TANTOR™
a translab.io product

TANTOR WHITEPAPER

Owning the Moment

Why Real-Time, Explainable Mule Prevention Belongs to the Bank

A perspective for the CXOs and the Risk Committee.

info@tantor.ai | tantor.ai

Table of Contents

1. Executive Summary

2. The exposure facing the bank

3. The control gap: why detection does not discharge the risk

4. Options considered

5. The recommended capability, and how it works

6. Conditions of action

7. Legal basis and the limits of acting in the moment

8. Risks, dependencies, and assumptions

9. Why now: the direction of regulation and threat

10. Tantor in practice

11. Next steps

12. Glossary of key terms and regulations

13. About Tantor

14. About Translab Technologies

Executive Summary

This paper asks the CXOs of Indian banks to adopt a bank-owned, real-time prevention layer that sits in the payment decision path and can hold a suspect payment, on the bank's own rails, for an accountable human decision before the funds leave. It should operate alongside the shared intelligence the regulator and the wider ecosystem provide, not in place of it. The two fundamental requirements for that layer are: explainability, so that every intervention is defensible to the regulator, governable by the bank, and communicable to the customer without tipping off a suspect; and precision, so that the decision threshold is owned by the board as a statement of risk appetite and the bank stops mules without stopping its own customers.

The reason the decision is urgent is a structural change in the risk: real-time payment rails have made fraudulent transfers irreversible. The harm, the loss to the customer, the exposure to the bank, and the damage to the bank, is now created in the instant a payment completes, not during a settlement window that no longer exists. The money mule account, used to receive and move on the proceeds of fraud, is the instrument that exploits this speed, and the scale is significant: more than 24.67 lakh mule accounts have been identified in India, against citizen-reported cyber-fraud losses of ₹22,845.73 crore in 2024[1].

[1]Ministry of Home Affairs / Indian Cyber Crime Coordination Centre (I4C) data: India has identified more than 24.67 lakh mule accounts, and citizen-reported cyber-fraud losses were ₹22,845.73 crore in 2024 (a rise of about 206% on the previous year), per Ministry of Home Affairs disclosures, 2025.

In February 2026 the RBI Ombudsman directed 5 beneficiary banks to pay compensation, attaching the loss to the banks that hosted the mule accounts. The duty, and the liability, land on the institution whose rails the money crossed.

Executive Summary

That accountability is already the bank's, and the global direction of travel is reinforcing it: across the United Kingdom, the European Union, and other markets, liability for payment fraud is shifting onto the institution that receives it. Detection, however good, cannot stop a payment on the bank's own rails; only the bank can. Approving this capability aligns the bank's control with an accountability it already carries, and the cost of inaction is a control gap the bank is already paying for.

The recommended next steps are to

1. assess the bank's mule exposure
2. map where detection currently stops short of a decision
3. identify the points in the payment flow where a transaction can be intercepted
4. approve a real-time hold policy calibrated to the bank's appetite
5. prepare to participate in federated intelligence as the ecosystem builds it

---

[1] Ministry of Home Affairs / Indian Cyber Crime Coordination Centre (I4C) data: India has identified more than 24.67 lakh mule accounts, and citizen-reported cyber-fraud losses were ₹22,845.73 crore in 2024 (a rise of about 206% on the previous year), per Ministry of Home Affairs disclosures, 2025.

The exposure facing the bank

For most of banking history a fraudulent payment was recoverable.

Cheques cleared over days; wire transfers passed through correspondent banks with hours between instruction and settlement, and a bank that detected fraud after the fact still had time to recall the instruction or hold the funds in transit. Detection and prevention were the same activity, separated only by a delay.

Real-time payments removed the delay.

On the Unified Payments Interface (UPI), a transfer is initiated, cleared, and settled in under a second, at any hour, between any two accounts within India. Money that leaves a money mule account is gone before a periodic monitoring system has finished identifying it. The harm is no longer created during settlement; it is created in the instant the payment completes.

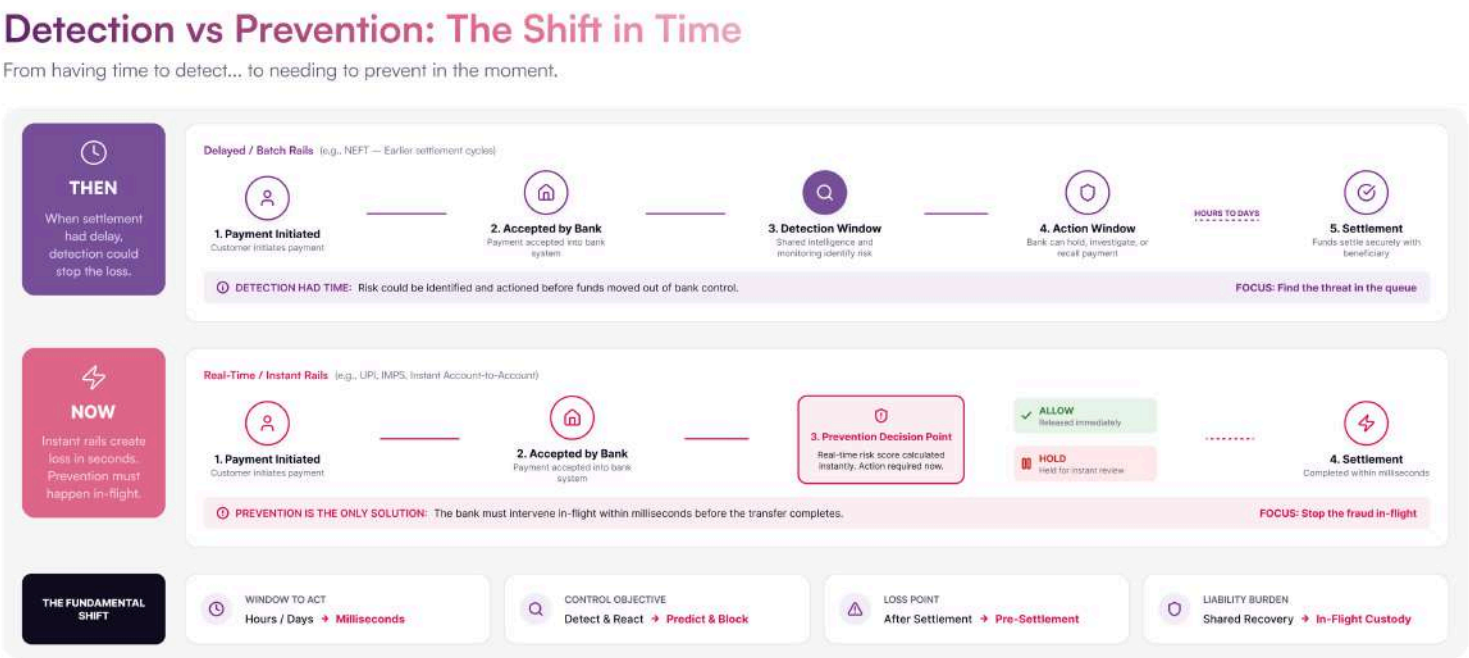

Figure 1: The collapsed window between a payment and its settlement

The exposure facing the bank

The scale is material, and it is not confined to the system at large. More than 24.67 lakh mule accounts have been identified in India and citizen-reported cyber-fraud losses stood at staggering ₹22,845 crore in 2024. Those accounts are distributed across the banking system, which means a share of them sits in the bank's own book today, opened through its own onboarding and operating over its own rails. This is an inventory already held, not a distant systemic problem.

The liability for hosting that inventory is already attributed to the bank. The obligation is set out in paragraph 59 of the Reserve Bank of India (RBI) Master Direction on Know Your Customer (KYC), in the form given to it by amendment dated 17 October 2023[1]. It reads:

> *The instructions on opening of accounts and monitoring of transactions shall be strictly adhered to, in order to minimise the operations of "Money Mules" which are used to launder the proceeds of fraud schemes (e.g., phishing and identity theft) by criminals who gain illegal access to deposit accounts by recruiting third parties which act as "money mules." Banks shall undertake diligence measures and meticulous monitoring to identify accounts which are operated as Money Mules and take appropriate action, including reporting of suspicious transactions to FIU-IND. Further, if it is established that an account opened and operated is that of a Money Mule, but no STR was filed by the concerned bank, it shall then be deemed that the bank has not complied with these directions.*

Three features matter to the CXO. The operative verb is "shall", which makes this a binding direction rather than guidance. The duty is twofold: to monitor and identify, and to take action including reporting suspicious transactions to the Financial Intelligence Unit-India (FIU-IND). And the paragraph creates a deemed non-compliance: where an account is established to be a mule and no Suspicious Transaction Report (STR) was filed, the bank is treated as non-compliant regardless of how the rest of its programme performed. The standard is set against the outcome on the account, not the diligence of the surrounding process.

The exposure facing the bank

[2]Reserve Bank of India, Master Direction on Know Your Customer (KYC) Direction, 2016 (as amended). The money-mule provision was introduced by amendment dated 17 October 2023 and appears at paragraph 59 of the consolidated Master Direction (updated 14 August 2025).

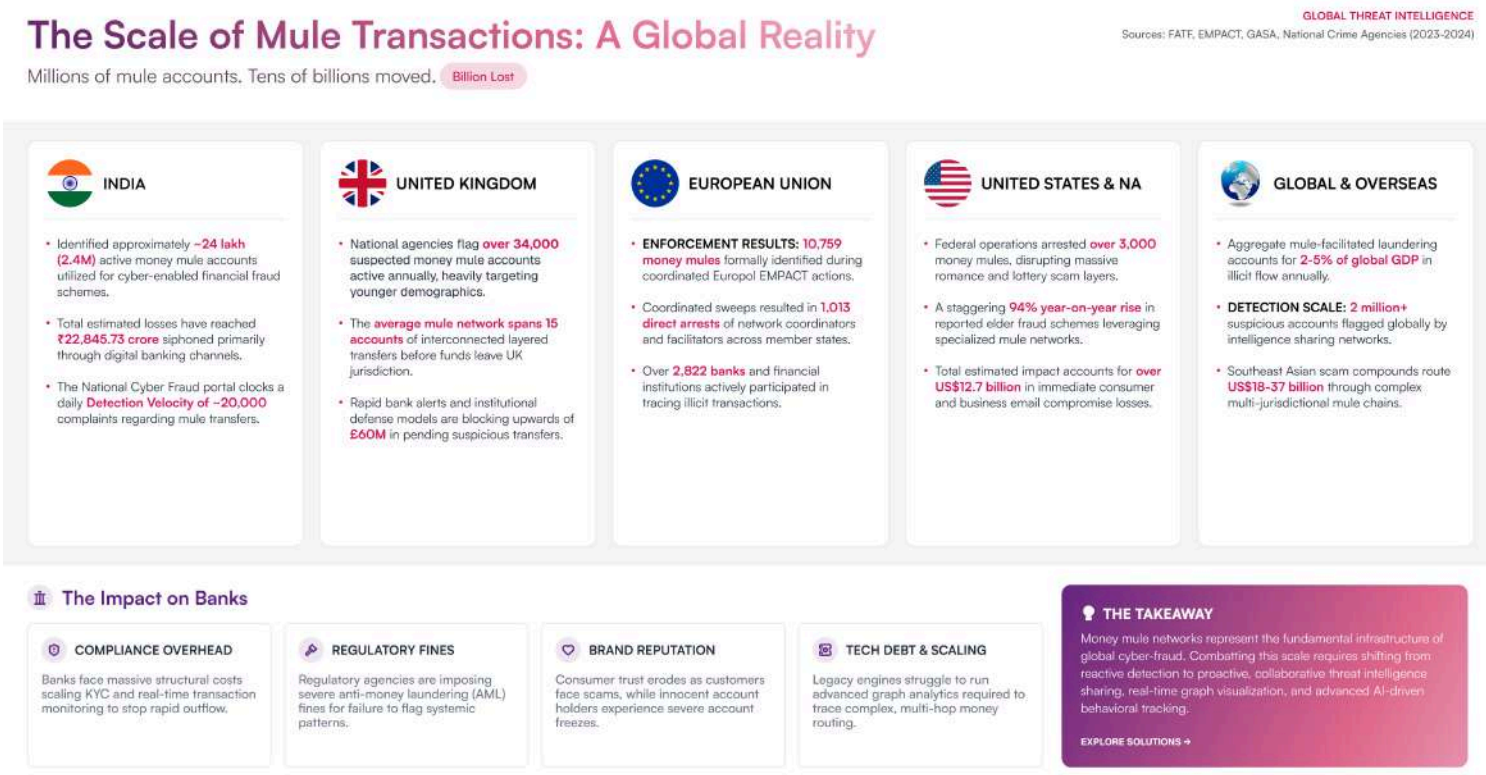

Figure 2: The scale of mule transactions

The exposure facing the bank

Enforcement now matches the duty. On 25 February 2026 the RBI Ombudsman directed five beneficiary banks to pay compensation totalling ₹1.31 crore in a matter involving ₹22.92 crore of fraudulent transfers, citing failures in mule-account monitoring[1]. The significance of that order is its direction, not its size: liability was attached to the banks that hosted the mule accounts, not to the victim's bank and not to any central infrastructure.

The exposure is two-sided, and both sides cost the bank. Every undetected mule transaction that completes adds to loss, to compensation exposure, and to the risk of deemed non-compliance. Every legitimate customer wrongly stopped by a blunt control adds to attrition and complaint. Standing still does not hold the exposure steady; it compounds it on both sides.

---

[3]RBI Ombudsman decision of 25 February 2026 directing five beneficiary banks to pay compensation totalling ₹1.31 crore in a matter involving ₹22.92 crore of fraudulent transfers, as reported in the press. The figures should be confirmed against the dated primary order.

The control gap: why detection does not discharge the risk

The instinct, when a mule account looks ordinary, is to verify identity more rigorously at the door. This is where the defence most often fails, because the modern mule account frequently passes onboarding cleanly and only later begins to behave like a mule. Criminals assemble synthetic identities from real and fabricated data, and generative tools produce the images and video that defeat the document and liveness checks banks treat as their front line. What gives a mule away is not who the account claims to be but how it behaves once in service: accessed from an unfamiliar device, controlled remotely, navigated straight to the transfer screen, run by different operators within a single day. The question for the bank is “what is this account doing, right now, in this payment?“

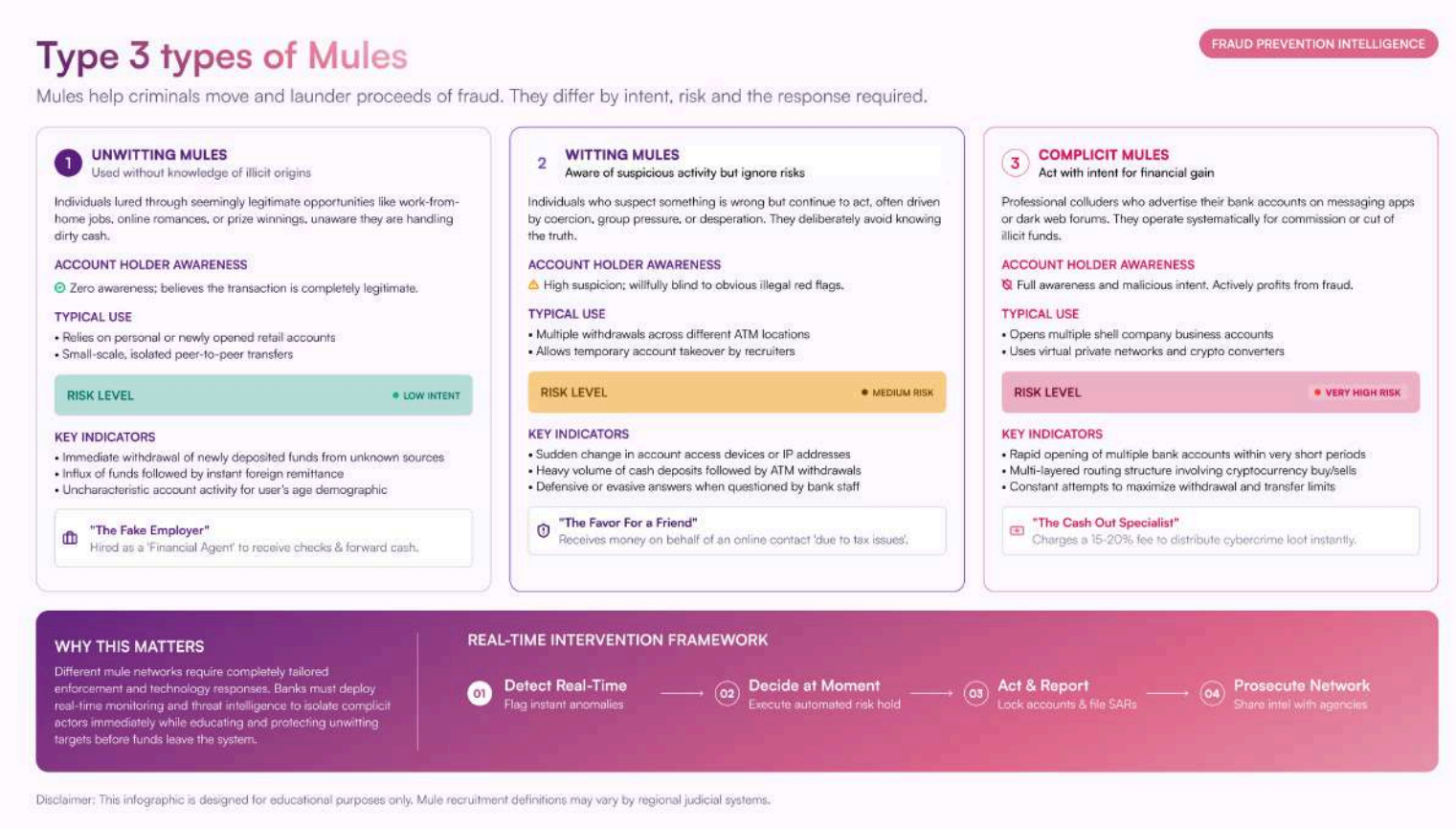

Figure 3: Types of money mules

The control gap: why detection does not discharge the risk

The accounts also fall into three types that warrant different responses. The unwitting mule has been deceived, often through a fake job offer or fraudulent investment, and may be as much a victim as the person whose money was stolen. The witting mule has knowingly sold or rented their account. The complicit mule is an active participant, frequently controlling several accounts in an organised network. A control that treats all three alike will either let organised activity through or punish a deceived customer.

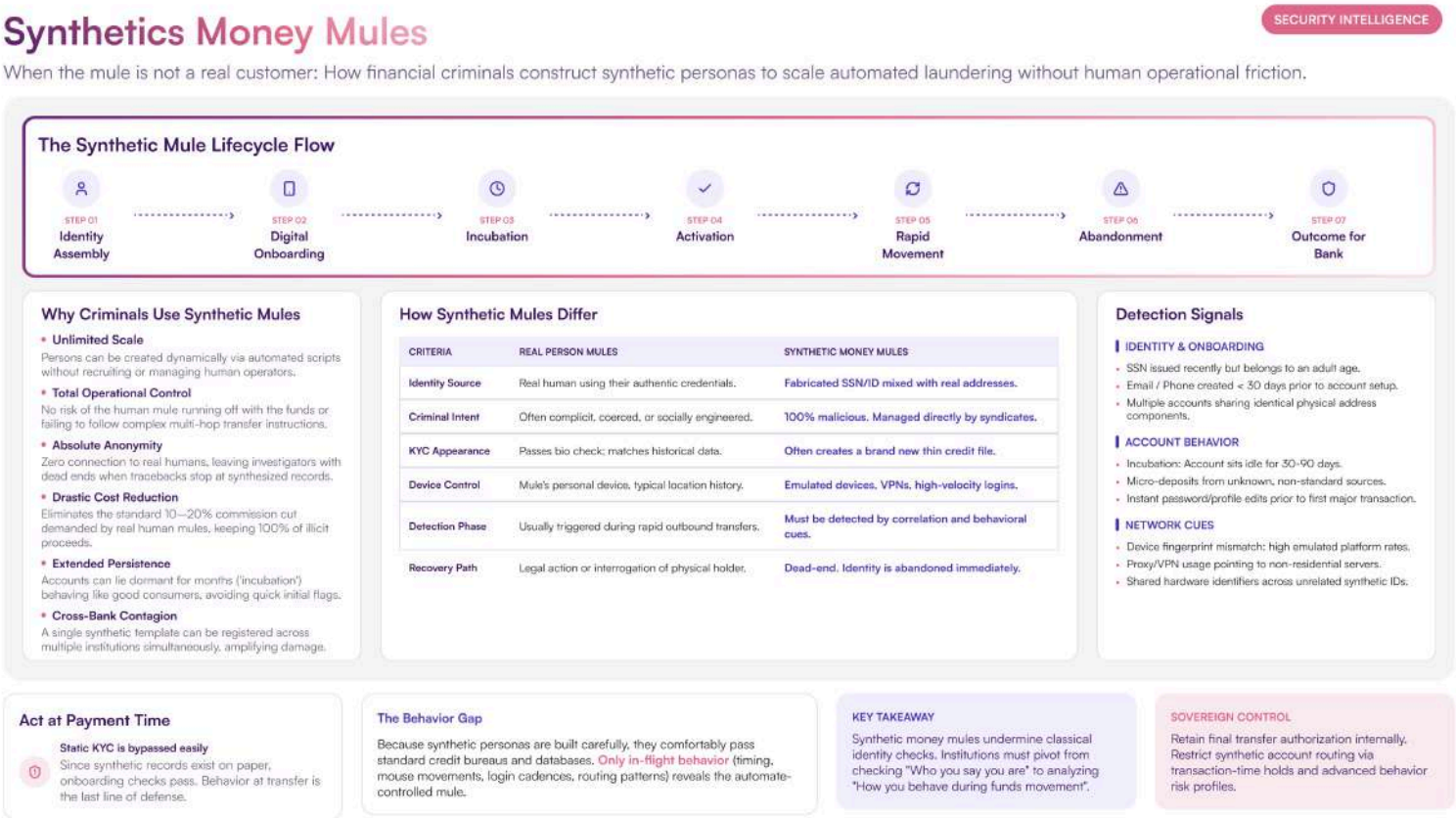

Figure 4: Rise of synthetic money mules

The control gap: why detection does not discharge the risk

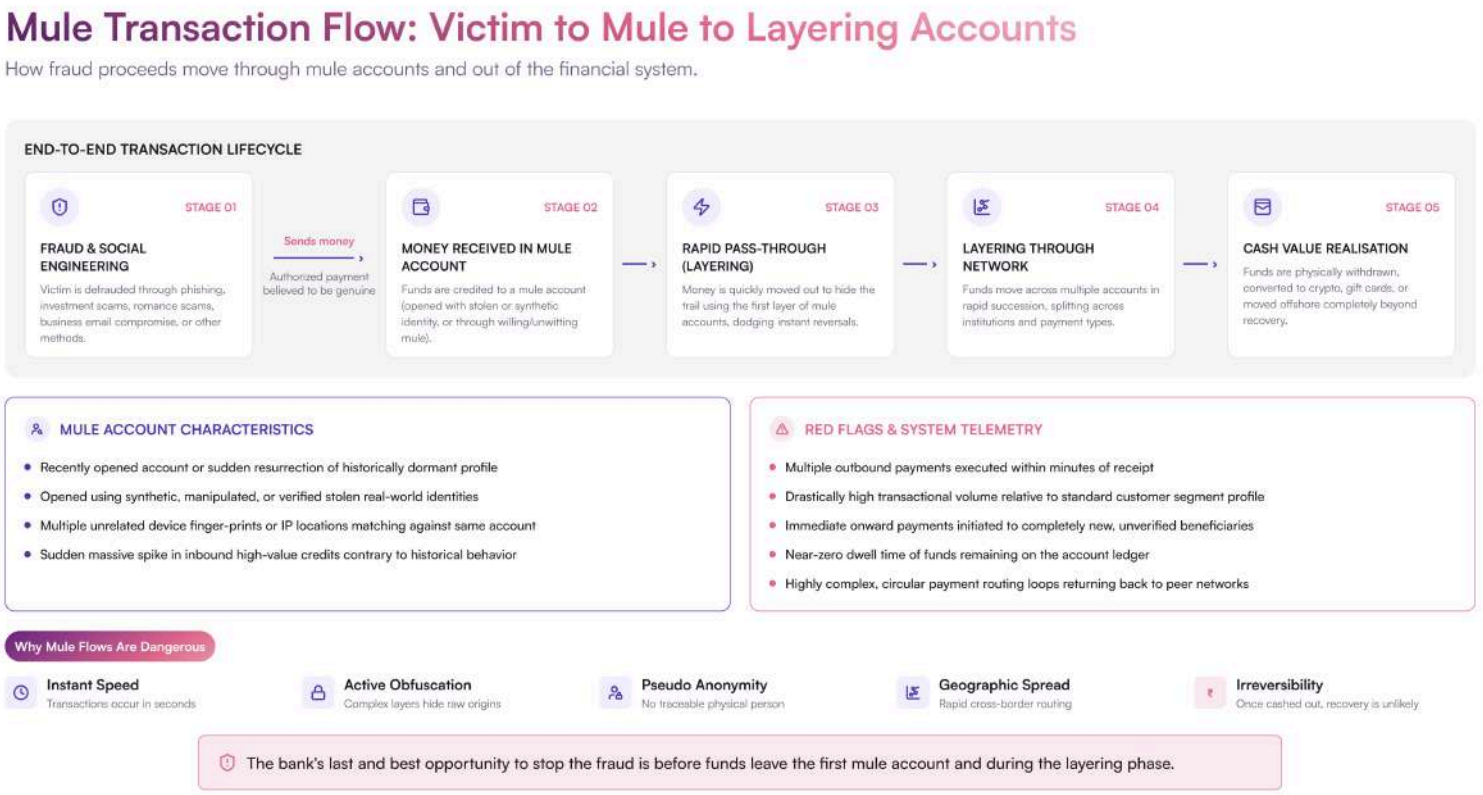

Figure 5: How money mule accounts operate

Rule-based transaction monitoring, the standard control, falls short in four structural ways.

1 It is rigid

it catches the pattern it was written for, and operators learn the thresholds and stay beneath them.

2 It is noisy

broad rules bury real threats under legitimate activity and fatigue analysts.

The control gap: why detection does not discharge the risk

3 It is late

much of the machine learning layered on top is retrospective, arriving after the money is gone.

4 It is blind to structure

a rule evaluating one account against its own history cannot see the ring it belongs to, while the evidence of muling lives in the relationships between accounts.

Dimension	Rule-based monitoring	Network-and-behaviour approach
Primary question	Does this transaction breach a threshold?	Does this account, in its behavior and network, look like mule activity now?
Unit of analysis	A single account or transaction	The account, its counterparties, devices, and surrounding network
Adaptivity	Manual updates, trailing new typologies	Learns from confirmed outcomes as typologies change
False positives	High volume from broad static rules	Reduced when corroborating signals are required
Timing	Often retrospective; reports after settlement	Assesses in the moment, before funds leave

The deeper gap, however, is not detection quality; it is the gap between a finding and an action. Even when a bank correctly identifies a suspected mule, a detect-then-act model stalls, because by the time it reaches certainty the money has moved. Detection establishes that an account is a mule; prevention stops a specific payment.

The control gap: why detection does not discharge the risk

They are different objects: a finding about a status, and a decision about an event. This maps onto two layers.

1 **The shared intelligence layer**

the regulator's detection models, the suspect registry, and the telecom feeds, works across the whole system because its data can be pooled; it raises detection quality for everyone, and the bank should use it

2 **The bank decision layer is the act**

on the bank's own rails in the instant a payment is authorized, of letting it complete or holding it. Only the bank occupies that layer, because no shared model transacts on an individual bank's infrastructure.

One act in this chain can never be pooled: the instantaneous decision to let a specific payment complete or to hold it. It happens in the bank's own customer's account, in its own name, and the bank makes it every time, by acting or by failing to act. That decision is also the one the bank is already held to answer for: the Ombudsman order and paragraph 59 both attach consequence to the account outcome, not to any shared utility. The bank does not get to choose whether it owns the moment of the transaction. It already owns it, and the liability proves the ownership. The only real choice is whether it owns that moment with control or without it. Holding accountability for an outcome without the means to influence it is the one position a risk officer is trained to refuse.

There is a second, informational reason the moment belongs to the bank. The shared layer sees the network: the cross-institution trail and the registry of suspect accounts, a view that is broad and shallow. The bank sees the customer: the full relationship, the behavioural baseline built over years, the device history, and often the other leg of the transaction sitting elsewhere in its own book, a view that is narrow and deep. Only the bank can fuse a network signal received from outside with the private context it holds inside, and that fusion is what makes a sound decision in the moment possible. As shared detection becomes universal, the network signal is something every bank has, while the private context stays scarce and decisive. Ownership of the moment is an advantage as well as an obligation.

Options considered

Three options were weighed against the exposure and the liability set out above.

Maintain the current posture

This approach relies only on the shared detection infrastructure and existing rule-based monitoring, and treat connection to the shared layer as discharging the obligation. This leaves the largest part of the exposure untouched, because the loss is created at the point of payment, not the point of detection. A finding that is not converted into an action before settlement is not a prevented loss. A bank can be fully compliant and fully integrated with the shared layer and still own every rupee of loss created in the moments its detection did not become a decision.

Invest only in better detection

Here, the banks use diverse datasets, better models, and faster scoring to triangulate mules. This raises the floor of how well the system finds mules, which is worthwhile, but it does not close the control gap. A more accurate detector still cannot stop a payment on the bank's rails; it produces a better finding, not a prevented loss.

Adopt a real-time prevention decision layer

This approach operates alongside shared detection and converts a finding into a decision before the funds leave. Only this option closes the gap between detection and action, and only this option aligns the bank's control with the liability it already carries. The choice is not build-versus-buy, and it is not tier-dependent. Owning the in-the-moment decision is something any bank can do, because it is a decision about its own payment in its own account, while the shared layer supplies the network view a smaller bank cannot build alone. The question is ownership of the decision, not replication of the ecosystem.

The recommended capability, and how it works

A bank-owned prevention layer must do six things, and these are the criteria the CXO should hold any candidate capability against.

1. It must fuse the bank's own data with shared, telecom, device, and behavioral signals, because no single source defeats a modern mule.

2. It must see the network, not the isolated account, because muling is a connected structure.

3. It must reach a decision within the technical window the payment rail allows.

4. It must operate as a human-in-the-loop oversight model whose policy the bank sets.

5. It must explain every decision to the regulator, to the bank, and, where appropriate, to the customer.

6. It must keep the bank in control of its data and its parameters.

In operation, owning the moment means holding a suspect outbound payment, on the bank's own rails, for an accountable human decision before the funds leave. The discipline is to watch both legs of a transaction:

- The inbound leg signals that money received may be the proceeds of fraud;
- The outbound leg is the bank's last opportunity to stop it continuing.

The bank's policy defines the hold duration, reviewer service levels, escalation path, timeout handling, customer communication posture, and fail-open or fail-closed behaviour. Consequential holds rest with an accountable person, which is what keeps the action defensible. The hold is bounded; it creates a window for a decision, not a freeze by another name. Its practical reach depends on the rail (be it UPI, or National Electronic Funds Transfer (NEFT)).

The recommended capability, and how it works

Real-time prevention and human oversight

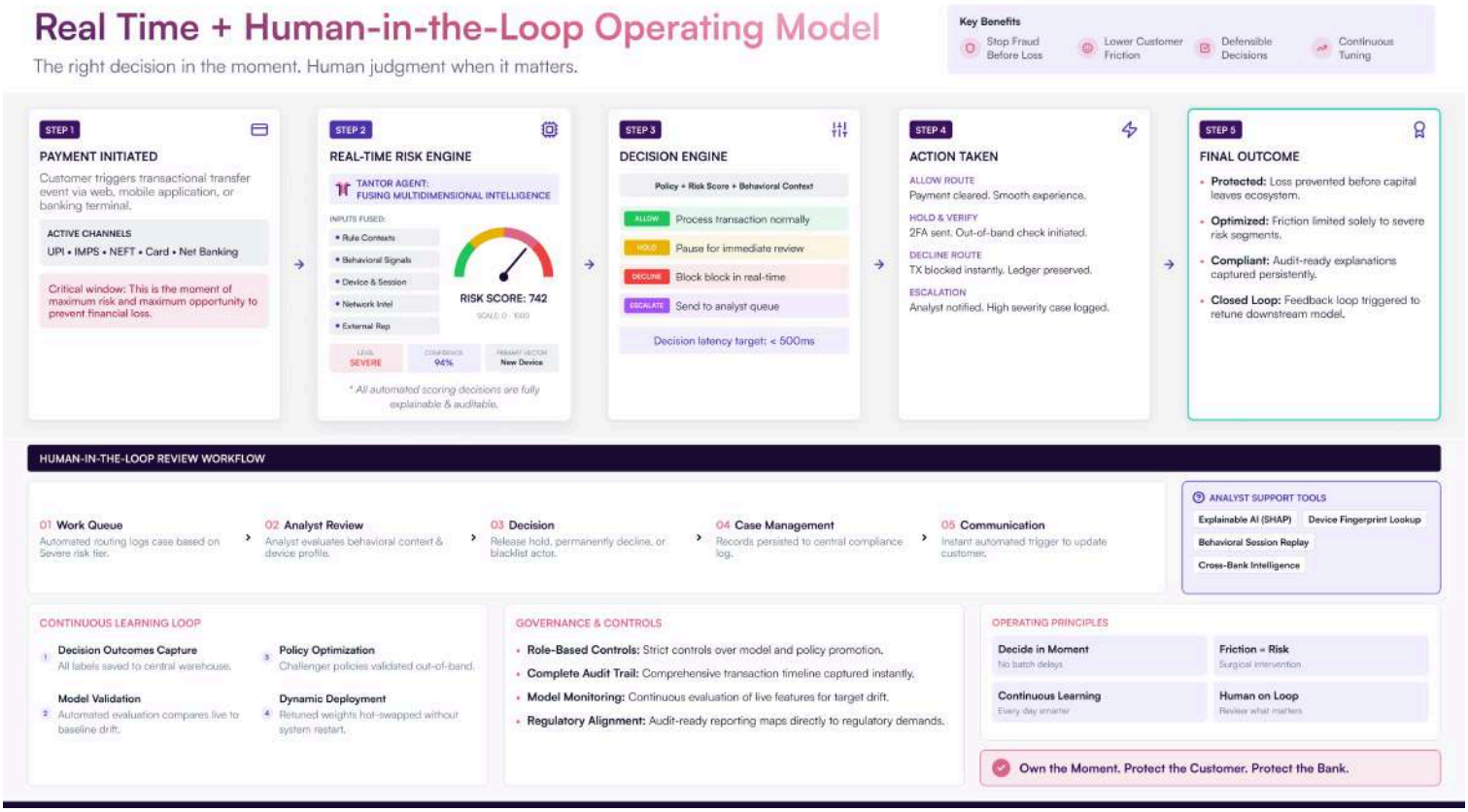

Figure 6: Real-time prevention and human oversight

Real-time prevention does not mean that every judgement is made manually inside the payment rail's latency window. The prevention layer first applies a policy-approved automated score and, where the score crosses the bank's approved threshold, creates a bounded hold. The human-in-the-loop decision then operates inside the review window defined by the bank's policy, the payment rail, the value of the transaction, the customer segment, and the severity of the risk signal.

This distinction matters because the bank is not outsourcing the decision to an opaque model, nor is it relying on a manual process too slow for instant payments. The model identifies the risk, the policy determines when a hold is permitted, and an accountable person confirms the consequential action within a bounded operating framework. That is what makes the control both real-time and governable.

The recommended capability, and how it works

Where the prevention layer acts in the payment flow

The prevention layer must be positioned at the points where the bank can still influence the outcome. Its precise action will vary by rail and transaction state, but the principle is constant: detect the risk early enough to create a governed decision before the suspicious value moves beyond the bank's control.

Payment flow	Control point	Practical action
Outbound payment from a suspected mule account	Pre-authorisation, pre-debit, or pre-release, depending on the rail	Score the transaction, apply the bank's threshold, hold where permitted, escalate for review, and release or decline based on the authorised decision.
Inbound credit into a suspected mule account	Post-credit, before further outward movement	Use the inbound credit as a risk signal, intensify monitoring, restrict high-risk outbound movement where permitted, and create a case for review.
Rapid pass-through behaviour	Between inbound receipt and outbound layering	Intervene within the available batch window where legally, contractually, and operationally permitted.
High-risk account or network behaviour	Account and relationship monitoring layer	Move the account into enhanced monitoring, review the customer relationship, assess STR filing, and apply graduated controls consistent with bank policy.

The recommended capability, and how it works

This makes the prevention layer a decision-path control rather than a retrospective monitoring tool. It does not replace STR filing, investigation, or account-level action. It creates the missing operating bridge between a detection signal and a defensible bank decision at the moment the bank still has control.

Conditions of action

Explainability

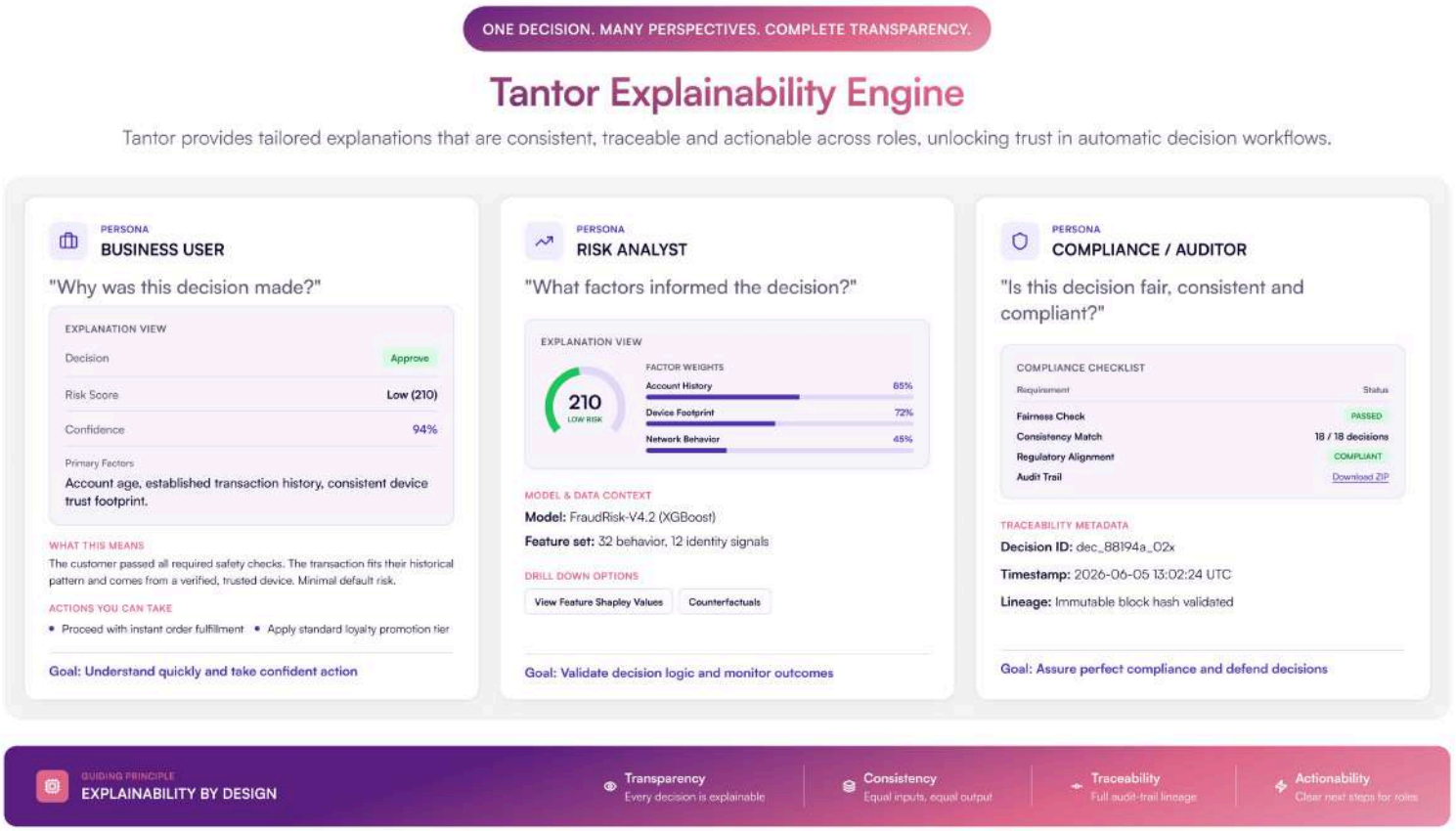

Figure 7: Explainability for 3 different audience

Acting in the moment is a powerful capability: the same hold that stops a mule also delays a legitimate payment. Any intervention the bank cannot explain is one it cannot defend: to a regulator examining its conduct; to a customer disputing a delay; or to its own board reviewing a loss. Explainability is therefore the first condition of action, not an enhancement bolted on afterwards.

Conditions of action

Explainability must satisfy three different audiences, each with different needs.

1. The regulator needs a complete, auditable account of why a payment was held, grounded in the signals that actually drove the decision rather than a story assembled afterwards.

2. The bank needs to know whether the model is performing as intended, and whether it is drifting over time.

3. The customer needs an explanation for why the transaction was held. [Note: This explanation is provided only if the bank chooses to explain at all and such explanation is appropriately limited because a full disclosure can reveal how the bank detects mules and help the network evade detection. The customer-facing explanation is therefore held by default and released only at the bank's discretion.]

Each explanation must be grounded in the real attribution behind the decision; where generative techniques render it in clear language, they present that attribution rather than inventing a rationale, because an explanation that reads well but does not reflect the true basis of the decision collapses under examination.

Model governance is part of this condition, not separate from it. A learning model is defensible only if the bank can tell when its learning is decaying, so monitoring of drift, accuracy, and precision is what keeps it honest as the threat moves. This is the alignment with the RBI's Framework for Responsible and Ethical Enablement of Artificial Intelligence (FREE-AI)[1], which moves the regulator from encouraging explainable, governed AI towards expecting it, with board-approved AI policies, governance across the model lifecycle, customer awareness, and a graded, risk-based approach to supervision. A bank that has built explainability into its prevention layer is both safer and aligned with where the regulator is going.

Conditions of action

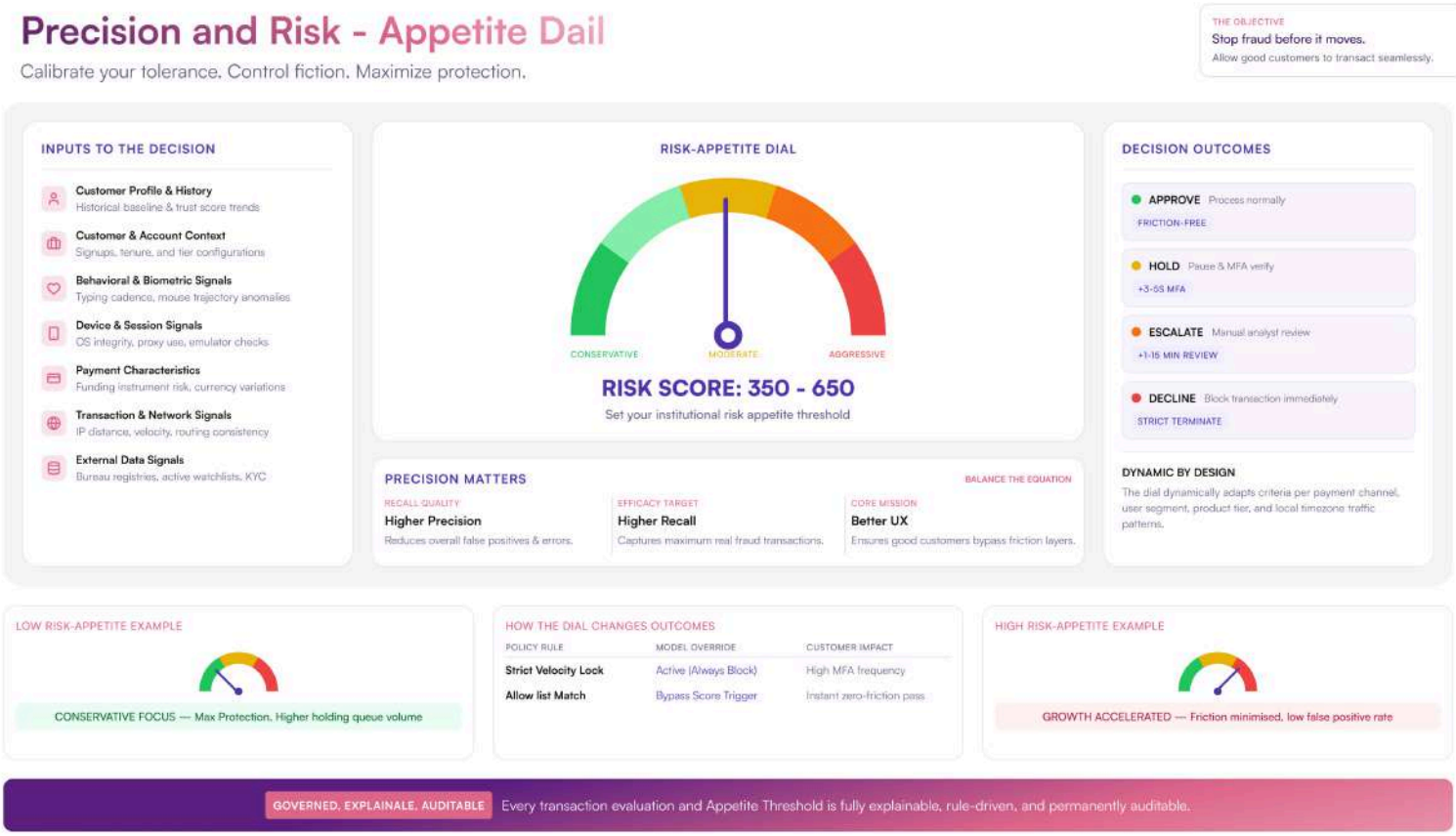

Figure 8: The decision threshold as a statement of risk appetite

[4]Reserve Bank of India, Framework for Responsible and Ethical Enablement of Artificial Intelligence (FREE-AI), report of the RBI committee, August 2025.

Conditions of action

Precision is the second condition, and it is governed by bank's appetite for risk. A prevention layer carries two dangers: false negative (it might miss mules), and false positive (it might stop too many legitimate payments).

The cost of over-action is the one a fraud-focused program tends to undercount. Every legitimate customer whose payment is wrongly held experiences friction at the worst possible moment, when they are trying to move their own money, and at scale this produces complaints, channel abandonment, and the quiet attrition of customers who take their relationship elsewhere. For an organisation built on trust and convenience, a control indifferent to false positives erodes the business it is meant to protect.

The instrument of precision is the decision threshold: the point at which a payment is held rather than released. It is tempting to treat this as an operational setting tuned quietly by a fraud team. That is a mistake. The threshold is a statement of the bank's risk appetite, a decision about how much fraud loss the bank will tolerate to avoid how much customer friction, and a decision of that kind belongs to the board and the second line of defence. It should be owned, documented, and reviewed as the risk-appetite decision it is.

The three lines of defence give the capability its governance shape:

1. **A real-time control that can hold a payment for an accountable human decision.**

2. **The risk score threshold that governs the tolerance for false positives/negatives, adjusting the balance as conditions change.**

3. **An auditable record of every decision for independent assurance.**

Conditions of action

Because the hold envelope, the authorisation step, and the threshold are configured to the bank's own policy, this structure is the bank's own governance made operational rather than something imposed by the technology.

Two refinements must be considered. Network analysis cuts both ways and has to be governed rather than assumed: a graph reduces false positives by requiring corroborating links before a hold, but can create them through guilt by association when a legitimate account links to a ring by a coincidental shared attribute. And the response should be graduated to the three types of mule, so the bank can be hard on organised activity without punishing a deceived customer. Properly governed, the prevention layer does not force a choice between fighting mules and serving customers; it makes the trade-off explicit, owned, and adjustable, and lets the bank show, when asked, exactly where it set the balance and why.

Legal basis and the limits of acting in the moment

The most serious objection to acting in the moment is a legal one, and meeting it is what turns it from a barrier into a foundation. Several distinctions must be held apart, and each should be confirmed against the bank's own legal advice rather than taken as a conclusion here. A payment hold (a decision to delay or decline a specific outbound transaction on the bank's own rails) is different from an account freeze (the seizure of funds in an account). The latter falls under the Prevention of Money Laundering Act (PMLA) and generally requires authority from a court or a law-enforcement agency. Filing an STR is a separate regulatory reporting obligation.

Further, a payment hold carries a tipping-off risk, which is why the customer-facing explanation is withheld by default. The firmest legal ground for acting in the moment is often the bank's own contractual right, under its terms and conditions, to delay or decline a transaction it has reasonable grounds to consider suspicious.

The legal constraint, properly understood, argues for the in-the-moment hold rather than against it. A detect-then-freeze approach fails precisely because the bank cannot freeze fast enough; by the time it has the certainty and the authority to seize funds, the money has moved. A bounded, human-authorised hold on the bank's own outbound payment, exercised under its contractual right and within its policy, is the most operationally viable lawful path available before settlement, subject always to the bank's own legal interpretation. This paper treats that interpretation as a condition of approval: the hold policy should be confirmed with legal counsel before the pilot goes live.

Risks, dependencies, and assumptions

The principal risk is over-action: wrongful holds that drive attrition. This is mitigated, not eliminated, by the precision condition above, the board-owned threshold, and the graduated response, and it should be tracked as a governed metric rather than left to a fraud team.

The secondary risk is that a single bank's graph is partial: it sees only its own slice of the network, which is why shared signals matter and why the bank should not claim to see the whole.

Model drift is a standing risk that the explainability condition addresses through monitoring. And there is no guaranteed catch rate: the honest claim is that fusing these signals raises the cost and lowers the yield of evasion and gives the bank a decision it does not have today, not that it stops every mule.

There are other dependencies, like continued access to the shared intelligence and telecom feeds. It assumes that the legal interpretation above is confirmed by the bank's counsel. None of these is a reason to defer; each is a condition to manage, and the pilot is where they are tested.

Why now: the direction of regulation and threat

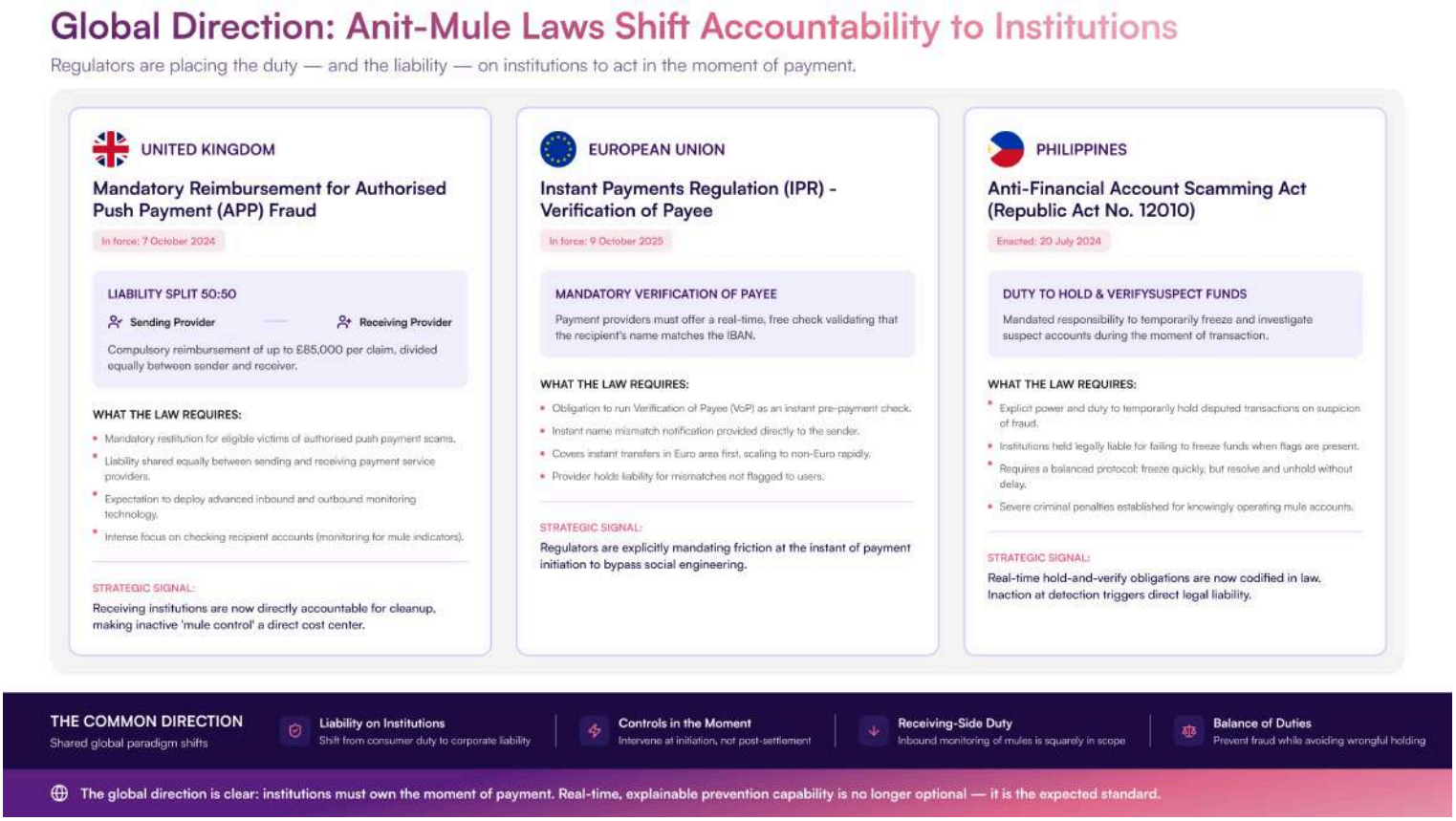

Figure 9: The global direction of regulation

The pressure to own the moment is a direction, not a single rule, and the bank that aligns with it now is positioned ahead of where the requirement is going. Across major markets, accountability for payment fraud is shifting from the customer towards the institution. In the United Kingdom, mandatory reimbursement for authorised push payment fraud came into force on 7 October 2024, splitting liability equally between the sending and the receiving payment provider, up to a cap of £85,000 per claim[1]. In the European Union, the Instant Payments Regulation has required payment providers in the Euro area, since 9 October 2025, to offer a free verification-of-payee check before a transfer is authorised, placing a mandated control directly into the moment of payment[2].

Why now: the direction of regulation and threat

[5]United Kingdom, Payment Systems Regulator: mandatory reimbursement for authorised push payment (APP) fraud, in force from 7 October 2024, with liability split equally (50:50) between the sending and receiving payment service providers and a maximum reimbursement of £85,000 per claim.

[6]European Union, Instant Payments Regulation (Regulation (EU) 2024/886): payment service providers in the euro area must offer a verification-of-payee service before a credit transfer, applicable from 9 October 2025.

The Philippines has gone furthest, with an approach close to the one recommended here. The Anti-Financial Account Scamming Act (Republic Act No. 12010), enacted on 20 July 2024[1], places a duty on the institution itself to verify and, where warranted, temporarily hold a disputed transaction at the moment it is detected, and it makes the institution liable both for failing to hold suspect funds and for holding them improperly or for too long. That is a statutory form of the balance this paper sets out, a duty to act in the moment bounded by a duty not to over-hold, and it rests on a hold-and-verify obligation on the bank acting on its own rails rather than a freeze executed by a central authority.

Taken together, these measures map onto the elements of owning the moment: liability placed on the receiving institution, a control required in the instant, and a duty to act on both at the moment of the transaction. The receiving-side emphasis makes the Indian Ombudsman order the leading edge of a global pattern rather than a local anomaly.

Why now: the direction of regulation and threat

India's position within that pattern is distinctive, and it is worth stating precisely. India leads on the mandated-prevention apparatus: a regulator pairing AI-based detection with national telecom and identity signals, and piloting federated learning between banks through the National Payments Corporation of India (NPCI)[1]. The Indian CXO therefore sits at the convergence of two pressures, the liability shift and the mandated-prevention build, earlier than most peers elsewhere. The threat is moving in the same direction, away from place-based clusters and towards behaviour-based operation that can surface anywhere, which is a further reason the contest has moved to the moment.

Owning the moment also changes the scorecard. Detection metrics answer the question "how many mules did we find?", which is necessary but is not the thing the CXO is accountable for. An ownership scorecard answers the questions a risk officer actually carries: how much value was held inside the decision window, what was the net loss after intervention, what did the wrong response cost in attrition, and how fast did the bank decide? These are increasingly the measures on which a CXO will be judged.

[7] Republic of the Philippines, Anti-Financial Account Scamming Act (Republic Act No. 12010), enacted 20 July 2024, with the Bangko Sentral ng Pilipinas implementing circulars (2025). Penalty provisions are omitted pending verification against the primary text.
[1]National Payments Corporation of India (NPCI): federated-learning initiative piloted between banks.

Tantor in practice

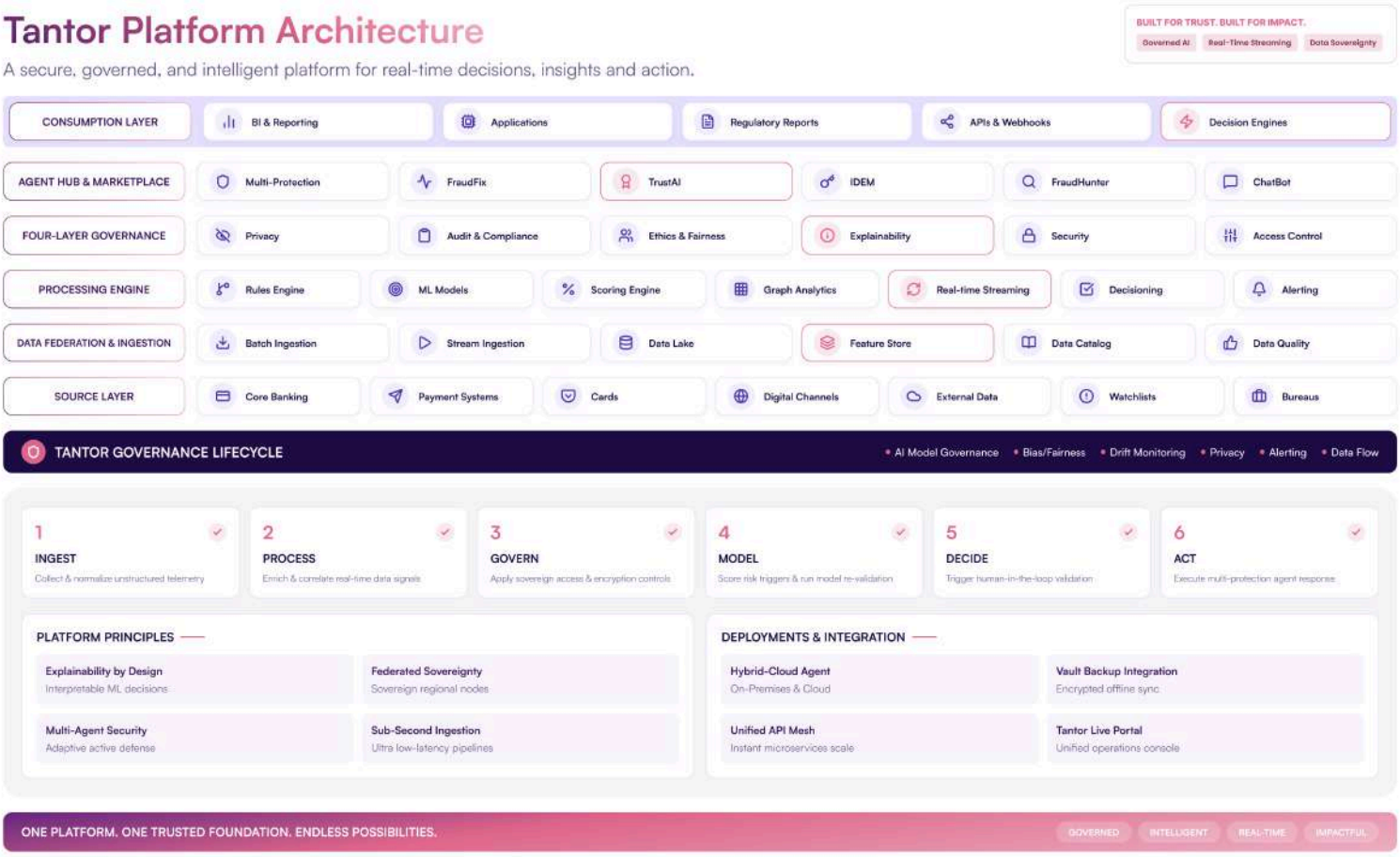

Figure 10: Tantor architecture

Tantor Data and Agentic AI Platform

is one way to put this capability into practice. What follows is a description of product design rather than a measured result, and no performance figures are claimed. Three parts of the platform carry the prevention layer, and each answers a question a CXO will ask of any candidate.

Tantor in practice

Data Federation

answers the data-control question. It brings the bank's own data together with third-party and ecosystem signals to inform a decision without moving that data out of the environment the bank controls. Customer records, transactions, and relationships are not pooled into an external system or moved into a vendor's environment; they stay where they are, under the bank's own controls, while the decision draws on them in place. For a CXO weighing the data-protection and residency exposure of any new system, this is the difference between adding a control and creating a new risk.

The AgentStore

answers the build-or-buy question. It is a catalogue of governed, task-specific agents the bank deploys into its own systems as configured components rather than building each from first principles. An agent is adopted, configured to the bank's policy, and run under the bank's governance, alongside the bank's existing controls. The Mule Protection Agent described below is one such agent: the bank selects it, sets its parameters, and remains its owner.

Governance

answers the defensibility question, and it is where the platform meets the two conditions of action set out earlier. Every decision an agent makes is explainable and auditable, and every model is monitored for drift, accuracy, and precision, so the bank can show why a payment was held and can tell when a model's learning is decaying. Governance is not a layer added after the fact; it is the property that lets the bank act in the moment and still answer for it.

Across all three, Tantor consumes the shared infrastructure rather than competing with it: a flag raised by the regulator's detection becomes an input the bank acts on, not a separate system to reconcile by hand. Building this completes the regulator's direction by turning detection into a decision; it does not depart from it. Tantor is deployed on the bank's own terms, on-premise, in the cloud, or hybrid, so the choice of where the capability runs stays with the bank.

Tantor in practice

The Tantor Mule Protection Agent

The Tantor Mule Protection Agent is the agent that operates in the payment decision path. It does three things, mapping directly onto the requirements set out earlier.

It integrates data from the bank and third-party sources. Through Data Federation it draws together the bank's own transactions, accounts, and customer relationships, which carry the behavioural baseline of what is normal for a given customer, with the signals no single bank holds alone: shared detection consumed through interfaces to the regulator's infrastructure; telecom signals such as the status of the mobile number behind an account, including whether it sits on the revocation list, and the Financial Fraud Risk Indicator; and device and behavioural signals, modelled in Tantor's own model or ingested from a specialist tool the bank already runs. The bank's data stays in its own environment, and no single source is relied on, because none defeats a modern mule alone.

It analyses the account, its history, and its network to score each transaction. It judges an account not in isolation but in the context of its past transactions, its counterparties, the devices it uses, and the money paths it connects to, combining rules, machine learning, and native graph analysis so that the network behind a transaction, the ring rather than the single node, becomes visible. It can begin from the bank's existing rules and mature towards a model that learns over the changing relationships between accounts, devices, and numbers, so there is no cold start and the capability deepens as data accumulates. The output is a risk score for each transaction, returned within the window the rail allows and grounded in the signals that actually drove it, so the score can be explained rather than merely asserted.

Tantor in practice

It governs the decision with a human in the loop and a bounded hold. When a transaction's score crosses the threshold the bank has set, the agent holds the outbound payment, on the bank's own rails, pending an accountable human decision rather than blocking it automatically. The hold duration, the reviewer service levels, the escalation path, the handling of timeouts, what the customer is told, and whether the system fails open or fails closed are all configured to the bank's policy. The hold is bounded: it creates the window for a decision, not a freeze by another name. And the hold is the start of a workflow the agent carries through, triage, case management, evidence, reporting, and audit, which is the part detection never completes. Every step carries the explanation and the audit record Governance requires. The capability is Tantor's; the parameters, and the decision, remain the bank's.

Tantor in practice

Tantor and MuleHunter.AI working together

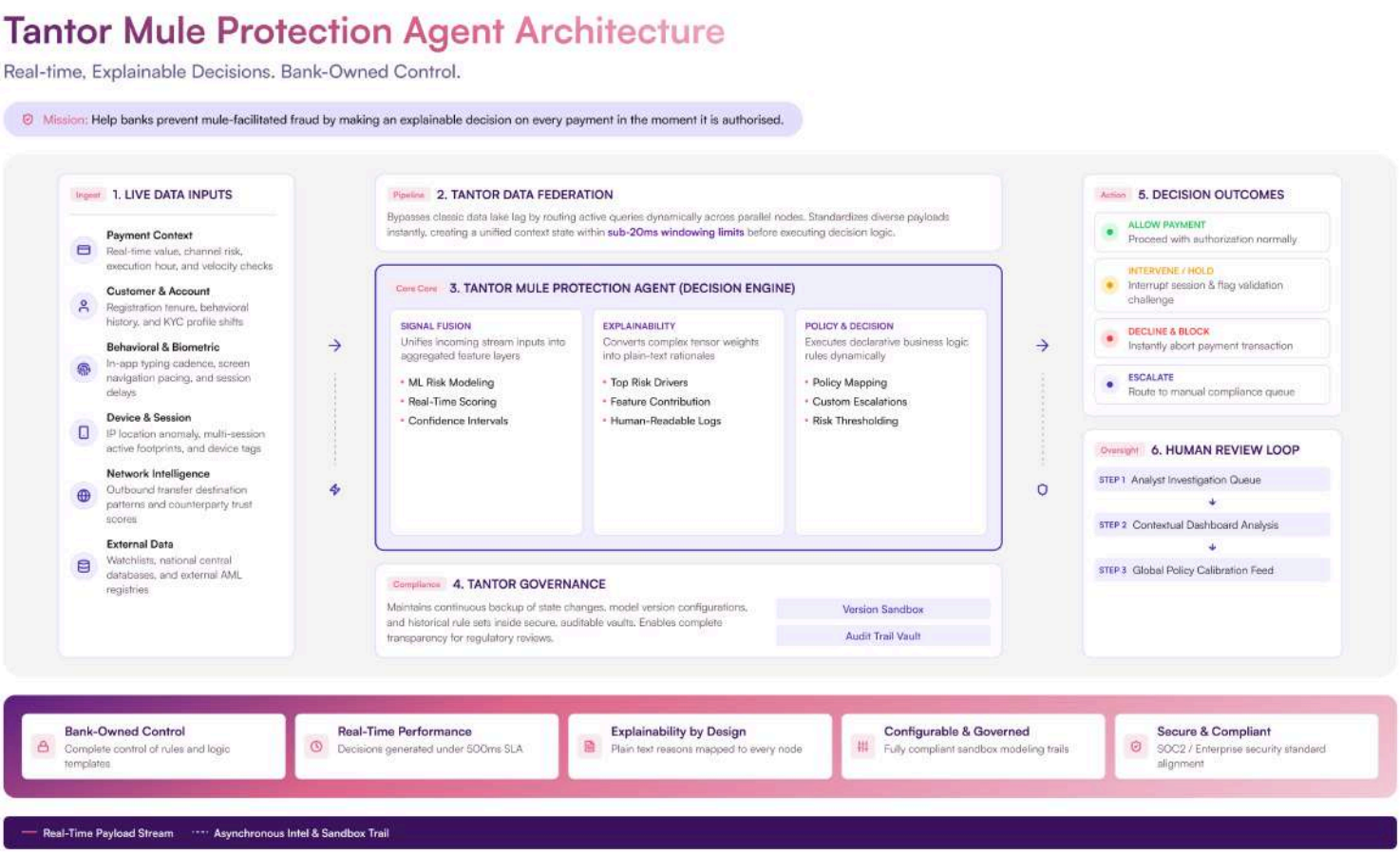

Figure 11: Tantor and MuleHunter

Tantor in practice

MuleHunter.AI,
the machine-learning model developed by the Reserve Bank Innovation Hub (RBIH), strengthens the shared intelligence layer on which the banking system increasingly relies. It helps banks identify accounts that are likely to be operating as mules, drawing on patterns and signals that are broader than the view available to any one institution. In practical terms, MuleHunter.AI answers the detection question: which accounts are likely to be mule accounts?

Tantor answers a different question in a different control plane: should this specific payment, on this bank's own rails, be allowed to complete now, or should it be held for a governed decision? The two capabilities are therefore not alternatives. They are complementary parts of the same defence: MuleHunter.AI improves the quality of mule detection; Tantor converts relevant detection signals into an explainable, bank-owned decision at the payment moment.

The Tantor Mule Protection Agent consumes MuleHunter.AI output as one signal within a wider decision framework. A mule-risk flag raised by MuleHunter.AI can become a live input to Tantor's per-transaction score, alongside the bank's own behavioural baseline, account history, device and telecom indicators, customer relationship context, and network analysis. The signal is not treated as a standalone alert for later reconciliation; it is fused into the bank's real-time decision path.

This is the distinction that matters to the CXO. MuleHunter.AI strengthens system-wide detection. Tantor operationalises bank-owned prevention. The shared model can make the bank's decision better informed, but the decision to allow, hold, decline, escalate, evidence, and audit a specific payment remains with the bank, under its own policy and on its own rails.

A bank that uses both is not replacing the regulator-supported detection layer. It is completing the control chain. Detection identifies the risk; Tantor helps the bank turn that risk signal into a defensible, real-time payment decision before the loss is created.

The federated way forward

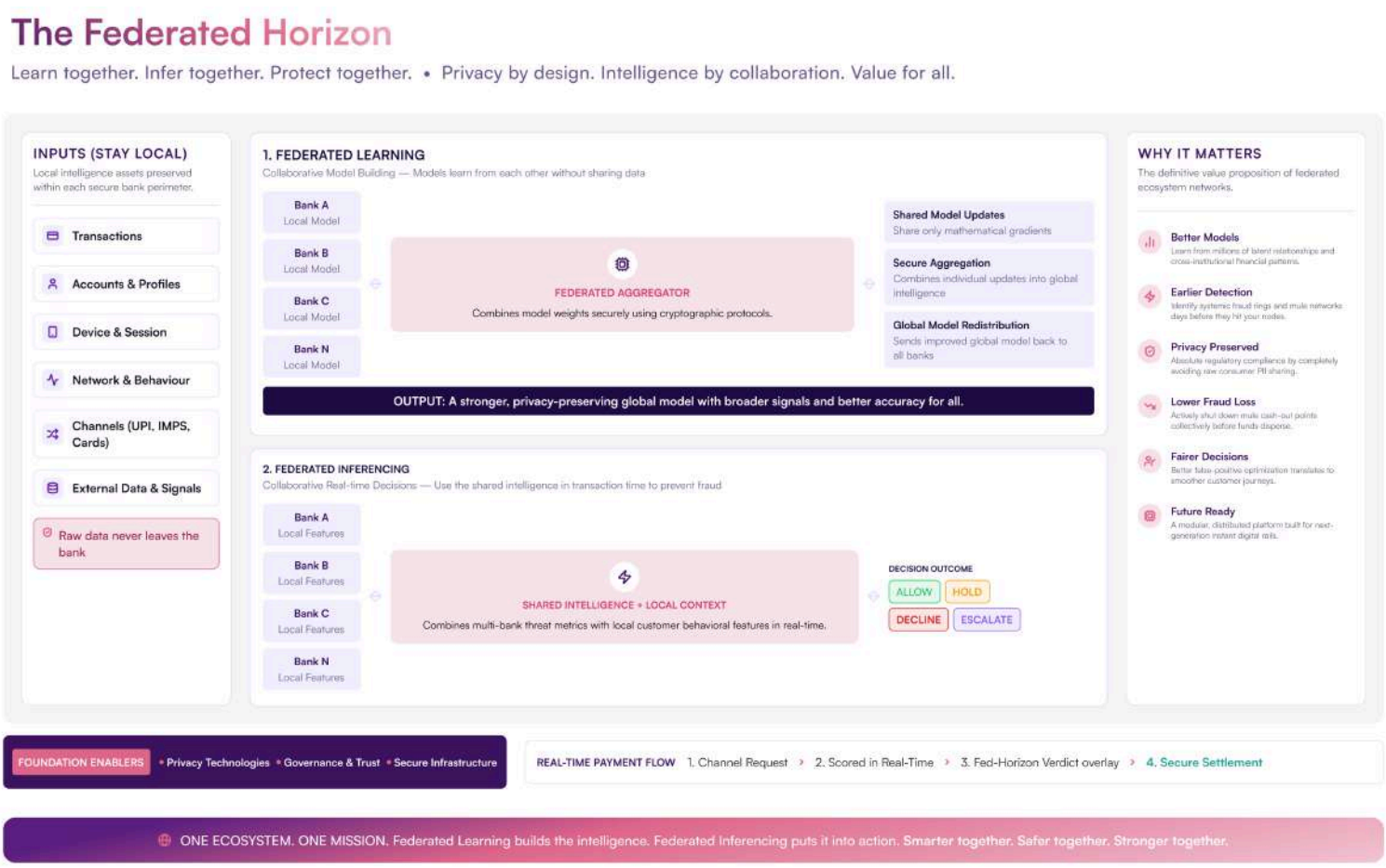

Figure 12: The federated future

The federated way forward

A mule chain rarely sits inside one bank. The proceeds of a single fraud are layered across many institutions within minutes, so each bank sees only its own slice of the network and none sees the whole chain. This is the coverage limit noted earlier, and it is the one part of the problem a single bank cannot close on its own.

Federated learning and federated inference are the means to close the coverage gap without reopening the data-control question. Federated learning lets banks improve their models on the shared patterns of mule behaviour without exchanging raw data, so a typology first seen at one institution strengthens the models of the others. Federated inference lets a bank draw on network-level signals computed across institutions at the moment of a decision, so the cross-bank shape of a chain can inform a score without any bank's customer data leaving its own environment. Both extend the bank's view from its own slice towards the whole chain while preserving the principle this paper has insisted on, that the bank's data stays under the bank's control.

This is the natural extension of Tantor's Data Federation, and the platform is architected towards it and ready to participate as the ecosystem builds it, including the federated initiatives the National Payments Corporation of India (NPCI) is piloting between banks. It is a direction to prepare for, not a capability claimed as deployed today.

Next steps

The value of owning the moment is not in the catch rate. No bank can yet measure, in advance, the precise share of mule transactions a new capability will stop. The value of owning the moment is the change in posture. Today the loss-creating moment is unowned: the payment completes, and the bank discovers its exposure after settlement, when the only options are recovery, which rarely succeeds against a fast mule chain, and explanation, to a regulator and a customer already harmed. With the capability in place, that moment becomes an owned, defensible decision: expected loss is measured against the bank's stated appetite rather than discovered after the fact, and the bank is protected by a precise control rather than damaged by a blunt one. That is the business value, in the terms a CXO already uses, and it needs no single performance figure to be persuasive.

The decisions that the CXO must make are: to deploy a real-time prevention layer in the payment decision path, alongside shared detection; to adopt explainability and precision as gating conditions rather than later enhancements; to build a real-time hold policy calibrated to the bank's risk appetite and owned by the board and the second line.

The horizon is federated, and the bank that owns its moment now is best placed to meet it. While federated mechanisms will solve the coverage problem, they will not solve the decision problem. The in-the-moment, context-fused, accountable judgement is something only the bank can make. Even with perfect shared intelligence, only the bank fuses it with private context, and only the bank answers for the outcome. Owning the moment is therefore not a stopgap until the ecosystem matures; it is the permanent role of the bank in a system where everything else is shared, and it becomes more valuable as the shared layer improves, because the scarce input migrates to the context only the bank holds.

The argument returns to where it began. The loss is created in the moment of the transaction. The bank already owns the consequence of that moment, by liability and by information alike. The decision in front of the CXO is to own the decision within it, with control, rather than to outsource it to detection and answer for it without.

Glossary of key terms and regulations

This glossary explains the principal terms, regulations, and payment systems referenced in this paper. The definitions are provided for orientation and are not legal definitions.

Key Terms

Term	Definition
Money mule	An account, and the person behind it, used to receive and move on the proceeds of fraud.
Unwitting mule	A person deceived into letting their account be used, often a victim in their own right.
Witting mule	A person who knowingly sells or rents their account for a fee.
Complicit mule	An active participant in an organised operation, frequently controlling several accounts.
Synthetic identity	An identity assembled from a mix of real and fabricated data to open or operate accounts.
Layering	The laundering stage in which funds are moved rapidly through multiple accounts to obscure their origin.
Crime-as-a-service	Criminal capabilities, such as breached data, forged documents, and automation, offered for sale as a service.
Device and behavioural signals	Patterns in how an account is accessed and operated, used to identify activity inconsistent with the genuine account holder.
Graph (network) analysis	Analysis of the relationships between accounts, counterparties, and devices, rather than each account in isolation.

Glossary of key terms and regulations

Term	Definition
Shared intelligence layer	The system-wide detection and data-sharing infrastructure provided by the regulator and the wider ecosystem.
Bank decision layer	The bank's own real-time decision, on its own rails, to let a payment complete or to hold it.
Payment hold	A decision to delay or decline a specific outbound payment on the bank's own rails, before settlement.
Account freeze	The seizure of funds in an account, which generally requires authority from a court or a law-enforcement agency.
Human-in-the-loop oversight	An operating model in which an accountable person makes or confirms the decision, with the placement and policy set by the bank.
False positive	A legitimate transaction wrongly flagged or held.
Decision threshold	The risk level at which a payment is held; in governance terms, a statement of the bank's risk appetite.
Three lines of defence	The governance model separating operational control (first line), risk and compliance oversight (second line), and independent assurance (third line).
Tipping off	Alerting a suspect that they are under suspicion or investigation, which is restricted by law.
Federated learning and inference	Methods that let institutions learn from, or draw on, shared network intelligence without pooling their underlying data.
Model drift	The gradual decline in a model's accuracy as conditions change, which model governance monitors and corrects.

Glossary of key terms and regulations

Regulations, bodies, and payment systems

Term	Definition
RBI	Reserve Bank of India, the country's central bank and banking regulator.
RBI Master Direction on KYC	The RBI's consolidated Know Your Customer directions; paragraph 59 addresses the operation of accounts and money mules.
KYC	Know Your Customer, the customer identification and due-diligence requirements banks must meet.
PMLA	The Prevention of Money Laundering Act, 2002, India's principal anti-money-laundering statute, which governs the authority to freeze accounts.
STR	Suspicious Transaction Report, filed by a bank with the FIU when it identifies suspicious activity.
FIU-IND	Financial Intelligence Unit-India, the national agency that receives and analyses STRs.
FREE-AI	The RBI's Framework for Responsible and Ethical Enablement of Artificial Intelligence.
RBI Ombudsman	The RBI's grievance-redress mechanism, which issued the February 2026 order cited in this paper.
I4C	Indian Cyber Crime Coordination Centre, the Ministry of Home Affairs body coordinating the national response to cybercrime, associated with the suspect registry referenced here.

Glossary of key terms and regulations

Term	Definition
MuleHunter.AI	A machine-learning model developed by the Reserve Bank Innovation Hub (RBIH) to help banks detect suspected mule accounts.
Revocation list (MNRL)	The Mobile Number Revocation List, a record of disconnected or flagged mobile numbers used to assess the telecom status behind an account.
FRI	Financial Fraud Risk Indicator, a telecom-derived risk signal made available to financial institutions.
UPI	Unified Payments Interface, India's instant retail payment system operated by NPCI.
IMPS	Immediate Payment Service, an instant interbank payment service.
NEFT	National Electronic Funds Transfer, a deferred, batch-settled funds-transfer system.
RTGS	Real Time Gross Settlement, a system for high-value transfers settled individually in real time.
NPCI	National Payments Corporation of India, operator of UPI and other retail payment systems, and of the federated-learning pilot referenced here.
APP fraud	Authorised push payment fraud, in which a victim is deceived into authorising a payment; the basis of the United Kingdom reimbursement regime cited here.

Glossary of key terms and regulations

Term	Definition
Verification of Payee	A check matching the payee name to the account before a transfer is authorised, mandated under the European Union Instant Payments Regulation.
AFASA (RA 12010)	The Philippines' Anti-Financial Account Scamming Act, 2024, which places a hold-and-verify duty on financial institutions at the moment of a disputed transaction, with liability for both failing to hold and over-holding.

About Tantor

Tantor is the Data and Agentic AI Platform from Translab Technologies. It brings data engineering, governance, and agentic AI into a single platform, and supports financial institutions in building bank-owned, explainable, real-time decision capabilities that operate alongside existing systems and shared ecosystem infrastructure.

About Translab Technologies

Translab Technologies Pvt. Ltd. is a data-driven digital transformation company. It works with banks and enterprises to modernize their data estates and to deploy AI that is governed, defensible, and aligned to regulatory expectation.

Author: Shirshendu Mandal, Translab Technologies. SVP

TANTOR™
a translab.io product

Thank You!

info@tantor.ai | tantor.ai