THE STUDIO · VOLUME IV

Governing Judgment

A First-Principles Framework for Decision Provenance

in the Deflationary Economy

A Discussion Paper

Richard Leclezio
Leclezio Consulting Corporation

May 2026

REFERENCE IMPLEMENTATION · KEEL v0.2 · LIVE IN PRODUCTION

About this paper

Governing Judgment is Volume IV in the Studio series at richardleclezio.com. It is a working paper, not a marketing document. It proposes an original framework — **Decision Provenance** — for protecting strategic judgment in an economy where the cost of executing an idea is collapsing faster than the discipline to choose the right ideas. The framework is presented vendor-neutral; the reference implementation, **KEEL**, is now in production and is described in Section 07.

> **STATUS — PHASE 0 COMPLETE**
>
> This is not a thought experiment. Since the first draft of this framework, the reference implementation has been built and deployed to production. KEEL v0.2 runs live, with five operating engines, a full CEO operating system, and a nine-table append-only ledger, on real Claude. What does not yet exist is outcome-confirmed data: no design partner has yet run a live commitment through to a recorded outcome. The framework is therefore deployed but pre-validation — built, honest about what it cannot yet prove, and ready for one design partner. That distinction is maintained throughout this paper.

Suggested citation

Leclezio, R. (2026). Governing Judgment: A First-Principles Framework for Decision Provenance in the Deflationary Economy. Discussion Paper. Leclezio Consulting Corporation, May 2026.

Position in the Studio series

This volume sits alongside three earlier working papers that have addressed adjacent layers of the same problem — governing an AI-transformed enterprise:

- **Volume I — The Token Economy:** a first-principles playbook for governing the cost of Claude in production.
- **Volume II — Managing Latency and Hallucination in Enterprise AI Agents:** a framework for governing the reliability of agentic systems.
- **Volume III — Becoming Discoverable:** how interactive AI-powered profiles are rewriting professional identity in the age of conversational machines.

Volume IV completes the arc. Volumes I–III govern, in turn, the *spend*, the *reliability*, and the *identity* layers of AI transformation. This volume addresses the layer those three depend on but none of them solve: the *judgment* of the human operator deciding what to do with the cheapened, faster, more legible world the first three describe.

Contents

From the Author

The first three volumes of the Studio series were written, in different ways, about the same thing: the parts of an enterprise that AI changes mechanically — what it costs, how reliably it runs, how identity is presented through it. Each is a tractable problem. Each has an architecture. Each is solvable.

This volume is harder, because it is about the part of the enterprise that AI does not touch directly and yet quietly threatens most: the judgment of the person at the top. The CEO. The operating partner. The leader making the bets that determine whether the cheaper, faster, more legible world the first three volumes describe becomes a compounding advantage or a faster route to value destruction.

I have spent the last several years working at the operating layer of AI inside financial services — first as a project manager governing generative-AI programs, then as an advisor on data and risk regimes, then as a builder of agentic systems. The pattern that surfaced across those engagements is the one this paper is about. The technology kept getting better. The discipline of the humans deciding what to do with it did not. The gap between those two curves is where most of the value is being lost right now, and where the next decade of operational alpha will be earned or surrendered.

This paper is the framework I wish had existed when I first encountered the gap. It is offered as a working draft, not a final word. Comments, critique, and counterexamples are welcomed.

Richard Leclezio · New York, May 2026

Cold Open

It is eleven o'clock on a Tuesday night. A CEO is about to destroy three million dollars of value, and no one in the building will stop her.

Six weeks ago she committed her company to an AI automation pilot. The thesis was clean: automating forty percent of invoice reconciliation removes one-point-two million dollars of annual headcount cost on an eighteen-month payback, with a ninety-three-percent accuracy threshold she declared non-negotiable before Phase 2. Three-point-two million dollars of capital. The board approved it. She wrote the kill-criteria herself: automation accuracy below ninety-three percent in production, payback projection past twenty-four months, board approval of Phase 2 by the third quarter. The investment committee signed off at eighty-five percent conviction.

Tonight, none of those criteria have triggered. Automation accuracy is at ninety-three point four percent. Payback is tracking at nineteen months — comfortably inside the twenty-four-month ceiling. The reallocation of freed headcount, the part of the thesis that compounds, has not even begun.

What has changed is the conditions around her. A board member asked pointed questions on this afternoon's call. The sector is jittery. A competitor is circulating doubt in the trade press. Her phone has not stopped vibrating since six.

At eleven thirty she will pull the program. She will tell herself, and the board, and her CFO, and eventually her therapist, that the conditions had changed and she was being prudent. None of that will be true. What changed was not the data. What changed was the room.

She is not stupid. She is not weak. She is operating without an instrument that, until recently, did not exist: a system that remembers what she committed to and why, that knows the difference between her data changing and her feelings changing, and that is empowered to confront her with the gap in the moment she is about to step across it.

> *This paper is about the design of that instrument — and, as of this volume, its arrival.*

01 The Judgment Deficit

The cost of executing an idea is collapsing. The cost of choosing the wrong idea is not.

1.1 Two curves, moving in opposite directions

Volume I of the Studio series — *The Token Economy* — argued that the price of inference is collapsing in a way that fundamentally re-rates which workloads can be afforded inside an enterprise. Volume II made the parallel case for reliability: the engineering effort to make an agentic system production-grade has fallen by an order of magnitude in eighteen months. The combined effect is what this volume calls the **deflationary execution cycle**: the era in which any clearly-specified idea can be built, tested, and shipped for a fraction of what it cost two years ago.

This is, on its face, a wonderful thing. It is also, on second inspection, the most dangerous environment for decision-making the modern enterprise has ever encountered. When execution is cheap, the constraint on the firm shifts from the question can we build it to the question are we making the right bets, fast enough. The marginal return on disciplined judgment rises. The marginal cost of undisciplined judgment rises with it, because mistakes that used to take six months to manifest now appear in six weeks, and the speed of competitive response leaves no room to recover.

> **THE JUDGMENT DEFICIT, DEFINED**
>
> The gap between the speed and quality of decisions a firm could make if it were operating with disciplined judgment, and the speed and quality of decisions it is actually making. In a slow-execution economy the deficit was a tax on returns. In the deflationary economy it is the difference between compounding and being compressed out.

1.2 Why judgment does not scale with technology

The dashboards have proliferated. The models have improved. The advisory firms have published their AI playbooks. None of this has appreciably improved the quality of strategic decisions at the top of the firm, for three structural reasons.

First, the leader is drowning, not starving. The scarce resource is not information; it is synthesis, dissent, and the will to act on uncomfortable conclusions. More dashboards make the deficit worse, not better.

Second, judgment fails predictably under pressure. The literature on cognitive bias is forty years old. The professional class has read it, nodded at it, and continues to make the same errors at the same rates because reading about bias does not, as it turns out, materially reduce bias in the moment it counts.

Third, the human systems surrounding a leader are structurally incapable of correcting her in the moment. The chief of staff serves her. The board meets quarterly. The consultants are paid by the engagement. The people closest to the leader are also, by design, the people least likely to confront her at the moment of reversal.

1.3 The shape of the failure

Across hundreds of post-mortems published by private equity operating teams, management consultancies, and corporate strategy groups in the last decade, the same two failure modes appear with monotonous regularity.

- **Capitulation under pressure.** A bet was made at high conviction, the thesis was sound, the kill-criteria did not trigger, but a non-evidence pressure — a bad quarter, a hostile board call, a sector dislocation, a competitor announcement — produced a reversal. The firm paid the carrying cost of the original bet and the reversal cost, then watched a competitor who held position compound the gains.
- **Commitment escalation under sunk cost.** A bet was made at moderate conviction, the kill-criteria triggered, the thesis was visibly dead, but the firm continued investing because the cost of admitting the error exceeded — emotionally, politically, careerally — the cost of continuing the error.

The deficit, in other words, is two-sided. Firms lose money by quitting good bets at the wrong moment, and they lose money by refusing to quit bad bets at the right moment. Every existing solution in the market addresses one side at most. This paper argues that the only system worth building addresses both — and that the only credible architecture for doing so begins with a discipline that, until this reference implementation, did not exist in the enterprise stack: **Decision Provenance**.

1.4 The deficit, in working figures

The figures below are the framework's working estimates, drawn from post-mortem literature and adjacent-domain studies. They are illustrative, not yet validated against the reference implementation's own outcome data — which by design does not yet exist. They are offered to size the problem, and to be argued against.

THE JUDGMENT DEFICIT — WORKING FIGURES

Term	Definition
~68%	Estimated share of material strategic reversals driven primarily by non-evidence pressure rather than triggered kill-criteria.

Term	Definition
~2.1×	Illustrative exit-multiple gap between leaders who held high-conviction theses through pressure and those who reversed under it.
0	Existing enterprise tools that record the decision state — thesis, conviction, pre-committed exit criteria — rather than only the decision.
1	Labeled training row generated by every single intercept, forever. The dataset compounds; see Section 06.

02 Why traditional advisory cannot close it

Each of the three incumbents — consultants, dashboards, generic AI — addresses some symptom of the deficit. None addresses the mechanism.

2.1 Consultants

Strategy consulting is the oldest professional response to the judgment deficit. It works well in the static problem domains it was designed for: a discrete decision, a defined window, a deliverable that lands and the engagement closes. It does not work for the continuous discipline this paper is about. Three structural reasons:

- The economic model is project-based. The consultant is paid for the engagement, not the long-tail outcome of the decisions made downstream of it. There is no contractual or financial mechanism by which the consultant is accountable for a reversal made twelve months after the deck was delivered.
- The cadence is wrong. Strategy is a continuous sensing-and-deciding loop; consulting is an episodic intervention. The most consequential reversals tend to happen between engagements, exactly when the consultant is not in the room.
- The incentive structure is wrong. The consultant's next engagement depends on the current client being happy. Confronting a CEO mid-reversal is not a happiness-producing act. The system, set up correctly, never confronts.

2.2 Dashboards and business intelligence

Dashboards solve a different problem: visibility. They are necessary infrastructure, but they do not address the deficit. A dashboard tells the leader what is happening. The deficit is about what to do about it, under what conditions, with what discipline against what prior thesis. Dashboards have no memory of intent, no record of conviction, and no opinion. They are arrows; the leader is the bow.

2.3 Generic conversational AI

It is tempting to assume that the third wave of generic AI — the chat assistants now in every office — will close the gap. They will not, for reasons that map closely to the failures of consulting and dashboarding combined. A general-purpose assistant has no persistent memory of the firm's commitments. It has no obligation to confront, because confrontation is an unsafe behavior for a product whose retention depends on user satisfaction. And it has no labeled history of which kinds of reversals turn out, with

hindsight, to have been right and which turn out to have been wrong. It is, structurally, a sophisticated agreement engine. The judgment deficit is not closed by sophisticated agreement.

> **A STRUCTURAL TEST FOR ANY PROPOSED SOLUTION**
>
> Three questions surface whether a candidate system can close the deficit. First: does it record the commitment in a form it can confront the leader with later? Second: is it empowered to dissent in both directions — against hubris and against capitulation? Third: does its accuracy improve with use, in a way no competitor can replicate? A no to any of the three disqualifies the candidate.

2.4 What an adequate solution must do

The remainder of this paper builds a system that answers yes to all three. Section 03 introduces the discipline of *Decision Provenance* — the recording of judgment in a form that can be re-confronted later. Section 04 introduces the *Dual Dissent* protocol, the two-sided red-team against both failure modes. Section 05 reframes strategy through the *Deflation Lens*, the rule-set for choosing bets under collapsing execution cost. Section 06 develops the *Portfolio Mirror*, the cross-firm labeled-judgment dataset that makes the system unreplicable. Section 07 specifies the reference architecture, now deployed. Sections 08 and 09 address deployment and the assumptions on which the whole edifice rests.

2.5 Where the categories stand

The clearest way to see the gap is to set the existing tool categories — operating-system frameworks of the EOS / Traction school, OKR software, and board portals — against the structural test above. The reference implementation is the only column that records the decision state rather than only the decision.

CAPABILITY BY CATEGORY

Term	Definition
Record decision state at commitment	KEEL: yes (thesis + conviction + kill-criteria). EOS/OKR/board portals: no.
Intercept reversal attempts live	KEEL: yes (sub-second classification). EOS/OKR/board portals: no.
Classify fear vs. analysis reversal	KEEL: yes (three-class with confidence). EOS/OKR/board portals: no.
Dual adversarial review before a hold	KEEL: yes (Bull + Bear synthesis). EOS/OKR/board portals: no.
30/60/90-day outcome review	KEEL: yes (auto-scheduled). OKR tools: manual at best. Others: no.
Decision-quality labeling	KEEL: yes (immutable, per-outcome). EOS/OKR/board portals: no.

Term	Definition
Board sentiment as structured data	KEEL: yes (signal timeline per member). Board portals: unstructured only. Others: no.
Cross-portfolio query in one place	KEEL: yes (full-context chief-of-staff chat). EOS/OKR/board portals: no.
Labeled dataset that trains a model	KEEL: yes (grows with every company). EOS/OKR/board portals: no.

03 Decision Provenance — the original framework

Audit trails for data are a forty-year-old discipline. The same discipline, applied to human judgment, has never been built.

3.1 The borrowed metaphor

In data governance — the field the author spent several years in, governing risk data under regimes like BCBS 239 — the central insight is that data without provenance is data that cannot be trusted. A number on a report is meaningless unless the lineage is known: where it originated, what transformations were applied, who signed off at each step, what assumptions were embedded. Provenance is the audit infrastructure that makes the data legible to its consumers and accountable across its life.

Strategic judgment, made by humans at the top of firms, is treated today as if it required no such infrastructure. The decision is announced; the discussion that produced it is partial and unrecorded; the assumptions under which it was made are remembered selectively if at all; the conditions under which it should be revisited are rarely articulated and almost never written down. When the decision is later reversed, the reversal is treated as a fresh event, evaluated against the new conditions only — not against the conditions of the original commitment.

This paper proposes a different posture. Every material strategic decision should be recorded with the same rigor as a material data point. The recording is what we call **Decision Provenance**.

3.2 The Decision Provenance schema

A Decision Provenance record is the minimum viable description of a judgment such that the judgment can be re-confronted later. It contains, at a minimum, the following nine fields. Schema is illustrative; the canonical reference appears in Appendix B.

DECISION PROVENANCE — THE NINE FIELDS

Term	Definition
Decision name	A short, retrievable identifier for the bet or commitment.
Thesis	The leader’s reasoning, in her own words, for why this is the right bet.
Conviction	A numeric score (0–100) of the leader’s confidence at commitment, supplied by the leader.
Capital committed	The financial and non-financial resources allocated. Non-financial matters: attention is a budget.

Term	Definition
Commitment data snapshot	The state of the world — internal metrics and external signals — at the moment of commitment.
Kill-criteria	Pre-committed, falsifiable conditions under which the bet will be reversed without further debate.
Expected J-curve	The leader's explicit acknowledgment of the period of discomfort she will hold through.
Reviewers	The named individuals who will be re-shown the original thesis if a reversal is attempted.
Provenance log	An append-only record of every subsequent event affecting the decision, including reversal attempts.

3.3 Pre-mortem commitment: the load-bearing innovation

Of the nine fields, one is doing more work than the others. The *kill-criteria* field is what transforms Decision Provenance from documentation into discipline. It is the act of writing down, in advance of pressure, the exact conditions under which the leader has authorized her future self to reverse the bet — and, by inversion, the exact conditions under which her future self will not be authorized to reverse it.

The intuition is the same one that powers Ulysses tying himself to the mast. The leader at the moment of commitment is the most lucid agent the firm has access to in relation to this decision. The leader eight weeks later, under pressure, is a different and noticeably less lucid agent. Pre-mortem commitment binds the second agent to the judgment of the first. It is the cheapest, most powerful intervention in the entire framework.

> **THE PRE-MORTEM COMMITMENT**
> The disciplined act, at the moment of high-conviction commitment, of writing down the falsifiable conditions under which the bet will be reversed. The criteria must be observable, time-bound, and specific enough that a future reader can determine, without further debate, whether they have triggered.

3.4 Reversal classification: the verdict states

Every reversal attempt against a Decision Provenance record is classified into one of three states. The classification is not the reversal itself — the leader retains decision authority — but it is the verdict on whether the reversal is evidence-driven or condition-driven.

- **Analysis reversal.** One or more pre-committed kill-criteria have triggered, or material new adverse facts have emerged that were not anticipated at commitment. The reversal is evidence-driven; the discipline is to reverse, and the reversal is recorded as such.
- **Fear reversal.** No kill-criterion has triggered. The data underneath the bet has not materially changed. What has changed is the conditions around the leader — sentiment, pressure, recency, herd behavior. The reversal pattern-matches to capitulation under non-evidence pressure, and the discipline is to confront.
- **Mixed signal.** The picture is genuinely ambiguous; some kill-criteria are stressed but not triggered, some new facts are adverse but not decisive. The discipline is to revisit the thesis explicitly rather than reverse silently. A mixed-signal verdict is a request for a deliberate re-commitment, not a reversal.

The classification step is the load-bearing artifact. It is what converts a continuous stream of moods, pressures, and opinions into a discrete, auditable judgment about the nature of the reversal — and gives the leader something to hold herself against. In the reference architecture (Section 07), the classification is initially produced by a large language model operating against a structured rubric, and over time graduates to a learned classifier as outcome data accumulates.

3.5 The honesty constraint

A system that classifies every reversal as a fear reversal becomes, immediately and permanently, a manipulation tool. The leader stops trusting it; the firm stops using it; the value disappears. The honesty constraint is therefore the load-bearing ethical commitment of the entire framework. When a kill-criterion has genuinely triggered, the system must return an analysis-reversal verdict and recommend reversing, even if doing so contradicts the leader's prior conviction. A system that protects the leader from her own past judgment when her judgment was wrong is as destructive as a system that fails to protect her from her own current panic when her panic is unjustified.

> *Capitulation and hubris are equal failure modes. A system that guards against only one becomes the instrument of the other.*

04 The Dual Dissent — guarding hubris and capitulation

Most red-team practices in the enterprise are one-directional. They guard against bad bets being made. They do nothing to guard against good bets being abandoned.

4.1 The asymmetry in existing practice

Pre-mortems, war-gaming, and devil's-advocate protocols are all standard tools in mature strategy organizations. They share a single direction of fire: they all attack the leading recommendation at the moment of commitment. They ask, with varying degrees of sophistication: *what could go wrong with this bet?* This is useful. It is also half the job.

The other half — and, by the evidence of post-mortems across two decades of strategy outcomes, the more expensive half — is the dissent against reversal. When a leader proposes to abandon an existing bet, no symmetrical practice exists. There is no formal protocol that asks: *what could go wrong with the reversal?* The board nods, the CFO updates the model, the announcement is drafted. The cost of abandoned bets that should have been held is never measured against the cost of pursued bets that should have been killed, because the discipline that would measure it has never been built.

4.2 The Dual Dissent protocol

The protocol introduced in this paper applies symmetric, structured dissent to both events.

- **At commitment:** a Hubris Dissent is run. The protocol attacks the leading recommendation, surfaces the assumptions that must hold for the thesis to be true, and registers a confidence-weighted bear case. The Hubris Dissent becomes part of the Decision Provenance record.
- **At reversal:** a Capitulation Dissent is run. The protocol attacks the proposed reversal, surfaces the original thesis, the kill-criteria status, and the evidence-versus-condition decomposition of the drivers, and registers a confidence-weighted hold case.

Both dissents are written, not verbal. Both are timestamped and attached to the Decision Provenance record. Both are read aloud, in their entirety, before the corresponding action is taken. The friction is the feature.

> **DUAL DISSENT IN ONE SENTENCE**
>
> A bet is committed only after a written argument against making it. A bet is reversed only after a written argument against reversing it. Both arguments live in the Decision Provenance record forever.

4.3 What changes when both directions are guarded

The empirical effect, observed across mature deployments of similar two-sided protocols in adjacent domains — trading desks running formal kill-and-hold protocols, clinical-trial committees with explicit stopping and continuation rules — is not that fewer reversals happen. It is that the reversals that do happen are concentrated in the cases where reversal is correct. The signal-to-noise ratio of the firm's strategic motion improves. Bets are killed faster when the kill-criteria trigger; bets are held longer when the pressure is non-evidence; the carrying cost of indecision falls toward zero.

This is the operational alpha of disciplined judgment. It does not require the leader to be smarter, calmer, or more contrarian. It requires only that the system around her be structured to convert her judgment into auditable artifacts, and to confront her, in writing, in the moment.

05 The Deflation Lens — strategy under collapsing execution cost

Most strategy frameworks were built for an era in which execution was the constraint. That era is ending. The frameworks should follow.

5.1 The inherited frameworks

The dominant strategy frameworks taught in business schools and applied by management consultancies — Porter's Five Forces, the BCG growth-share matrix, the McKinsey horizons, where-to-play and how-to-win — were all developed in a world where execution was slow, costly, and risky. The implicit assumption is that the firm's scarce resource is the ability to do things, and the role of strategy is to allocate that scarce ability to the highest-return arenas.

In the deflationary execution cycle, the implicit assumption is no longer true. Execution is becoming abundant. The scarce resource is now the discipline to choose, and the speed to choose well. A strategy framework that does not foreground the collapse of execution cost is, by construction, mis-aimed.

5.2 The Deflation Lens

The Deflation Lens is not a replacement for the inherited frameworks. It is a reframing of their inputs. Every strategic question is rewritten through three sub-questions, asked in order.

THE THREE DEFLATION LENS QUESTIONS

Term	Definition
Where does it collapse?	Which parts of our cost structure will be subject to a step-change reduction in unit cost over the next thirty-six months, driven by AI, automation, and the broader deflationary execution cycle?
What do we reallocate to?	When the cost in those parts of the structure falls by 60–90%, where does the freed capital and attention go? Which adjacencies, which growth bets, which moat-building investments?
How fast must we move?	What is the window before competitors run the same compression, and what is our schedule for executing the reallocation inside that window?

The Lens does not produce the answers. It produces the correct questions. A strategy conversation that has not addressed all three of these is, by the standard advanced here, incomplete.

5.3 The error pattern the Lens corrects

In the absence of the Lens, the most common strategic error of the deflationary cycle is what this paper calls *symmetrical savings* — the firm cuts cost in the deflating part of the stack, books the savings as

margin, and reports the move as efficiency. The error is the failure to reallocate. The competitor who runs the same compression and ploughs the freed capital into growth compounds at the same time as the symmetrical-savings firm flattens, and within two to four years the gap is structural.

> **THE REALLOCATION IMPERATIVE**
>
> In a deflationary execution cycle, the question is never how much cost you cut. The question is what you do with the dollars and the people the cut frees up. Firms that book the savings get a one-time margin lift. Firms that reallocate the savings compound.

5.4 Why the Lens belongs in the same framework as Decision Provenance

The Deflation Lens is the strategic content of the framework; Decision Provenance is the governance scaffold around it. They are complementary by design. The Lens generates the bets worth making. Provenance ensures those bets are committed deliberately, recorded auditably, and reversed only on evidence. Either without the other is half the system. A firm that uses the Lens without Provenance generates good ideas and then capitulates on them at the first sign of pressure. A firm that uses Provenance without the Lens disciplines the wrong bets. The combination is the operating posture this paper is proposing.

06 The Portfolio Mirror — why the moat is unreplicable

The defensibility of the system proposed here does not rest on any single feature. It rests on a dataset that, by construction, no entrant can buy.

6.1 The three moats — state, data, trust

The defensibility of a system built around this framework rests on three moats, each stronger than the last. They are named for what they protect: the decision *state*, the labeled *data*, and the *trust* that makes the whole thing usable.

- **The state moat.** No other tool records the state in which a decision was made — the thesis in the leader's exact words, the conviction score, the pre-committed exit criteria, the emotional context. This is the precondition for everything else: you cannot confront a decision-maker with their prior reasoning if you never captured it. The state moat is what makes confrontation possible at all.
- **The data moat.** The labeled dataset — decision, state, outcome, and quality label — is the second moat. No competitor has it, because no competitor captures the state needed to create it. At roughly a thousand outcome-confirmed rows the dataset trains a real classifier; the longer the system runs, the more accurate it becomes. This is the compounding advantage developed in the rest of this section under its framework name, the *Portfolio Mirror*.
- **The trust moat.** The system is honest. When kill-criteria trigger, it recommends reversing; when fear is the driver, it says so plainly. A system that manufactured fear narratives to keep leaders committed would be noticed and discarded within weeks. Honesty at the classification layer is what makes the confrontation credible — and credibility, once earned across an organization, is itself slow and costly for a competitor to displace.

6.2 Why no new entrant can replicate the data moat

The Portfolio Mirror has three properties that make it functionally impossible to replicate.

1. **It is not for sale.** There is no broker market in labeled strategic decisions with outcome data. Firms do not sell, and would not sell, their commitment histories, their reversals, and the verdicts attached to each. The dataset has to be produced by the act of operating the framework; it cannot be purchased after the fact.
2. **It cannot be synthesized.** A foundation model can be trained on vast public text; it cannot be trained on the private history of how leaders at hundreds of firms actually committed, hesitated, capitulated, or held. The Mirror lives in the gap between what is observable from the outside and what is recordable only from inside the moment of decision.

3. **It compounds with every customer.** Each new firm that adopts the framework adds its labeled records to the aggregate. The system becomes more accurate at classifying reversal types, more useful at benchmarking conviction levels against peers, and more discriminating in its dissent — for every customer simultaneously. New entrants begin at zero, with no path to catch up except to wait for the incumbent to fail.

6.3 What the Mirror surfaces

Once enough outcomes accrue, the Portfolio Mirror benchmarks how leaders who held high-conviction theses under pressure fared against those who reversed. The figures below are the framework's illustrative target benchmark — the shape of the gap the dataset is expected to reveal, not yet a measured result, because no design partner has yet run a commitment through to outcome. They are presented to show what the Mirror computes, not to claim it has been validated.

PORTFOLIO MIRROR — ILLUSTRATIVE PEER-COHORT BENCHMARK

Term	Definition
Revenue, holders	+18% (illustrative) — leaders who held high-conviction theses through pressure.
Revenue, reversers	+3% (illustrative) — leaders who reversed under non-evidence pressure.
Margin recovery	Holders +8.2 pts vs. reversers −1.1 pts (illustrative).
Exit-multiple gap	~2.1× median (illustrative) — the spread the framework is built to protect.

6.4 The four decision-quality labels

Every outcome review writes one of four labels against the original verdict. These four classes are the ground-truth supervision signal for the eventual classifier — the thing that converts the Conviction Ledger from a record into a training set.

DECISION-QUALITY LABELS

Term	Definition
right_held	Held course; thesis validated. A fear reversal correctly avoided; exit multiple protected.
wrong_held	Held course; thesis failed. The signals were present and the system should have caught the inflection. The label the system must learn from most.
right_reversed	Reversed on evidence; kill-criteria had genuinely triggered. The correct, disciplined kill.

Term	Definition
wrong_reversed	Reversed on fear; the thesis would have validated. This is the precise loss the framework exists to prevent.

THE LABELED-JUDGMENT MONOPOLY

A trained fear-versus-analysis classifier requires three things: decisions, the cognitive state in which they were made, and the outcomes that followed. Only a system that operates the Decision Provenance discipline at scale, across many firms, with outcome tracking, can produce all three. The classifier that emerges is the monopoly.

6.5 The honest sequencing

The classifier described above does not exist yet. It cannot exist yet, because no outcome-confirmed labeled data exists yet — the reference implementation is deployed but pre-design-partner. The system's current classifier is therefore a large language model operating against a structured rubric, judging the evidence-versus-condition decomposition of each reversal with reasoning rather than learning. As Decision Provenance records accumulate and outcomes are tracked, the labeled set grows. At a threshold — on the order of a thousand outcome-confirmed records across a heterogeneous customer base — the rubric is augmented and then partly replaced by a learned classifier whose accuracy can be reported with statistical confidence.

This sequencing is not a weakness. It is the feature of the moat that distinguishes this system from anything a foundation-model vendor or a consulting incumbent could plausibly stand up in response. The first mover, having deployed the framework first, controls the only meaningful supply of training data in the category for as long as the framework is operating.

07 Reference architecture — the KEEL implementation

What follows is no longer a proposal. The reference implementation is built, deployed, and running in production. It is offered vendor-neutral; substitutions are possible at every layer.

7.1 Design philosophy

The reference implementation, called **KEEL**, is named for the part of a ship below the waterline that no one sees and that keeps the vessel from capsizing in heavy weather. The naming choice is deliberate. The system is meant to be invisible in calm conditions and load-bearing in storms. Four non-negotiable principles are encoded directly in the system, not merely aspired to.

- **The engine must stay honest.** When kill-criteria trigger, the verdict is analysis reversal and the recommendation is reverse. A system that always cries fear is a manipulation tool and destroys the trust premise entirely.
- **Decision support, never decision-maker.** No endpoint, copy, or surface implies KEEL decides. It classifies, confronts, and logs. The human chooses. This line is legally and ethically load-bearing and does not move.
- **The Conviction Ledger is the moat — protect its integrity.** Every reversal attempt, its classification, the emotional-state label, and the decision are recorded immutably. No code path lets a decision happen without a provenance row.
- **Provisional things stay labeled provisional.** The classifier is LLM-as-judge against a rubric, because no outcome-confirmed data exists yet. This is stated plainly wherever the verdict appears. With roughly a thousand outcome rows it trains a real model; until then, provisional.

7.2 The agent fleet — five engines

KEEL is built as a fleet of specialist engines, in the pattern Volume II of the Studio series argued was the right shape for governed agentic systems. Five engines are deployed and running on real Claude. Each has a narrow job, a defined interface, and a logged output.

THE DEPLOYED KEEL ENGINES

Term	Definition
01 Decision Provenance Engine	The spine. node_retrieve → node_check_kill_criteria → node_detect_drivers → node_classify. Strict JSON out. Mock fallback always on so the system never fails to run.

Term	Definition
02 Dual Dissent Engine	Two adversarial agents — Bull and Bear — run simultaneously against the same commitment, producing a synthesized verdict with an explicit key risk. Forces structured disagreement before any hold decision.
03 Chief-of-Staff Chat	Streaming responses with full portfolio context injected: all portfolio companies, all commitments, all intercept history, all board signals. Reasons across companies in a single turn.
04 Decision Memo Engine	A structured four-section memo — Situation, Provenance Classification, Bull/Bear, Recommendation — board-presentable and PDF-ready in roughly two seconds.
05 Outcome Review Engine	Auto-schedules 30/60/90-day review checkpoints on every decision and captures the verdict plus a decision-quality label. Each row is a labeled training sample for the eventual classifier.

The strategy-synthesis, external-sensing, and internal-performance agents described earlier in this framework (Section 05 and the data-lineage discussion) are the roadmap layer — partially realized through the Chief-of-Staff Chat's context injection and slated for full build-out after the first design partner. The five engines above are the production reality today.

7.3 The provenance pipeline

When a reversal is attempted against a commitment in the Conviction Ledger, the Decision Provenance Engine runs in fixed-order stages, each producing an auditable artifact. The current example commitment — Meridian Logistics' invoice-reconciliation pilot — classifies in under one second.

4. **Retrieve commitment.** The original thesis in the leader's own words, the conviction score at commitment, the capital deployed, and the kill-criteria as pre-committed and unmodified are loaded from the ledger.
5. **Audit kill-criteria.** Each pre-committed criterion is evaluated against live data. If any has triggered, the engine leans toward analysis reversal. If none have, that is the first signal of fear.
6. **Detect drivers.** The leader's stated reason is decomposed into evidence-based and condition-based drivers. The latter — recency bias, social/board pressure, loss aversion, herd behavior, fatigue — are named, not to shame but to surface.
7. **Classify.** claude-sonnet-4-6 issues a strict-JSON verdict: classification (fear_reversal, analysis_reversal, or mixed), confidence 0–100, a reasoning paragraph, and a recommendation (hold, reverse, or revisit_thesis). The prompt has never been a yes-machine.
8. **Confront.** The classification, reasoning, and recommendation are surfaced. The leader chooses to hold or override. Either way the decision is logged immutably and 30/60/90-day reviews are auto-scheduled.

7.4 The CEO operating system

Around the engines sits a complete operating console. Three supporting modules each feed signals back into the Decision Provenance Engine, so that the system's reading of a reversal attempt is informed by the accumulated context that preceded it.

THE CEO OPERATING SYSTEM MODULES

Term	Definition
My Week	Weekly priority history with avoidance detection — tasks moved consecutively without progress are flagged with a times-moved counter. The question it forces: what are you avoiding, and why?
My Team	Direct reports with their committed words from 1:1s, tracked on-track / at-risk / missed. Surfaces accountability patterns across the leadership team.
Board Room	CEO-to-board commitments with exact words quoted, board composition with real-time stance, countdown timers to each commitment, and a per-member signal log. The board never hears "I think I said" again.

The Board Signals module deserves particular note: board sentiment is captured as structured data — Comfortable, Neutral, Concerned, Escalated, Approved, Rejected — per member, per commitment. Sentiment is a leading indicator for reversal pressure. A Concerned signal today is frequently the precursor to a fear-reversal attempt thirty days out; capturing it in structured form means that when the reversal attempt arrives, the engine can reference the exact sequence of board signals that produced the pressure.

7.5 The decision arc and outcome reviews

Every commitment has a full lifecycle: committed, decided, outcome recorded. KEEL auto-schedules 30/60/90-day checkpoints, and each review captures what actually happened versus the thesis, scored against the original decision. The verdict plus the decision-quality label is the labeled training data. This is the moat in motion — and the reason the Outcome Review Engine, unglamorous as it is, is the most strategically important of the five.

7.6 The Conviction Ledger and the stack

The Conviction Ledger is the append-only data store in which Decision Provenance lives, now spanning nine tables (full schema in Appendix B). Three properties are worth emphasizing: it is append-only, so no record is altered or deleted; overrides are logged with full context, so the system never adapts its rubric to flatter a past override; and it is outcome-linked, so the reconciliation cadence produces the labels that train the eventual classifier.

DEPLOYED STACK

Term	Definition
Backend	FastAPI · Python · SQLite, Postgres-ready.
LLM	claude-sonnet-4-6 via the Anthropic API.
Frontend	Single-file HTML application, zero build step.
Deploy	Railway (nixpacks), auto-deploy on push. Running live in production at v0.2.
Orchestration	Structured as discrete nodes, designed 1:1 for a LangGraph StateGraph migration.

08 Where to deploy first — the operator wedge

The hardest decision in commercializing this framework is not how to build it. It is who to sell it to first.

8.1 Why not the public-company CEO

It is tempting to position the system as a tool for the CEO of a large public company. The reasoning is intuitive: that is where the largest decisions are made, that is where the value at stake is most material, that is where the failure modes this paper describes most visibly destroy value.

The intuition is correct. The go-to-market that follows from it is wrong. A sitting public-company CEO has the highest trust bar of any buyer in enterprise software, the smallest time budget, a human chief of staff already in place, and a relationship with at least one major consultancy on speed-dial. Selling a brand-new category to that buyer, cold, is the slowest and most expensive motion available.

8.2 The PE operating-partner wedge

The right wedge is a structurally different buyer with the same problem at higher concentration: the operating partner inside a private-equity firm. Five properties make this buyer the right first customer.

9. **Concentrated portfolio.** A mid-market PE firm holds fifteen to forty portfolio companies, each running an aggressive value-creation plan, most with no real strategy staff. The judgment deficit per company is high; the buyer-side aggregation is also high.
10. **Single buyer, many seats.** The operating partner commits once and deploys across the portfolio. The unit economics of the sale invert relative to selling individual CEOs.
11. **Mandate and access.** The operating partner has a contractual right to portfolio company data and a mandate to impose tools. The two hardest objections in the CEO sale — trust and data access — collapse.
12. **Aligned incentives.** Operating partners are paid on exit multiples. A system that catches one fear-driven reversal across a portfolio pays for itself many times over. The ROI narrative is short.
13. **Network expansion.** Once the framework is deployed across the portfolio, the same engine retargets, with minor modification, to the PE firm's own deal-screening and diligence function. The land-and-expand path is structural rather than improvised.

8.3 The land motion

The first product surface inside a portfolio company is not the full framework. It is an *AI-readiness diagnostic* — a structured assessment of where the company sits on the deflationary execution cycle, how

many of its operators are AI-fluent, and where the next thirty-six months of cost compression are likely to occur. The diagnostic is fast, defensible, and lands a sponsor inside the company. The full Decision Provenance framework follows the diagnostic by ninety to one hundred eighty days.

8.4 The expand motion

Once the framework is operating inside the portfolio, the second product surface is the cross-portfolio benchmark — the Portfolio Mirror — sold to the operating partner directly. This is the surface that is unreplicable. The third surface, eighteen to thirty-six months in, is the application of the same engine to the PE firm's own deal pipeline: every prospective acquisition is screened through the Decision Provenance frame before the investment committee, and post-acquisition the operating playbook is wired into the framework from day one.

8.5 What deployment looks like — one partner, three companies

The deployed system illustrates the wedge concretely. A single operating partner covers three portfolio companies simultaneously — each running a multi-million-dollar transformation, each CEO feeling board pressure at a different inflection point. KEEL intercepts every reversal attempt across all three, flags the fear-driven ones, validates the evidence-driven ones, and builds a provenance record that travels with the investment to exit.

THE DEPLOYED THREE-COMPANY PORTFOLIO

Term	Definition
Meridian Logistics	AI invoice-reconciliation pilot in the J-curve. Fundamentals tracking (93.4% accuracy, 19-month payback) but board confidence eroding. The watch item: the June board call, before Phase 1 completes.
Apex Manufacturing	Supplier adoption running nine points below plan at 71% against a 75% threshold and already over budget. If adoption misses by quarter-end, the inventory-improvement thesis genuinely breaks — an analysis reversal, not a fear one.
Vertex Software	Market expansion on track. No flags. The system suppresses rather than surfaces — editorial restraint is half the value.

The contrast across the three is the point: KEEL holds the line at Meridian (fear), would recommend reversing at Apex if the threshold breaks (analysis), and stays quiet at Vertex (no signal). One operating partner, three inflection points, one console — the bandwidth problem the wedge exists to solve.

8.6 The single ask

THE ASK

> One portfolio company. Thirty days. No integration required — KEEL runs in a browser. No IT dependency, no upfront data migration, no commitment beyond access to one CEO. The CEO gets a structured system for their single biggest active commitment; the operating partner gets visibility into every reversal attempt before it happens; and the Conviction Ledger begins building from day one. The ask is specific, the timeframe is fixed, and the downside is bounded.

8.7 Status and sequence

The framework is past the proposal stage. Where the earlier draft described a system to be built, the reference implementation is now deployed.

- **Phase 0 — complete.** The full operating system is built and live: five engines on real Claude — Decision Provenance, Dual Dissent, Chief-of-Staff Chat, Decision Memo, Outcome Review — plus the CEO operating system and a nine-table ledger. Production-deployed, not a pitch deck.
- **Phase 1 — now.** One PE design partner, one portfolio company, thirty days. No integration, access to one CEO, thirty days of live provenance data. This is the gate the entire venture turns on.
- **Phase 2 — after.** Multi-company deployment. Cross-portfolio intelligence activates, the Portfolio Mirror benchmark becomes real, the labeled dataset begins growing, and the moat becomes visible and measurable.

09 What must be true — the falsifiable assumptions

Every framework should publish the conditions under which it would be wrong. This one publishes three.

9.1 Governed data access at sufficient depth

The framework assumes that, with a PE sponsor's contractual right, governed access to a portfolio company's financials, board materials, and operational metrics can be obtained quickly enough that the system's outputs are non-generic. If this assumption fails — if the data access in practice is too shallow, too slow, or too contested — the system's outputs degrade to advisory boilerplate and the framework loses its claim to differentiation. The author's view is that this assumption holds in the PE-portco wedge and would not hold in a direct sale to public-company CEOs, which is one of the reasons the wedge is structured as it is.

9.2 Operating-partner willingness to act on the output

The framework assumes that the operating partner will not route around the system with her own judgment once it is in place — that the Capitulation Dissent will be read, the verdict will be weighed, and the override decisions will be made deliberately rather than reflexively. If the operating partner treats the framework as theatre, the dataset that powers the moat never accumulates and the system collapses to a vanity layer over existing dashboards. The Conviction Ledger's visibility of overrides is the principal mechanism against this failure mode.

9.3 The honesty constraint survives commercial pressure

The framework's value depends on the verdict being honest. Commercial pressure — to flatter the customer, to soften the dissent, to avoid the awkward analysis-reversal verdict that tells the leader her bet is dead — will be continuous. If the honesty constraint is weakened in response to that pressure, the framework converts into the manipulation tool it was designed to prevent. The governance of the honesty constraint must therefore be structural — written into the system's rubric, audited externally, and treated as the load-bearing ethical commitment of the firm.

> **THE PUBLICATION TEST**
>
> A framework that cannot publish the conditions under which it would be wrong is not a framework. It is a brochure. The three assumptions above are the conditions under which this framework would be wrong. They are offered to be argued against.

Closing — the discipline that compounds

The argument of this paper is, at heart, narrow. The technology of running a company is being rewritten faster than the discipline of running a company. Volumes I through III of the Studio series have addressed parts of the technology side. This volume addresses the discipline side.

The proposition is that the gap between those two curves is where the next decade's operational alpha is concentrated, and that closing the gap requires a specific architecture: pre-committed kill-criteria written at the moment of conviction, symmetric dissent at both commitment and reversal, classification of every reversal attempt against the leader's own prior reasoning, and an append-only ledger of judgment that — across many firms, over time — becomes the unreplicable training data for an instrument no incumbent can stand up in response.

A leader operating with this discipline is not smarter, calmer, or more contrarian than the leader operating without it. She is, instead, *legible to herself across time* — her past judgment is preserved in a form her future self can be confronted with, and her future judgment is therefore harder to mistake for her past judgment, easier to honor when honoring is right, and easier to revise when revising is warranted.

Multiplied across the bets a leader makes in a year, and compounded across the years a firm operates, the gap between those two postures is the gap between compounding and being compressed out. That is the gap this paper has been about.

Appendix A: The Conviction Diagnostic

A structured self-assessment for surfacing where a firm's current judgment practice sits against the framework. The diagnostic is a series of statements; the operator rates each on a scale of 1 (strongly disagree / not in place) to 5 (strongly agree / fully in place). Total scores map to a posture rating that surfaces the highest-leverage area for the next quarter's investment.

Section 1 — Pre-mortem commitment

- Every active strategic bet in our firm has a written thesis, in the decision-maker's own words, accessible to the senior team.
- Every active strategic bet has at least three falsifiable, observable kill-criteria recorded at the moment of commitment.
- The kill-criteria for every active bet have been reviewed for triggering status within the last sixty days.
- No bet in our active portfolio has been informally extended past its kill-criteria without an explicit, written re-commitment.

Section 2 — Dual Dissent

- Every material commitment in the last twelve months had a written Hubris Dissent attached at the time of decision.
- Every material reversal in the last twelve months had a written Capitulation Dissent attached at the time of decision.
- Both dissents are filed in a location that is accessible to the next reviewer at any subsequent decision point.
- No reversal in our active portfolio has occurred without a written argument against the reversal first being read aloud.

Section 3 — Decision Provenance

- Every material decision in our firm has an append-only provenance record that captures thesis, conviction, kill-criteria, and subsequent events.
- When a leader proposes a reversal, the original thesis and kill-criteria are surfaced as a precondition to the decision being made.

- Reversal attempts in our firm are explicitly classified into evidence-driven and condition-driven categories, with the classification preserved in the record.
- The override of a system-rendered verdict is logged with the same rigor as the decision it overrides.

Section 4 — Deflation Lens

- Every active bet in our portfolio has been examined against the three Deflation Lens questions (where collapse, what reallocation, what window).
- We have written reallocation plans for the savings produced by every active cost-compression initiative in our firm.
- Our strategy reviews explicitly compare our reallocation pace against the inferred pace of relevant competitors.
- We have, in the last quarter, killed at least one initiative whose savings were not paired with a reallocation plan.

Section 5 — Portfolio Mirror posture

- We track outcomes against every material commitment for at least eight quarters following the decision.
- We re-evaluate the original verdict (analysis vs. fear vs. mixed) against the realized outcome, and revise our internal rubric accordingly.
- We participate in, or operate, a cross-firm benchmark that compares our judgment patterns against anonymized peers.
- Our internal review cadence treats the labeled-judgment dataset as a strategic asset, not a compliance byproduct.

Scoring

INTERPRETATION

A score of three or fewer questions answered with documented evidence is a firm operating with no judgment discipline. A score of seven to twelve is a firm with partial discipline, typically concentrated in commitment but absent at reversal. A score of seventeen or more is a firm operating with full Decision Provenance discipline — and, by the argument of this paper, accumulating the labeled judgment data that becomes its unreplicable moat. The gap between three and seventeen, multiplied across the bets a firm makes and compounded across the years it operates, is the gap this entire framework is designed to close.

Appendix B: Reference schema for the Conviction Ledger

The schema for the Decision Provenance store, as deployed — nine append-only tables. SQL-flavored for portability; the identical structure works in a document store. Field names follow the live implementation.

B.1 — commitments

FIELDS

Term	Definition
decision_id	Stable, append-only identifier for the commitment.
firm_id	The operating entity. Permits cross-portfolio aggregation under appropriate anonymization.
decision_name	Short retrievable label.
thesis	The decision-maker's reasoning at the moment of commitment, in her own words.
conviction	Integer 0–100, supplied by the decision-maker.
capital_committed	Financial and non-financial commitments. Attention is a tracked resource.
commitment_data	Structured snapshot of internal and external signals at commitment time.
kill_criteria	Array of falsifiable, observable conditions for reversal. Required.
expected_j_curve	Explicit acknowledgment of the discomfort window the decision-maker has authorized.
committed_at	Timestamp.
status	active | held | reversed. State machine; transitions logged in provenance.

B.2 — provenance [sacred]

FIELDS

Term	Definition
provenance_id	Append-only record identifier.
decision_id	Foreign key to commitments.

Term	Definition
event	commitment | reversal_attempt | held | reversed | outcome_recorded.
classification	analysis_reversal | fear_reversal | mixed | null. Populated on reversal attempts.
confidence	Integer 0–100. The system's confidence in the classification.
drivers	Array of named cognitive/condition drivers (recency, loss aversion, social pressure, herd, fatigue).
evidence_for_reversal	Array of evidence-based reasons for reversal. Empty for clean fear reversals.
reasoning	The Capitulation Dissent rendered for this attempt.
decision	held | overridden | null. The decision-maker's resolution.
pressure_signals	Array of non-evidence conditions present at the moment of the attempt.
created_at	Timestamp.
outcome_link	Foreign key to outcomes, populated at the reconciliation cadence.

B.3 — outcomes

FIELDS

Term	Definition
outcome_id	Append-only.
decision_id	Foreign key to commitments.
verdict	achieved | partial | failed.
thesis_held	Boolean — did the original thesis hold?
actual_result	The realized outcome value against the original thesis.
vs_thesis	The delta between realized result and the committed thesis.
decision_quality	right_held | wrong_held | right_reversed | wrong_reversed. The label that trains the Portfolio Mirror classifier.
recorded_at	Timestamp. Immutable once written.

B.4 — outcome_reviews

FIELDS

Term	Definition
review_id	Append-only.
decision_id	Foreign key to commitments.
checkpoint	30d | 60d | 90d. Auto-scheduled at decision time.
due_at	Scheduled date; surfaced on the dashboard when due or overdue.
status	scheduled | due | overdue | complete.
captured	The thesis-holding read recorded at the checkpoint.

B.5 — kill_criteria_updates

FIELDS

Term	Definition
update_id	Append-only. No silent edits; no retroactive rationalization.
decision_id	Foreign key to commitments.
criterion	The criterion being changed.
old_value	Prior threshold and triggered-state.
new_value	New threshold and triggered-state.
triggered	Whether the criterion was triggered at the time of change.
updated_at	Timestamp.

B.6 — board_signals

FIELDS

Term	Definition
signal_id	Append-only.
decision_id	Foreign key to the commitment under discussion.
member_id	Foreign key to board_members.
signal	comfortable | neutral | concerned | escalated | approved | rejected.
note	Free-text context captured after the board interaction.
created_at	Timestamp. Forms a per-member sentiment timeline.

B.7 — board_members

FIELDS

Term	Definition
member_id	Stable identifier.
firm_id	The portfolio company.
name	Board member name.
stance	Current rolled-up stance, updated by the latest signal.

B.8 — weekly_priorities

FIELDS

Term	Definition
priority_id	Append-only.
firm_id	The portfolio company.
label	The priority text.
times_moved	Avoidance counter — increments each week the priority is carried without progress.
week_of	The week this snapshot covers.

B.9 — direct_reports

FIELDS

Term	Definition
report_id	Stable identifier.
firm_id	The portfolio company.
name	Direct report name.
committed_word	The commitment made in a 1:1, quoted.
delivery	on_track | at_risk | missed.

About The Studio

The Studio is the intelligence imprint at richardleclezio.com. It publishes frameworks, blueprints, and working papers on AI transformation, regulatory intelligence, enterprise governance, and — with this volume — the governance of human judgment in an AI-deflated economy. No takes. Only original thinking. The Studio publishes on an irregular schedule.

Contact: richardleclezio.com · linkedin.com/in/richard-leclezio · richard.leclezio@gmail.com

About KEEL

KEEL is the reference implementation of the Decision Provenance framework introduced in this paper — built, deployed, and running in production at v0.2. The system is named for the part of a ship below the waterline that no one sees and that keeps the vessel from capsizing in heavy weather. KEEL is decision support, not the decision-maker: it classifies, confronts, and logs; the human chooses, always. Built by the author, operated by Leclezio Consulting Corporation, powered by Claude (Anthropic).

Architecture: Five live engines — Decision Provenance, Dual Dissent, Chief-of-Staff Chat, Decision Memo, Outcome Review — plus a full CEO operating system (My Week, My Team, Board Room) and a nine-table append-only Conviction Ledger. FastAPI · Python · SQLite (Postgres-ready); claude-sonnet-4-6; single-file frontend; deployed on Railway; designed 1:1 for LangGraph migration. Specification in Section 07; schema in Appendix B.

Live: keel-production-c7c5.up.railway.app · richard.leclezio@gmail.com

Status: Phase 0 complete (built and live). Phase 1 — the single ask — is one PE design partner, one portfolio company, thirty days. The classifier is LLM-as-judge against a rubric until outcome-confirmed data accrues; provisional things stay labeled provisional.