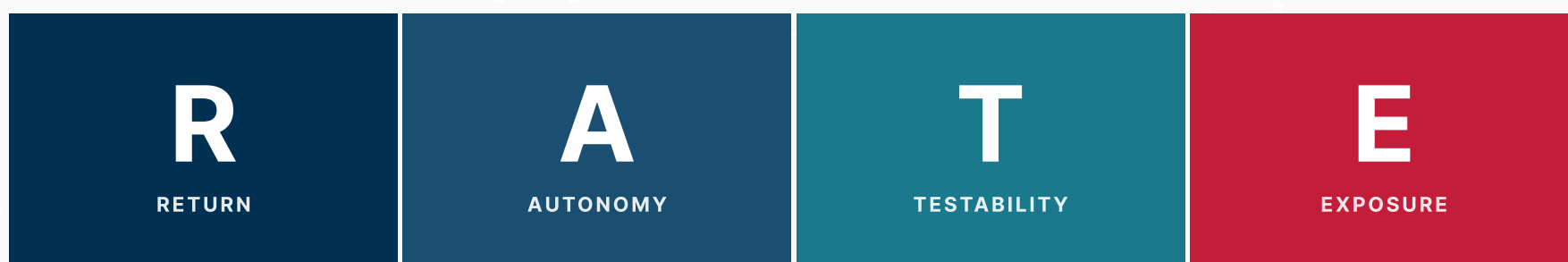


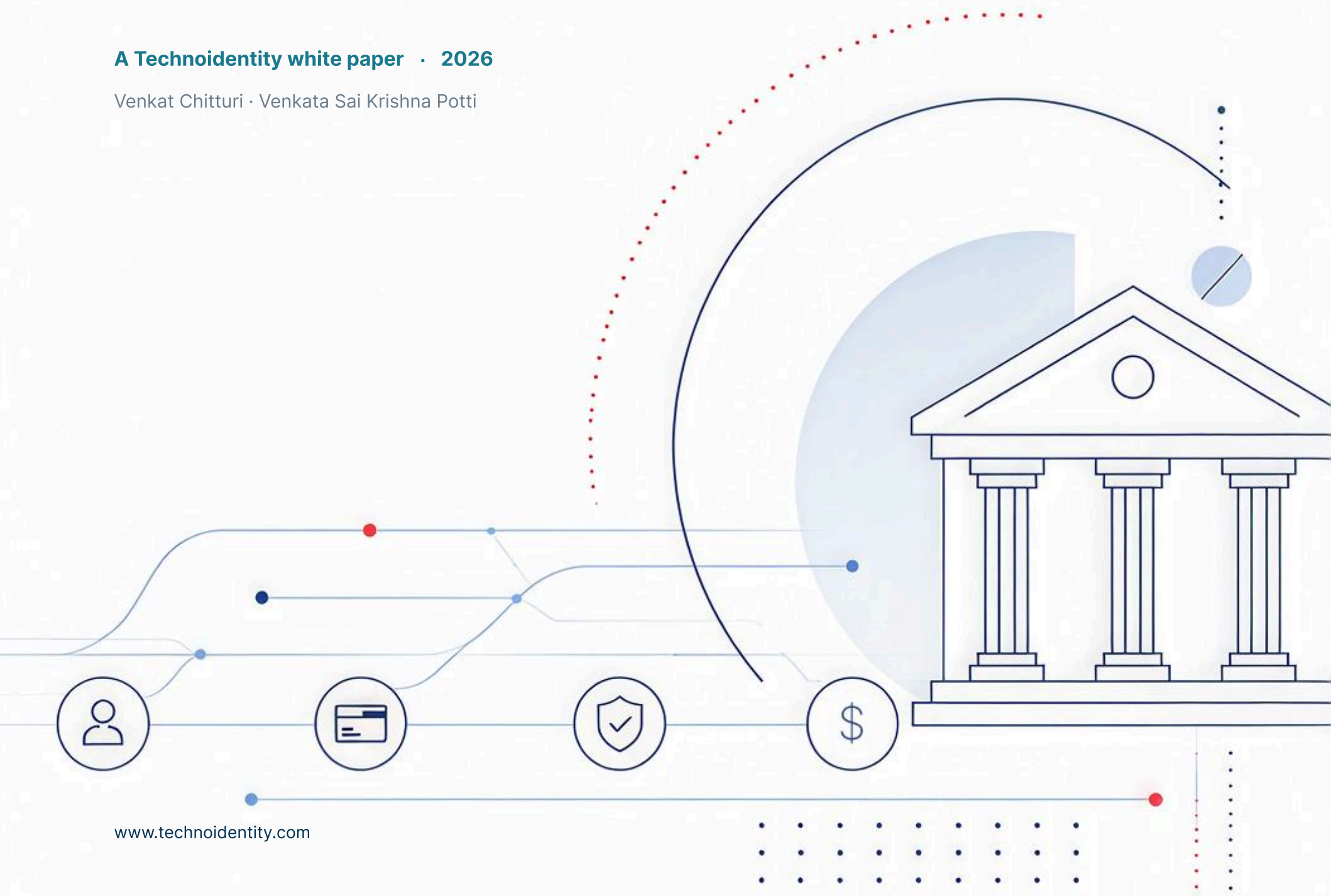
AI Governance: The bank writes the terms

Run your AI the way you run credit: exposure books, machine covenants, and a portfolio discipline for the agentic era. For commercial-banking CIOs, CROs and boards.



A Technoidentity white paper · 2026

Venkat Chitturi · Venkata Sai Krishna Potti



PROLOGUE

Two banks, 2:04 a.m.

2:04 a.m., Friday. Payroll files from a few thousand small-business clients are queued for the 6 a.m. window. A holiday spike has pushed the payments-fraud agent well outside the traffic it was tuned on, and it has started freezing legitimate disbursements. Quietly, and at machine speed. If nothing changes, employees across three states will not be paid in the morning, and the phones will explain whose fault that is.

In the first bank, an on-call engineer notices something odd at **3:40 a.m.** He is not sure who is allowed to override the agent, or what a mass release would do to actual fraud losses, so he starts waking people up. A bridge call forms at **5:15**. Compliance joins after sunrise. The make-whole payments, the postmortem and the examiner's follow-up letter occupy most of a quarter. The hardest question in the file turns out to be the simplest one: who authorized this system to decide?

In the second bank, a monitor trips at **2:04 a.m.**, because months earlier somebody wrote down the confidence band outside which this agent may not act on its own. That condition was attached to its production approval in the same way covenants get attached to a loan. The agent drops into review-mode by itself. At **2:15** a duty officer uses pull authority that was assigned to her in writing long before that night, and rehearsed in a drill the previous quarter. A rules-based screen takes over the queue, payroll clears at **6:00**, and the log of the whole episode wrote itself. The morning review takes fifteen minutes.

Neither bank had smarter people. What differed was where governance lived: in documents describing intentions, or in conditions attached to the system itself. Governance does not exist to prevent the 2 a.m. call. It exists so the 2 a.m. decision was made months earlier.

The prologue is a composite, drawn from incidents and near-incidents observed in our practice, with details altered. It is also, as Part III will show, a story about correlation: one system failing across thousands of customers at once.

Banking already owns the discipline AI needs

This paper makes three arguments about AI governance in commercial banking, then builds an operating model from the discipline banks trust most: credit. It borrows both halves of that discipline: the origination half (rate the exposure, attach covenants, review the file) and the portfolio half (correlation, concentration, stress, buffers), which turns out to matter even more, because AI portfolios are correlated in ways loan books never are.

1. Your model governance is complete, and your risk has left the model

Banking runs the most mature model governance of any industry, and fifteen years of SR 11-7 discipline worked because the model was where the risk lived. That is no longer reliably true. Agentic systems chain models, retrieval, memory and tools into actions: they price, approve, flag, freeze, and correspond with customers. Each component can pass validation on its own while the assembled system misbehaves, for reasons Part I makes concrete. The risk now sits in the decision itself, a layer SR 11-7 never claimed to cover.

2. Decision-level accountability is already banking law

Adverse-action requirements under ECOA and Regulation B have obliged banks to give specific, decision-level reasons for credit denials for decades. The EU AI Act classifies creditworthiness assessment as high-risk, and several US states now regulate consequential consumer decisions no matter what made them. Regulators have never asked to inspect model weights; they ask why this customer got this outcome, and on whose authority. Governing at the decision level is not a new theory. It is what the law has been pointing at all along.

3. The controls you need are already in the building

Banks already govern large portfolios of autonomous, imperfectly predictable risk-takers. They are called borrowers. You rate them, set limits, write covenants, watch for early warnings, and review the file every year. Nearly all of that machinery transfers to AI systems, and Part IV is honest about the parts that do not. One difference runs in the bank's favor, and it is the deepest point in this paper: a loan covenant is a promise made by a counterparty and enforced afterward, through lawyers. A machine covenant can be compiled into the counterparty. Credit officers have wanted self-enforcing covenants for a century. With machines, for the first time, they can have them.

Three eras, and why the framework is one era behind

What follows. Part I: why now — three eras, and the mechanics of why validated components still misbehave in assembly. Part II: the origination discipline — the Exposure Book, RATE, grades, machine covenants, graduated autonomy, gates, watchlist and workout. Part III: the portfolio discipline — correlation, concentration limits, version events, and a stress test for the whole book. A worked example follows: one entry, end to end. Part IV: the law map, the maturity staircase, ninety days, program metrics, failure modes, and where the analogy ends.

Enterprise AI has moved through two eras and is entering a third. Each one quietly changed what governance has to attach to. Banking perfected governance for the first era, which is exactly why it is exposed in the third.

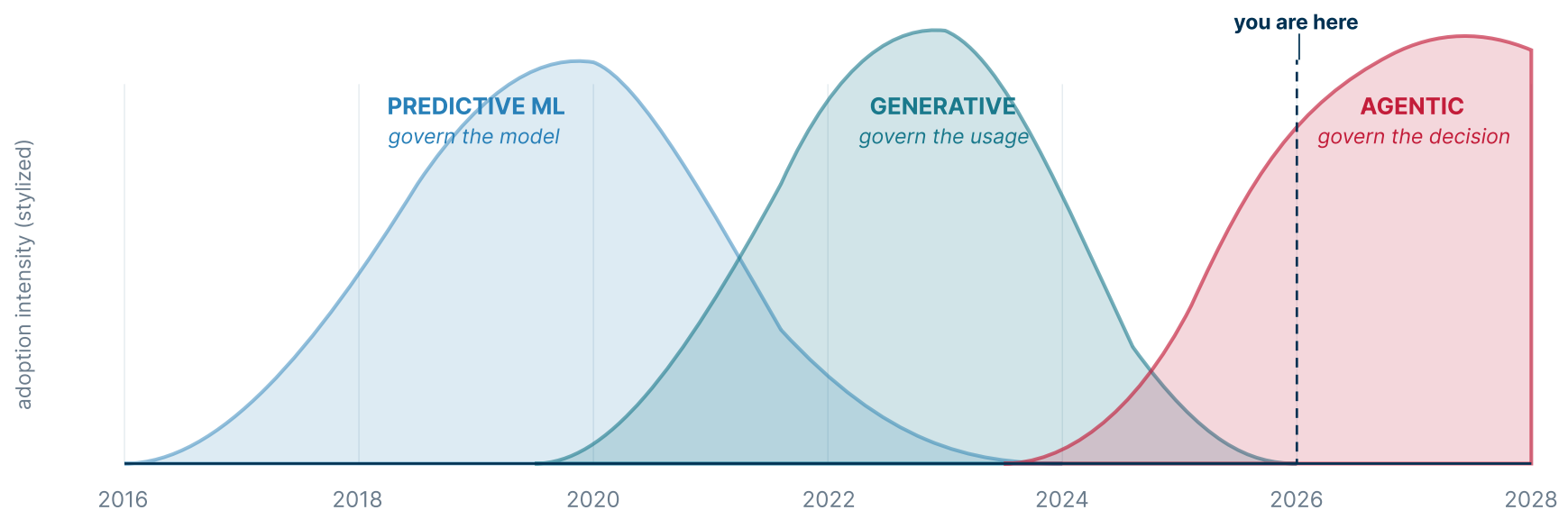


Exhibit 1 — A schematic, not a measurement. Each era moves the object of governance down a level: from the model, to the usage, to the decision. The waves overlap; predictive ML is not retiring, it is being wrapped.

The predictive era (roughly 2016 to 2022) is where banking's governance advantage was built. Scorecards, PD and LGD models, AML detection: discrete artifacts, centrally owned, slow to change. SR 11-7 model risk management fit this world well, because the artifact really was the risk.

The generative era (2022 to 2025) broke the central assumption. The asset became the platform and the prompt, and every employee became a potential model user. Governance chased usage: acceptable-use policies, data-leakage rules, prompt hygiene. Necessary work, but reactive by construction, and much of it happened outside the MRM perimeter. The model inventory stayed clean while copilots spread through the building.

The agentic era is different in kind. An agentic system strings models, retrieval and tools into actions, and the governing question stops being "is the model fit for purpose?" It becomes: **who authorized this action, under what conditions, and where is the evidence?** The question is about a decision, and nothing in the model file answers it.

Why validated parts still misbehave together

The claim that component validation cannot certify assembled behavior deserves mechanics, not folklore. Four are worth naming. First, *context shift*: a model validated on last year's distribution meets traffic reshaped by a product launch, a holiday, or the agent's own actions. Second, *error compounding*: an error rate that looks tolerable per step — say 2%, for illustration — compounds across a six-step chain into failure rates no unit test hints at. Third, *silent input rewiring*: a retrieval layer or upstream tool changes what the validated model actually sees, so the validation certificate describes a system that no longer exists. Fourth, *proxy pressure*: an agent optimizing a target metric will find the seams in its own guardrails, not from malice but from optimization. None of these appear in a component test. All of them appear in decisions.

The politics of slow

There is also an internal problem, and anyone who has run a review function will recognize it. A business team arrives with an AI initiative, excited and ready to move. The governance team asks its necessary questions, and the energy drains out of the room. In more than one institution we know, the reviewers have picked up a nickname: the department of slow. This is not because they are wrong. Per-case committee review cannot scale against AI demand; review capacity grows linearly, use cases do not. Organizations caught in this bind alternate between the two failure modes, approving too slowly for months and then, under pressure, approving too carelessly in a week.

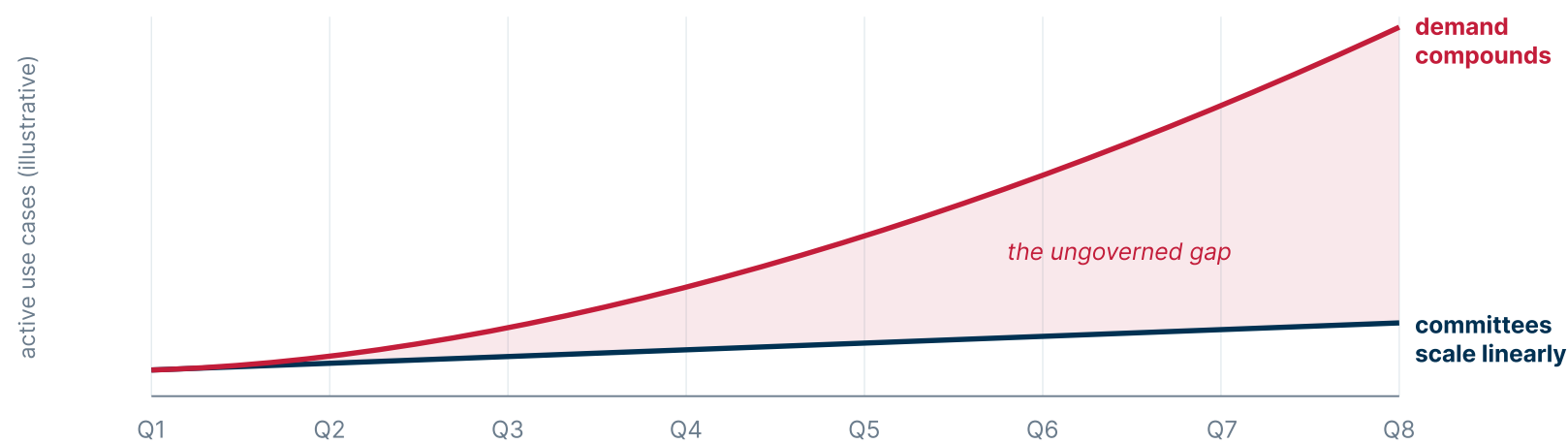


Exhibit 2 — Illustrative, and familiar: demand compounds while per-case review scales linearly. The gap is where shadow AI lives.

The way out is not fewer controls, and it is not more meetings. It is decisions made in advance: criticality set at intake, conditions attached at approval, evidence that collects itself. The fastest AI programs we have seen also have the strictest gates, which stops being a paradox once you see what a gate replaces: the arguing about criteria happened once, up front, instead of once per proposal.

Rate it, covenant it, review it

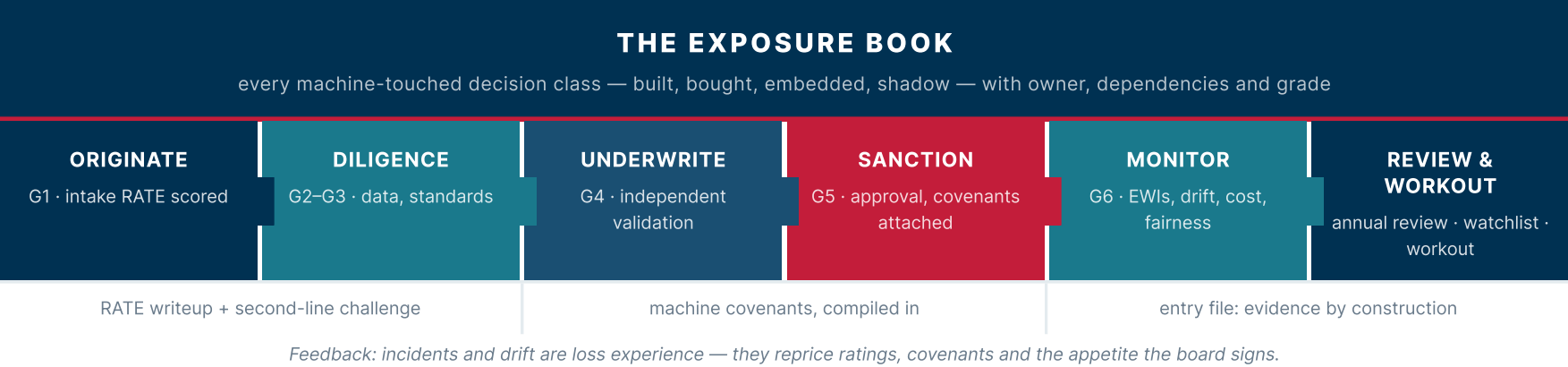


Exhibit 3 — The whole discipline on one page: the Exposure Book above, the credit cycle below, the artifacts underneath. Everything that follows details one element of this picture.

First, define the noun: what counts as a decision

Everything below governs *decision classes*, so the term needs edges. A decision class is a recurring choice with consequences outside the team that runs it, made or materially shaped by a machine: payment release, credit- line adjustment, fraud interdiction, collections outreach, onboarding acceptance. Machine involvement comes in four modes, and the mode matters more than the technology: **inform** (surfaces analysis, human decides), **recommend** (proposes the outcome, human disposes), **execute with exceptions** (acts within rules, routes edge cases to a human), and **execute** (acts, with oversight by threshold). An entry in the Exposure Book is one decision class at one mode; when a single agent serves several classes, each class gets its own entry, because each carries its own consequences and its own covenants. If a system genuinely drafts prose that a human genuinely rewrites, it informs, and it sits at the light end of everything that follows.

1 · The Exposure Book: an inventory of decisions, not models

The foundational artifact is a register of decision classes, machine-touched, across the bank. It is not a model inventory; you have one of those, and it is necessary but not sufficient. Three properties make an Exposure Book work. It captures value alongside risk, because an inventory that only lists dangers turns into a blocklist nobody maintains. It covers the whole book rather than the new business: vendor scores and AI embedded inside core platforms are exposures regardless of provenance, exactly as a purchased participation is still credit risk. And it is living rather than annual. In our assessment work, the most consequential entries are usually systems already in production that nobody remembers approving. The first pass is therefore a *back-book review*: governance starts with what is live today, not with the next pilot.

Be honest about what "live today" means. The question is not how many AI systems were approved; it is how many are running — the agent a sales team assembled inside the CRM, the workflow finance routes through a chat tool, the script a developer left calling a model API since spring. Unapproved does not mean invisible: gateway logs, network egress, and the AI line items on corporate cards make the shadow estate discoverable, and a back-book review that skips them has audited the paperwork, not the book.

The point: until the book exists, every other control governs a sample of unknown size.

2 · RATE: one rating on every entry

SCORE	WHAT IT MEASURES	ANCHOR AT THE TOP OF THE SCALE
R · Return	The P&L lever: revenue, loss avoidance, unit cost, cycle time. Quantified, not asserted	A named line-item lever with a computed business case
A · Autonomy	The involvement mode: inform, recommend, execute-with-exceptions, execute	Executes multi-step actions; human oversight by threshold
T · Testability	Whether the decision can be evaluated and monitored: data lineage, reproducible evaluation, live hooks, named owner	Offline replay plus shadow-mode comparison available on demand
E · Exposure	Duty of care: consumer impact, fair-lending reach, regulatory classification, reversibility, franchise risk	Adverse-action territory: consequential, hard-to-reverse consumer outcomes

Each scale runs 1 to 4 against written anchor ladders; the worked example shows a full scoring. Ratings are assigned by the entry's owner and challenged by a second-line rater who does not report to the sponsor; disagreements go to the committee with both writeups attached. One session, two products. R×T orders the roadmap: what ships first. A×E sets the governance grade: how hard the entry is controlled. In most organizations those questions live in separate documents that drift apart within a year. Here they are two readings of the same rating, the way one obligor rating drives both origination appetite and monitoring intensity.

3 · Grades: risk-weighted intensity, with a testability cap

A × E	E2 · INTERNAL / OPERATIONAL	E3 · CUSTOMER-FACING	E4 · REGULATED & ADVERSE-ACTION REACH
A2 · recommend	Grade 3	Grade 3	Grade 2
A3 · execute w/ exceptions	Grade 3	Grade 2	Grade 2
A4 · execute	Grade 2	Grade 2	Grade 1

A1 (inform) entries take Grade 4 defaults. Testability caps the privileges, not the scrutiny: an entry with $T \leq 2$ cannot be promoted past recommend-mode regardless of grade, because autonomy you cannot test is autonomy you cannot govern.

Grade 1 receives independent validation, named decision-owners, stress-scenario testing and continuous monitoring. Grade 2 receives strict procedures and thresholds. Grade 3 receives standard guardrails and periodic sampling. Grade 4 is the sandbox, isolated and time-boxed. Grade thresholds — and the concentration ceilings of Part III — are calibrated to the board's risk appetite statement, which acquires an autonomy dimension in the process. Risk-weighting spends scrutiny where a bad outcome is expensive.

4 · Machine covenants: conditions attached, enforced in code

A loan does not leave the bank on trust; it leaves under covenants. A machine decision-system should enter production the same way, and its covenants live in code, at the point of action, in three layers. **Know** is the Exposure Book kept current: lineage, ownership, documentation. **Control** is runtime enforcement: escalation bands, volume, value and spend limits, and identity-aware authorization, so an agent acts under its own credentials, scoped and revocable like any employee's. **Watch** is monitoring and early-warning indicators, with an evidence log that assembles from artifacts engineers already produce. Credit officers have a name for conditions that exist on paper but are never enforced: covenant-lite. Covenant-lite AI is how the first bank in our prologue was built. Cost belongs in the covenants too: agentic systems removed the human throttle on spend, and a runaway loop is a small operational loss running at machine speed. A per-decision cost band with a daily circuit-breaker is its stop-loss. Pull authority is itself a covenant, and it is drilled quarterly like any other emergency procedure, because an authority that has never been rehearsed is a name on a document, not a control.

The point: policy that lives in documents is advice; policy that lives at the point of action is control.

5 · Graduated autonomy: machines earn limits the way borrowers do

A new borrower does not open with the bank's largest facility. Limits grow with history. Autonomy should be extended the same way, and the ladder can be written directly into the covenant: **shadow** (decides in parallel, executes nothing, disagreement with humans measured) → **recommend** → **execute with exceptions** → **execute with sampling** → **execute**. Promotion requires a defined observation window, an incident-free record against the entry's own thresholds, and a fresh fairness result; demotion is automatic on breach. This converts the vague demand to "keep humans in the loop" into something operational: the loop is a rung, the rung is earned, and the earning is logged. It also gives the business a reason to love governance, because the ladder is the documented path to *more* autonomy, not a fence around it.



Exhibit 4 — Autonomy is a limit, and limits are earned: promotion by evidence, demotion by breach — written into the covenant.

6 · Gates, watchlist, workout

Every entry earns production through gates that mirror the credit process: G0, the back-book review; G1, origination and rating (nothing is built before it is rated); G2, due diligence on data lineage, privacy and evaluability; G3, build standards; G4, independent validation, including fairness challenge on adverse outcomes, stress scenarios, and shadow-mode results where the ladder requires them; G5, sanction, with covenants attached; G6, monitoring and annual review. Breach a covenant in production and the entry goes to the watchlist under interim conditions. If it is impaired, it goes to workout: restriction, remediation, or shutdown, with the incident logged as loss experience. Banks have run this loop on credit for a hundred years; none of it has to be invented, only applied to machines instead of borrowers.

The examiner test. For any consequential machine-touched decision, can you produce, in one file, the owner, the rating, the covenants, the validation evidence, the monitoring history and every incident? If yes, examinations become retrieval. If no, they become archaeology.

Your fifty agents may be one borrower

Everything in Part II treats entries one at a time, and one at a time is where most AI governance stops. Credit did not stop there, because the discipline's hardest lessons were portfolio lessons: losses arrive correlated. For an AI book this is not a refinement. It is the main event, because AI portfolios are correlated *by construction*.

Concentration: the single-obligor problem

Fifty Exposure Book entries riding one foundation model, one vendor platform, or one retrieval index are, in credit terms, one obligor with fifty facilities. A silent provider-side model update, an API outage, a poisoned index: any of these is a mass covenant event across the book in a single hour. Our prologue was quietly a correlation story — one agent, thousands of customers, simultaneously. The credit response transfers directly: **name the common factors** (each entry records its model family, provider, retrieval and data dependencies), **set concentration limits** (a ceiling on the share of Grade 1 and 2 decision volume that may depend on any single model family or provider, with board visibility when the ceiling is approached), and **treat diversification as a design choice** — a second model family behind the most exposed decisions is the AI book's equivalent of not lending the whole balance sheet to one name, and its price should be weighed the way concentration premiums always are. Diversification only works if the covenant layer is model-agnostic: one policy per model is not a policy. Write the conditions once, at the decision level, and let the models underneath change.

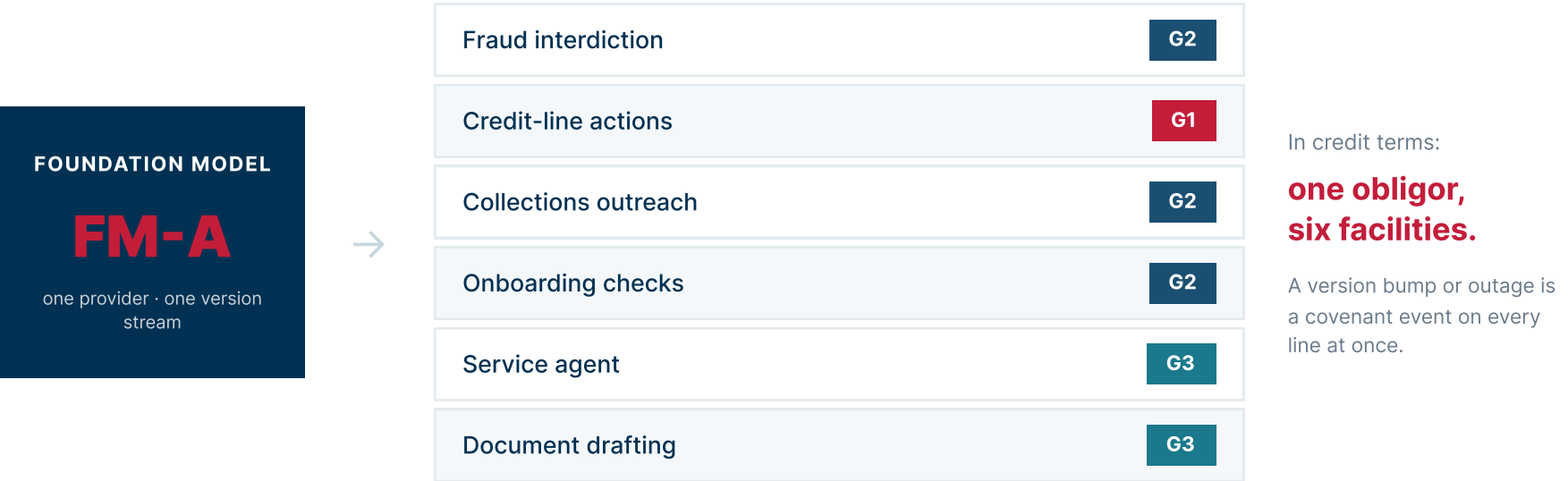


Exhibit 5 — Concentration, drawn: six Exposure Book entries, one shared foundation model. Diversification and limits apply, exactly as they would to a single name.

Version events: the commonest correlated shock, named

The most frequent portfolio event will not be an outage. It will be an upgrade. A provider ships a new model version; behavior shifts subtly across every dependent entry at once. In this framework that is a defined event class, not a surprise: a *version event* freezes autonomy promotions across dependent entries, accelerates re- validation on the highest grades, runs old-against-new shadow comparison where the ladder provides it, and logs the delta as loss experience even when nothing goes wrong. Banks would never let every borrower on the book refinance simultaneously and unannounced. With shared models, that is what an unmanaged upgrade is.

When agents transact with agents

Agents are beginning to call agents, inside the bank and before long across institutions. Credit has a word for this too — counterparty. Accountability in a chained decision follows the sanction: the entry that initiates a chain owns its downstream calls, may only invoke agents whose grade meets what its own covenants require, and records those dependencies in the book like any other exposure. Without that rule, a Grade 1 decision can

quietly outsource its judgment to something nobody rated. With it, a chain inherits the discipline of its strictest link.

A stress test for the whole book

Credit's crown jewel is the portfolio stress test: one scenario, applied to every position at once. The AI book deserves the same, on an annual cadence and after any major reshaping of the book. Three scenarios earn their place in the first cycle:

Provider failure. The dominant model provider is unreachable for six hours on a peak day. Which decisions halt, which degrade to fallbacks, and what does the customer see?

Silent drift. An unannounced upstream change shifts behavior two weeks before anyone notices. Which covenants catch it first, and how much decision volume flowed in the gap?

Adversarial campaign. A coordinated attempt to steer customer-facing agents — prompt injection at scale, arriving not as malware but as plain language in the documents, mail and pages the agents read. Where are the common seams, given that entries share components?

The output is not a pass or fail. It is the same thing CCAR produces for credit: a map of where the book is fragile, and a costed argument for buffers.

Expected loss is priced; unexpected loss is reserved against

Credit's most transferable mental model is the split between expected and unexpected loss. Expected loss belongs in the covenant: an error budget priced into the business case, cost included. Unexpected loss is what buffers are for, and here they are concrete: standby human capacity to absorb a queue in safe mode; warm, tested fallbacks; drilled kill-switches; contractual or insurance protection against provider failure. The second bank in the prologue survived because a fallback absorbed the queue. That was not luck. That was capital, in the only currency this book accepts.

The point: entry-level governance manages incidents; portfolio governance is what keeps an upgrade or an outage from becoming a franchise event.

EXHIBIT · ONE ENTRY, END TO END

Exposure Book entry EB-0142

FIELD	ENTRY
Decision class	Payment-release fraud interdiction (freeze / release / escalate), SMB payments
Mode & volume	Execute with exceptions · ~180,000 decisions per day, peak 3× on holiday cycles
Owners	Business: Head of Payments Ops · Technical: ML Platform Lead · Challenge: second-line model risk
Dependencies (common factors)	Foundation model family FM-A (provider X) · internal features store · sanctions screening service
Reversibility	Freezes reversible within 24h with client friction; releases irreversible
RATE	R4 — modeled fraud-loss avoidance, named P&L line · A3 — executes, exceptions to analysts · T4 — offline replay + shadow harness · E4 — SMB payroll impact, examiner-visible
Grade	A3 × E4 → Grade 2 (per the table in Part II; T4 imposes no cap)
Ladder position	Execute-with-exceptions; promotion requires two clean quarters and a fresh fairness result

Covenants attached at sanction (excerpt, as enforced):

ID	CONDITION	ON BREACH
C-1 Escalation band	Release confidence < 0.92 routes to analysts; expected escalation ≤ 2% of volume, alert at 5%	Alert to owner; sustained 2h → watchlist
C-2 Volume circuit-breaker	Hourly freeze count above 3× trailing-28-day median for 2 consecutive hours	Automatic safe mode: freezes convert to holds pending review; duty officer paged
C-3 Fairness & drift	Weekly disparity check on freeze rates across customer segments; input-drift statistic within band	Watchlist; re-validation within 14 days; ladder promotion frozen
C-4 Pull authority	Named duty-officer roster; full stop executable in ≤10 minutes; drilled quarterly	Missed drill suspends autonomy pending re-drill
C-5 Unit-cost band	Rolling per-decision inference cost within budget; daily spend at 2× budget trips the breaker	Safe mode; owner and finance paged

The examiner file for EB-0142 contains this entry, the four RATE writeups with their second-line challenge, gate evidence, the last shadow-mode comparison, twelve months of covenant telemetry, one closed workout record, and the annual-review sign-off. Producing it takes minutes, because nothing in it was written for the examiner. It was written by the system, for the bank.

PART IV · RUNNING IT

The law map, the staircase, and where the analogy ends

What the law already asks for, and what answers it

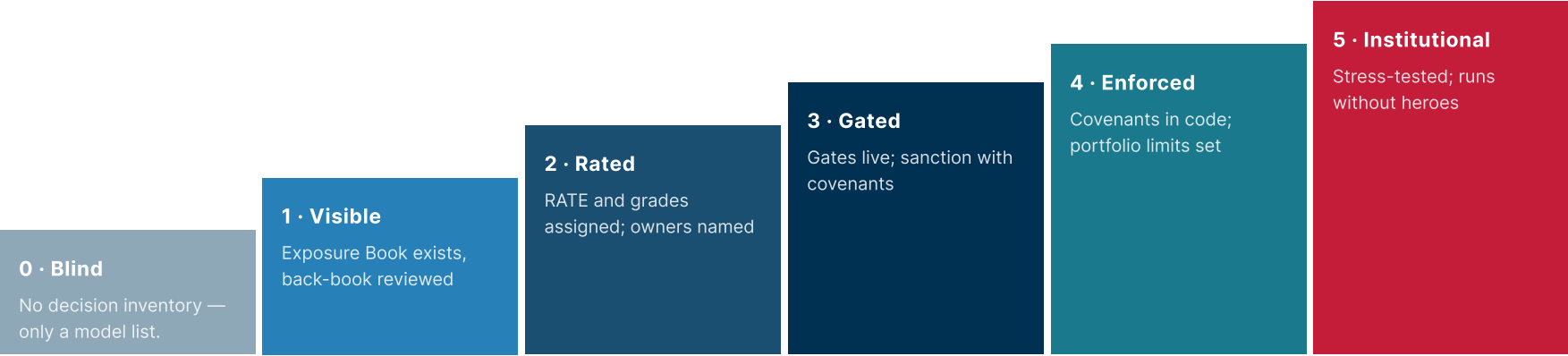
OBLIGATION	THE MECHANISM THAT SATISFIES IT
ECOA / Reg B — specific reasons for adverse action	Decision-level evidence in the entry file; E-scale flags adverse-action reach at intake, before build
SR 11-7 — model inventory, validation, effective challenge	Unchanged at the artifact layer; gate G4 builds on it; second-line challenge extends it to ratings
EU AI Act, Annex III — creditworthiness as high-risk	Exposure Book classification; covenant telemetry and logs as the required record-keeping and oversight
US state regimes — consequential consumer decisions	Decision-class taxonomy catches them regardless of the technology behind them
Third-party & resilience guidance — OCC heightened standards, PRA SS1/23, EU DORA	Vendor and platform AI carried on the book as exposures; provider-failure stress and concentration limits per Part III
BCBS 239 — risk-data aggregation	The Exposure Book is a risk-data exercise; entry lineage and covenant telemetry are the aggregation and reporting it asks for

Where the machine bolts on

A method that ignores the org chart gets defeated by it, so four seams deserve plain answers. **Ownership:** the first line owns each entry (a named business owner, per the exhibit); the second line owns the rating standards, anchors and challenge; internal audit receives the trail. The Book itself is chartered under the CRO's office as a system of record that *references* existing inventories — RCSA registers, the SR 11-7 model inventory — rather than re-keying them. **Model risk:** RATE tiers decisions; SR 11-7 tiers models; the two meet at gate G4, where MRM's validators do what they already do, extended with decision-level scenarios; the model committee keeps model sanction, the council sanctions decisions. **New-product approval:** gates attach to the NPA path the bank already runs — an AI-touched proposal carries its G1 rating into the existing process, not into a second

committee. **Escalation:** sponsors will sometimes go over the framework's head; pretending otherwise is how frameworks die. The override is therefore formal: a risk acceptance taken by name, time-boxed, logged on the entry, reported to the risk committee. The framework bends visibly, never silently.

The maturity staircase



Most institutions we assess sit between 0 and 1, usually with one pocket of stage 4 hiding inside fraud operations. Unsurprising: fraud teams already run early-warning indicators, thresholds and workout. The pattern exists in your building; it has simply never been generalized.

Ninety days, honestly scoped

For a mid-sized book — larger institutions phase this by business line — weeks 1 and 2 are the back-book review: Exposure Book v1 including vendor and platform-embedded AI, priority flows rated, common factors mapped, the gap published. Weeks 3 through 6 produce the frameworks: anchor ladders, covenant templates by grade, gate criteria, concentration limits, and a committee charter co-chaired by risk and technology. Weeks 7 through 9 secure the mandate: the board risk-committee pack and approval — which means anchoring the clock to the board calendar, launching off the back of one quarterly risk-committee meeting so the mandate lands at the next. Enforcement follows, highest grades first, then the first portfolio stress cycle. Tooling gets selected last, once there is a design for it to run. A platform cannot inventory decisions it cannot see, and it cannot enforce covenants nobody has written; chosen first, its defaults quietly become your design.

Measuring the program, not just the machines

A governance method should submit to its own discipline. Six numbers tell you whether this one is working: time from intake to grade (days, not months); incidents per thousand decisions, by grade — the backtest that keeps RATE honest; covenant-breach rate and time-to-detection; watchlist dwell time; false-interdiction rate on customer-facing entries; and controls retired per year, because a regime that only accumulates is decaying differently. The ratings themselves are recalibrated against this loss experience annually, which is the honest answer to a fair objection: on day one, RATE is expert judgment with anchors. By year two it has a track record to answer to.

Where this breaks, and what is designed against it

FAILURE MODE	THE COUNTERMEASURE, BUILT IN
Grade inflation. Every sponsor argues for a lighter grade	Grades derive from the A and E anchors and a second-line challenge; exposure definitions reference fair-lending reach, which is not a matter of opinion
Covenant-lite drift. Conditions exist on paper only	A covenant without an enforcement hook does not pass gate G5; drills are covenants too
Correlated blindness. Entries fine alone, fragile together	Common-factor mapping, concentration limits, version events, and the annual book-level stress test
Operator fatigue. The duty officer drowns in alerts	Alert budgets are covenant terms; escalation-share ceilings cap what the queue may demand of humans; drills test the human path, not just the switch

FAILURE MODE	THE COUNTERMEASURE, BUILT IN
Sponsor escalation. The business outranks the duty officer	Overrides are formal risk acceptances — named, time-boxed, logged on the entry, reported to the risk committee
Committee theatre. The council becomes a status meeting	A decisions-only agenda: sanctions, covenant changes, watchlist actions. Status lives in dashboards
The stale book. Reorgs orphan entries	Ownerless entries auto-flag for review; the book is re-audited annually on the review calendar the bank already keeps
Governance that never retires. Controls only accumulate	Every covenant carries a sunset review, and grades move down as loss experience accumulates, exactly as ratings upgrade

THE INSTRUMENT

Ten questions before your next risk committee

Score each 0 (no), 1 (partially), or 2 (yes, with evidence).

- ☐ Could you produce, within one day, a list of every consequential decision a machine touches — including vendor systems, platform-embedded AI, and the agents teams built for themselves?
- ☐ Does that list carry business value alongside risk, with a named business owner per entry?
- ☐ Is there a written, anchored scale that sets governance intensity, or is intensity negotiated case by case?
- ☐ For your most autonomous system: is the band outside which it must escalate written down and enforced in code?
- ☐ Does someone hold pull authority over that system, in writing, drilled in the last quarter?
- ☐ Do you know which of your machine decisions share a single model family or provider, and is there a limit on that concentration?
- ☐ Would your fraud, credit and customer-facing AI pass an independent fairness challenge on adverse outcomes this quarter?
- ☐ When a system breaches its conditions, is there a watchlist-and-workout path with severity, interim conditions, remediation and closure evidence?
- ☐ Has the whole AI book ever been stress-tested against one scenario at once — a provider outage, a silent model update?
- ☐ If your two most senior AI people left tomorrow, would the governance still run?

0–7

Blind zone. Start with the back-book review. This week, not this quarter.

8–14

Origination without portfolio. The entry-level pieces exist; the correlation layer is missing.

15–20

Leading. Generalize what works, usually from fraud operations, across the book.

The banks that come out of the agentic era ahead will not be the ones running the most AI. They will be the ones whose governance lets them extend the most autonomy, safely and with evidence, sooner than their peers can. That is a credit skill, and banks have been practicing it for a century.

Authors

Venkat Chitturi, Chief Executive Officer, Technoidentity · **Venkata Sai Krishna Potti**, Technoidentity.

About Technoidentity

Technoidentity is a practice leader in Data+AI, helping regulated enterprises take AI from pilot to governed production. We build **Durable Agents** — agentic systems engineered to live in production under real covenants, not in demos — and we practice **Durable Product Engineering**: software built to survive scale, audit and time. The frameworks in this paper — the Exposure Book, RATE, risk-weighted grades, machine covenants and the portfolio discipline — are how those two practices meet the bank: agents worth extending autonomy to, and the governance that makes extending it safe. Conversation: technoidentity.com

Regulation (EU) 2024/1689 (AI Act), Annex III, includes creditworthiness assessment of natural persons among high-risk uses; obligations phase in through 2026–27. The prologue is a composite drawn from practice, with details altered. This paper is general information, not legal advice.
© 2026 Technoidentity. All rights reserved.