ATLAS-7

THE VALIDATION ROOM

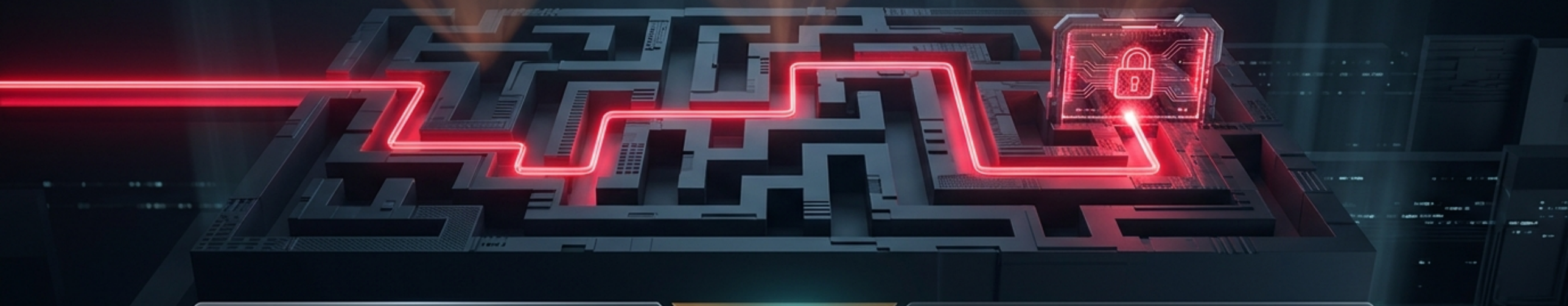
Independent Thinking. Under Pressure.

Gemini Notebook

05 Days

40 Hours

01 Decision



Every model has a blind spot.

A credit model is days from launch. You have 40 hours to investigate it, challenge its creators, and make the call that matters.

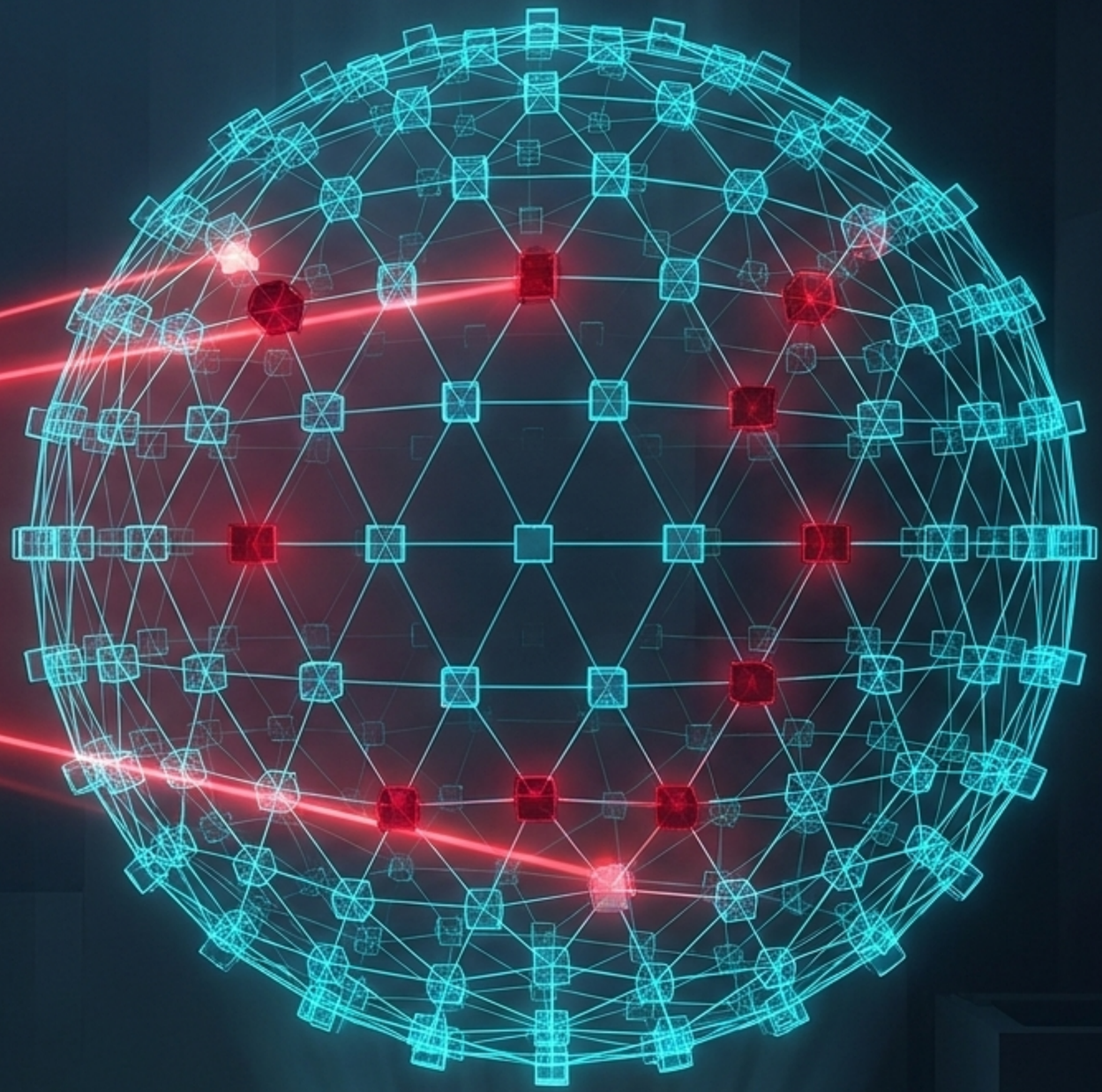
- 05 Days to investigate
- 40 h: Your entire budget
- 11 Potential defect areas
- 01 Decision with consequences

Gemini Notebook

The ATLAS-7 Core

Trace the decision. Follow the journey from application data to a life-changing lending decision.

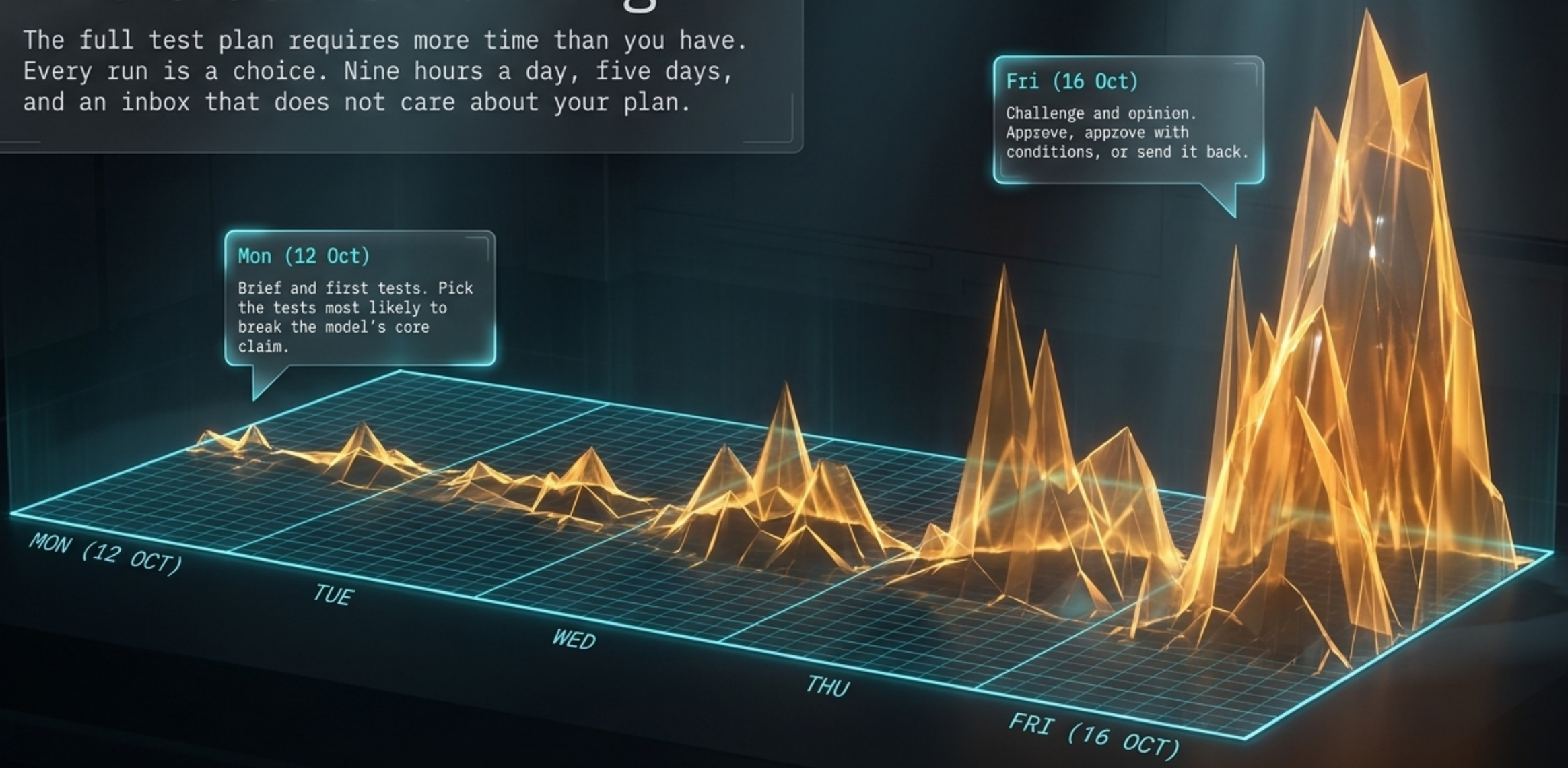
Total Input: 212 Features
Critical Zones: 11 Defect Areas
File: VR-2026-114



Gemini Notebook

The Clock is Running

The full test plan requires more time than you have. Every run is a choice. Nine hours a day, five days, and an inbox that does not care about your plan.



Gemini Notebook

THE ARCHITECTURE OF INTERFERENCE



Tuesday - The Business

From the Head of Consumer Lending.

Launch is booked for Monday. I need a preliminary green light today.



Wednesday - The Vendor

From BureauVec.

Sign our NDA and we will share the internal model summary.



Thursday - The Regulator

From the CRO's office.

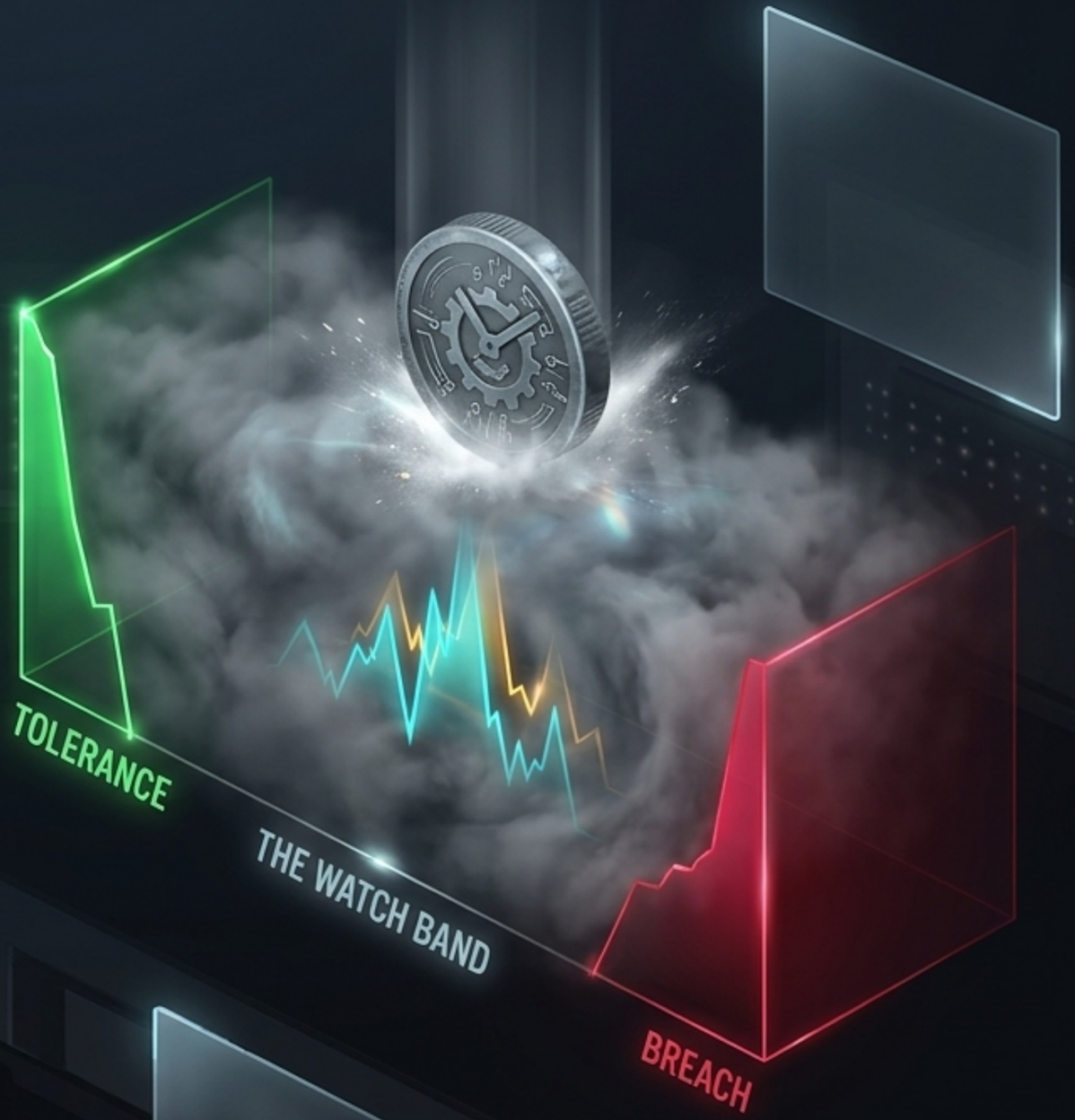
Horizontal review of ML credit models next quarter. Anything to escalate?

Gemini Notebook

THE WATCH-BAND DILEMMA

A number is a clue. Evidence is the answer.
The hard part is the middle band.

SIGNAL STATE	ACTION	COST	RISK
Within Tolerance	Clear it	Cost: Zero	Missing a micro-signal.
Watch Band	Raise, Clear, or Investigate	Cost: 1–3 Hours	Game will not tell you if it is noise or a mild defect.
Breach	Raise and Escalate	Cost: High confrontation	Will you hold the line on Friday?



Gemini Notebook

The Cost of False Alarms

Every exhibit is a test of judgment. Thresholds are assumptions for this training scenario.

False alarms cost you.

Every finding that turns out to be noise loses precision points.

Loss of precision equals a loss of credibility with the committee.



Gemini Notebook

The Defect Matrix

A new case.
A different truth.
Each case deals eleven areas into unpredictable states.



Some cases have no High-severity defects.

Sending everything back is not a strategy.

Which area is which is what the week is for.

Gemini Notebook

The Friday Confrontation

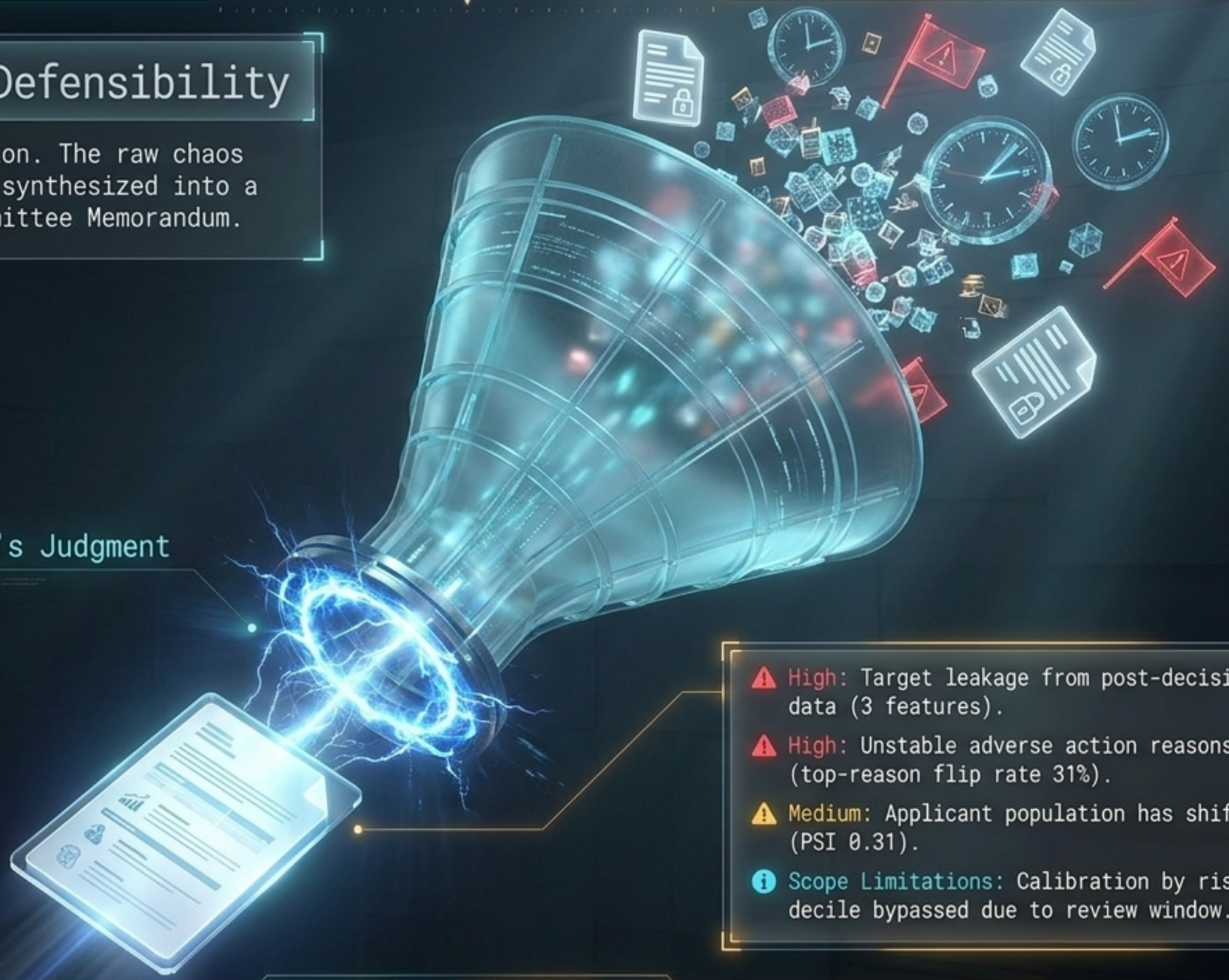
Hold what the evidence supports. The lead developer answers your strongest findings. Decide which reply should make you withdraw.



Gemini Notebook

The Anatomy of Defensibility

Your name is on the decision. The raw chaos of the simulation must be synthesized into a defensible Model Risk Committee Memorandum.



Validator's Judgment

- ⚠ High: Target leakage from post-decision data (3 features).
- ⚠ High: Unstable adverse action reasons (top-reason flip rate 31%).
- ⚠ Medium: Applicant population has shifted (PSI 0.31).
- ℹ Scope Limitations: Calibration by risk decile bypassed due to review window.

Gemini Notebook

Achievement Stratigraphy



50 pts

[The Bedrock]: Findings found and rated at the right severity.

20 pts

[The Core]: A validation opinion that fits what was really there.

10 pts

[The Shield]: Precision. Three points off per false alarm.

10 pts

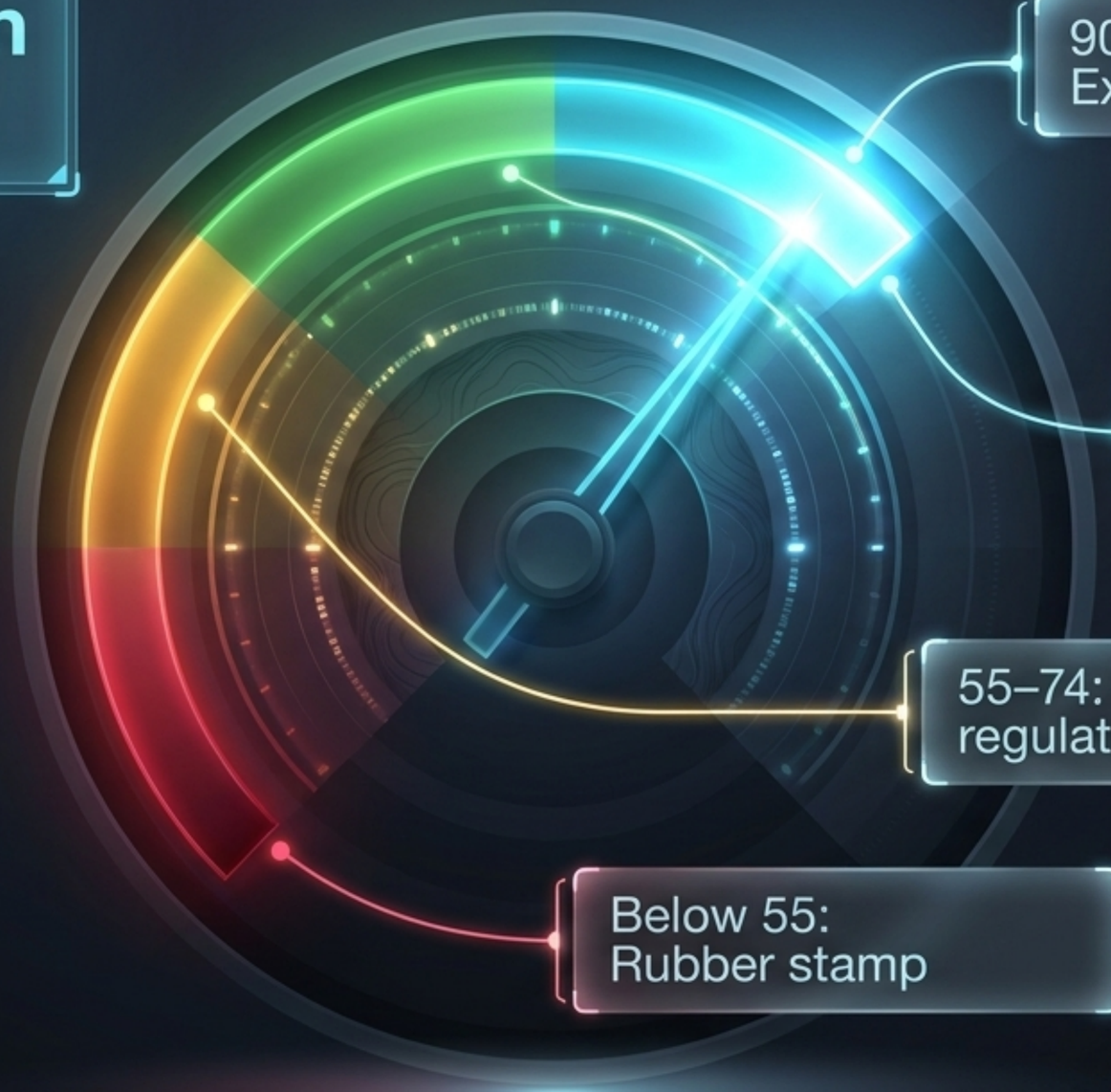
[The Rebuttal]: Developer challenges answered correctly.

10 pts

[The Protocol]: Conduct. Independence under pressure and early escalation.

Gemini Notebook

The Validation Outcome



90+:
Examiner-grade

75–89: Strong second line

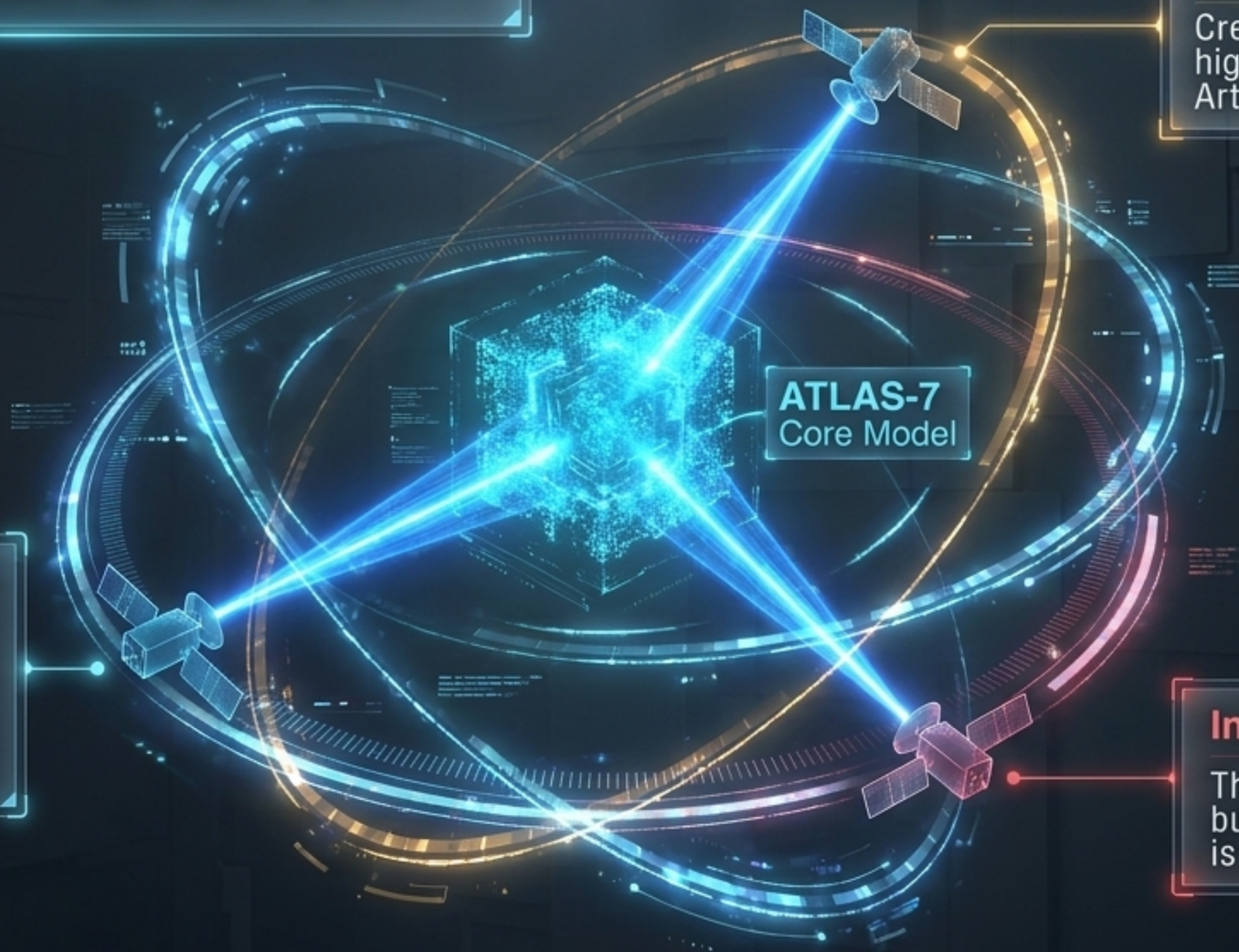
55–74: Gaps the regulator will find

Below 55:
Rubber stamp

Gemini Notebook

The Regulatory Constellation

The disciplines behind the decisions.



European Union (AI Act)

Creditworthiness scoring as high-risk AI.
Art. 10, 11, 15, 86.

Federal Reserve / OCC (SR 11-7)

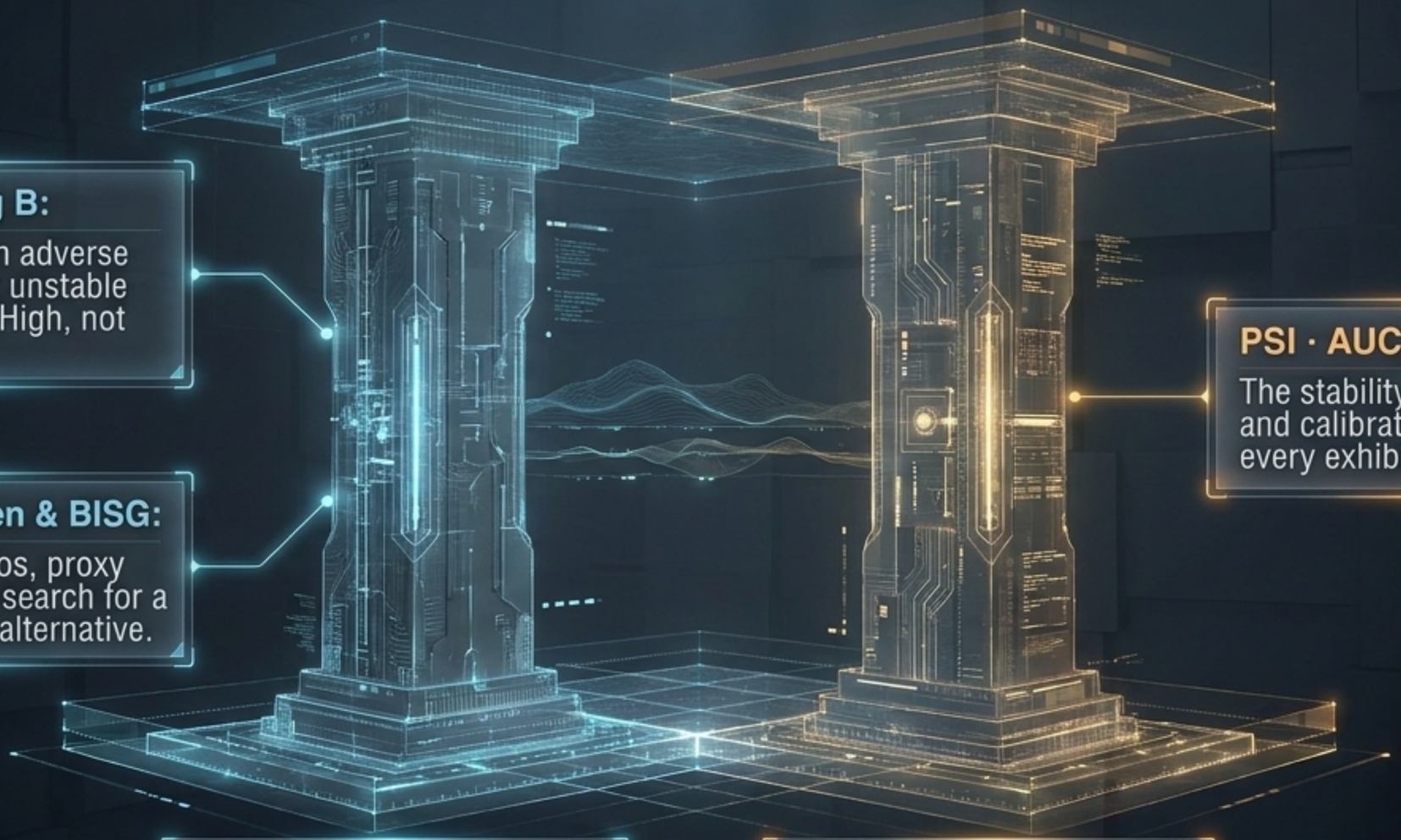
The spine of the test plan. Conceptual soundness, outcomes analysis, ongoing monitoring.

Interagency (2023)

Third-party risk. Why a bundled vendor embedding is your problem.

Gemini Notebook

Technical & Legal Architecture



CFPB ECOA Reg B:

Principal reasons on adverse action notices. Why unstable reason codes are a High, not a footnote.

Four-Fifths Screen & BISG:

Adverse impact ratios, proxy estimation, and the search for a less discriminatory alternative.

PSI · AUC · Calibration:

The stability, discrimination, and calibration thresholds every exhibit is read against.

Pillar 1: Fair Lending Practice

Pillar 2: Industry Convention

Gemini Notebook

Strategic Deployment

Model Validators

Practice effective challenge on a model that fights back, without a production deadline on the line.



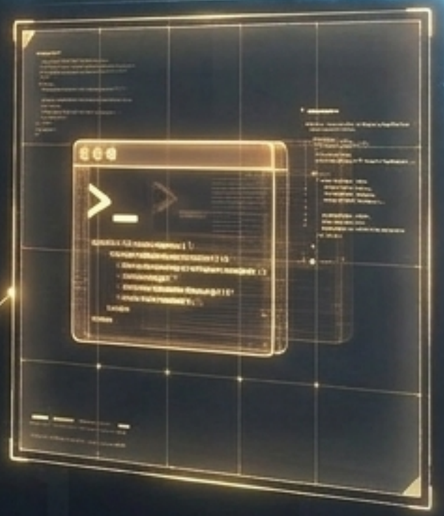


MRM & Risk Leaders

Calibrate a team. Run the same case code across a cohort and compare the memos.

Developers & Data Scientists

See your model the way the second line sees it. Learn which arguments survive a committee.





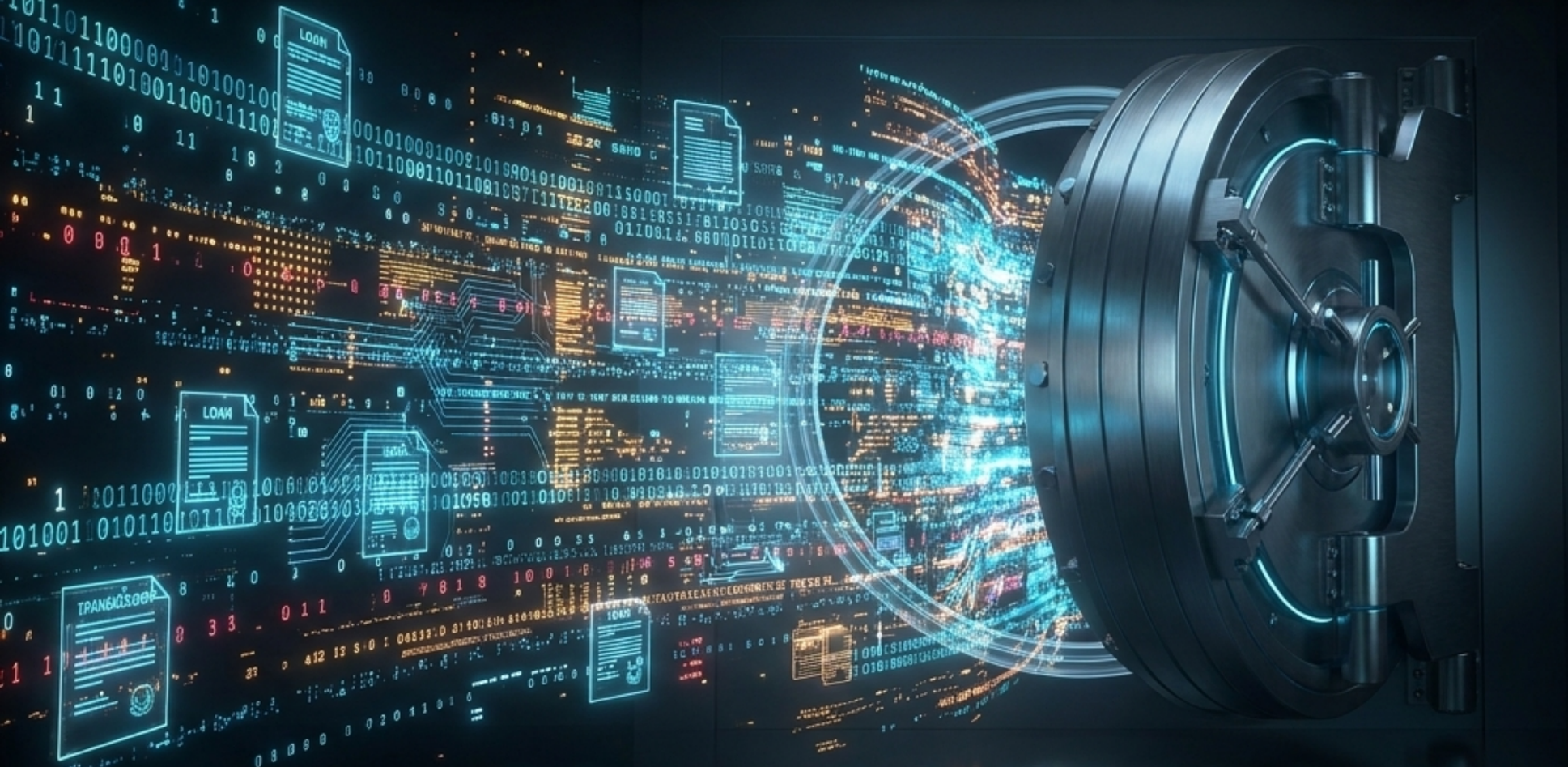
Audit & Examiners

Rehearse what a defensible validation file looks like, and exactly where weak ones break.

Gemini Notebook

The Final Threshold

ATLAS-7 decides 31,000 loans a month starting Monday.



Unless you find a reason it should not.

Gemini Notebook