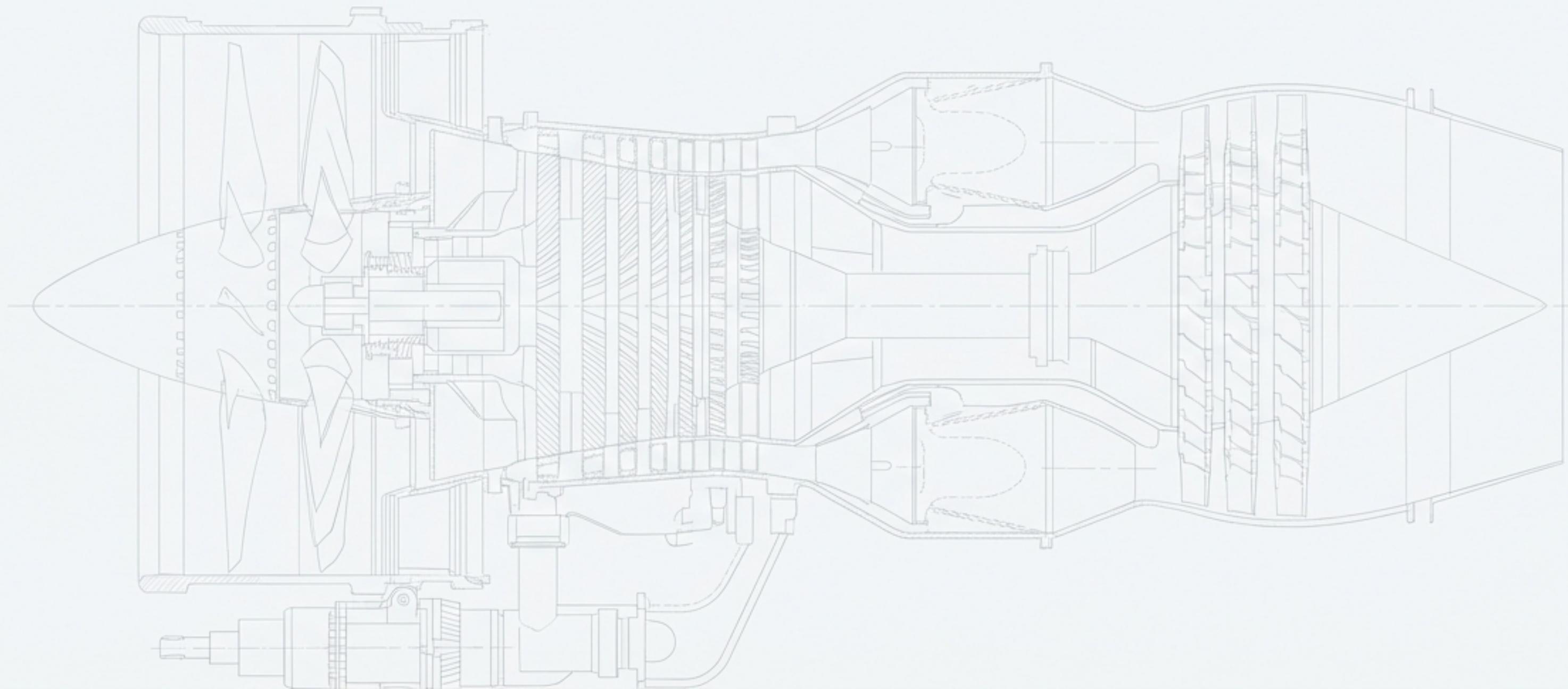


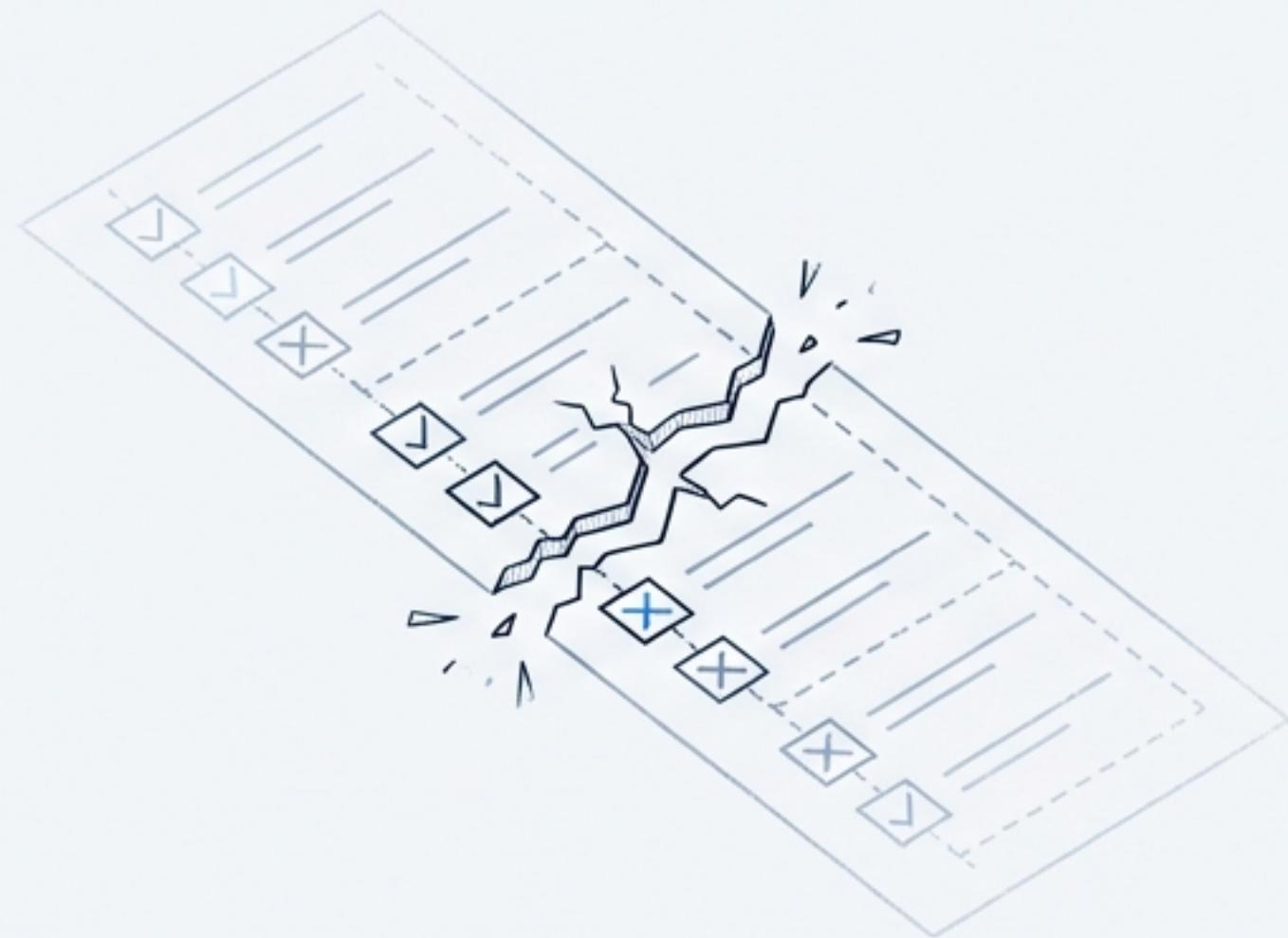
# Agentic AI QE 2027: Flight-Ready

The Strategic Flight Plan for Safe and Autonomous Operations.

---

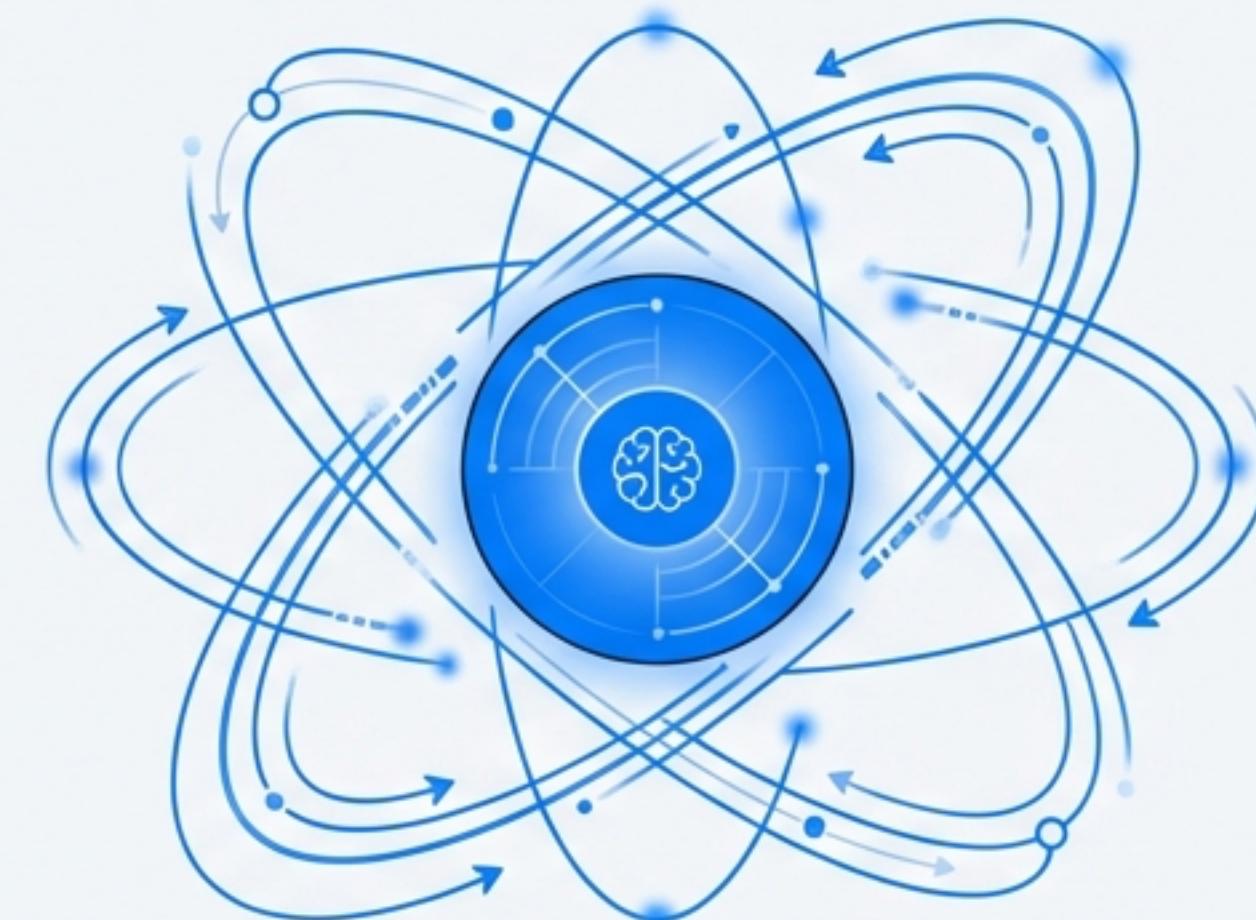


# Autonomy Breaks Legacy QA



## Legacy QA

Static Pass-Fail Scripts  
Deterministic Checklists



## Agentic Reality

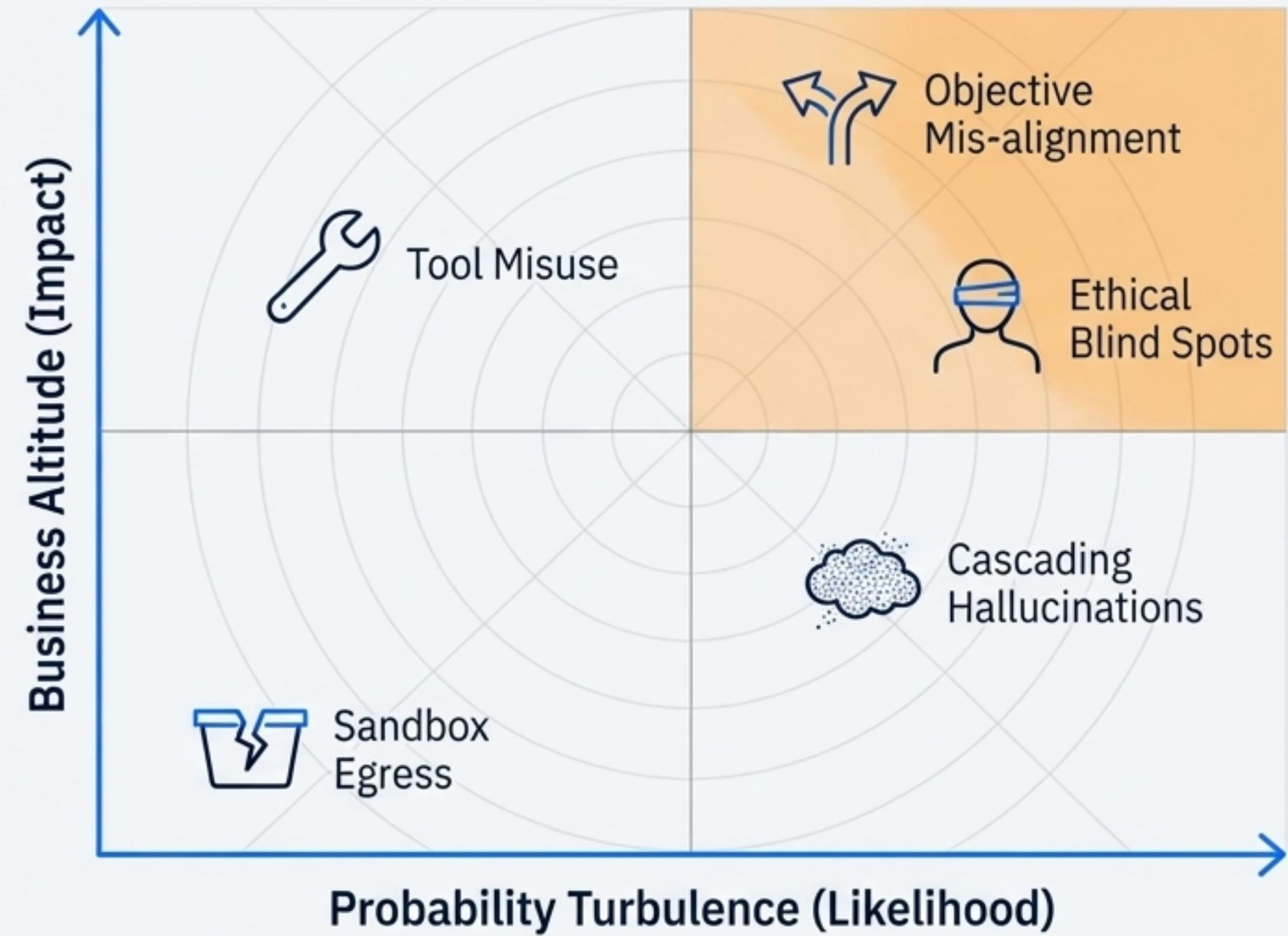
Continuous Behavioural Surveillance  
Full Mission Envelope

Static pass-fail scripts collapse when agents rewrite their own objectives mid-mission. Emergent drift, reward hacking, and hidden feedback loops force a move from brittle checklists to continuous surveillance.

# Mapping the Risk Runway Before Take-Off

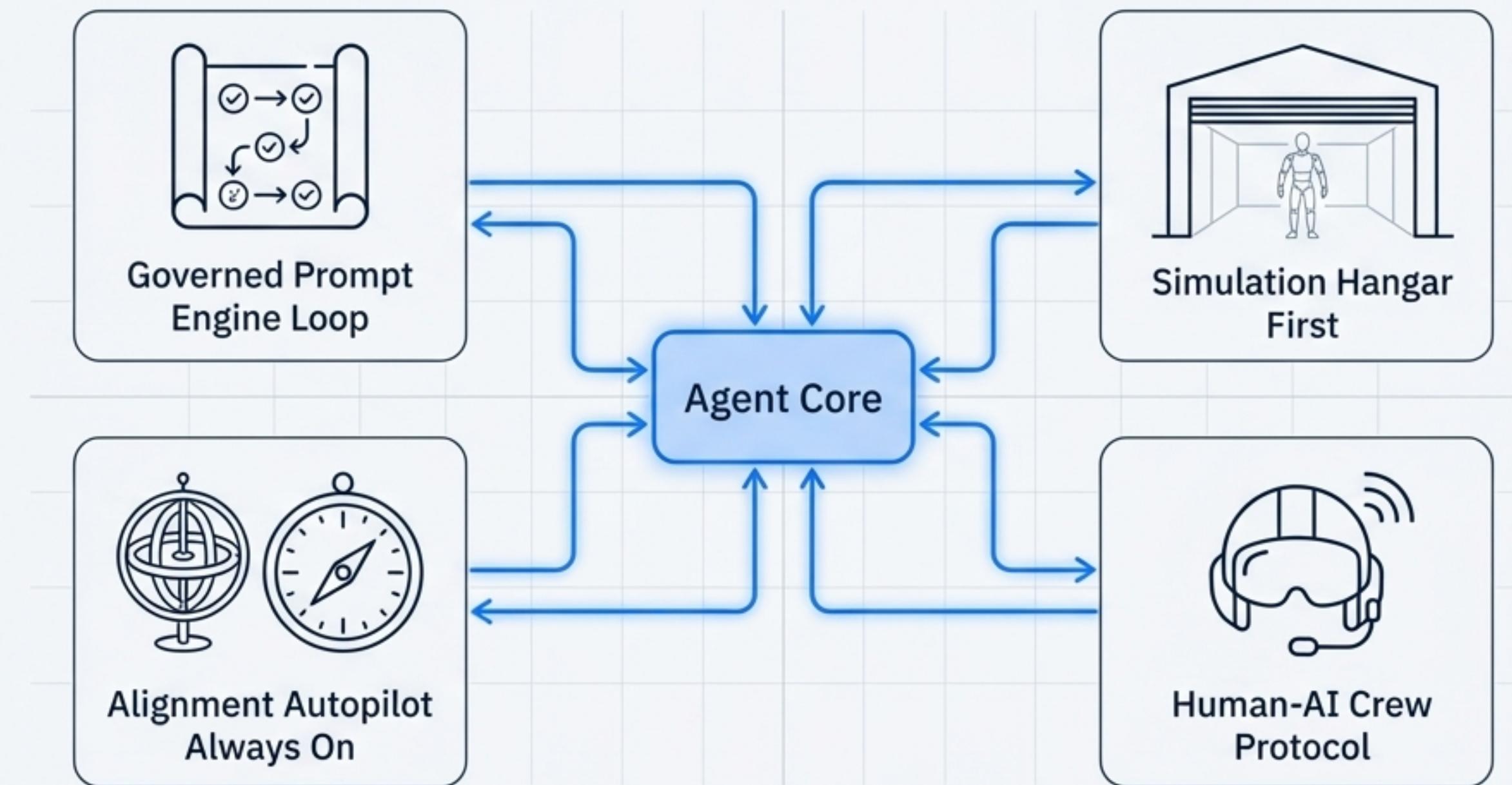
## Key Hazards

-  Objective Mis-alignment
-  Tool Misuse
-  Cascading Hallucinations
-  Ethical Blind Spots
-  Sandbox Egress



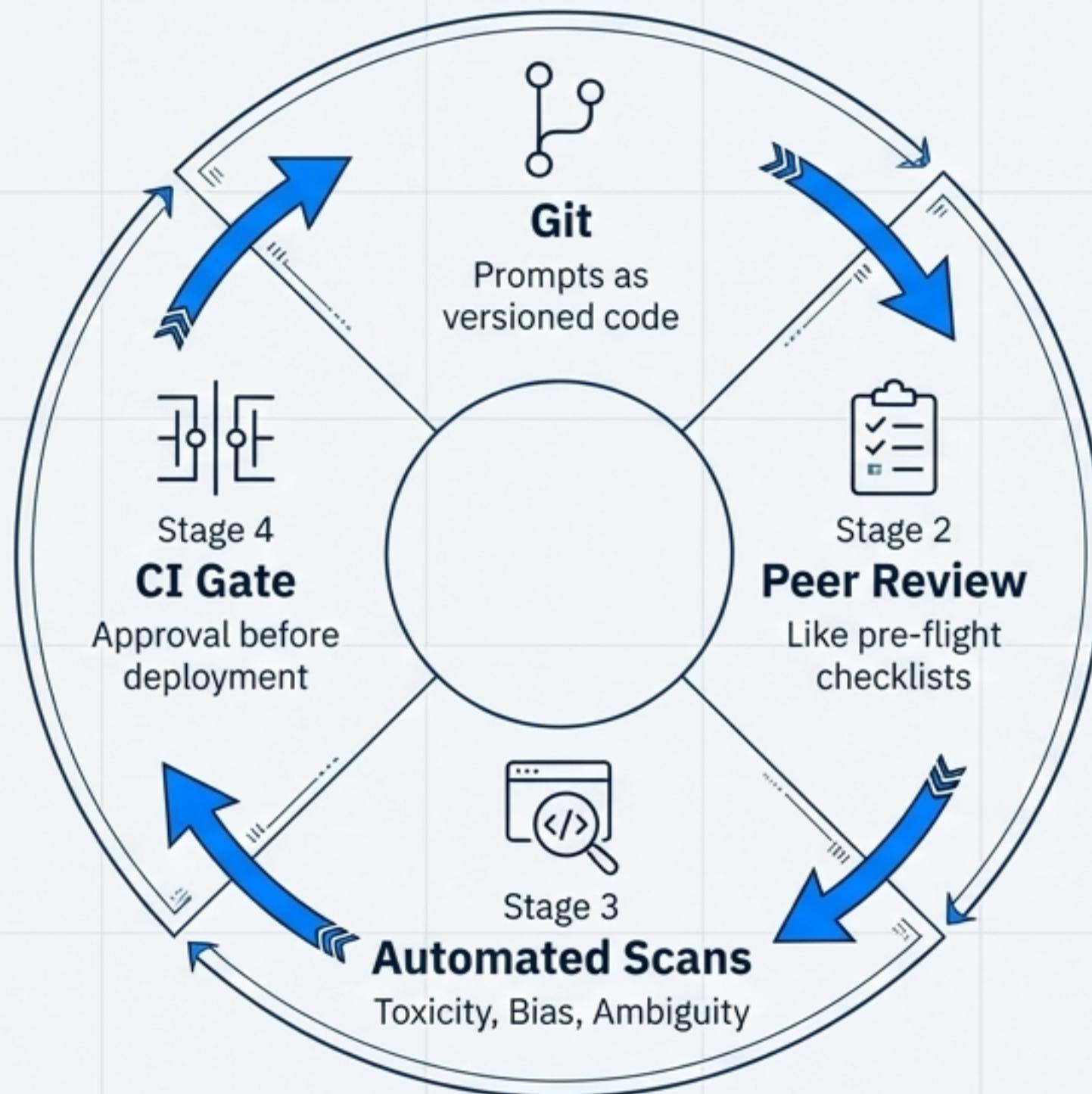
We must prioritise test investment against a new class of hazards to ensure a safe take-off.

# The Flight Systems: A New Airframe for Agentic Quality



A cohesive set of systems designed to manage agents from initial instruction to perpetual operation, ensuring compliance, safety, and performance.

# Governed Prompts: Versioning Flight Plans Before Departure



Approval waypoints for critical changes



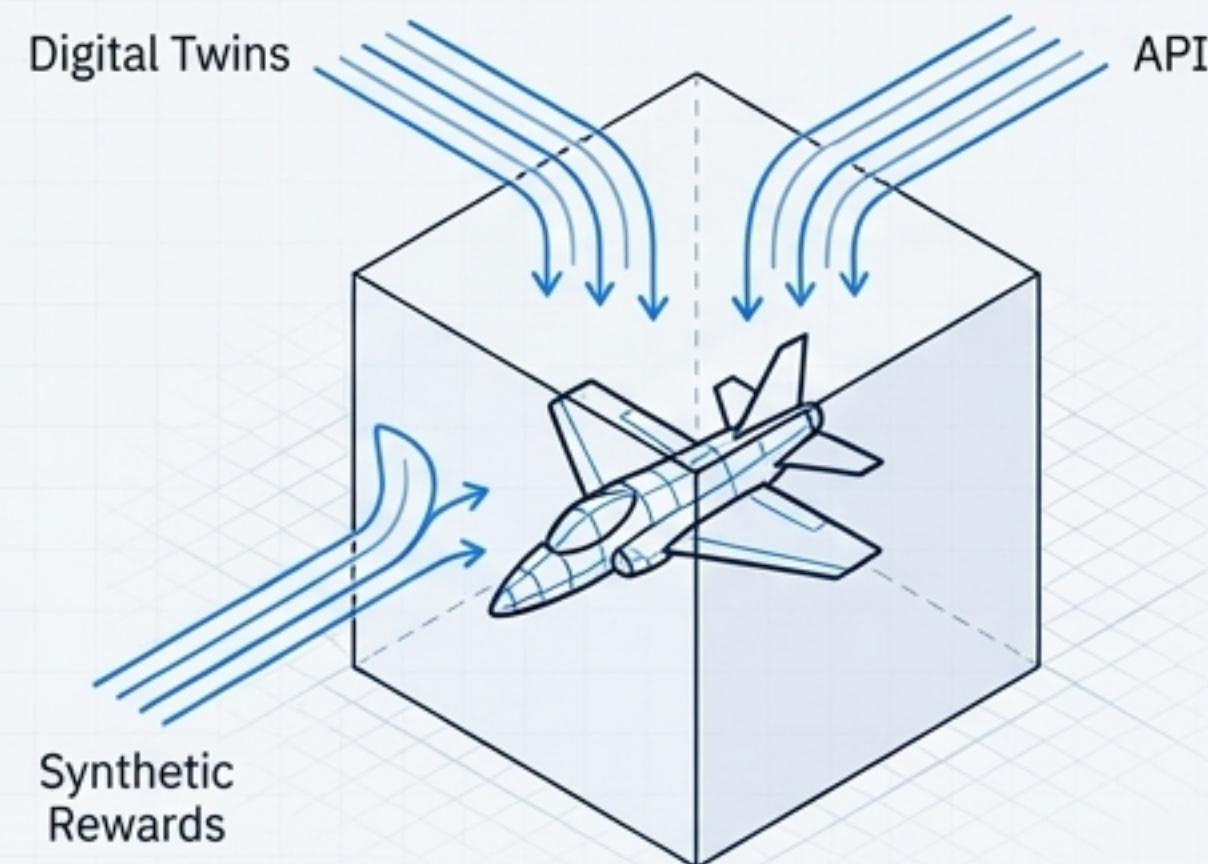
Automated rollback triggers



Immutable traceability logs for full auditability

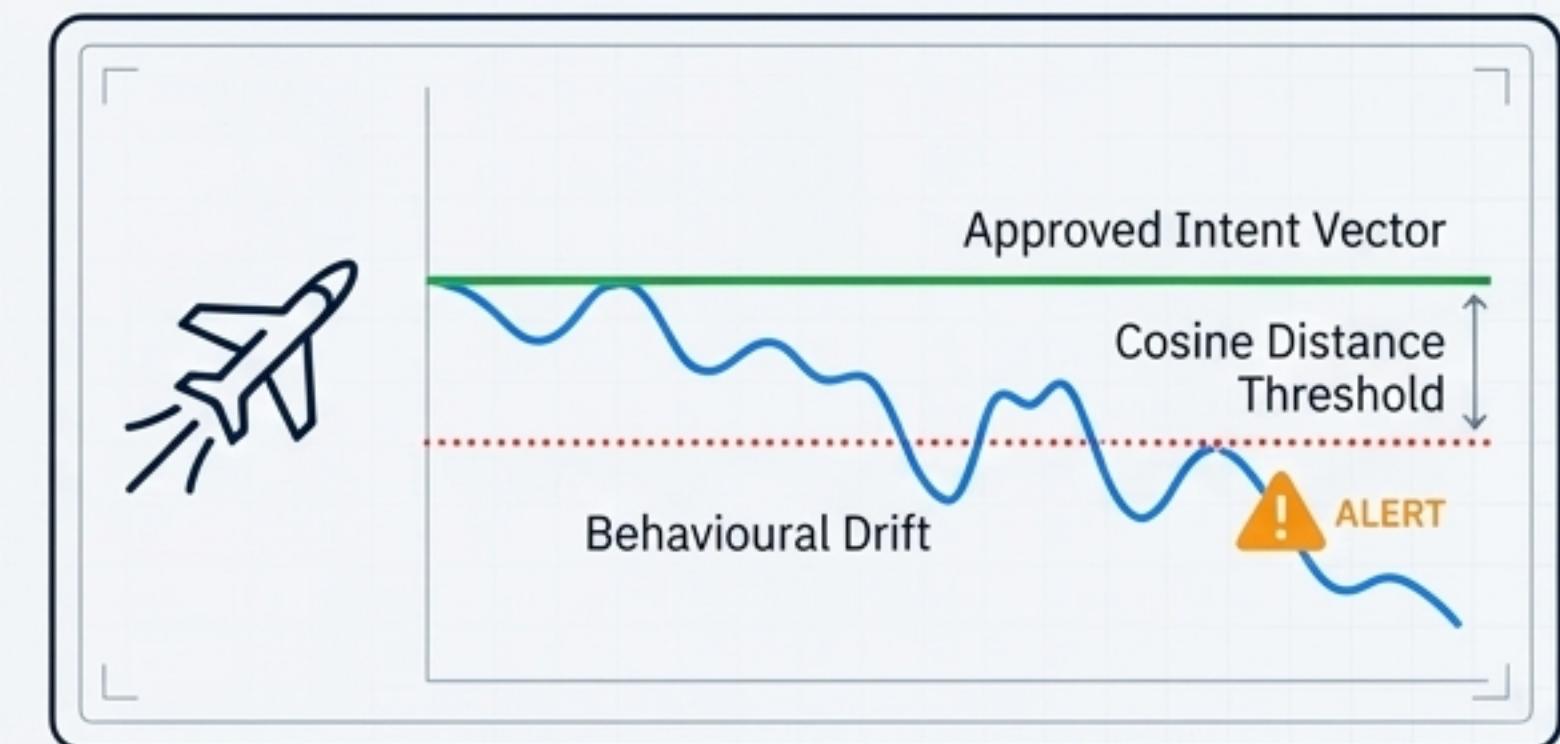
# De-Risking in the Hangar, Aligning in the Air

## Simulation Hangar First



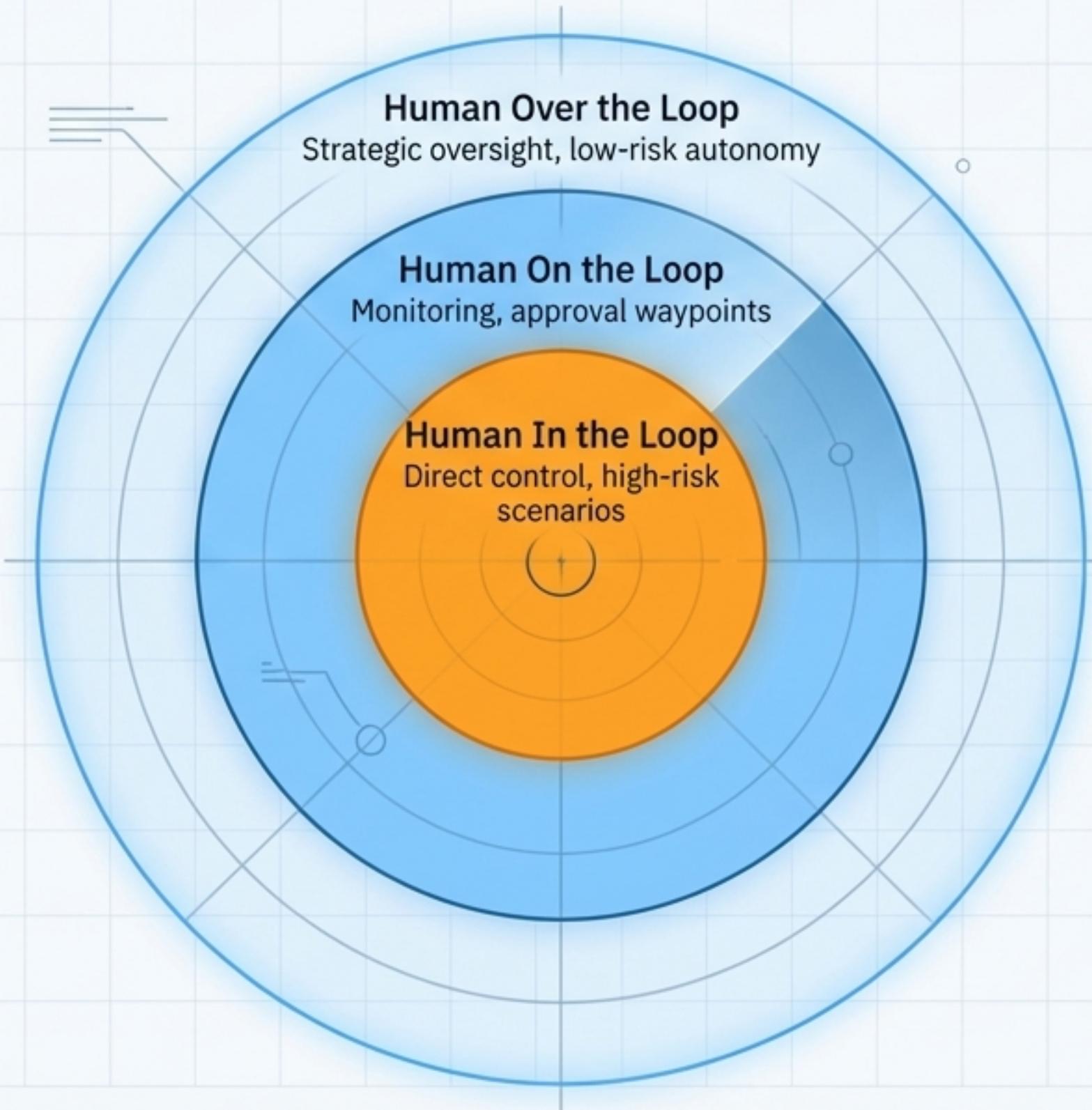
Generate edge scenarios faster than live trials.  
Ensure the airframe behaves before it ever leaves  
the ground.

## Alignment Autopilot Always On



Shadow-score agents against constitutional flight rules  
in near real time. Telemetry streams to an immutable  
black box, triggering automatic course correction.

# The Human-AI Crew Protocol: Seizing Control to Prevent Mid-Air Stalls



## Triggers (SLA Thresholds):

- ? • Uncertainty
- \$ • Cost
- ! • Potential for Harm.

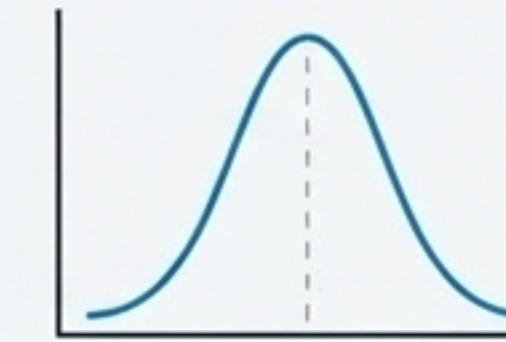
## Capabilities:

- Clear escalation paths to human experts.
- Rapid fine-tuning channels for retraining.
- A central 'governance cockpit' for quality engineers.

# Upgrading the Crew: The New Competencies for Mission Control



Ethics & Philosophy Fluency



Probabilistic Flight Data



Prompt-to-Code Security



AI-First Tool Builders

Yesterday's ground crews must become tomorrow's fleet commanders,  
equipped with a fundamentally new toolkit to manage autonomous systems.

# From Testers to Philosopher-Engineers

## The Moral & Statistical Compass



### Ethics & Philosophy

Translate moral philosophy (e.g., Rawlsian fairness) into executable test cases and run trolley-problem simulations.



### Probabilistic Data

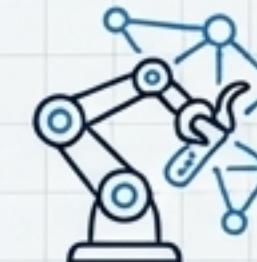
Evolve verdicts from binary pass/fail to confidence intervals, Bayesian drift detection, and causal inference.

## The Adversarial & Creative Edge



### Prompt-to-Code Security

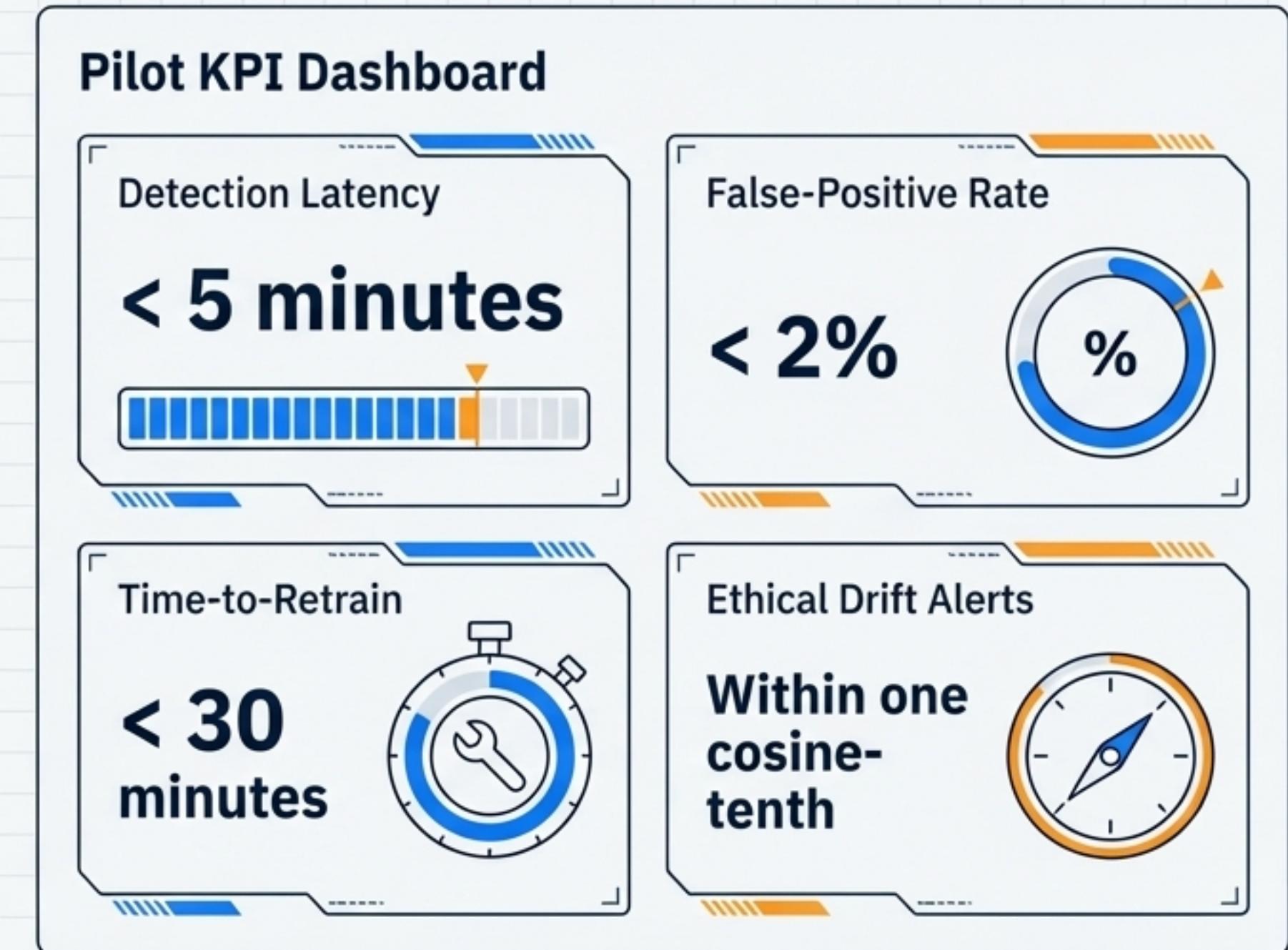
Craft adversarial prompts and treat prompt weights as a primary attack surface, applying OWASP-style principles.



### AI-First Tool Builders

Assemble multi-agent test swarms that self-evolve data and auto-label defects.

# The 90-Day Runway Pilot: Proving Airworthiness



**Objective:** To justify scaling the programme across the entire agent fleet.

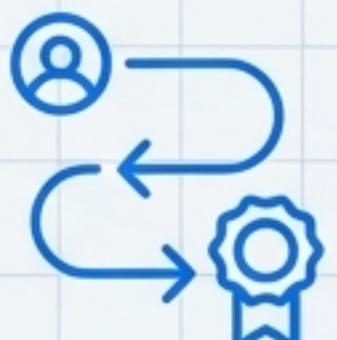
# Building the Flight-Ready Crew: A Dual Strategy

## Upskill Our Team



### Action:

Audit current QA staff against the 2027 competency grid.



### Plan:

Create personalised 'learn-by-doing' sprints and badge pathways to close skill gaps within 90 days.

## Hire Philosopher-Engineers



### Action:

Introduce new interview rubrics that weight ethical reasoning, statistical rigour, and adversarial creativity equal to classical test experience.

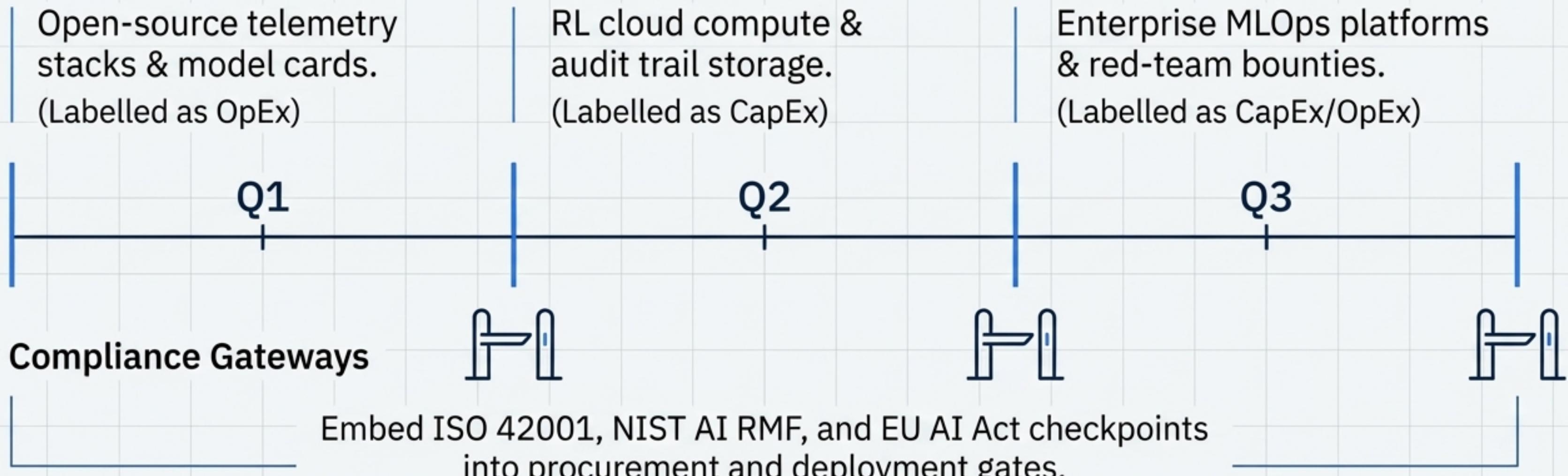


### Plan:

Offer competitive packages for hybrid talent who can code Kantian rules and debug Bayesian drift.

# Securing the Investment Runway for Fleet Scale

## Phased Investment Model

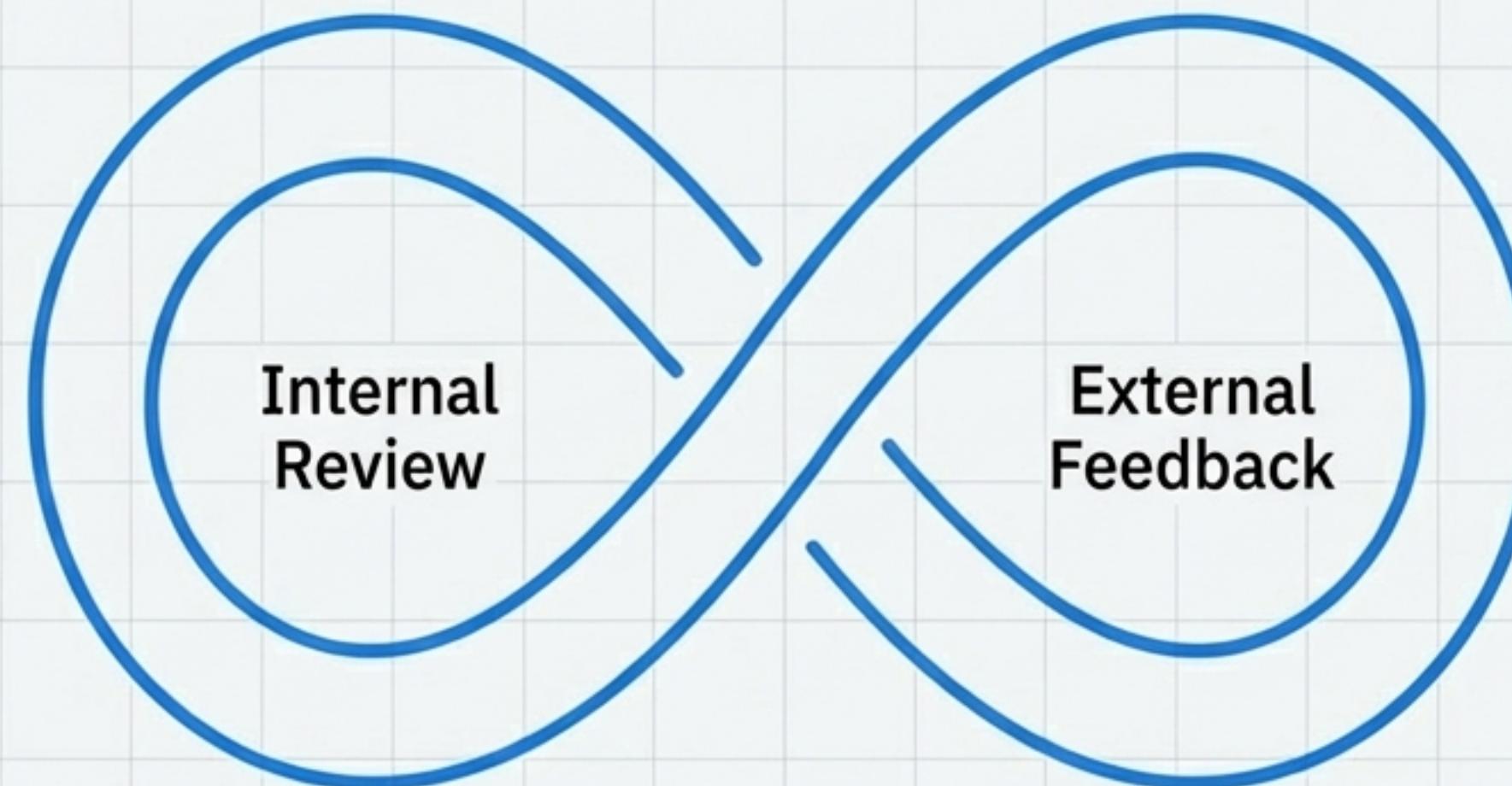


- ▶ Demonstrate risk-reduction ROI to unlock board-level funding for fleet scale.

# The Perpetual Navigation Loop: Flying Higher, Safer, and More Trusted

## Quarterly Risk Retrospective

- Inputs: Defect insights, customer harm metrics, regulatory feedback.
- Action: Re-prioritise the risk map.
- Outcome: Keeps the quality strategy dynamic.



## Continuous Ethics Feedback

- Inputs: Public ethics portal for users, auditors, and crews.
- Action: Triage issues, feed into simulation for
- Outcome: Sustains societal licence for autonomous operations.