



From Prediction Machine to Decision Machine

Autonomy Platform · Public Blog

CONFIDENTIAL

© 2026 Azirella Ltd. All rights reserved worldwide.

Strictly confidential and proprietary — do not distribute.

From Prediction Machine to Decision Machine

**LLMs are still just Prediction Machines.
Supply Chain Planning needs Decision Machines.**

Ajay Agrawal, Joshua Gans, and Avi Goldfarb published *Prediction Machines* in 2018. The thesis was elegant: AI makes prediction cheap, and complements to prediction — data, judgment, action — become valuable as a result. Their AI Canvas decomposes any task into seven slots: Prediction, Judgment, Action, Outcome (the top row), and Input, Training, Feedback (the bottom row). **In their framing, AI sits only in the Prediction slot.** Humans (and deterministic systems) do everything else.

Autonomy provides full-scope coverage of the canvas. Each slot has a substrate behind it, by design.

**AI Canvas
slot**

Autonomy substrate

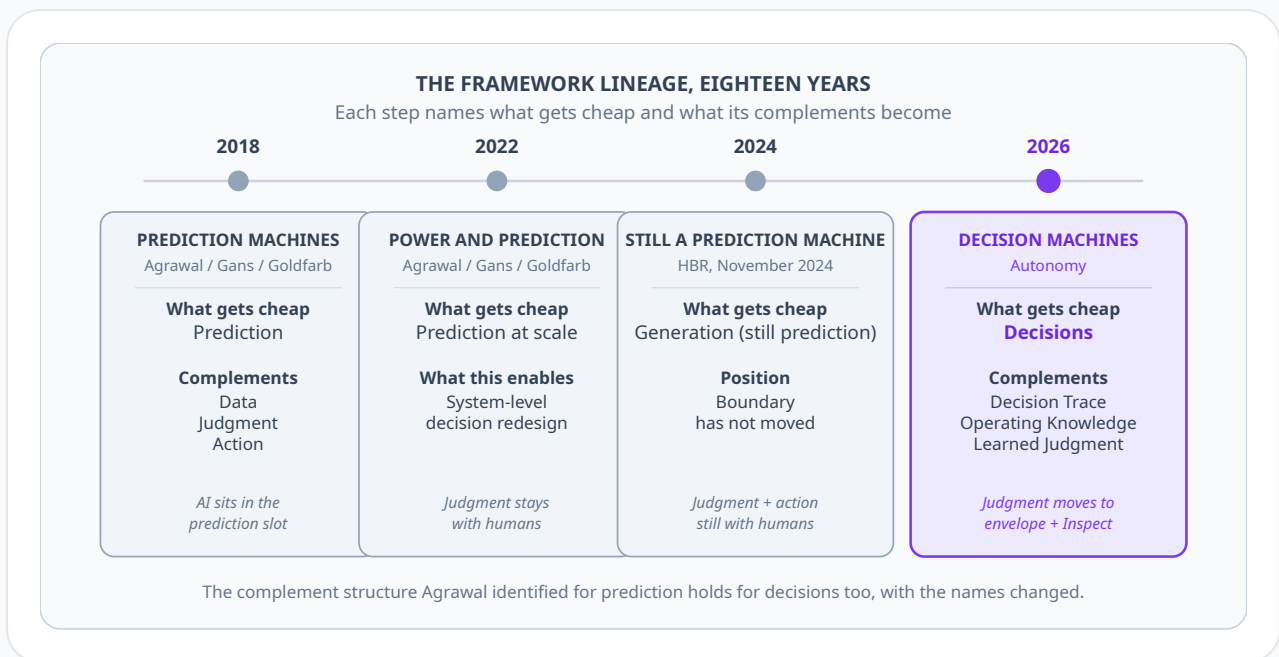
Prediction	NN forecasting agents wrapped in a conformal layer that emits P10 / P50 / P90 intervals on every prediction.
Judgment	BSC weights, governance pipeline thresholds, the Automate envelope, plus elicited Operating Knowledge priors.
Action	A trained decision policy emits an action; AIIO routes it to Automate, Inform, or Inspect; canonical state is written.
Outcome	The Decision Trace ledger captures Prompt, Decision, Expected, Likelihood; the BSC reward function scores realised outcomes.
Input	Canonical state via the AWS SC Data Model; the Context Engine routes external signals to the right tier.
Training	RL on the digital twin against a BSC reward, plus conformal recalibration and causal posterior updates on cadence.
Feedback	Operating Knowledge interviews elicit planner priors; the override classifier and causal layer score each override.

That is what makes a decision machine cheap to operate. The complements Agrawal, Gans, and Goldfarb correctly identified as the bottleneck on cheap prediction become part of the substrate when the goal is cheap decisions.

In 2022 they published *Power and Prediction*, extending the framework to system-level redesign. In November 2024 they returned to HBR with a piece titled “Generative AI Is Still Just a Prediction Machine.” The word *still* is doing a lot of work. Their position is that even genAI’s writing, drawing, coding, and summarising are reframed prediction tasks. Judgment stays

with humans. Action stays with humans (or deterministic systems). Their boundary has not moved.

The boundary is correct, as far as it goes. The economic argument they pioneered is one of the clearest pieces of thinking about AI's impact in the public canon. But the framework was named for what AI could do well in 2018, and the answer was prediction. **The question for 2026 is what gets cheap next, and what its complements become.**



What gets cheap next

In industrial supply chains, the answer is decisions.

Not predictions over decisions. Not recommendations dressed up as decisions. **Calibrated, traceable, inspectable, certifiable decisions** that the substrate can produce at the cadence the operation demands, with the audit surface the regulator and the insurer require.

That is not a vendor claim. It is what the four pillars of a working industrial AI substrate compose into when you put them next to one another:

- A **trained decision policy** (NN or solver) that maps state to action without an LLM in the path. Engineered, not generated.
- A **conformal layer** that puts a calibrated likelihood on every prediction the policy reads from, so the decision the policy emits carries a calibrated likelihood too.
- A **digital twin** that the policy was trained against, under RL on a Balanced Scorecard reward, before it ever touched canonical state.
- A **causal layer** that scores realised overrides against the substrate's own recommendations, attributes outcome to decision, and feeds the next training pass.

Put those four together and the decision is no longer a high-cost artefact requiring scarce human judgment for every emission. It is something the substrate produces with the same economic property prediction had in 2018: cheap, and reliable enough to redesign the workflow around.

Cheaper, faster, better, the standard test for tech adoption

The Creative Destruction Lab's Prediction Machines argument rested on the standard test for whether a technology actually displaces the prior way of doing things: it has to be **cheaper**, **faster**, or **better**, and ideally all three. Cheap prediction passed the test for one slot of the workflow. Cheap decisions have to pass it for the whole workflow.

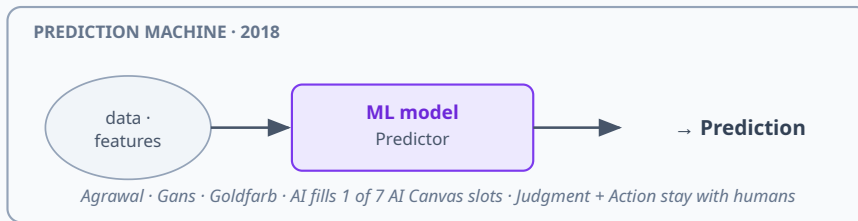
Autonomy hits all three for the full decision, with one important caveat.

- **Cheaper.** A decision produced by the substrate costs a fraction of a decision produced by a senior planner. The marginal cost of running one more cycle is compute, not headcount. Continuous optimisation that previously cost too much engineering time to run more than quarterly becomes a continuous default.
- **Faster.** A decision lands in seconds, not in the planning meeting two days from now. The two operational quantities that matter most, **time to detect** (when an exception surfaces) and **time to correct** (when the substrate has acted on it), both compress from days to seconds.
- **Better, with a caveat.** We do **not** claim that every atomic decision the substrate makes is better than what a senior planner would have made on her best day. We do claim that *on aggregate*, the combination of cheaper and faster on detection and correction makes the overall outcome materially better. The failure mode of most supply chains is not “the wrong decision on a clean signal.” It is “the right decision after the signal sat in someone’s inbox for three days.” A substrate that detects in seconds and corrects in seconds avoids the major problems that compound while humans are scarce, and over hundreds or thousands of decisions a week, the aggregate of detect-and-correct is what moves the BSC.

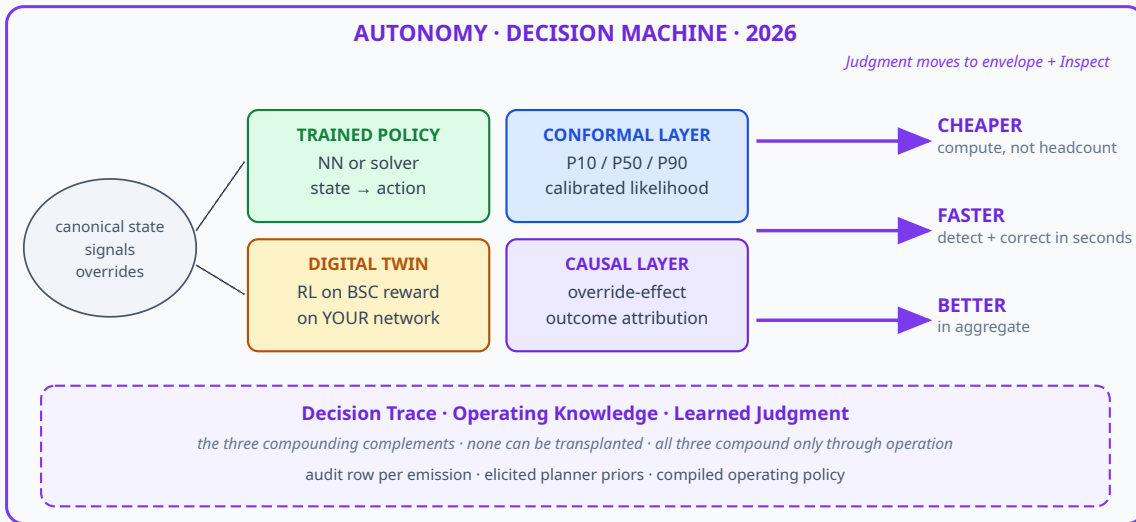
This is the practical sense in which cheap decisions matter. Atomic-decision quality is contested case-by-case; aggregate outcome is not. Cheaper plus faster on detection and correction is the mechanism that delivers the better.

PREDICTION MACHINE → DECISION MACHINE

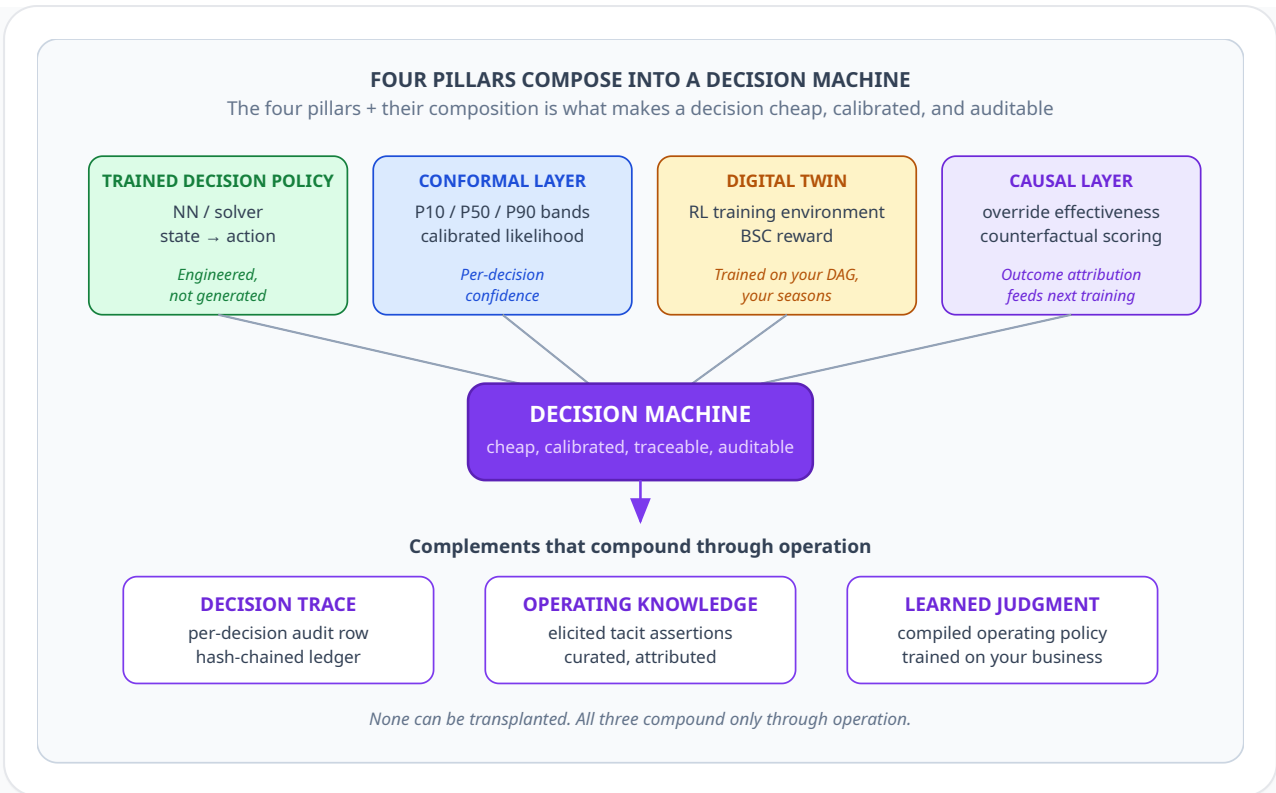
Cheap prediction (2018) → cheap decisions (2026). Same economic shape, one level up.



↓ *what AI got cheap next →*



The model is the commodity. The substrate is the differentiator.



What the complements become

Agrawal, Gans, and Goldfarb’s economic argument said cheap prediction makes its complements valuable. The same shape applies one level up.

When decisions get cheap, the complements to decisions become valuable:

- The **Decision Trace**, the per-decision audit row, hash-chained into a verifiable ledger. The aviation, pharma, and nuclear industries already use *trace* as the noun for an immutable per-event record. The supply chain industry will too.
- **Operating Knowledge**, the tacit assertions experienced planners carry, captured before they retire, structured with provenance, curated under the two-person rule, surfaced to the substrate as active priors.

The knowledge that took a senior planner thirty years to internalise no longer walks out the door at retirement.

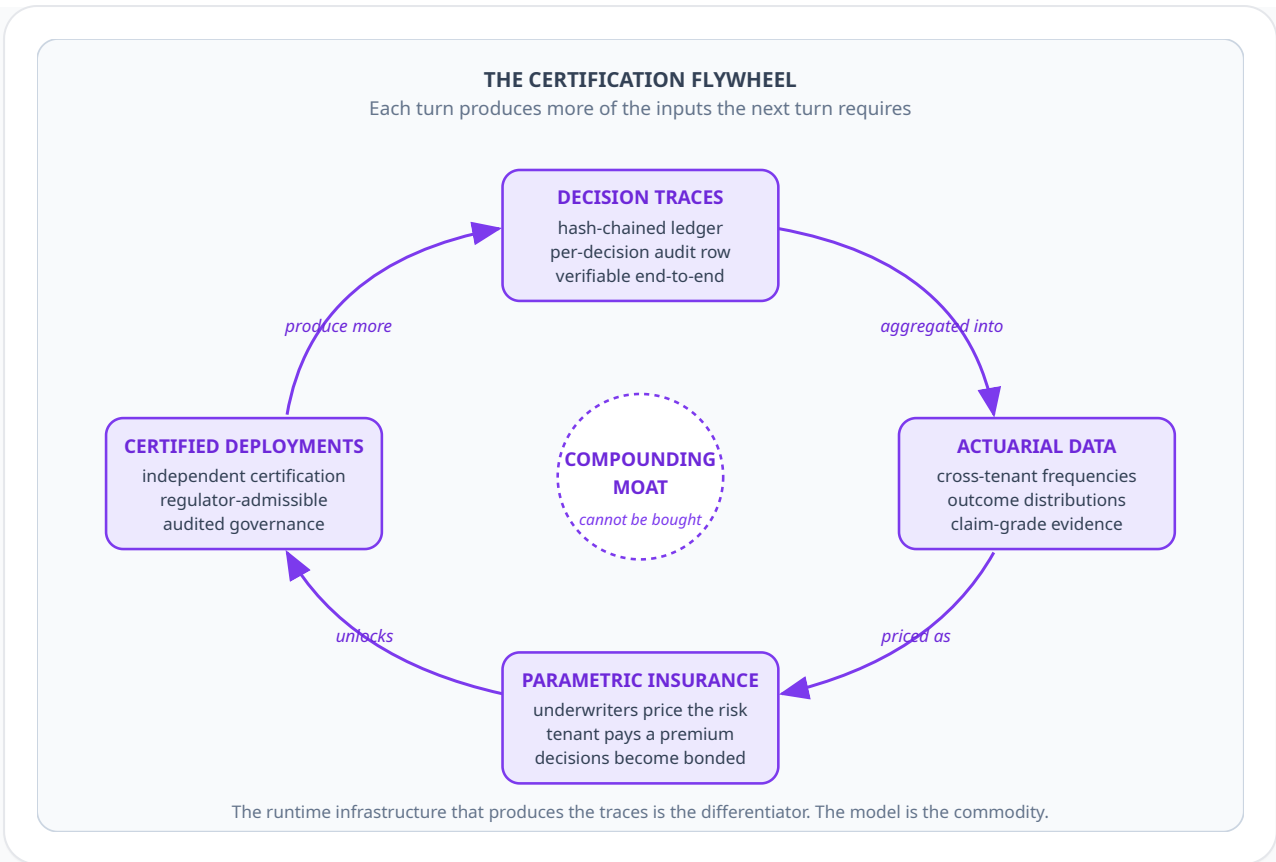
- **Learned Judgment**, the substrate's compiled operating policy. Trained model weights that have seen every season your business goes through. Calibrated conformal intervals that know your lead-time and demand distributions. Causal estimands that have scored every override your planners have made.

None of these can be transplanted. All three compound only through operation. *You cannot buy three years of decision traces. You cannot buy thirty years of operating knowledge. You cannot buy a learned policy trained on every season your network has run through.* Each one is the trained-on-your-business specificity that a competitor on the same platform does not get for free.

The certification flywheel

There is one more thing that becomes valuable when decisions get cheap, and it is the property that turns Autonomy from a software vendor into something else.

Hash-chained decision traces produce actuarial-quality data. Actuarial-quality data lets a parametric insurer underwrite autonomous decisions. Insurable autonomous decisions get certified by independent bodies. Certified deployments produce more decision traces. The flywheel turns.



This is the property the [XMPPro CEO Pieter van Schalkwyk](#) calls the industrial AI moonshot. He is right that it matters. He is also right that the runtime infrastructure that produces those traces is what differentiates. The substrate is the differentiator. The model is the commodity.

Where this leaves the framework

Agrawal, Gans, and Goldfarb’s framework is not wrong. It is the cleanest description of where AI’s economic contribution sat between 2018 and 2024. Their refusal to extend the framework themselves into 2026, holding the prediction line even as genAI’s writing and reasoning capabilities have surged, is intellectually honest. They are economists, not industrial AI architects.

The extension is the supply chain industry’s to make:

- **2018** (Agrawal / Gans / Goldfarb): AI makes prediction cheap. Complements to prediction become valuable.
- **2022** (Agrawal / Gans / Goldfarb): Cheap prediction enables system-level decision redesign. Judgment stays with humans.
- **2024** (Agrawal / Gans / Goldfarb): Even generative AI is still just prediction. Judgment still stays with humans.
- **2026**: Cheap decisions, in industrial supply chains. Complements to decisions, the trace, the operating knowledge, the learned judgment, become valuable. Judgment moves from making each decision to specifying the envelope and inspecting the trace.

The framework deserves the respect of being extended, not contradicted. Cheap decisions are the next economic step in the same lineage Agrawal, Gans, and Goldfarb drew. The complement structure they identified for prediction holds for decisions too, with the names changed.

Industrial supply chains are where the next step has to happen first, because they are where the cost of getting a decision wrong is highest and where the architectural infrastructure to produce calibrated decisions at scale is closest to in place.

The model is replaceable on a six-month cycle. The substrate is licensable to every operator who wants it. The moat is the three-bucket compound of your operation's decisions, your planners' heuristics, and the substrate's compiled judgment.

Generative AI is still just a prediction machine. Industrial supply chains need decision machines. We are building one.

For the architectural depth behind this post, see the [How Agents Learn](#) *whitepaper*.

See Autonomy in action

Walk through how Autonomy models, executes, monitors, and governs supply chain decisions with autonomous AI agents.

[See It Live](#)