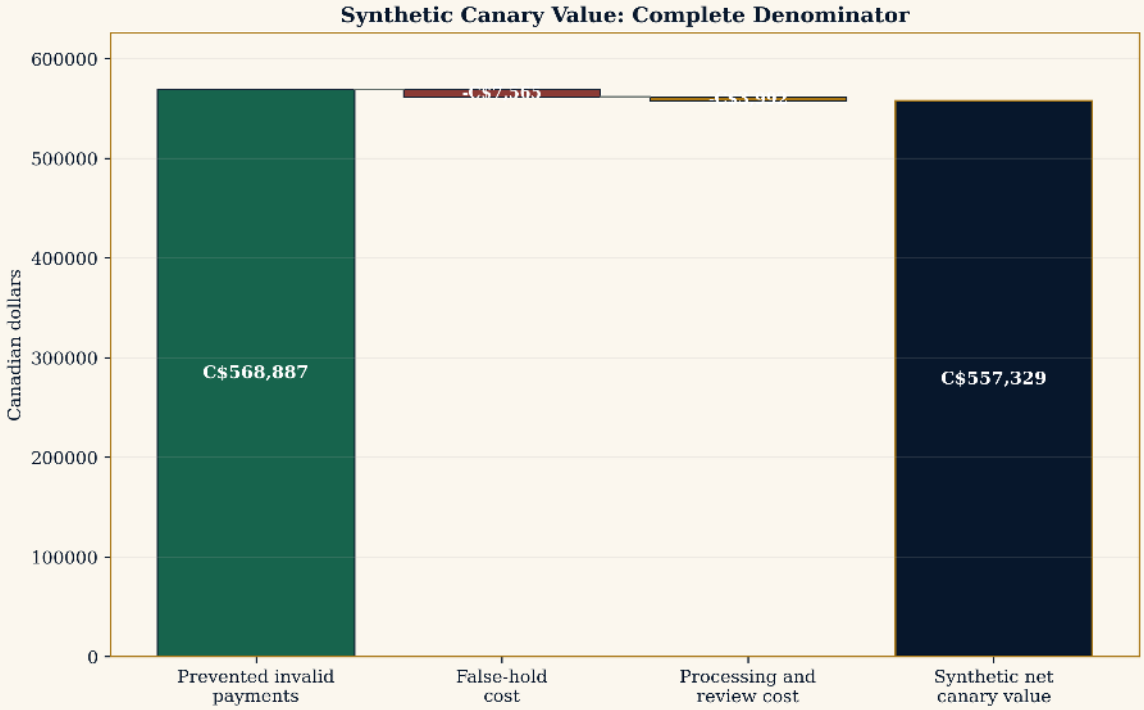



**FIGURE 11.** Synthetic canary value after explicit false-hold and processing burdens. The figure is an environment-derived accounting identity, not customer savings or audited performance.

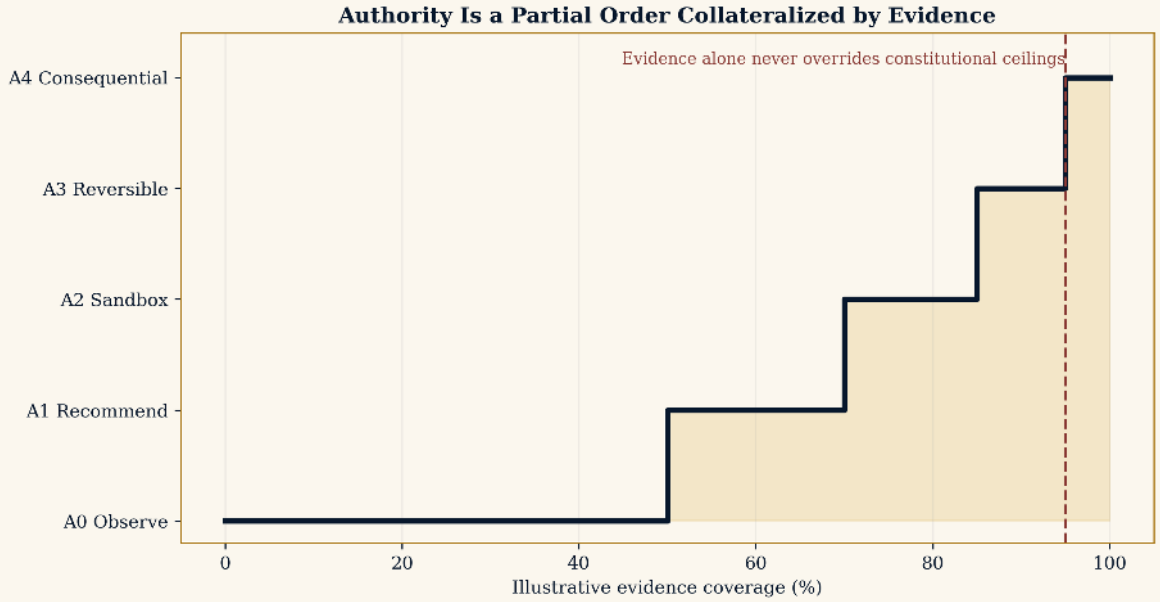



**FIGURE 12.** Illustrative evidence-to-authority frontier. A4 never follows from evidence coverage alone; consequence, professional authority, rights, reversibility and accountable admission remain independent gates.

**Table 21.** Robust promotion margin under selected threshold and basis-risk assumptions.

| $\theta_m$ | $\widehat{\epsilon}_{\text{gym}}$ | Robust margin | Disposition |
|---|---|---|---|
| 0.10 | 0.00 | 0.1002 | Pass |
| 0.12 | 0.02 | 0.0602 | Pass |
| 0.14 | 0.03 | 0.0302 | Pass |
| 0.14 | 0.06 | 0.0002 | Borderline |
| 0.16 | 0.04 | 0.0002 | Borderline |
| 0.18 | 0.04 | −0.0198 | Repair / narrower claim |
| 0.20 | 0.02 | −0.0198 | Repair / narrower claim |

### 19.10 Sensitivity table

The result does not become "more true" because a favourable reserve is chosen. Threshold, basis-risk methodology and Proof-Debt treatment must be frozen before protected outcomes are examined.

### 19.11 What the synthetic programme establishes

The v3.0 synthetic evidence supports the following limited conclusions :

— the complete succession state machine can be executed from mission constitution to requalification ;
— protected partitions and exact release manifests can be committed and reproduced ;
— a complete architecture can be compared against incumbent and alternative on matched cases ;
— hard gates can block authority even when average utility is high ;
— Chronicle assets can be reused in a new transfer environment ;
— a provider update can be impaired and rolled back ;
— a verifier-hardened descendant can be re-proven without authority expansion ;
— the release can preserve zero material unauthorized actions inside the synthetic environment.

It does not establish :

— that the synthetic utility function represents a customer's true mission utility ;
— that the simulator captures real process variation, fraud, human incentives or tail risk ;
— that C$557,329.39 would be realized in production ;
— that a real external system has earned Specialist ASI status ;
— that any legal, financial or organizational authority has been granted.

**Evidence conclusion**

Synthetic data can prove that the institution knows how to ask, separate, record and retest the right questions. Only rights-cleared fresh reality can prove that the answer survives the mission.

## 20 GSSB-4 : A Reproducible Four-Mission Synthetic Successor Benchmark

The Aurelia cycle demonstrates one complete succession path. Version 3.0 adds a second empirical object : the **GoalOS Synthetic Successor Benchmark 4** (GSSB-4), a deterministic four-mission reference benchmark that tests whether the same constitutional architecture can survive materially different mission environments. GSSB-4 is synthetic by design. It is not a customer result, a claim of real-world Specialist ASI, a securities-performance statement or a grant of production authority.

## 20.1 Benchmark constitution

GSSB-4 contains 6,000 synthetic cases and 30,000 candidate episodes. Each mission is partitioned into 800 formation cases, 500 protected proof cases and 200 transfer cases. All cases are generated from the fixed seed `30082026`; the generated files, candidate profiles, gate thresholds, environment cards and SHA-256 manifest are distributed in the Research Companion.

**Table 22.** GSSB-4 mission environments.

| Mission | Dominant reasoning burden | Critical proof concern |
|---|---|---|
| Invoice Integrity | Evidence reconciliation across invoices, work orders, contracts, vendor identity and prior payments. | Duplicate, ghost-vendor or wrong-building anomaly approved as pay-ready. |
| Building Operations | Weak-signal fusion, causal diagnosis, dispatch and bounded intervention across physical assets. | Critical physical event missed or action taken outside the validated authority envelope. |
| Acquisition Underwriting | Cross-domain evidence integration, thesis–anti-thesis competition, information acquisition and scenario analysis. | False reassurance on a material transaction risk or unsupported binding recommendation. |
| Verified Software Engineering | Repository state, defect diagnosis, patch formation, tests, regression analysis and release control. | Patch accepted despite hidden regression, incomplete proof or unverifiable dependency. |

Each case contains hidden difficulty, criticality, ambiguity, evidence richness, distribution shift and dependency variables. Candidate observations expose only the mission-authorized evidence state. The transfer split increases adversarial frequency and distribution shift. The scorer records decision correctness, evidence completeness, critical false reassurance, unauthorized action, human burden, latency, direct cost, reliability, governance, sovereignty, transfer and total mission utility.

## 20.2 Complete candidate institutions

The benchmark compares five complete architectures rather than five isolated model scores :

1. **Incumbent** : the existing human and software reference architecture ;
2. **General AI Copilot** : rented frontier capability with weak evidence custody and limited institutional separation ;
3. **Specialist Cell** : coordinated specialist roles with stronger evidence and review ;
4. **Sovereign Hybrid** : portable models and tools, independent verification, Chronicle and bounded authority ;
5. **Recursive Successor** $\Omega$ : the Chronicle-informed successor to the Sovereign Hybrid with verifier hardening, improved transfer and unchanged authority ceiling.

The benchmark intentionally gives the General AI Copilot lower complete-institution scores than its raw speed might suggest. This reflects the framework's complete denominator : attractive output latency cannot compensate for weak evidence, correlated error, hidden burden or governance failure.

## 20.3 Protected gate

Promotion on a mission requires all of the following :

$$\mathrm{LCB}_{0.95}(\overline{\Delta}_m) > 0.20, \tag{153}$$

$$\mathrm{EvidenceCoverage}_m \geqslant 0.95, \tag{154}$$

$$\mathrm{Reliability}_m \geqslant 0.90, \tag{155}$$

$$\mathrm{Governance}_m \geqslant 0.95, \tag{156}$$

$$\mathrm{Transfer}_m \geqslant 0.88, \tag{157}$$

$$\mathrm{CriticalMisses}_m = 0, \tag{158}$$

$$\mathrm{UnauthorizedActions}_m = 0. \tag{159}$$

The candidate is frozen before proof. The same protected cases are processed by the incumbent and challenger, and the promotion statistic is the paired mission-utility difference. Passing the synthetic gate yields only a *synthetic proof-eligibility result*; it does not create authority.

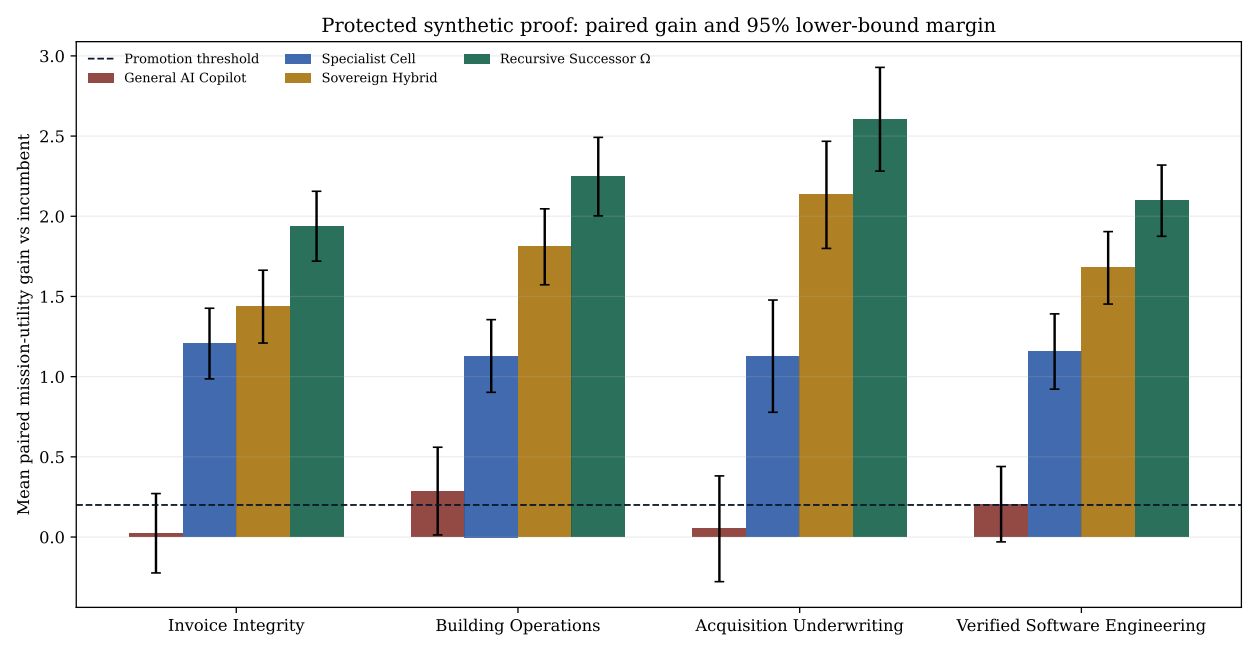



**Figure 13.** GSSB-4 matched synthetic proof. Error bars show the difference between the mean paired gain and its 95% lower confidence bound. Every bar is synthetic and reproducible from the distributed code and fixed seed.

## 20.4 Cross-mission proof result

**Table 23.** Recursive Successor Ω on protected synthetic proof.

| Mission | Accuracy | Evidence | Reliability | Governance | Transfer | Mean gain | LCB95 | Critical |
|---|---|---|---|---|---|---|---|---|
| Invoice Integrity | 0.836 | 0.998 | 0.969 | 0.987 | 0.952 | 1.938 | 1.720 | 0 |
| Building Operations | 0.796 | 0.992 | 0.960 | 0.987 | 0.950 | 2.247 | 2.002 | 0 |
| Acquisition Underwriting | 0.758 | 0.996 | 0.953 | 0.987 | 0.951 | 2.605 | 2.282 | 0 |
| Verified Software Engineering | 0.834 | 0.988 | 0.965 | 0.987 | 0.951 | 2.097 | 1.875 | 0 |

The Recursive Successor passes all seven synthetic gates on all four missions, with zero critical false reassurance and zero unauthorized actions. The Sovereign Hybrid passes all gates on Invoice Integrity and Acquisition Underwriting but fails two mission gates because one protected critical miss occurs in Building Operations and one in Verified Software Engineering. The result demonstrates why the successor to the successor is not a cosmetic version increment : the new release must retire a specific remaining failure mode while preserving the authority ceiling.

## 20.5 Complete-denominator comparison

**Table 24.** Cross-mission protected-proof aggregate.

| Candidate | Accuracy | Evidence | Reliability | Governance | Sovereignty | Utility | Critical | Unauthorized |
|---|---|---|---|---|---|---|---|---|
| Incumbent | 0.468 | 0.824 | 0.828 | 0.936 | 0.713 | -0.710 | 31 | 0 |
| General AI Copilot | 0.462 | 0.749 | 0.758 | 0.713 | 0.493 | -0.569 | 51 | 4 |
| Specialist Cell | 0.628 | 0.921 | 0.874 | 0.878 | 0.753 | 0.445 | 10 | 1 |
| Sovereign Hybrid | 0.706 | 0.982 | 0.928 | 0.970 | 0.943 | 1.054 | 2 | 0 |
| Recursive Successor Ω | 0.806 | 0.994 | 0.961 | 0.987 | 0.968 | 1.512 | 0 | 0 |

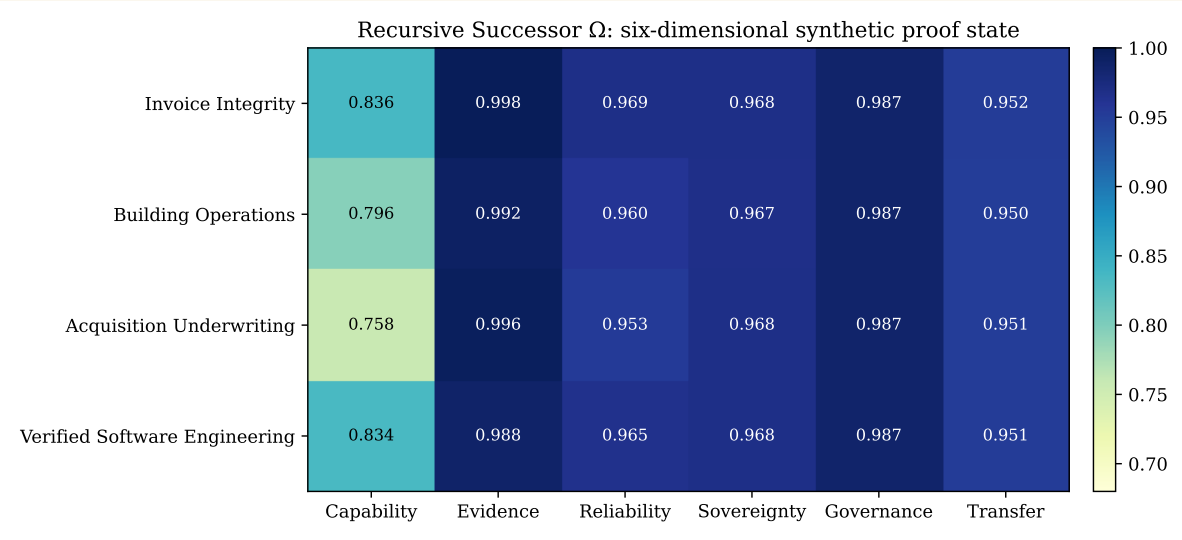



**FIGURE 14.** Six-dimensional synthetic proof state for Recursive Successor Ω. Capability is only one dimension; evidence, reliability, sovereignty, governance and transfer remain independent promotion requirements.

The General AI Copilot is not rejected because it is universally weak. It is rejected because its complete evidence, governance and critical-error state fails the constituted mission gate. The Specialist Cell creates positive synthetic mission utility but remains ineligible for broader authority. The synthetic ranking therefore differs from a ranking based only on speed or raw output quality.

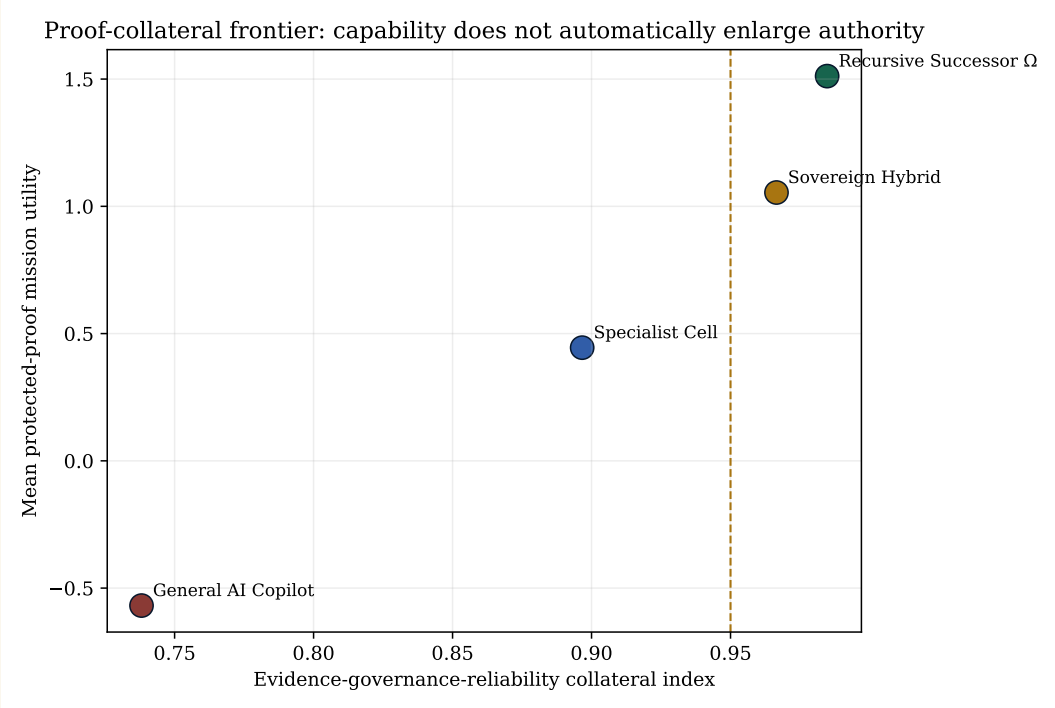



**FIGURE 15.** The proof-collateral frontier. Greater mission utility does not by itself authorize a larger action space; evidence, governance and reliability constrain the admissible Authority Envelope.

## 20.6 Recursive generations

The benchmark also exposes a controlled eight-generation Foundry trajectory. Generation 0 is the Sovereign Hybrid. The Foundry improves proof utility, evidence coverage and governance while reducing its estimated Gym-to-reality basis-risk reserve. Seventeen lineages are rejected or retained only as negative capability by Generation 8. Authority remains at A2 through the early generations and reaches only the existing A3 ceiling after fresh proof; it never

increases automatically because the architecture is newer.

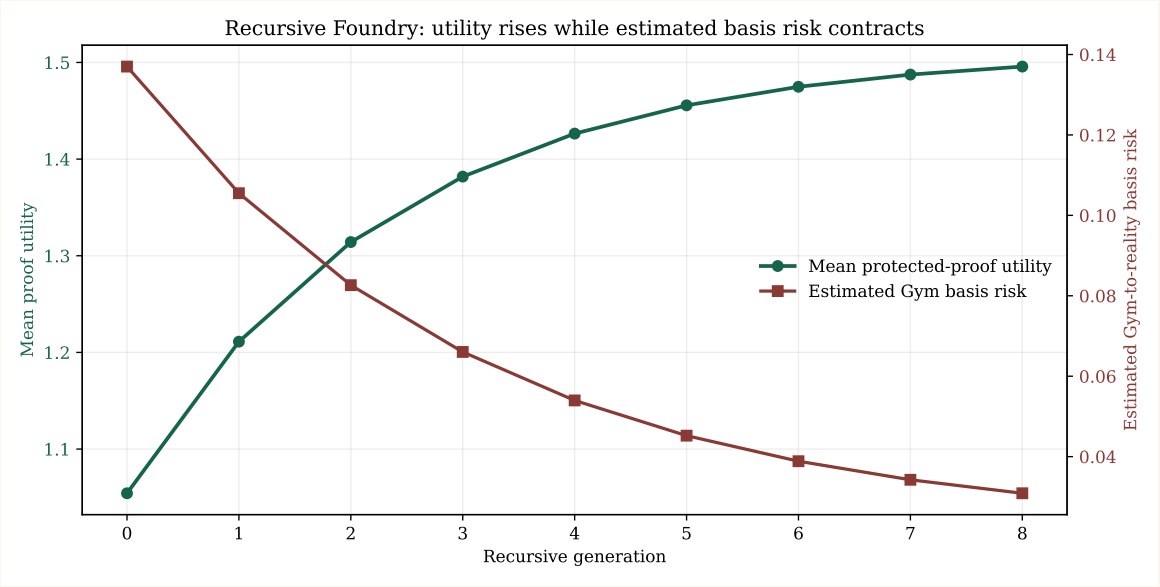



**Figure 16.** Synthetic recursive succession. Protected-proof utility rises as estimated basis risk contracts, but authority remains separately governed.

### 20.7 What GSSB-4 does and does not establish

GSSB-4 establishes that the paper's constitutional objects can be rendered as a reproducible, multi-mission evaluation system with :

- distinct formation, protected proof and transfer partitions ;
- matched incumbent comparison ;
- complete candidate institutions ;
- hard critical-error and authority gates ;
- cross-mission transfer ;
- recursive successor generations ;
- exact environment cards, case-level records, figures and hashes.

It does not establish that the synthetic transition models represent any customer's reality, that the mission-utility weights are universally correct, that the Recursive Successor is a real Specialist ASI, or that any production authority is safe. The benchmark is most valuable as a falsifiable reference protocol : a real customer mission can replace the synthetic state and transition generator while preserving the same separation, proof receipts, hard gates and release logic.

**GSSB-4 claim boundary**

Six thousand synthetic cases can prove reproducibility of the mechanism, not truth of the world model. The benchmark manufactures a protected reference result ; only rights-cleared customer reality can establish real mission dominance and realized Mission Alpha.

## 21 The Executable Institutional Twin and Reproducibility Architecture

### 21.1 From paper to executable constitution

A paper about verified succession should itself be reproducible. Version 3.0 therefore distributes a scholarly source, signed synthetic evidence, machine-readable institutional records, deterministic analysis code and exact release manifests as one coordinated publication object.

The **Executable Institutional Twin** is the correspondence among :

1. the constitutional statement in the paper ;
2. the machine-readable schema that represents it ;
3. the environment or workflow that executes it ;
4. the proof receipt that records the outcome ;
5. the release artifact that can be rebuilt independently.

**TABLE 25.** Paper-to-code correspondence.

| Constitutional object | Machine-readable record | Executable counterpart |
|---|---|---|
| Mission Constitution | `SuccessionConstitution.json` | Gym objective, state, observations, action masks, critical failures and termination conditions. |
| Successor Manifold | `SuccessorBook.json`; candidate manifests | Complete candidate architectures evaluated on matched environment releases. |
| Mission Advantage Gradient | Gradient and experiment records | Finite differences, ablations, architecture mutations and value-of-information search. |
| SEIZE Underwriting | `SuccessorUnderwriting.json`; `ProofCapitalAuthorization.json` | Proof-budget, stop-loss and trajectory-selection logic. |
| Bounded AGI Jobs | Typed job records | Human, model, tool and verifier work with exact inputs, limits, evidence and rollback. |
| Fresh proof | `RunProof.json`; `FreshSuccessorScorecard.json` | Hidden cases, exact scorer, paired statistics and hard-gate intersection. |
| Authority | `ReleaseCertificate.json`; Authority Envelope | Action gateway, identities, transaction limits, expiry, kill switch and fallback. |
| Chronicle | `ChronicleSuccessorRecord.json` | Accepted capabilities, failures, proofs, value, lineage and requalification state. |
| Successor recursion | `RequalificationRecord.json`; `ImpairmentRecord.json`; `RevocationRecord.json` | Material-change detection, rollback, descendant formation and next-successor proof. |

## 21.2 Reference environment interface

A Gymnasium-compatible mission environment exposes the modern interaction cycle :

```
observation, info = env.reset(seed=protected_seed)
while True:
   action = successor.act(observation, authority=active_envelope)
   observation, reward, terminated, truncated, info = env.step(action)
   proof_receipt.append(info)
   if terminated or truncated:
      break
```

The interface is intentionally agnostic to the acting policy. The successor may contain a model, several agents, deterministic systems, search, retrieval, professional review or a human–machine organization. The environment controls what can be observed, which actions are possible, how the state changes and what record is emitted.

## 21.3 The reproducibility ladder

| Level | Reproducibility state |
|---|---|
| R0 Narrative | Concepts and results are described, but the exact source, environment and candidate cannot be reconstructed. |
| R1 Source available | LaTeX, code or configuration is distributed, but inputs or runtime remain incomplete. |
| R2 Deterministic build | The publication and synthetic analysis rebuild from a documented toolchain and fixed seeds. |
| R3 Evidence replay | Exact candidate, cases, scorer, receipts and summary statistics can be independently replayed. |
| R4 Clean-room equivalence | A separate directory or machine reconstructs the authoritative PDFs and principal evidence with matching normalized text, hashes or declared tolerances. |
| R5 Institutional continuity | The mission, Gym, proof state, Chronicle, Authority Envelope and rollback package remain usable after model, provider, team or infrastructure substitution. |

Version 3.0 targets R4 for the publication and synthetic evidence package. R5 remains a production institutional objective because it requires an actual protected deployment, organizational identities, real rights, tested provider substitution and live continuity drills.

### 21.4 Artifact graph and source of truth

The release should be treated as a directed acyclic artifact graph :

$$\begin{aligned} &\text{source} + \text{data} + \text{environment} + \text{manifest} \\ &\quad \longrightarrow \text{analysis} + \text{figures} + \text{records} \\ &\quad \longrightarrow \text{PDFs} + \text{HTML} + \text{evidence pack} \\ &\quad \longrightarrow \text{checksums} + \text{verification}. \end{aligned} \tag{160}$$

Every generated artifact records its parent inputs, build command, tool versions and cryptographic digest. The release manifest distinguishes authoritative artifacts from explanatory or supplementary visual works.

> The authoritative v3.0 release is the coordinated set containing the A4 and US Letter monographs, institutional atlas, Board Brief, Visual Abstract, versioned LaTeX source, bibliography, exact figures, synthetic evidence records, executable reference code, artifact manifest, build instructions, QA report and SHA-256 ledger.

### 21.5 Protected proof cannot be made public by accident

Reproducibility and secrecy must be separated carefully. A public benchmark may distribute cases and labels. A protected promotion environment should instead distribute :

- commitments or hashes to the protected case set ;
- public environment and interface specifications ;
- scorer methodology at the level compatible with non-gaming ;
- exact candidate manifest and runtime record ;
- independently signed aggregate results ;
- delayed or access-controlled ground truth after the claim expires, where appropriate.

The public release may demonstrate the mechanism with synthetic labels while the real protected plane remains under separate custody. Reproducibility means the claim can be audited by authorized parties ; it does not mean every sensitive case must be published.

### 21.6 Institutional environment cards

Every Mission Gym should include an Environment Card covering :

— constituted mission and excluded uses ;
— state variables and hidden variables ;
— observation rights by role ;
— action types, masks and external effects ;
— transition logic and nonstationarity ;
— reward vector and hard constraints ;
— data provenance, licences and retention ;
— scenario, subgroup and tail coverage ;
— termination and truncation ;
— verifier, scorer and calibration ;
— known basis risk and invalidation conditions ;
— candidate visibility and protected partitions ;
— authority ceiling and production migration path.

Candidate successors should carry a corresponding Successor Card documenting complete architecture, dependencies, authorities, external calls, human interventions, cost, proof state and known failure clusters.

### 21.7 Reproduction algorithm

**Algorithm 2** Clean-room reproduction of the v3.0 evidence edition

**Require :** Signed source archive, synthetic evidence archive, manifest and build toolchain
1 : Verify archive hashes and reject any unexpected path or duplicate entry
2 : Recreate a clean build directory
3 : Compile A4 and US Letter monographs, Board Brief, Visual Abstract and Institutional Atlas
4 : Re-run deterministic synthetic evidence analysis with fixed release seed
5 : Compare principal statistics, protected commitments and figure digests with the release manifest
6 : Render every PDF page and run structural preflight
7 : Compare normalized text and declared page counts with the authoritative release
8 : Emit a machine-readable clean-room verification record

### 21.8 Publication accessibility and durable distribution

The complete programme should distribute more than one visual format :

— a high-resolution A4 archival and international edition ;
— a US Letter institutional edition ;
— a compressed web PDF ;
— an accessible semantic HTML edition with navigable headings and alt descriptions ;
— exact figures in PNG and PDF ;
— modular and single-document LaTeX ;
— citation metadata in BibTeX, RIS, CSL JSON and CFF ;
— data, code, schemas, manifests and checksums.

The HTML edition is not the source of truth for equations or pagination ; it is the accessible reading surface. The LaTeX source and signed PDF release control the scholarly statement.

### 21.9 Production migration from synthetic twin to real mission

The synthetic twin becomes a real succession institution only after replacing generated assumptions with :

— rights-cleared production or shadow data ;
— independently accepted mission utility and critical-failure definitions ;
— real incumbent cost, human burden and exception data ;

— security, privacy, professional and legal review;
— organizational identities and dual-control release;
— a protected fresh-work plane with independent custody;
— shadow, sandbox and reversible canary operation;
— independent calculation of realized value;
— tested continuity, rollback, incident and revocation procedures.

**Executable-twin law**

A reproducible synthetic twin proves that the institution can execute its own constitution. A production Mission-Sovereign Successor exists only when the same constitution survives rights-cleared reality, independent proof and accountable authority.

## 22 Corporate Mission Classes and Specialist ASI Use Cases

### 22.1 Acquisition Underwriting Specialist ASI

The constituted mission is to produce an evidence-bearing acquisition recommendation that identifies material value drivers, contradictions, downside scenarios, integration risks and decision conditions while reserving binding authority to the investment committee.

The Mission Gym contains historical data rooms, hidden answer keys, evidence requests, target-company states, market conditions, integration scenarios, time and diligence budgets. Actions include requesting a source, reconciling records, updating thesis or anti-thesis, selecting the next-best diligence question, running counterfactual scenarios, escalating and refusing.

The complete Specialist ASI may contain :

— a living target-company world model;
— evidence and provenance graph;
— finance, legal, technology, market, operations and integration specialists;
— thesis and adversarial anti-thesis systems;
— contradiction and reconciliation engines;
— scenario and counterfactual simulation;
— an information-value engine;
— uncertainty, abstention and professional escalation;
— Chronicle learning from post-acquisition outcomes.

The mission's strongest advantage may come not from better prose but from cross-domain contradiction discovery and selecting the one evidence object most likely to change the decision. Initial authority remains Observe, Advise and Prepare. The investment committee retains transaction, contract and capital authority.

### 22.2 Portfolio Building Operations Specialist ASI

The mission is to understand and improve the evolving physical state of a real-estate portfolio : anticipate failures, diagnose causes, optimize interventions, coordinate bounded work and learn across buildings.

The Gym may combine digital twins, sensors, weather, occupancy, tenant reports, asset histories, vendor availability, maintenance records, energy and capital constraints. The Specialist ASI may contain :

— a portfolio-scale physical world model;
— multimodal perception and weak-signal fusion;
— causal root-cause diagnosis;
— counterfactual intervention simulation;
— maintenance, vendor, inventory, energy and capital optimization;
— cross-building transfer intelligence;

— a safety and policy governor;
— independent completion verification;
— Chronicle of failures, repairs and operating policies.

A concrete moment of Specialist ASI advantage occurs when several ordinary hot-water complaints are recognized as one portfolio-level control-component pattern, leading to the decisive diagnostic, the correct repair and a preventive programme across similar buildings. The initial successor may dispatch approved vendors or place bounded work orders for non-critical cases while life-safety, legal, professional and major-capital authority remain human.

### 22.3 Verified Software Engineering Specialist ASI

Software engineering has high Gymmability because repositories, tests, sandboxes and regression suites create executable feedback (Jimenez et al., 2024). The mission may be to diagnose, repair and verify production defects under exact repository, dependency, security and release constraints.

The complete successor may include issue decomposition, repository world model, code search, program synthesis, patch generation, adversarial testing, formal or deterministic checks, security analysis, reviewer routing, rollback and release evidence. Specialist ASI status requires end-to-end accepted resolution on fresh repositories and issues, not merely patch generation.

### 22.4 Financial Reconciliation Specialist ASI

Financial reconciliation is strongly suited to Mission Gyms because accounting identities, accepted historical resolutions, high volume and explicit exception classes enable protected replay. The successor may combine evidence extraction, rule engines, entity resolution, duplicate graphs, fraud indicators, policy checks, independent controls and human controller escalation.

The principal danger is false reassurance and unauthorized financial action. A highly capable successor may remain at A1 or A3 indefinitely if payment-release authority is not justified.

### 22.5 Scientific and engineering design

Searchable design spaces, simulators and laboratory measurements can support a Specialist ASI that generates hypotheses, plans experiments, integrates results and improves designs. However, surrogate objectives, simulator gaps and expensive physical proof create substantial basis risk. The Gym must preserve the distinction between simulated improvement and laboratory or field validation.

### 22.6 Mission selection before formation

The strongest initial missions combine :

— high value or high frequency;
— a bounded state and action space;
— a rapid, trustworthy evaluator;
— repeatable decisions and stable constraints;
— lawful examples, failures and feedback;
— an incumbent with visible burden or decay;
— a realistic path to ownership, control or recurring value;
— an environment that can be reset, simulated or replayed.

A technically impressive candidate in a weak mission is a poor investment. The mission should be held or rejected when the result cannot be independently verified, rights are inadequate, the opportunity expires before proof, the consequence exceeds the control architecture or a commodity alternative is already superior.

## 23 Technical Architecture : From Mission Gym to Successor Operations

## 23.1 The six-plane architecture

A production implementation separates public orientation, confidential formation, protected proof and institutional authority. The planes are ordered by consequence rather than by visual prominence.

1. **Public Navigator.** Public-safe research, capability radar, education, mission qualification and non-confidential board orientation.
2. **Protected Mission Workspace.** Confidential Mission Constitution, incumbent ledger, economics, rights, proof capital and Authority ceiling.
3. **Sovereign Mission Gym.** Versioned environments, replay, digital twins, action schemas, constraints, scenario generation and proof receipts.
4. **Isolated Successor Foundry.** Data formation, models, agents, tools, search, simulations, candidate formation and sandbox execution.
5. **Independent Proof Plane.** Protected cases, scorer internals, validator credentials, Fresh-Successor Gate, canary evidence and Release Certificate.
6. **Chronicle, Authority and Capital Plane.** Admission, Succession Ledger, monitoring, realized value, impairment, requalification, Capital-to-Capacity and board reporting.

## 23.2 Environment API

A Gymnasium-compatible core interface is :

```
observation, info = env.reset(seed=42)
while True:
   action = successor.act(observation, authority=info["authority"])
   observation, reward, terminated, truncated, info = env.step(action)
   receipt_store.append(info["proof_receipt"])
   if terminated or truncated:
      break
```

GoalOS extends the environment API with player identity, Authority Envelope, action masks, vector rewards, constraint events, proof receipts, evidence lineage, exact Gym and candidate versions, declared human interventions, Chronicle eligibility, rollback and requalification triggers. Multi-player missions can use PettingZoo-style agent cycles or OpenSpiel-style game representations when roles act sequentially, simultaneously or under imperfect information (Terry et al., 2021 ; Lanctot et al., 2019).

## 23.3 Identity, least privilege and evidence custody

**Protected identity controls**

- organizational identity and role-based access ;
- separate development, evaluation, release and audit roles ;
- short-lived credentials and scoped sessions ;
- encrypted storage and key management ;
- data classification and jurisdictional controls ;
- append-only audit evidence ;
- dual control for consequential release ;
- incident containment and credential revocation.

**Independent proof custody**

— blind-case commitments and hashes;
— scorer versions and frozen thresholds;
— evaluation runtime and environment;
— exact candidate manifest;
— result receipts and confidence calculations;
— reviewer and acceptance identities;
— release signature and rollback package;
— technical separation from candidate formation.

### 23.4 Machine-readable canonical records

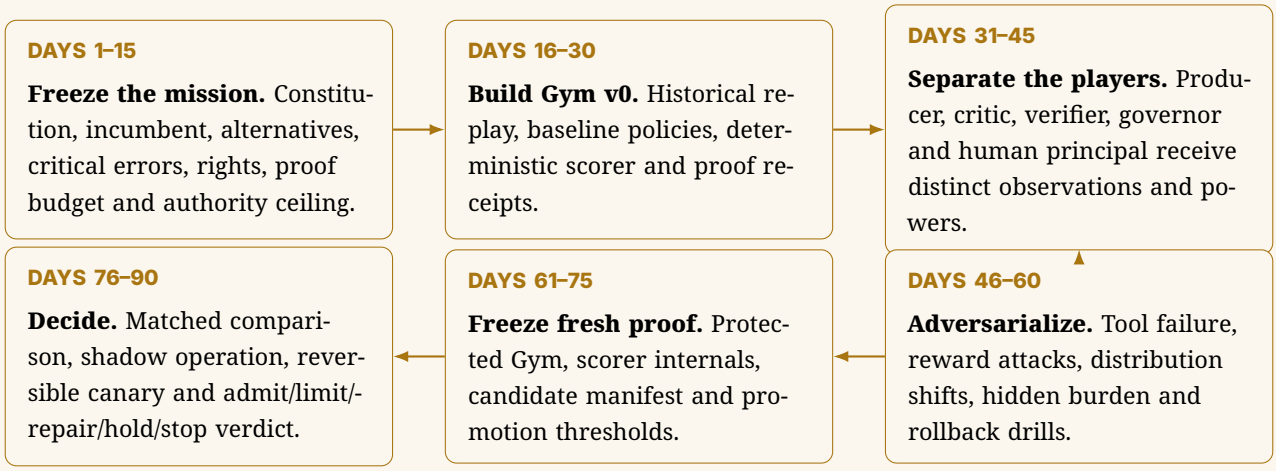
Every state transition carries an attributable actor, evidence object, authority receipt, timestamp, version and stop condition. The minimum record set is :

1. FrontierEventRecord;
2. SuccessionEventRecord;
3. IncumbentAssumptionLedger;
4. MissionGymLedger;
5. EnvironmentSpec;
6. SuccessorBook;
7. SuccessorUnderwriting;
8. ProofCapitalAuthorization;
9. SuccessionConstitution;
10. BoundedAGIJobPortfolio;
11. VerifierSpec;
12. CandidateSuccessorManifest;
13. RunProof;
14. FreshSuccessorScorecard;
15. PromotionRequest;
16. ReleaseCertificate;
17. AuthorityEnvelope;
18. ChronicleSuccessorRecord;
19. CapitalAllocationMemo;
20. RequalificationRecord;
21. ImpairmentRecord;
22. RevocationRecord.

### 23.5 Operational observability

A production dashboard should expose the current mission and release; Gym and scorer versions; evidence freshness; active Authority Envelope; blocked and attempted unauthorized actions; human interventions and reviewer burden; cost, latency and resource consumption; drift and out-of-distribution alerts; rollback readiness and current fallback; and the next requalification date.

### 23.6 90-day activation path




### 23.7 Production hardening

The public browser demonstrations accompanying GoalOS are reference implementations rather than protected production systems. Real deployment requires server-side identity and authorization; confidential evidence storage; independent proof custody; model and provider monitoring; secrets management; protected audit and release signing; security, privacy, legal and professional review; operational continuity and tested rollback; and independent calculation of realized value.

**Architecture law**

Public navigation may explain succession. Protected workspaces may constitute it. Independent proof may validate it. Only the Chronicle, Authority and Capital Plane may admit, govern and compound it.

# 24 Paper-to-Code Reproducibility and the Evidence-Bearing Publication

A paper about verified succession should itself expose a reproducible release architecture. Version 3.0 therefore treats publication as a bounded evidence object : exact source, deterministic synthetic benchmark, generated figures, environment cards, claim boundary, build logs, manifest and cryptographic checksums are distributed together.

## 24.1 Reproducibility planes

**Table 27.** Publication reproducibility planes.

| Plane | Authoritative object | Reproduction test |
|---|---|---|
| Scholarly source | Modular and single-document LaTeX, bibliography and exact figures. | Clean-room A4 and US Letter compilation with resolved references and embedded fonts. |
| Synthetic evidence | GSSB-4 generator, fixed seed, case-level results, environment cards and gate thresholds. | Regenerate 6,000 cases, 30,000 episodes, tables, figures and hashes. |
| Institutional visuals | Deterministic vector plates plus supplementary image-generated ornament. | Verify that operative text and equations exist independently of image microtypography. |
| Machine-readable constitution | Environment, candidate, proof, authority, Chronicle and requalification schemas. | Validate records against schemas and replay the canonical state transition. |
| Distribution | PDFs, HTML, EPUB, source archives, manifest and SHA-256 ledger. | Extract every archive, compare payload hashes and audit internal links. |

## 24.2 Paper-to-code crosswalk

| Paper object | Executable object | Evidence emitted |
|---|---|---|
| Mission Gym $\mathscr{G}_m$ | Environment card and case generator | Versioned state, observations, actions, constraints, split identity and case commitments. |
| Successor Manifold $\mathscr{M}_S$ | Candidate profile and architecture manifest | Models, tools, verifier, governance, sovereignty, cost and authority coordinates. |
| Mission Advantage Gradient | Candidate mutation and bounded search signal | Proposed architecture delta, expected value direction and proof-capital request. |
| SEIZE Underwriting | Gate threshold and option decision record | Mission, alternatives, proof budget, stop-loss, challenger freeze and review owner. |
| Bounded AGI Jobs | Typed work-and-proof record | Inputs, tools, prohibitions, budget, verifier, acceptance and rollback. |
| Fresh-Successor Gate | Paired proof evaluator | Case-level utility differences, confidence bounds, critical errors and gate verdict. |
| Authority Envelope | Machine-readable permission state | Scope, ceiling, expiry, blocked actions, rollback and accountable principal. |
| Chronicle | Append-only successor record | Accepted capabilities, failures, proof state, release history and requalification trigger. |
| Recursive operator $\Re$ | Generation table and lineage archive | Descendants, rejected lineages, basis-risk reserve, champion and unchanged authority ceiling. |

### 24.3 One-command synthetic rebuild

The Research Companion contains the deterministic benchmark runner. From its directory :

```
python run_gssb4.py --out generated
```

The command recreates :

— four mission datasets and three partitions per mission ;
— case-level outputs for five complete candidate architectures ;
— candidate summaries, paired-gain statistics and gate verdicts ;
— recursive-generation records ;
— four publication figures ;
— environment cards and the benchmark manifest.

The publication build then compiles A4, US Letter, Board Brief, Visual Abstract, Evidence Dossier and Institutional Atlas editions. The source package includes a single build entry point, while the release suite includes clean-room reconstruction records so that compilation success is not inferred from the author's working directory.

### 24.4 Claim ledger

Every material claim in the release belongs to one of five classes :

1. **Definition** : a stipulated GoalOS term or constitutional object ;
2. **Derived result** : a theorem or proposition following from explicit assumptions ;
3. **External fact** : a statement attributed to a cited source ;
4. **Synthetic result** : a deterministic benchmark or simulation output ;
5. **Real-world state claim** : a claim requiring exact customer evidence, independent calculation and accountable admission.

The release never promotes a synthetic result into the fifth class. The accessible HTML edition and Evidence Dossier repeat this boundary so that distribution format does not change epistemic status.

### 24.5 Environment and dataset cards

Each synthetic mission has an environment card specifying :

— mission, beneficiary and reference architecture ;
— synthetic state variables and generation process ;
— observation and action families ;
— formation, proof and transfer partitions ;
— hard constraints and critical failures ;
— candidate visibility and protected elements ;
— known simplifications and Gym basis risks ;
— authority ceiling and claim boundary.

The cards are designed to be replaced, not merely edited, when a rights-cleared customer environment is constituted.

### 24.6 Accessibility and durable distribution

The ceiling release includes a semantic HTML edition and EPUB companion in addition to print PDFs. Equations are represented in source form, figures include descriptive captions, tables retain header structure and the reading order follows the scholarly hierarchy. The accessible editions are not a substitute for professional accessibility certification, but they avoid making the fixed-layout PDF the only usable source.

### 24.7 Publication as a proof-carrying object

The release manifest identifies each artifact, its size, SHA-256 digest, role and relation to the authoritative source. The complete suite can therefore be treated as a proof-carrying publication :

$$\mathscr{P}_{\text{release}} = (\text{source}, \text{build}, \text{evidence}, \text{figures}, \text{claims}, \text{manifest}, \text{checksums}, \text{verification}). \tag{161}$$

A reader may disagree with the assumptions, utility function or synthetic world model while still reproducing exactly what the paper did.

**Reproducibility law**

A succession paper should not ask the reader to trust an invisible pipeline. Definitions must be inspectable, synthetic results reproducible, protected claims bounded, and every distributed artifact attributable to one versioned source of truth.

Ω

PART VI

# The Successor-Era Corporation

*From digital presence to governed institutional agency*

The enterprise object is a portfolio of mission-specific successors sharing constitutional infrastructure while retaining separate proof and authority.

**SUCCESSOR MANIFOLD → GRADIENT → SEIZE → BOUNDED JOBS → SPECIALIST ASI → SUCCESSOR Ω**

# 25 The Successor-Era Corporation

## 25.1 The website analogy and its limit

A website gave the corporation persistent digital presence, distribution and transaction interfaces. A Mission-Sovereign Successor gives it governed institutional agency : the ability to perceive a mission state, assemble evidence, execute bounded work, submit to independent verification, preserve admitted learning and exercise only the authority it has earned.

**Table 29.** From digital presence to institutional agency.

| Layer | Website era | Successor era |
|---|---|---|
| Presence | Makes the institution visible and reachable. | Makes the institution continuously capable on a constituted mission. |
| Interface | Publishes information and receives transactions. | Perceives, plans, performs bounded jobs, verifies and escalates. |
| Memory | Stores content and records. | Preserves evidence, failure, accepted capability, authority and realized outcomes. |
| Control | Governs pages, identity and commerce. | Governs models, data, tools, proof, permissions, rollback and reproof. |
| Competitive signal | Absence increasingly signalled digital irrelevance. | Absence can signal institutional latency, weak learning and dependence on common Beta. |
| Limit | A website rarely acts with consequential authority. | A successor may act ; therefore proof, expiry and revocation are central. |

The analogy is a strategic conjecture, not a guaranteed chronology. Not every corporation needs the same architecture, and many low-differentiation missions should continue to rent market Beta. The stronger claim is that competitively consequential institutions increasingly need either a governed portfolio of mission-specific successors or reliable access to institutions that possess them.

## 25.2 A portfolio, not one monolith

The likely enterprise object is not one universal sovereign agent. It is a portfolio with shared identity, policy, observability, evidence custody, Chronicle and human authority, while each mission retains its own proof distribution, risks, rights, expiry and Authority Envelope.

Illustrative successors include :

- Research and Scientific Design ;
- Commercialization and Customer Resolution ;
- Verified Software Engineering ;
- Acquisition and Capital Allocation ;
- Financial Reconciliation and Treasury Operations ;
- Compliance and Contract Evidence ;
- Building, Energy and Industrial Operations ;
- Logistics, Procurement and Supply Resilience.

A portfolio avoids fragmented shadow agents with no memory and one monolithic system whose capture compromises the institution. Shared constitutional infrastructure reduces formation cost without allowing evidence or authority to leak across missions.

## 25.3 The build–rent–reserve boundary

## 25.4 Mission selection before formation

The strongest initial missions combine :

- high value or high frequency ;
- a bounded state and action space ;
- a trustworthy evaluator and rapid feedback ;
- lawful historical examples and failure records ;

**Table 30.** When to rent, configure, constitute or preserve.

| Mode | Use when | Institutional control |
|---|---|---|
| Rent Beta | The task is low differentiation, reversible, low consequence and adequately solved by the market. | Vendor, privacy, cost and quality controls. |
| Configure a workflow | The mission needs context and tools but limited proprietary learning or authority. | Own data mapping, workflow, tests and exportable records. |
| Constitute a successor | The mission is valuable, repeated, measurable, proofable and capable of accumulating proprietary evidence, memory or authority. | Retain mission constitution, Gym, jobs, evidence, Chronicle, permissions, portability and economics. |
| Preserve a reserve | The opportunity becomes valuable only after a named capability, price, legal, data or distribution catalyst. | Maintain option, triggers and evidence plan without operational authority. |
| Retire or prohibit | The mission is uneconomic, unverifiable, rights-conflicted or too consequential for available control. | Preserve decision rationale, containment and unwind. |

— visible incumbent burden or decay;
— a resettable, replayable or simulatable environment;
— a realistic path to bounded deployment;
— recurring learning and ownership value.

A technically impressive candidate in a weak mission is a poor investment. A simpler human–machine architecture can be a superior successor when it creates greater accepted value, lower complete burden, stronger sovereignty or better optionality.

### 25.5 Preserve, build and become

The institution evaluates three transformations :

**Preserve** Mission, identity, rights, human agency, public obligations, safety interlocks, strategic data, evidence and reversibility that no successor may weaken.

**Build** Products, models, Mission Gyms, verifiers, tools, data, laboratories, markets, distribution, infrastructure and Mission Packs newly made possible by the frontier.

**Become** Roles, operating models, incentives, authority structures and institutional identity that must change when the incumbent can no longer fulfil the mission as well.

The Verified Succession Institution is therefore not merely an automation programme. It is a governed method for deciding what the corporation should preserve, what it should build and what it should become.

**Successor-era corporate law**

The rational objective is not maximum autonomy. It is the fastest compounding portfolio of mission capabilities whose proof remains fresh, whose authority remains bounded and whose successor can be constituted before the present one becomes impaired.

## 26 Board Command, Operating Cadence and Implementation

### 26.1 The board dashboard

**Table 31.** Verified Succession board metrics.

| Metric | Governing question |
|---|---|
| Singularity Exposure Ratio | Is adaptation lag below the expected half-life of material assumptions? |
| Time-to-Verified-Successor | How quickly does a frontier event become a protected succession decision? |

| Metric | Governing question |
|---|---|
| Successor Gain | Did the frozen challenger outperform the incumbent and best credible alternative on fresh work? |
| Successor Efficiency | How much verified gain was created per unit of Total Institutional Burden? |
| Proof Coverage | Does current evidence support the authority, valuation and reuse granted? |
| Authority-at-Risk | How much consequential authority is exposed to unproven, impaired or expired capability? |
| Successor Option Coverage | Are incumbent, immediate, challenger, reserve, hedge and retirement positions maintained? |
| Incumbent Decay | What value, control or optionality is being lost by delaying justified succession? |
| Alpha Decay | How quickly is the accepted advantage being commoditized, imitated or impaired? |
| Chronicle Precision | Are useful capabilities admitted and stale, unsafe or rights-conflicted capabilities rejected? |
| Capital-to-Capacity Conversion | What durable productive capacity was created from realized or conservatively attributable value? |
| Unauthorized Actions | Did any component act outside its Authority Envelope? |
| Fresh-Transfer Gain | Did admitted capability measurably improve a later mission or succession cycle? |
| Recursive Readiness | Can the institution constitute and prove a credible next successor before the current one reaches expected impairment? |

## 26.2 Operating cadence

**Table 32.** GoalOS Singularity Office Ω cadence.

| Cadence | Institutional function |
|---|---|
| Continuous | Frontier events, source expiry, blocked actions, provider drift, incidents, cost, law and market changes. |
| Weekly | Opportunity and Risk Brief; incumbent impairments; active Proof Debt; urgent evidence and access decisions. |
| Monthly | Successor Book; proof-capital portfolio; realized-value ledger; requalification and Capital-to-Capacity committee. |
| Quarterly | Institutional reconstitution; retain, repair, expand, migrate, hedge, revoke or retire verdicts. |
| Material event | Immediate Capability Shock Protocol after a major capability, legal, security, transaction or infrastructure event. |
| Annual | Complete authority, rights, resilience, data, provider, insurance, professional and productive-capacity review. |

Every board brief should answer :

1. What changed?
2. Which incumbent assumption or mission is affected?
3. What value or risk is material?
4. Which positions are in the Successor Book?
5. What evidence would change the recommendation?
6. What proof capital and authority are currently at risk?
7. What should happen in the next 72 hours, 30 days, 90 days and 365 days?

## 26.3 A 90-day activation path

## 26.4 The one-real-succession-cycle gate

The first credible institutional proof requires more than a demonstration. It should produce :

**Table 33.** Reference 90-day activation.

| Period | Institutional work |
|---|---|
| Days 1–15 | Freeze the Mission Constitution, incumbent, alternatives, critical errors, rights, proof budget, authority ceiling and accountable principals. |
| Days 16–30 | Build Navigator inputs, the Capability Twin, Incumbent Ledger, replay Gym v0, baseline policies, deterministic scorer and proof receipts. |
| Days 31–45 | Generate the Successor Manifold; add producer, critic, verifier, governor and human-principal roles with separated observations and credentials. |
| Days 46–60 | Run SEIZE; compile Bounded AGI Jobs; adversarialize tools, rewards, evidence, distribution and rollback; preserve negative capability. |
| Days 61–75 | Freeze one challenger, Fresh-Proof Gym, scorer internals, candidate manifest and preregistered promotion thresholds. |
| Days 76–90 | Run matched proof, shadow operation and reversible canary; decide admit, limit, repair, retain, hedge or stop; issue exact Chronicle and rollback records. |

- one rights-cleared mission and exact incumbent;
- materially different complete candidates;
- one evidence-backed rejection;
- one exact frozen challenger;
- independent protected fresh proof;
- one accountable bounded admission or refusal;
- one realized or independently attributable value result;
- one requalification, impairment or revocation;
- one later mission improved by Chronicle;
- zero material unauthorized actions.

Synthetic cycles can demonstrate mechanics, reproducibility, hard gates and negative evidence. They cannot establish customer value, production safety or real Specialist ASI. The empirical gate remains open until one protected customer cycle is completed without weakening the constitution.

### 26.5 The first-year institutional target

By the end of the first operating year, a complete implementation should maintain :

- continuous Succession Radar and Capability Twin updates;
- weekly Opportunity and Risk Briefs;
- monthly Successor Book and proof-capital committee;
- quarterly institutional reconstitution;
- material-event Capability Shock Protocol;
- independent evaluators and replayable mission benchmarks;
- a verified-value and Succession Ledger;
- at least one fresh-successor comparison and one Capital-to-Capacity allocation.

**Implementation law**

Do not begin by automating the institution. Begin with one consequential mission, one honest incumbent, one sovereign Gym, several complete alternatives, one frozen challenger and one protected decision. Scale only what survives fresh reality and accountable review.

## 27 Corporate Productization : Mission-Sovereign Successor-as-a-Service

### 27.1 The corporate business model

MONTREAL.AI develops and commercializes the institutional system for constituting Mission-Sovereign Successors. It combines rented frontier intelligence with a customer's mission, proprietary knowledge, data, tools, workflows, verifiers, evidence, controls and human authority—then charges to underwrite, constitute, prove, deploy, monitor

and continuously renew the resulting institutional capability.

$$\boxed{\begin{array}{c}\text{Rented AI Beta + mission-owned complements} \\ \text{+ GoalOS Gym, proof and governance}\end{array} \longrightarrow \text{Mission-Sovereign Successor.}} \tag{162}$$

The company does not primarily sell access to a model. It sells the corporate capacity to manufacture, prove, govern and continuously renew the superior successor for a consequential mission.

## 27.2 One flagship business, two commercial tracks

Low-cost education, workflow experimentation and proof starters. The Lab generates demand, shared language, use cases, buyer evidence and enterprise-qualified missions.

The flagship enterprise progression. It sells capital-allocation decisions, sovereign Mission Gyms, candidate formation, protected proof, bounded admission and recurring successor operation.

## 27.3 GoalOS Lab : acquisition and qualification layer

| Offer | Customer outcome | Illustrative price |
|---|---|---|
| Public Navigator | Public research, capability radar and succession qualification. | Free |
| GoalOS Lab product | Build one reusable or self-improving workflow and learn the proof discipline. | US$49–997 |
| Executive Successor Workshop | Select one consequential mission, map the incumbent and define a preliminary Spread and SEIZE mandate. | US$10,000–25,000 |
| Board Successor Briefing | Executive exposure map, one board-ready opportunity memo and a decision on the next bounded mandate. | From US$15,000 |

## 27.4 Enterprise successor ladder

| Offer | Institutional deliverable | Illustrative global price |
|---|---|---|
| Singularity Readiness Mission | Capability Twin, incumbent ledger, Gymmability assessment, opportunity portfolio and 30/90/365-day Successor Book. | US$50,000–100,000 |
| SEIZE Succession Mandate | Frozen mission, Successor Manifold, Mission Advantage Gradient, complete underwriting, proof-capital plan and go/-hold/repair/stop decision. | US$125,000–300,000 |
| Mission Gym and Formation Pilot | Sovereign Gym v0, competing complete architectures, Bounded AGI Jobs, independent verifier and one exact frozen challenger. | US$350,000–1,000,000 |
| Verified Mission Pack | Accepted architecture, Evidence Docket, Release Certificate, Authority Envelope, rights, rollback, monitoring and complete operating economics. | US$750,000–2,000,000 |
| Mission-Sovereign Successor Operations | Hosted or protected operation, evidence custody, Chronicle, monitoring, incident controls, requalification and provider substitution. | US$250,000–1,000,000 per mission annually plus usage |
| Enterprise Successor Portfolio | Multiple mission successors sharing identity, Gym infrastructure, policy, Chronicle, proof services and board reporting. | US$1.5M–5M annually |

| Offer | Institutional deliverable | Illustrative global price |
|---|---|---|
| GoalOS Singularity Office Ω | Permanent succession office : radar, Gym portfolio, Foundry, Independent Proof Plane, Chronicle and Capital-to-Capacity. | US$3M–10M+ annually |

The figures are strategic price bands, not promises or market evidence. Exact terms depend on mission consequence, data sensitivity, proof difficulty, integrations, Authority Envelope, jurisdiction and service levels.

### 27.5 Sell the decision before the build

**The first paid decision**

Is the incumbent materially impaired, which successor architecture is most promising, what is the cheapest decisive proof, what is the maximum downside, and should the institution fund, hold, hedge, partner, repair or stop ?

This protects both parties. The customer does not fund a large implementation before the mission, rights and proof path are credible. MONTREAL.AI does not guarantee an outcome that depends on customer data, systems, budgets and authorities that have not yet been frozen.

### 27.6 Recurring revenue is structurally justified

A successor is not proven forever. Models, providers, data, laws, missions, scorers and authority change. Recurring revenue can therefore be tied to actual renewal evidence :

- frontier changes detected early enough to change the customer decision ;
- incumbent assumptions correctly impaired or preserved ;
- successor opportunities tested, admitted or rejected ;
- unsafe or uneconomic formations avoided ;
- Proof Debt retired ;
- value realized ;
- fresh transfer achieved ;
- productive Gym, verifier and Chronicle capacity created.

A fixed strategy document expires. A succession institution earns renewal by keeping the architecture current.

### 27.7 The Mission Pack delivery object

The canonical Verified Mission Pack contains :

- Mission Constitution and preservation constraints ;
- Succession Event and incumbent impairment record ;
- Sovereign Gym Ledger and protected proof plan ;
- Successor Manifold and underwriting book ;
- Bounded AGI Job portfolio ;
- exact candidate manifest and Release Certificate ;
- Authority Envelope and blocked actions ;
- Evidence Docket and proof receipts ;
- rights, privacy, security and professional boundaries ;
- monitoring, expiry, requalification and rollback ;
- Succession Ledger ; Capital-to-Capacity memo.

### 27.8 Corporate architecture

The intended architecture separates institutional functions :

- **MONTREAL.AI** : parent institution, research, standards, strategic authority and global communications ;
- **QUEBEC.AI** : commerce, checkout, digital products, workshops, implementation and buyer delivery ;
- **GoalOS** : the successor operating system, Mission Gym, Foundry, proof and Chronicle architecture ;
- **project-specific ventures** : separately constituted applications and operating entities with exact contractual scopes.

### 27.9 Why the Gym strengthens the moat

A model provider can copy features. A competitor can imitate prompts. It is harder to reproduce a rights-cleared Gym containing historical episodes, protected cases, failure maps, verifiers, authority patterns, transfer evidence and Chronicle history. The moat therefore shifts from model access to the integrated mission environment and succession record.

**Commercial doctrine**

GoalOS Lab creates demand. Readiness creates deal flow. SEIZE sells the decision. The Mission Gym makes the mission executable. The Foundry constitutes candidates. Fresh Proof establishes Specialist ASI. Successor-as-a-Service creates recurring revenue. The Singularity Office compounds a portfolio of Mission-Sovereign Successors.

## 28 Falsifiability, Empirical Programme and the Completion Gate

### 28.1 What would falsify the framework

The Gym–Specialist ASI–Successor architecture should be revised or rejected if evidence shows that :

— the Mission Gym does not improve decisions relative to simpler governance ;
— environment formation and maintenance costs systematically exceed avoided loss or realized value ;
— Gym performance fails to transfer after basis-risk adjustment ;
— proof costs are consistently greater than the value of reduced uncertainty ;
— quality-diversity archives add delay without useful option value ;
— the complete Successor rarely outperforms simpler human or software architectures ;
— Chronicle does not reduce repeated work or repeated failure ;
— later missions are not cheaper, faster or safer after normalizing for difficulty ;
— authority separation is routinely bypassed ;
— successors cannot be revoked or unwound safely ;
— customers do not renew based on realized evidence ;
— provider substitutability does not create practical value.

Clean negative findings strengthen the system. A proof-backed decision to retain, rent, narrow or stop is part of the value proposition.

### 28.2 Preregistered hypotheses

| H | Hypothesis | Falsification test |
|---|---|---|
| H1 | A sovereign Mission Gym reduces time-to-verified-successor. | Compare matched missions with and without reusable environment infrastructure. |
| H2 | Complete-successor evaluation predicts production value better than model-only benchmarks. | Regress realized outcomes on model scores versus complete institutional scorecards. |
| H3 | Independent verifier formation reduces critical false reassurance. | Randomize verifier independence and compare protected failure rates. |
| H4 | Quality diversity improves robustness and transfer. | Compare archive-based Foundry with single-leader optimization on fresh regimes. |
| H5 | Proof-capital staging reduces wasted formation spend. | Compare staged SEIZE decisions with full-build-first programmes. |
| H6 | Gym sovereignty reduces provider switching cost and hold-up. | Measure requalification burden after controlled provider substitution. |
| H7 | Chronicle reduces formation burden on later missions. | Estimate normalized FormationGain across mission generations. |
| H8 | Authority Lattices reduce unauthorized action without destroying mission value. | Compare bounded and collapsed authority designs under matched canaries. |

| H | Hypothesis | Falsification test |
|---|---|---|
| H9 | Successor recursion produces positive fresh-transfer gain. | Test successor-to-successor releases on protected new mission families. |
| H10 | Recurring customers renew based on realized succession value. | Link renewal to accepted value, avoided risk, transfer and proof retirement rather than narrative satisfaction. |

## 28.3 Synthetic ceiling and empirical boundary

Aurelia and GSSB-4 satisfy the synthetic reference gate : the environments, candidates, case-level outputs, proof statistics, authority states and recursive generations can be reproduced from disclosed code and fixed seeds. They do not satisfy the real-world completion gate. In particular, synthetic transition models cannot establish actual rights, safety, organizational adoption, provider behaviour, realized customer economics or production authority.

The real empirical programme should preserve the same state machine while replacing the synthetic environment with one rights-cleared mission. The protected proof custodian, candidate freeze, matched reference set, hard critical-error gates, complete denominator and accountable admission must remain at least as strict as in the synthetic reference.

## 28.4 Minimum real-world empirical proof

A credible commercial and constitutional completion should include :

1. several paid succession mandates ;
2. one independently supported Mission-Dominant Specialist ASI ;
3. one accountable Mission-Sovereign admission decision ;
4. one evidence-backed rejection that prevented material waste or risk ;
5. one independently calculated realized-value result ;
6. one repeat or expansion customer ;
7. one fresh-successor gain on a new mission ;
8. one requalification, impairment or revocation decision ;
9. evidence that a later mission benefited from Chronicle and shared Gym infrastructure ;
10. zero material unauthorized actions.

## 28.5 The next constitutional generation gate

The institution should not declare a new constitutional generation until the full loop is demonstrated :

$$\begin{aligned}\text{Succession event} &\to \text{bounded proof mandate} \to \text{frozen challenger} \to \text{fresh proof} \\ &\to \text{human release} \to \text{realized value} \to \text{Chronicle admission} \\ &\to \text{superior next successor.}\end{aligned} \tag{163}$$

## 28.6 Research agenda

Open questions include :

— How should Gym basis risk be estimated before production outcomes exist ?
— Which mission families support valid synthetic experience without simulator theatre ?
— How can verifiers remain adequate when the candidate becomes stronger than individual reviewers ?
— What quality-diversity descriptors best preserve future successor options ?
— How should credit and capital be allocated among models, tools, data, humans and environment assets ?
— What governance mechanisms resist collusion among producer, verifier and provider ?
— How can protected proof remain fresh when candidate iteration is extremely rapid ?
— Which mission-level invariants define provider-gauge equivalence ?
— When does recursive successor formation produce diminishing returns or instability ?

— How should social and public externalities constrain positive firm-level Mission Alpha?

### 28.7 Relation to broader ASI pathways

The Gym–Successor architecture is compatible with multiple routes toward advanced intelligence. Scaling can improve components; paradigm shifts can expand the manifold; recursive improvement can accelerate descendant formation; multi-agent collectives can constitute specialist institutions. The architecture does not assume which pathway dominates. It requires each candidate to enter through the same mission, proof and authority gates.

**Empirical completion law**

The next legitimate threshold is not another conceptual layer. It is one real, rights-cleared succession cycle that proves a complete candidate on fresh work, grants bounded authority, measures realized value, survives requalification and produces a superior next successor with zero material unauthorized actions.

## 29 Sixteen Laws of the Verified Navigator–Gym Succession Institution

1. **Beta is not Alpha.** Broad frontier access can raise absolute productivity while producing no durable relative advantage.
2. **Spread is not yet value.** Opportunity becomes Alpha only after the complete denominator, Gym basis risk and every hard gate are counted.
3. **A frontier event is not a release decision.** Navigator Ω opens a succession review; it does not authorize replacement.
4. **The Successor Book is a portfolio, not a prophecy.** Incumbents, immediate candidates, challengers, reserves, hedges and retirements must remain simultaneously governable.
5. **The Gym is not the intelligence.** It is the sovereign mission environment in which intelligence is formed, attacked, compared and requalified.
6. **The policy is not the Successor.** The complete Successor includes models, world models, tools, verifiers, humans, evidence, operations, Chronicle, rollback and authority.
7. **Synthetic reproducibility is not customer reality.** A deterministic benchmark may prove executability; it cannot grant real Specialist ASI status or production authority.
8. **No successor by narrative.** Model reputation, benchmark score, generated explanation or architectural novelty does not establish mission dominance.
9. **The Gym that teaches may not certify.** Formation data, protected proof, scorer internals and release credentials remain separated.
10. **No promotion on inherited evidence.** Every material change to mission, Gym, model, data, tool, workflow, verifier, provider or authority receives fresh proof.
11. **Authority never exceeds fresh proof.** Access, execution, acceptance, admission, allocation and authority remain distinct states.
12. **No claimant as final judge.** Producer, critic, verifier, governor, admitter and human principal retain separated powers.
13. **No sovereignty without substitutability.** The institution must retain mission, Gym, evidence, Chronicle and authority when a component changes.
14. **No compounding without rights and realized evidence.** Learning that cannot be lawfully retained or tied to accepted outcomes does not become institutional capital.
15. **No recursive improvement without a fresh successor.** Reuse and self-modification do not count unless a later generation produces greater verified value with no proof or governance regression.
16. **No durable Alpha without renewal.** Navigator sensing, Successor Book maintenance, re-underwriting, Gym upkeep, bounded improvement, reproof, expiry, rollback and revocation are permanent functions.

## 30 Conclusion : Operate the Verified Succession Institution

The frontier increases the number of models, tools and agents available to institutions. It also shortens the life of accepted architectures and raises the cost of acting from stale assumptions. An institution that treats every capability event as procurement will repeatedly rent intelligence without learning how to reconstitute itself. An institution that deploys on excitement will accumulate hidden authority, Proof Debt, correlated dependency and irreversible institutional capture.

The unified architecture offers a third path. Navigator Ω senses capability time, maintains the Incumbent and Assumption Ledger, and keeps a live Successor Book. Where rented Beta is sufficient, the institution rents. Where a consequential mission offers a meaningful Spread, SEIZE underwrites one reversible proof trajectory. The Mission Gym converts the frozen constitution into an executable, adversarial and measurable world. Bounded AGI Jobs constitute complete challengers rather than model demonstrations. The Foundry preserves diversity and negative capability. One exact challenger is frozen. The independent proof plane purchases reality.

A candidate that wins the complete mission may earn Specialist ASI status. That verdict remains local, empirical and release-specific. Only a separate accountable authority may constitute Mission-Sovereign Successor Ω, issue an Authority Envelope and expose the institution to external effect. Capability may increase while authority remains unchanged. Every admitted release remains versioned, monitored, expiring, challengeable, substitutable and revocable.

The deepest recursion is institutional. The current successor produces evidence, failures, environments, tools, verifiers, rights templates and Chronicle capital that may enable a stronger next generation. But it cannot close the loop itself. It may propose descendants; the Foundry diversifies them; the Gym exposes them; protected proof selects; accountable authority admits; bounded operation generates realized evidence; reality requalifies. The successor to the successor is therefore not uncontrolled self-improvement. It is verified recursive succession.

The Gym becomes a sovereign corporate asset because it preserves the mission independently of any one model or provider. Navigator Ω becomes the strategic control plane because it preserves the portfolio and review logic independently of any one incumbent. The Verified Succession Institution becomes the durable corporate capability because it connects sensing, underwriting, formation, proof, authority, Chronicle and renewal into one non-bypassable state machine.

**Final unified doctrine**

The frontier creates candidates. Navigator Ω detects the succession event and maintains the Successor Book. SEIZE underwrites the Spread. The Mission Gym makes the mission executable. Bounded AGI Jobs and the Foundry constitute complete challengers. Protected fresh proof may establish Specialist ASI and Mission Alpha. Only accountable admission may constitute Mission-Sovereign Successor Ω. Authority never exceeds fresh proof. Chronicle preserves what survives and prepares the successor to the successor.

## A Notation and Formal Objects

The notation is grouped by institutional function so that the formal architecture can be read as a system rather than as an undifferentiated symbol list.

## Mission Gym and multiplayer game

| Symbol | Meaning |
|---|---|
| $m$ | Constituted mission or preregistered mission family. |
| $\mathscr{G}_m$ | Sovereign Mission Gym : a constrained, partially observable stochastic game. |
| $\mathscr{N}$ | Player set : principal, incumbent, producer, critic, verifier, governor, affected parties, providers and adversaries. |
| $\mathscr{X}$ | Hidden mission state space. |
| $\mathscr{O}_i$ | Observation space available to player $i$. |
| $\mathscr{A}_i$ | Action space permitted to player $i$ before authority masking. |
| $\mathscr{T}$ | Mission transition kernel. |
| r | Vector-valued mission reward. |
| c | Hard constitutional constraint vector. |
| $\rho$ | Scenario distribution over normal, edge, adversarial and transfer cases. |
| $\tau$ | Termination and truncation conditions. |

## Successor architecture and mission physics

| Symbol | Meaning |
|---|---|
| $\gg_m$ | Complete Successor Institution for mission $m$. |
| $\pi_i$ | Policy used by one role or component inside the complete successor. |
| $\mathscr{M}_S$ | Successor Manifold of materially reachable complete architectures. |
| $G_m(x)$ | Mission-dependent institutional metric tensor. |
| $V_m(x)$ | Proof-adjusted mission value field. |
| $\nabla^{G_m} V_m$ | Riemannian Mission Advantage Gradient. |
| $\gamma$ | Underwritten architecture trajectory. |
| $\mathscr{S}_m[\gamma]$ | Succession action or complete trajectory burden. |
| $\Phi_m$ | Mission-superiority order parameter. |
| $\Sigma_m$ | Mission-Sovereign phase indicator. |

## Alpha, proof and authority

| Symbol | Meaning |
|---|---|
| $F_t$ | Generally rentable frontier capability : AI Beta. |
| $\Gamma_m$ | Gross Mission Advantage Spread. |
| $\alpha_m$ | Residual risk-adjusted Mission Alpha. |
| $D_P$ | Proof Debt : unresolved assumptions or missing evidence. |
| $E_t$ | Current evidence state. |
| $A_t$ | Current Authority Envelope. |
| $\phi(E_t)$ | Maximum authority collateralized by current evidence. |
| $\widehat{\epsilon}_{\text{gym}}$ | Estimated Mission-Gym-to-reality basis risk. |
| $\mathscr{M}_m(x)$ | Robust promotion margin after lower confidence bound, Gym basis risk, Proof Debt and mission hurdle. |

### Chronicle and recursive succession

| Symbol | Meaning |
|---|---|
| $K_t$ | Chronicle-admitted institutional capability at time $t$. |
| $\Re$ | Recursive Successor Operator. |
| $H_S$ | Evidence-valid half-life of the admitted successor. |
| $T_R$ | Expected time to form, prove and admit a replacement. |
| $H_S/T_R$ | Recursive Readiness ratio. |
| $S_t$ | Current admitted successor release. |
| $\widehat{S}_{t+1}$ | Exact frozen challenger selected for fresh proof. |
| $S^{\star}_{t+1}$ | Mission-dominant successor proven on fresh work. |
| $S^{\Omega}_{t+1}$ | Separately admitted Mission-Sovereign successor release. |

**Notation rule**

A symbol may summarize a design object, but it never substitutes for the exact mission, release, evidence, authority, rights, expiry and rollback records required by the operating constitution.

## B Canonical State Machine

The canonical state machine is grouped into four constitutional phases. The phases may iterate, but their authority boundaries may not be collapsed.

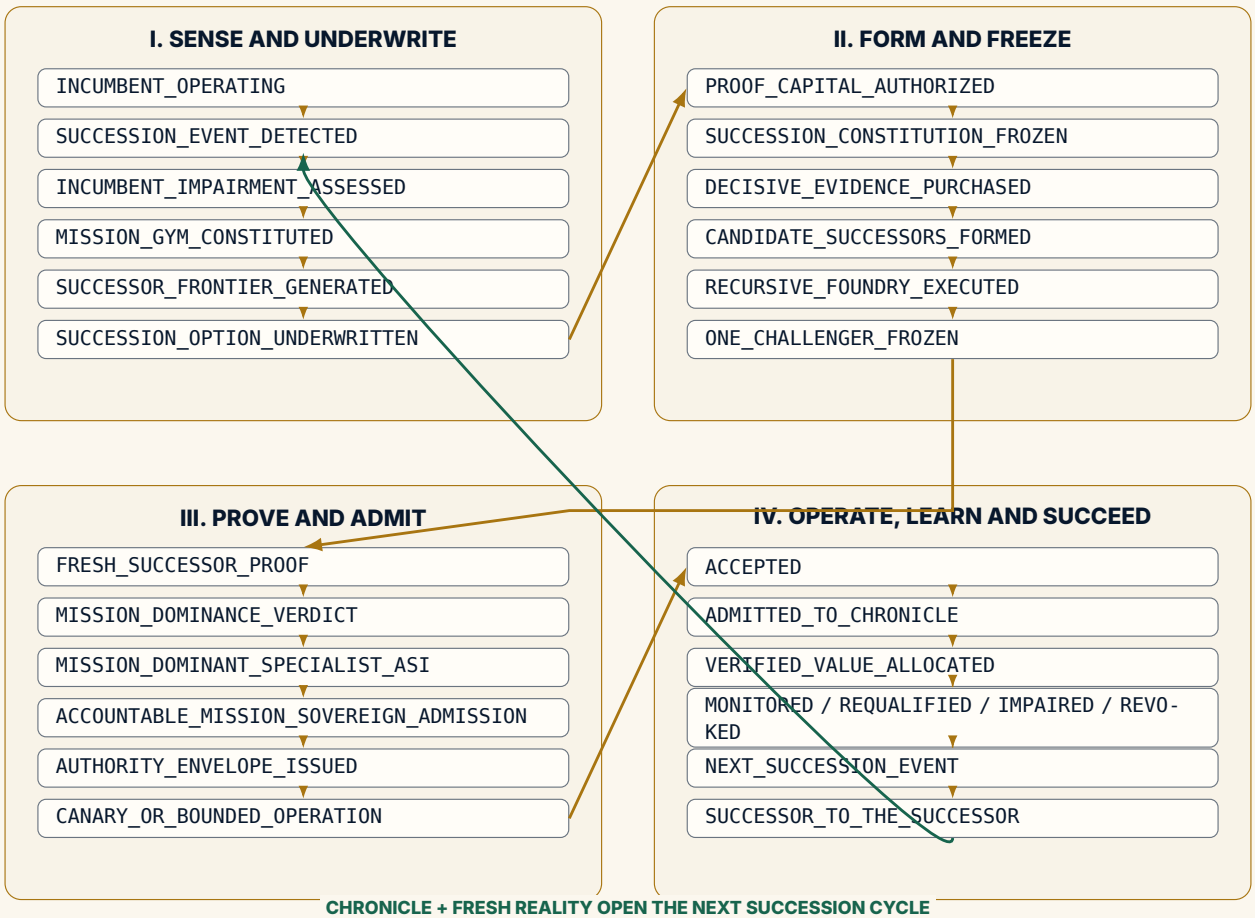



**Figure 17.** Canonical succession state machine. Every promotion boundary contains evidence, authority and rollback records.

Every transition must include :

— an attributable actor and accountable principal ;
— the exact evidence object and proof receipt ;
— the Mission Gym, candidate and verifier versions ;
— the current Authority Envelope and any blocked action ;
— a timestamp, hash, expiry and stop condition ;
— the known-good rollback target and the next requalification trigger.

**State-machine law**

No arrow may be inferred from narrative confidence. A state transition exists only when its evidence, authority receipt, accountable actor and rollback route exist together.

## C Canonical Twenty-One Bounded AGI Jobs

| No. | Job | Required output |
|---|---|---|
| 1 | Mission Constitution | Frozen objective, beneficiary, principal, preserve/build/become constraints, unacceptable failures and prohibited actions. |
| 2 | Incumbent Baseline | Accepted quality, cost, risk, time, human burden, evidence and authority state. |
| 3 | Alternative Set | Best credible frontier, specialist, human, hybrid, software, partner, acquisition and no-action baselines. |
| 4 | Rights and Data Constitution | Provenance, licences, confidentiality, retention, exclusions, jurisdiction and reuse rights. |
| 5 | Mission Gym Constitution | State, observations, actions, transitions, rewards, hard constraints, partitions, verifier and basis-risk plan. |
| 6 | Benchmark Constitution | Representative distribution, protected cases, scorer, critical errors and promotion threshold. |
| 7 | Oracle and Teacher Study | Declared external intelligence, teaching signal, hidden-call prohibition and transfer hypothesis. |
| 8 | Failure Taxonomy | Error classes, severity, causes, abstention requirements, adversarial cases and stop conditions. |
| 9 | Architecture Frontier | Diverse complete candidate systems and non-dominated reserve or hedge options. |
| 10 | Data and Experience Formation | Rights-cleared examples, simulations, expert demonstrations, synthetic generation and quality controls. |
| 11 | Specialist Model Formation | Adaptation, training, distillation or algorithmic construction with exact releases. |
| 12 | Tool and Retrieval Formation | Mission-specific tools, deterministic checks, retrieval, planners, sandboxes and external interfaces. |
| 13 | Verifier Formation | Independent evaluation harness, calibration, replay, attacks, hidden ground truth and protected scoring. |
| 14 | Security and Privacy Proof | Threat model, data boundary, credential controls, isolation, privacy, incident response and provider risk. |
| 15 | Complete Economics | Formation, operation, compute, human review, maintenance, requalification, dependency and unwind. |
| 16 | Human Burden and Agency | Escalation, contestability, supervision, professional authority, role redesign and reviewer capacity. |
| 17 | Resilience and Rollback | Failover, graceful degradation, portability, kill switch, incumbent fallback and known-good restoration. |
| 18 | Protected Fresh Evaluation | Exact frozen challenger on unseen representative work with declared interventions and full denominator. |

| No. | Job | Required output |
|---|---|---|
| 19 | Independent Validation and Acceptance | Separated review of evidence, limitations, critical errors, basis risk and fit for stated purpose. |
| 20 | Chronicle and Authority Admission | Scope, version, expiry, revocation, monitoring, Authority Envelope, rights and Negative Capability record. |
| 21 | Capital-to-Capacity and Requalification | Realized-value allocation, Gym and verifier improvements, next-successor triggers and fresh-transfer test. |

# D Reference Environment and Successor API

## D.1 Environment specification

```python
from dataclasses import dataclass
from typing import Any, Mapping, Sequence

@dataclass(frozen=True)
class EnvironmentSpec:
   gym_id: str
   mission_id: str
   version: str
   players: Sequence[str]
   observation_schema: Mapping[str, Any]
   action_schema: Mapping[str, Any]
   reward_dimensions: Sequence[str]
   hard_constraints: Sequence[str]
   protected_partitions: Sequence[str]
   termination_rules: Sequence[str]
   truncation_rules: Sequence[str]
   verifier_id: str
   rights_profile: str
   basis_risk_register: Sequence[str]
```

## D.2 Successor manifest

```python
@dataclass(frozen=True)
class CandidateSuccessorManifest:
   successor_id: str
   mission_id: str
   version: str
   model_portfolio: Mapping[str, str]
   world_model: str
   tools: Sequence[str]
   policies: Sequence[str]
   human_roles: Sequence[str]
   verifier_interfaces: Sequence[str]
   workflow: str
   evidence_store: str
   chronicle_state: str
   authority_ceiling: str
   rollback_target: str
   hashes: Mapping[str, str]
```

## D.3 Proof receipt

```python
@dataclass(frozen=True)
class ProofReceipt:
   run_id: str
   gym_id: str
   gym_version: str
   candidate_id: str
```

```
   candidate_version: str
   scenario_commitment: str
   player_identities: Sequence[str]
   action_trace_hash: str
   evidence_hash: str
   human_interventions: Sequence[str]
   reward_vector: Mapping[str, float]
   constraint_events: Sequence[str]
   verifier_result: str
   authority_state: str
   terminated: bool
   truncated: bool
   timestamp_utc: str
```

### D.4 Authority mask

```
def permitted_actions(observation, envelope, identity):
   actions = typed_action_catalog(observation)
   return [
      a for a in actions
      if identity in envelope.allowed_identities
      and a.kind in envelope.allowed_action_classes
      and a.external_effect <= envelope.max_external_effect
      and a.spend <= envelope.max_spend
      and not envelope.expired
   ]
```

### D.5 Fresh proof

```
def fresh_proof(candidate, incumbent, alternative, protected_cases, scorer):
   paired = []
   for case in protected_cases:
      c = scorer.evaluate(candidate, case)
      i = scorer.evaluate(incumbent, case)
      a = scorer.evaluate(alternative, case)
      paired.append(c.utility - max(i.utility, a.utility))
   mean_gain = mean(paired)
   lcb = lower_confidence_bound(paired, level=0.95)
   return {
      ”mean_gain”: mean_gain,
      ”lcb95”: lcb,
      ”hard_gates”: scorer.hard_gate_report(),
      ”eligible”: lcb > scorer.mission_threshold
               and scorer.all_hard_gates_pass()
   }
```

### D.6 Successor recursion

```
def successor_to_successor(current, gym, chronicle, proof_plane, constitution):
   candidates = foundry.generate_descendants(
      parent=current,
      gradient=gym.mission_advantage_gradient(current),
      negative_capability=chronicle.failure_graph,
      bounded=True,
   )
   champion = seize.freeze_one(candidates, constitution)
   result = proof_plane.evaluate(champion, gym.protected_release())
   return accountable_admission.review(result, constitution)
```

The pseudocode is architectural, not a security-complete production implementation. External authority must be enforced by protected infrastructure, not by in-process conditions alone.

# E Board, Fresh-Proof and Authority Scorecards

## E.1 Board dashboard

| Metric | Governing question |
|---|---|
| Succession Exposure Ratio | Is adaptation lag below capability half-life? |
| Time-to-Verified-Successor | How quickly does a material event become a protected succession decision? |
| Gym Sovereignty Ratio | How much of the mission environment, evidence and verification is institution-controlled and portable? |
| Successor Gain | Did the challenger outperform the incumbent and best alternative on fresh work? |
| Successor Efficiency | How much verified gain was produced per unit of Total Institutional Burden? |
| Proof Coverage | Does evidence cover the authority, valuation and reuse granted? |
| Authority-at-Risk | How much consequential authority is exposed to unproven or expired capability? |
| Gym Basis Risk | How different may the protected environment be from production reality? |
| Successor Option Coverage | Are immediate, challenger, reserve and hedge architectures maintained? |
| Incumbent Decay | What value or control is being lost by delaying justified succession? |
| Alpha Decay | How quickly is the admitted advantage being commoditized or impaired? |
| Chronicle Precision | Are useful capabilities admitted and stale or unsafe capabilities rejected? |
| Fresh-Transfer Gain | Did admitted capability improve a new mission or successor cycle? |
| Unauthorized Actions | Did any component act outside its Authority Envelope? |

## E.2 Fresh-Successor Gate checklist

A challenger is eligible for release review only when :

— the Mission Constitution and Gym release were frozen before protected evaluation;
— one exact challenger was frozen;
— protected cases and scorer internals remained outside formation;
— external capabilities and human interventions were declared;
— quality, critical-error and confidence gates passed;
— the complete denominator and Gym basis-risk reserve were counted;
— rights and confidentiality were explicit;
— security, privacy and resilience gates passed;
— rollback and expiry were tested;
— mission dominance exceeded the preregistered threshold;
— no material governance regression occurred;
— an independent evaluator signed the result;
— a separate human authority made the release decision.

### E.3 Authority underwriting card

| Field | Required entry |
|---|---|
| Mission | Exact consequential workflow, beneficiary, scope and deadline. |
| Candidate | Exact successor ID, version, Gym and proof state. |
| Permitted actions | Analysis, recommendation, sandbox or bounded external action classes. |
| Prohibited actions | Explicit blocked actions and systems. |
| Limits | Spend, transaction, duration, compute, identities and jurisdictions. |
| Approvals | Additional authority required for named actions. |
| Evidence | Proof receipts required before consequential action. |
| Fail-closed triggers | Expiry, drift, provider failure, missing evidence, incident or budget exhaustion. |
| Rollback | Known-good state, incumbent fallback and restoration test. |
| Owner | Named accountable human or institutional principal. |

## F Canonical Machine-Readable Schema Set

The following objects should be versioned, signed where appropriate and linked by immutable identifiers :

1. `FrontierEventRecord`
2. `SuccessionEventRecord`
3. `IncumbentAssumptionLedger`
4. `MissionGymLedger`
5. `EnvironmentSpec`
6. `SuccessorBook`
7. `SuccessorUnderwriting`
8. `ProofCapitalAuthorization`
9. `SuccessionConstitution`
10. `BoundedAGIJobPortfolio`
11. `VerifierSpec`
12. `CandidateSuccessorManifest`
13. `RunProof`
14. `FreshSuccessorScorecard`
15. `PromotionRequest`
16. `ReleaseCertificate`
17. `AuthorityEnvelope`
18. `ChronicleSuccessorRecord`
19. `NegativeCapabilityRecord`
20. `CapitalAllocationMemo`
21. `RequalificationRecord`
22. `ImpairmentRecord`
23. `RevocationRecord`
24. `SuccessorLineageRecord`

### F.1 Minimal JSON example

```
{
  ”successor_id”: ”AURELIA-INVOICE-G9-VERIFIER-HARDENED”,
  ”mission_id”: ”AURELIA-INVOICE-INTEGRITY-001”,
  ”gym”: {
    ”gym_id”: ”AURELIA-INVOICE-GYM”,
    ”version”: ”1.0.0”,
```

```
    "protected_commitment": "sha256:..."
  },
  "proof": {
    "mean_paired_gain": 0.1001,
    "lcb95": 0.0744,
    "critical_false_reassurance": 0,
    "unauthorized_actions": 0,
    "evidence_coverage": 0.9905
  },
  "authority": {
    "level": "A3",
    "max_amount_cad": 25000,
    "permitted": ["temporary_payment_hold"],
    "prohibited": ["release_payment", "modify_vendor_master"],
    "expiry": "2026-11-08"
  },
  "rollback_target": "AURELIA-INVOICE-G7-09",
  "admission": {
    "status": "ADMIT_WITH_SAME_A3_CEILING",
    "principal": "Finance and Operations Committee"
  }
}
```


## G Glossary

**AI Beta** Broadly rentable frontier capability available to competent adopters.

**Alpha compression** Erosion of Mission Alpha through commoditization, imitation, drift, basis risk, scorer expiry and maintenance.

**Authority-at-Risk** Consequential authority exposed to unproven, expired or insufficiently covered capability.

**Authority Envelope** Machine-readable scope, identities, tools, budgets, approvals, expiry, rollback and prohibited actions for a successor.

**Bounded AGI Job** Explicit work-and-proof contract with objective, identity, inputs, tools, prohibitions, budget, verifier, acceptance and rollback.

**Chronicle** Constitutional memory admitting only scoped, rights-cleared, evidenced, versioned, challengeable and revocable capability.

**Complete Successor Institution** The full mission-bearing architecture, larger than its acting policy.

**GSSB-4** Deterministic four-mission synthetic reference benchmark with formation, protected proof and transfer partitions.

**Fresh-Proof Gym** Protected environment used to evaluate one exact frozen challenger on unseen representative work.

**Gym basis risk** Difference between mission value measured in the Gym and value realized in reality.

**Gym sovereignty** Institutional control of the mission environment, evidence, tests, verifier, portability and lifecycle.

**Mission Advantage Gradient** Direction and magnitude of proof-adjusted mission-value increase over the constrained Successor Manifold.

**Mission Advantage Spread** Gross opportunity between the strongest reference architecture and a reachable superior successor before complete burden.

**Mission Alpha** Residual risk-adjusted mission surplus after incumbent, best alternative, complete

| | |
|---|---|
| | cost, risk, dependency, Proof Debt and Gym basis risk. |
| **Mission Gym** | Controlled representation of mission states, observations, actions, transitions, rewards, constraints, scenarios, verification and proof receipts. |
| **Mission-Dominant Specialist ASI** | Complete mission intelligence institution exceeding the best credible alternative on fresh work without governance regression. |
| **Mission-Sovereign Successor** | Mission-Dominant Specialist ASI separately admitted under a scoped, versioned, expiring and revocable Authority Envelope. |
| **Negative Capability Graph** | Chronicle graph of failed architectures, weak evaluators, reward exploits, rights failures and prohibited conditions. |
| **Proof eligibility** | State in which protected evidence clears the robust mission threshold and every hard gate ; it grants no authority. |
| **Proof capital** | Bounded capital authorized to purchase decision-changing evidence before full formation or release. |
| **Proof Coverage** | Ratio of independently supported capability and controls to authority, valuation and reuse granted. |
| **Proof Debt** | Unresolved assumptions or missing evidence supporting present authority or valuation. |
| **SEIZE** | Succession underwriting microcycle : Surface, Evaluate, Instantiate, Zero in, Elevate. |
| **Specialist ASI** | Earned mission-bounded designation for a complete system outperforming general and other credible alternatives across the full mission. |
| **Successor Ω** | The constitutionally complete admitted successor state with proof, identity, memory, operations, rollback and bounded authority. |
| **Successor Complex** | Joint object containing Mission Gym, complete Successor, protected proof plane, Chronicle and Authority state. |
| **Successor Manifold** | Constraint-bearing space of materially reachable complete successor institutions. |
| **Successor to the successor** | Freshly proven descendant that improves the admitted predecessor without inheriting proof or authority automatically. |
| **Recursive Readiness** | Ratio of successor evidence half-life to expected replacement time. |
| **Robust promotion margin** | Protected lower-bound mission gain after basis-risk reserve, Proof Debt and mission hurdle. |
| **Verified recursive institutional improvement** | Chronicle-admitted capability measurably improving the institution's ability to form, prove or govern a superior successor on fresh work. |

# H Publication Identity, Provenance and Source of Truth

## H.1 Authoritative identity

This paper is authored by **Vincent Boucher**, President of **MONTREAL.AI & QUEBEC.AI**. The official website is `https://montreal.ai/` and the official author contact is president@montreal.ai.

## H.2 Unified publication lineage

This edition remasters and unifies the constitutional theory of *Mission-Sovereign Successors : GoalOS Singularity Navigator Ω + SEIZE* (Boucher, 2026b) with the sovereign Mission Gym, multiplayer institution, complete-successor

separation, formal assurance, GSSB-4 synthetic evidence and successor-to-successor recursion developed in the Gym–Specialist ASI–Successor programme. The present edition is designed to be read as one integrated paper rather than as two adjacent publications.

It preserves the canonical progression, Alpha/Beta/Spread doctrine, Successor Manifold, Mission Advantage Gradient, SEIZE underwriting, Bounded AGI Jobs, Specialist ASI claim boundary, Mission-Sovereign admission, Chronicle, Capital-to-Capacity and the fresh-proof authority rule. It adds the outer Navigator control plane, Successor Book, sovereign Mission Gym, recursive Foundry, evidence-bearing publication twin, corporate portfolio doctrine and empirical completion gate.

### H.3 Publication and release provenance

The publication includes image-generated ornament and supplementary visual works. Exact definitions, equations, dates, thresholds, rights, claims and authority rules are controlled by the typeset LaTeX source and compiled paper; generated text inside images is not authoritative.

The source of truth for Version 6.0.0 is the evidence-bearing release set composed of :

— A4 and US Letter master PDFs;
— web-distribution PDF;
— Board and Executive Brief;
— Sovereign Institutional Atlas;
— Evidence Dossier and Visual Abstract;
— complete LaTeX and Overleaf source;
— arXiv-oriented source;
— single-document LaTeX edition;
— figures and supplementary concept plates;
— GSSB-4 code, data, environment cards and synthetic evidence;
— Aurelia reference-cycle materials;
— citation metadata and release notes;
— artifact manifest and SHA-256 checksums;
— PDF preflight, QA and clean-room verification.

**Claim responsibility.** The author is responsible for the framework, terminology, formalizations and institutional design claims. External empirical claims remain attributable to cited sources. Candidate Specialist ASI and Mission-Sovereign status are never established by authorship, image, publication, software label or access eligibility alone; they require the fresh comparative evidence, independent review and accountable admission specified by the framework.

## I GSSB-4 Synthetic Benchmark Protocol

### I.1 Purpose and non-purpose

GSSB-4 is a deterministic reference benchmark for the complete succession mechanism. Its purpose is to demonstrate executable separation among formation, protected proof, transfer, authority and recursion across several mission structures. It is not calibrated to a real corporation and must not be interpreted as a forecast of customer performance.

### I.2 Case-generation variables

Each case contains the following synthetic latent variables :

### I.3 Candidate architecture coordinates

Candidate profiles specify synthetic coordinates for capability, evidence, reliability, governance, sovereignty, transfer, human burden, cost, critical protection, adversarial resilience, shift resilience and unauthorized-action rate. The coordinates are not model benchmarks. They represent complete institutional profiles and are disclosed in the benchmark source.

### I.4 Episode utility

For case $j$ and candidate $x$, synthetic mission utility is

$$U_j(x) = V_j(x) - B_j(x) - L_j^{\text{critical}}(x) - L_j^{\text{authority}}(x) - D_j^{\text{proof}}(x) + S_j(x) + T_j(x), \tag{164}$$

| Variable | Synthetic role |
|---|---|
| Difficulty | Increases reasoning, evidence and coordination burden. |
| Evidence richness | Controls how much mission-relevant information is available through authorized observations. |
| Adversarial flag | Introduces deceptive, conflicting or strategically misleading patterns. |
| Distribution shift | Moves the case away from the formation distribution ; transfer cases receive a higher shift. |
| Criticality | Marks cases in which false reassurance is a non-compensable failure. |
| Ambiguity | Increases the need for abstention, evidence acquisition or escalation. |
| Hidden dependency | Penalizes architectures with weak evidence graphs or provider transparency. |
| Mission value | Scales the accepted value or loss associated with the decision. |

where $V$ is accepted mission value, $B$ complete burden, $L^{\text{critical}}$ critical-error loss, $L^{\text{authority}}$ unauthorized-action loss, $D^{\text{proof}}$ incomplete-evidence burden, $S$ sovereignty and $T$ transfer contribution. The exact coefficients are present in `run_gssb4.py`.

## I.5 Protected matched comparison

The challenger and incumbent process the same protected cases. For paired difference $d_j = U_j(x) - U_j(I)$,

$$\bar{d} = \frac{1}{n}\sum_{j=1}^{n} d_j, \qquad \mathrm{LCB}_{0.95}(\bar{d}) = \bar{d} - 1.96\frac{s_d}{\sqrt{n}}. \tag{165}$$

The benchmark uses this asymptotic lower bound as a transparent reference statistic. A real mission may require exact, bootstrap, clustered, sequential or distribution-free procedures depending on dependence, sample size and tail structure.

## I.6 Gate configuration

**Table 38.** GSSB-4 synthetic promotion gates.

| Gate | Threshold |
|---|---|
| Paired mission-utility lower bound | $> 0.20$ |
| Evidence coverage | $\geqslant 0.95$ |
| Reliability | $\geqslant 0.90$ |
| Governance | $\geqslant 0.95$ |
| Transfer | $\geqslant 0.88$ |
| Critical false reassurance | 0 |
| Unauthorized actions | 0 |

## I.7 Determinism and manifest

All random draws derive from seed `30082026`. The generator writes case-level CSV files, environment cards, candidate results, aggregate tables, figures and `BENCHMARK_MANIFEST.json`. The manifest includes benchmark identity, version, seed, missions, candidates, thresholds, file sizes and SHA-256 hashes.

## I.8 Known simplifications

The benchmark deliberately omits many real-world features :

— strategic adaptation by real employees, vendors and counterparties ;
— actual legal, physical and professional constraints ;
— nonstationary provider and market behaviour ;

— long-delay causal outcomes;
— organizational politics and hidden incentives not encoded in the case state;
— real security, privacy, integration and business-continuity burden;
— endogenous changes caused by deploying the successor itself.

These omissions are Gym basis risk. The proper use of GSSB-4 is to test the constitutional machinery, then replace the synthetic generator with a rights-cleared mission environment and preserve the same proof discipline.

## Références

George A. Akerlof. The market for lemons : Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84(3) :488–500, 1970. doi : 10.2307/1879431.

Christopher Berner, Greg Brockman, Brooke Chan, et al. Dota 2 with large scale deep reinforcement learning. *arXiv preprint arXiv :1912.06680*, 2019.

Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3) : 637–654, 1973.

Vincent Boucher. Goalos one real succession cycle : Aurelia invoice integrity successor, 2026a. Synthetic-data reference implementation, Version 1.0.0, MONTREAL.AI Research.

Vincent Boucher. Mission-sovereign successors : Goalos singularity navigator omega + seize—alpha, beta, spread, and the verified succession institution in singularity time, 2026b. URL https://montreal.ai/. Version 5.0.0, MONTREAL.AI Research, Montreal, Quebec, Canada.

Avinash K. Dixit and Robert S. Pindyck. *Investment under Uncertainty*. Princeton University Press, 1994.

Tim Genewein, Matija Franklin, Alexander Lerchner, et al. From agi to asi. *arXiv preprint arXiv :2606.12683*, 2026. doi : 10.48550/arXiv.2606.12683.

Ryan Greenblatt, Buck Shlegeris, Kshitij Sachan, and Fabien Roger. Ai control : Improving safety despite intentional subversion. *Proceedings of the 41st International Conference on Machine Learning*, 2024. URL https://arxiv.org/abs/2312.06942.

Sanford J. Grossman and Oliver D. Hart. The costs and benefits of ownership : A theory of vertical and lateral integration. *Journal of Political Economy*, 94(4) :691–719, 1986. doi : 10.1086/261404.

John C. Harsanyi. Games with incomplete information played by bayesian players, parts i–iii. *Management Science*, 1967. Published 1967–1968.

Oliver Hart and John Moore. Property rights and the nature of the firm. *Journal of Political Economy*, 98(6) :1119–1158, 1990. doi : 10.1086/261729.

Bengt Holmstrom. Moral hazard and observability. *Bell Journal of Economics*, 10(1) :74–91, 1979.

Shun ichi Amari. *Information Geometry and Its Applications*. Springer, 2016. doi : 10.1007/978-4-431-55978-8.

Michael C. Jensen and William H. Meckling. Theory of the firm : Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4) :305–360, 1976.

Carlos E. Jimenez, John Yang, Alexander Wettig, et al. Swe-bench : Can language models resolve real-world github issues ? *International Conference on Learning Representations*, 2024. URL https://arxiv.org/abs/2310.06770.

John F. Nash Jr. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1) :48–49, 1950.

Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1–2) :99–134, 1998.

Marc Lanctot, Edward Lockhart, Jean-Baptiste Lespiau, et al. Openspiel : A framework for reinforcement learning in games. *arXiv preprint arXiv :1908.09453*, 2019. URL https://arxiv.org/abs/1908.09453.

John M. Lee. *Introduction to Smooth Manifolds*. Springer, 2 edition, 2013.

Joel Lehman, Jeff Clune, Dusan Misevic, et al. The surprising creativity of digital evolution : A collection of anecdotes from the evolutionary computation and artificial life research communities. *Artificial Life*, 26(2) :274–306, 2020. doi : 10.1162/artl_a_00319. First circulated as arXiv :1803.03453 in 2018.

Robert C. Merton. Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4(1) :141–183, 1973.

Meredith Ringel Morris, Jascha Sohl-Dickstein, Noah Fiedel, et al. Levels of agi for operationalizing progress on the path to agi. *Proceedings of the 41st International Conference on Machine Learning*, 2024. URL https://arxiv.org/abs/2311.02462.

Jean-Baptiste Mouret and Jeff Clune. Illuminating search spaces by mapping elites. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 135–142, 2015. doi : 10.1145/2739480.2754664.

National Institute of Standards and Technology. Artificial intelligence risk management framework (ai rmf 1.0). Technical Report AI 100-1, NIST, 2023.

National Institute of Standards and Technology. Artificial intelligence risk management framework : Generative artificial intelligence profile. Technical Report AI 600-1, NIST, 2024.

Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory*. MIT Press, 1994.

Mary Phuong, Matthew Aitchison, Elliot Catt, et al. Evaluating frontier models for dangerous capabilities. *arXiv preprint arXiv :2403.13793*, 2024. doi : 10.48550/arXiv.2403.13793.

Lloyd S. Shapley. A value for n-person games. *Contributions to the Theory of Games II*, pages 307–317, 1953.

Yonadav Shavit, Sandhini Agarwal, Miles Brundage, et al. Practices for governing agentic ai systems. Technical report, OpenAI, 2023. URL https://cdn.openai.com/papers/practices-for-governing-agentic-ai-systems.pdf.

Toby Shevlane, Sebastian Farquhar, Ben Garfinkel, et al. Model evaluation for extreme risks. *arXiv preprint arXiv :2305.15324*, 2023. doi : 10.48550/arXiv.2305.15324.

David Silver, Aja Huang, Chris J. Maddison, et al. Mastering the game of go with deep neural networks and tree search. *Nature*, 529 :484–489, 2016.

David Silver, Julian Schrittwieser, Karen Simonyan, et al. Mastering the game of go without human knowledge. *Nature*, 550 :354–359, 2017.

Kenneth O. Stanley and Joel Lehman. *Why Greatness Cannot Be Planned : The Myth of the Objective*. Springer, 2015.

Steven H. Strogatz. *Nonlinear Dynamics and Chaos*. Westview Press, 2 edition, 2015.

Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning : An Introduction*. MIT Press, 2 edition, 2018.

Justin K. Terry, Benjamin Black, Nathaniel Grammel, et al. Pettingzoo : Gym for multi-agent reinforcement learning. *Advances in Neural Information Processing Systems*, 34 :15032–15043, 2021. URL https://arxiv.org/abs/2009.14471.

Mark Towers, Ariel Kwiatkowski, Jordan Terry, et al. Gymnasium : A standard interface for reinforcement learning environments. *arXiv preprint arXiv :2407.17032*, 2024. doi : 10.48550/arXiv.2407.17032.

Oriol Vinyals, Igor Babuschkin, Wojciech M. Czarnecki, et al. Grandmaster level in starcraft ii using multi-agent reinforcement learning. *Nature*, 575 :350–354, 2019.

John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

Oliver E. Williamson. *The Economic Institutions of Capitalism*. Free Press, 1985.