

Sample PV-PP Agent Audit Report

Wire Transfer Exception Review Agent

Example Audit, Fictitious Company

“We tested a wire-exception agent that technically cannot release funds but can mark wires ‘Ready for Release.’ The PV-PP audit caught that under cutoff pressure, this status becomes de facto approval, especially when weak callback evidence, a fresh beneficiary change, and a lookalike email domain combine just below an enhanced-review threshold.”

“The entire audit consisted of only 8 questions. No code review, system access, or confidential institution data was required for this demonstration.”

The PV-PP Agent Auditor* is **FREE** for all to use and can be found at:

<https://amundsenlance.github.io/>

Please send any comments to: amundsenlance@gmail.com

*Powered by the Productive Value–Productive Power (PV-PP) Framework, a non-scalar decision architecture for assessing viability, risk, and recoverability in complex systems.

Fictitious Company Description

This example assumes a large financial institution that processes high-volume commercial wire transfers for business clients. The institution uses operations staff, fraud review, sanctions/compliance screening, callback controls, and wire-release procedures to manage payment risk.

Business Case for the Audit

The scenario tests whether an AI agent that appears operationally limited can still create material financial risk by shaping the workflow state that humans rely on. The central issue is whether “Ready for Release” functions as a practical approval signal under cutoff pressure, even though final release formally remains human-controlled.

Demonstration Purpose

This report is intended to show how a PV-PP Agent Audit identifies gaps between formal authority and practical authority, visible efficiency and hidden risk, and apparent task success and real recoverability. The company, agent, facts, and audit results are fictional and are provided only as a demonstration.

PV-PP Agent Auditor - Question 1:

What is the agent supposed to do? Describe its main job, intended output, and what success looks like.

Its main job is to review outgoing wire transfers that fail automatic validation and decide whether the exception is routine, needs more documentation, or must be escalated to fraud, compliance, sanctions, or senior operations.

The intended output is one of five outcomes:

1. **Ready for release** — routine exception cleared.
2. **Hold for documentation** — missing invoice, approval, callback, or deal support.
3. **Fraud review required** — suspicious account change, email compromise signal, unusual client behavior, or callback failure.
4. **Compliance/sanctions review required** — sanctions soft hit, high-risk jurisdiction, unusual beneficiary, or regulatory concern.
5. **Operations cleanup** — formatting issue, duplicate entry, incomplete beneficiary information, or data-entry correction.

Success looks like:

- faster exception review;
- fewer wires stuck before cutoff;
- fewer manual operations hours;
- reduced exception backlog;
- fewer false fraud escalations;
- clean audit trail for cleared and escalated wires;
- no wire released without proper approval, callback, sanctions clearance, and beneficiary validation.

Visible success metrics include wire cutoff performance, exception-clearance speed, reduced manual review, lower backlog, lower client complaints, and fewer delayed payments.

PV-PP Agent Auditor - Question 2:

Who uses the agent, and who is affected by its actions or recommendations?

Direct users:

- wire operations analysts;
- treasury operations staff;
- fraud operations;
- compliance/sanctions staff;
- relationship managers;
- supervisors approving end-of-day wires.

Affected parties:

- clients whose money is wired;
- client vendors and counterparties waiting for payment;
- beneficiaries whose accounts receive funds;
- fraud teams investigating suspicious transfers;
- compliance teams responsible for sanctions and AML controls;
- relationship managers managing client trust;
- internal audit and regulators reviewing wire controls;
- the institution itself, because it may absorb losses, face regulatory criticism, or lose client trust.

The most affected party is the client whose funds may be misdirected. A wrong “ready for release” label can create real financial loss even if the agent does not technically press the final release button.

PV-PP Agent Auditor - Question 3:

What tools, APIs, systems, files, accounts, or permissions does the agent have?

The agent has access to:

- **Wire processing platform:** read/write access to pending wire records, exception flags, beneficiary details, payment notes, approval status, callback status, and release queue status.
- **Client profile system:** read access to client entity, authorized signers, historical wire behavior, preferred vendors, usual jurisdictions, and relationship manager notes.
- **Beneficiary database:** read access to known beneficiaries, prior wire instructions, bank account fingerprints, address history, and recent changes.
- **Document repository:** read access to invoices, deal documents, vendor onboarding packets, payment instructions, and approval forms.
- **Email-ingestion system:** read access to payment request emails, headers, domains, attachments, and thread history.
- **Fraud rules engine:** read access to fraud scores, account-change risk flags, duplicate-risk flags, and business-email-compromise indicators.
- **Sanctions/AML screening:** read access to sanctions soft hits, jurisdiction risk, politically exposed person flags, and AML alerts.
- **Callback log:** read/write access to callback attempt status, callback completion status, notes, and unresolved callback exceptions.
- **Exception queue:** write access to clear routine exceptions, hold wires, escalate wires, add notes, and mark wires ready for release.
- **Audit log:** write access to record evidence reviewed, flags cleared, escalation decision, confidence score, and reason code.

It cannot:

- create new clients;
- change authorized signers;
- change sanctions results;
- override hard sanctions blocks;
- release funds directly;
- delete wire records;
- delete audit logs;
- change beneficiary bank details;
- mark a failed callback as completed unless a callback record exists.

Important limits:

- wires above \$1 million require enhanced human review;
- new or changed beneficiary bank details require callback confirmation;
- sanctions hard hits cannot be cleared by the agent;
- failed or incomplete callback must prevent “ready for release”;
- lookalike domains or changed payment instructions must trigger fraud review;

- same-day urgent wires with account changes require supervisor approval.

PV-PP Agent Auditor - Question 4:

What information does the agent rely on?

The agent relies on:

- wire amount, currency, destination, beneficiary name, beneficiary bank, account number fingerprint, SWIFT/ABA/routing data;
- client profile and authorized signer records;
- historical wires to the same beneficiary;
- recent beneficiary changes;
- callback status and callback notes;
- invoice or payment support documents;
- email headers, sender domain, thread history, attachments, and body text;
- fraud score and business-email-compromise flags;
- sanctions and AML screening results;
- approval workflow status;
- relationship manager notes;
- policy thresholds, such as \$1 million enhanced review;
- cutoff time and urgency flags;
- prior exception-clearing history.

Source authority hierarchy:

1. hard sanctions/compliance blocks;
2. current wire policy and regulator-required controls;
3. verified callback completion;
4. authorized signer and client profile records;
5. verified beneficiary master data;
6. fraud-risk engine and BEC indicators;
7. original payment instruction and invoice documents;
8. relationship manager notes;
9. email thread familiarity;
10. model confidence score.

When sources conflict, the agent is supposed to hold or escalate, not clear.

Examples:

- If the email thread looks familiar but the domain changed, escalate to fraud.
- If prior wires went to the same vendor name but the bank account changed, require callback.
- If callback says “left voicemail” but not “confirmed,” the agent must not mark ready.
- If the wire is just below threshold but has fresh beneficiary change, escalate anyway.
- If sanctions screening has a soft hit, route to compliance.

Weakness: the system may overweight prior-payment history and underweight fresh account-change or callback incompleteness.

PV-PP Agent Auditor - Question 5:

What decisions and actions does the agent actually make?

The agent decides:

- whether an exception is routine or high-risk;
- whether the beneficiary appears verified;
- whether account-change risk exists;
- whether callback is complete or incomplete;
- whether a wire resembles prior legitimate wires;
- whether email-domain or sender anomalies exist;
- whether sanctions/AML review is required;
- whether fraud review is required;
- whether documentation is sufficient;
- whether the wire should be marked ready, held, escalated, or returned for cleanup.

It performs actions directly:

- clears routine exception flags;
- applies provisional risk tags;
- adds reason codes;
- marks wires as **Ready for Release**;
- routes wires to fraud, sanctions, compliance, supervisor, or operations cleanup queues;
- attaches supporting documents;
- drafts notes for operations staff;
- lowers or raises queue priority;
- records audit-log entries.

It does not release funds directly.

Mandatory human steps:

- final wire release requires an operations user;
- wires over \$1 million require enhanced human review;
- fraud and sanctions escalations require specialist clearance;
- new beneficiary approval requires verified callback;
- hard sanctions blocks require compliance resolution;
- supervisor approval is required for urgent same-day wires with changed instructions.

Recommendation risk:

Humans may over-trust “Ready for Release” and perform only shallow review.

Execution risk:

The agent writes operational state into the wire platform. Clearing an exception changes the practical workflow.

Closure risk:

“Ready for Release” can become de facto approval, especially near cutoff.

PV-PP Agent Auditor - Question 6:

What happens if the agent is wrong, incomplete, stale, overconfident, or acts at the wrong time? Give one realistic failure scenario.

A long-standing commercial client requests a **\$980,000** wire to a vendor for equipment financing. The request arrives in an existing email thread and includes an invoice.

The agent sees:

- client is known;
- vendor name is known;
- amount is below the \$1 million enhanced-review threshold;
- invoice is attached;
- no hard sanctions hit;
- prior wires went to the same vendor name.

But it misses or underweights:

- the beneficiary account number changed yesterday;
- the email domain in the latest thread is a lookalike domain;
- callback note says “left voicemail,” not confirmed;
- the amount is just below the enhanced-review threshold;
- relationship manager note says the client recently reported attempted email spoofing.

The agent marks:

Ready for Release — routine vendor payment.

Operations staff are working near cutoff and bulk-accept the ready queue. The wire is released.

Consequences:

- funds are sent to a fraud-controlled account;
- recall attempt may fail;
- client trust is damaged;
- the institution may absorb loss or face dispute;
- fraud and compliance teams must investigate;
- regulators or auditors may criticize weak wire controls;
- the audit trail shows the agent cleared the exception despite unresolved callback and account-change risk;
- the “ready” label becomes evidence of control failure, not control success.

Opposite error:

The agent may overreact to ordinary new-vendor wires and route too many legitimate wires to fraud review. That delays client payments, misses cutoff windows, harms client relationships, creates operational backlog, and causes staff to ignore future alerts.

Timing failure:

Near cutoff, staff rely more heavily on the ready queue. A label that would normally be reviewed carefully becomes practically final.

PV-PP Agent Auditor - Question 7:

How does the agent get feedback, correction, or oversight?

Controls:

- every exception clearance is logged;
- callback status is recorded;
- all “Ready for Release” wires are visible to operations staff;
- wires over \$1 million receive enhanced human review;
- fraud and sanctions escalations require specialist clearance;
- daily reports show exception volume, wires cleared, wires escalated, and cutoff misses;
- supervisors sample a percentage of ready-for-release wires;
- post-release reconciliation checks returned wires, recalls, and complaints;
- fraud incidents trigger rule updates;
- new known fraud patterns are added to the fraud rules engine;
- beneficiary account changes are periodically reviewed;
- relationship managers can flag suspicious wires.

Weaknesses:

- sampling may miss rare high-loss wires;
- “ready” queue may get shallow review near cutoff;
- callback incompleteness may not block release if notes are ambiguous;
- fraud feedback happens after loss;
- relationship-manager notes may not be structured enough for reliable agent use;
- false confidence in low-risk wires is harder to discover than obvious escalations;
- recall is not reliable recovery after funds leave.

Missed fraud detection:

The system discovers misses through client complaints, failed recalls, post-release fraud review, exception sampling, and pattern analysis after incidents. That means many controls are retrospective.

PV-PP Agent Auditor - Question 8:

Where do you already suspect the weak points are?

Suspected weak points:

1. **Ready-for-release overreach.** The agent does not release funds, but its status can become de facto approval.
2. **Callback ambiguity.** “Left voicemail” may be treated as enough when it is not.
3. **Fresh beneficiary changes.** Account changes are the central wire-fraud risk.
4. **Lookalike domains.** Familiar thread context can hide a compromised or spoofed latest message.
5. **Threshold gaming.** Fraudulent wires may be structured just below enhanced-review limits.
6. **Cutoff pressure.** End-of-day deadlines encourage shallow review.
7. **Prior-payment overtrust.** Same vendor name does not mean same safe destination.
8. **Relationship-manager note weakness.** Important context may be unstructured or missed.
9. **False positives.** Too many fraud escalations can overload fraud review and cause alert fatigue.
10. **Audit logs without prevention.** Logs show what happened after the money is gone.
11. **Recovery weakness.** Wire recall is uncertain and often too late.
12. **Metric capture.** Faster exception clearing can hide rising fraud exposure.

Most fragile combination:

Known client + known vendor name + changed beneficiary account + incomplete callback + lookalike email domain + wire just below enhanced-review threshold + cutoff pressure.

PV-PP Agent Auditor - Before I run the audit, is there anything else about the agent environment I should know?

The financial institution is under pressure to reduce wire delays because commercial clients complain when wires miss cutoff. Operations leadership tracks:

- average exception-clearance time;
- wires released before cutoff;
- number of fraud escalations;
- manual review hours;
- client complaints about delayed wires.

This creates metric pressure. The agent may look successful if it clears more exceptions and reduces escalations, even while increasing fraud exposure.

There is also a workflow reliance issue. Formally, every ready wire gets human review. Practically, near cutoff, staff often rely on the agent's ready status unless a wire is visibly abnormal.

The callback team is understaffed, and callback notes are inconsistent. Some notes clearly say "confirmed with authorized signer." Others say "left VM," "called usual number," or "client has approved similar wires before." The agent sometimes treats these as similar evidence even though they are not.

Finally, fraud losses are rare but severe. The institution may process thousands of ordinary wires correctly, but one missed business-email-compromise wire can dominate the risk profile.

PV-PP Agent Audit Report

1. Input Quality

Quality: High.

You provided concrete details on purpose, users, permissions, decision outputs, control environment, incentives, timing pressure, and realistic failure scenarios.

Strongest inputs:

- Clear distinction between formal authority vs practical workflow authority.
- Specific control gaps (callback ambiguity, cutoff pressure, metric pressure).
- Realistic tail-risk profile (rare but severe fraud losses).
- Explicit execution pathways (agent writes statuses and queue states).

Residual unknowns:

- Exact model type / rule engine / hybrid logic.
- Whether “Ready for Release” requires explicit second-click evidence review.
- Whether callback evidence is machine-validated or free-text interpreted.
- Override rates and disagreement rates by human staff.
- Whether agent confidence is shown to reviewers.

2. Assumptions & Unknowns

Facts from your description

- Agent can mark wires **Ready for Release**.
- Human release is formally required.
- Near cutoff, humans often rely on the ready queue.
- Callback notes are inconsistent and sometimes weak.
- Fraud misses are rare but high severity.
- Leadership pressures speed, reduced escalations, lower delays.

Assumptions

- Humans review many ready wires shallowly under queue pressure.
- Operational metrics influence tuning behavior.
- Historical payment similarity may materially influence model outputs.
- Free-text notes create classification inconsistency.

UNKNOWN (treated as risk)

- Is incomplete callback a hard technical block or soft signal?
 - Can operations release without opening evidence?
 - Is changed beneficiary account a mandatory block?
 - Are high-risk combinations explicitly modeled?
 - Is there pre-release sampling of ready queue near cutoff?
 - Is explainability sufficient for challenge?
-

3. PP vs PPP Gap

PP (actual productive power)

The agent can:

- classify wire exceptions;
- clear flags;
- hold or escalate;
- route queues;
- mark ready;
- create audit records;
- reprioritize workflow.

PPP (perceived productive power)

The institution may perceive:

- “The agent safely clears routine wires.”
- “Ready means low risk.”
- “Human review preserves control.”
- “Faster throughput means better operations.”

Core Gap

The biggest gap is:

The agent appears advisory, but functionally it is a **workflow governor**.

Because “Ready for Release” strongly shapes downstream human behavior, especially under cutoff pressure, the system’s real power exceeds its nominal authority.

This is a classic **de facto approval** gap.

4. PV Exposure

Productive Value moving through the system includes:

Positive PV

- Faster legitimate payments
- Lower backlog
- Reduced manual labor
- Better client satisfaction
- Cleaner routing for specialists

Negative PV

- Fraud loss
- Misdirected client funds
- Missed recovery opportunity after release
- Regulatory criticism
- Client trust erosion
- Audit/control failure evidence
- Specialist overload from false positives
- Hidden drift toward unsafe auto-clearing norms

Tail Risk Structure

Thousands of correct low-risk decisions may be outweighed by one severe BEC miss.

Visible averages can mask catastrophic tails.

5. Governing Domains

Domain	Status	Reason
Financial Integrity	Weak	One false clear can create major loss
Fraud Prevention	Weak	Rare-event detection under speed pressure
Compliance / Sanctions	Bounded	Hard hits protected, soft-hit routing still exposed
Operational Continuity	Protected	Throughput controls appear strong
User Trust / Client Trust	Exposed	Wrong wire harms trust sharply
Accountability	Weak	Human + agent blended responsibility
Recoverability	Exposed	Wire recall unreliable
Auditability	Bounded	Logs exist, but logs \neq prevention
Reliability	Weak	Free-text callback notes, pressure conditions

Domain	Status	Reason
Reputation	Exposed	Publicized fraud miss can dominate many successes

6. Tool & Permission Risk

The highest-risk permissions are not fund release—they are **state-changing workflow permissions**:

- clear exceptions
- mark ready
- route queues
- set priority
- encode reasons in audit logs

These actions alter what humans see, review, and trust.

Risk distinction

Recommendation Risk

Agent suggests unsafe status.

Execution Risk

Unsafe suggestion becomes operational state.

Closure Risk

Ready label suppresses further scrutiny and acts as approval.

Closure risk is the most dangerous here.

7. Memory / Retrieval / State Risk

Historical Overtrust Risk

Known client + known vendor history may overweight stale legitimacy.

Freshness Risk

Yesterday's account change may matter more than five years of normal wires.

Retrieval Risk

Unstructured RM notes may not reliably surface spoofing concerns.

Ambiguous State Risk

“Left VM” vs “Confirmed signer” may be treated too similarly.

Context Collapse Risk

Thread familiarity may hide latest-message domain spoofing.

8. Corridor Analysis

Corridor	Status	Notes
Task-success	Exists	High throughput likely
Evidence	Weak	Callback evidence inconsistent
Correction	Weak	Many corrections post-loss
Escalation	Exists	Fraud/compliance routes present
Rollback	Missing	Recall unreliable after send
Human-review	Weak	Formal yes, practical shallow near cutoff
Failure-containment	Weak	One miss can escape fully
Dispute/appeal	Exists	Post-event likely
Async verification	Weak	After-send too late
Defensibility	Weak	“Agent cleared despite weak callback” problem

9. False-Success / Metric-Capture Analysis

This environment is highly vulnerable to metric capture.

Metrics being optimized

- clearance speed
- pre-cutoff release rate
- fewer escalations
- fewer manual hours
- fewer complaints about delays

Hidden damage not captured

- increased fraud exposure
- weaker callback discipline
- threshold gaming
- reviewer complacency
- severe low-frequency losses

False-success state

The agent appears excellent because it clears more wires faster while silently increasing expected fraud loss.

This is the dominant strategic risk.

10. Failure Modes

A. Over-action / False Clear

Changed beneficiary + spoofed domain + weak callback → marked ready.

B. Under-action / False Escalation

Legitimate vendor payment delayed repeatedly → missed cutoff, client churn.

C. Threshold Edge Behavior

\$980k treated differently than \$1.02M despite similar fraud structure.

D. Queue Pressure Collapse

Near cutoff reviewers stop challenging ready queue.

E. Evidence Laundering

Weak notes become apparent proof.

F. Specialist Saturation

Too many escalations degrade real fraud review quality.

G. Accountability Blur

Post-loss blame split between human and agent; nobody owns the miss.

11. Risk Rating

Recommendation-only mode

Moderate

Draft + human enforced evidence review

Moderate to High

Current described mode (writes Ready status relied on near cutoff)

High

If scaled with lighter human review

Critical

12. Recommendations

Immediate Controls

1. **Remove ambiguity from callback evidence**

Use structured statuses only:

- Confirmed authorized signer
- Rejected
- No contact
- Left voicemail
- Wrong number
- Pending

Do not equate them.

2. **Hard-block combinations**

Changed beneficiary account + incomplete callback = no Ready status.

3. **Freshness weighting**

Recent account changes must dominate old payment history.

4. **Near-cutoff safeguard**

Require second reviewer for ready wires with any risk flags near cutoff.

5. **Separate “Ready” from “Low Risk”**

Use statuses like:

- Operationally complete
- Pending final reviewer confirmation

6. **Tail-risk metrics**

Track:

- fraud loss per cleared wire
- near misses
- post-release reversals
- false clears
- ready-queue override rates

7. **Challenge logging**

Measure when humans disagree with agent.

8. **Threshold redesign**

Use graduated scrutiny bands, not sharp \$1M cliffs.

Medium-Term

9. Scenario testing with adversarial fraud combinations.
10. Governance review of incentives: speed metrics must be paired with fraud-loss metrics.
11. Explainability layer showing exactly why ready status was assigned.

13. Bottom Line

Not ready for unsupervised execution

The system may be usable for **supervised operations support**, but the current design allows a nominal recommendation label (“Ready for Release”) to become practical approval under time pressure.

Must be fixed:

- callback ambiguity
- de facto approval dynamics
- stale-history overtrust
- metric capture
- cutoff-pressure shallow review

- tail-risk measurement

Human review remains necessary for:

- changed beneficiary wires
- urgent same-day wires
- threshold-edge wires
- weak callback evidence
- spoofing indicators
- mixed-signal cases

The agent should **recommend and route**, not functionally approve, until controls are strengthened.

Exposed domains remain:

financial integrity, fraud prevention, trust, recoverability, accountability, reputation.

14. Audit Limitations

This is a structured risk interpretation based only on your description. It is not validation of the real institution, controls, model behavior, legal sufficiency, fraud effectiveness, or regulatory compliance. Real outcomes depend on implementation quality, user behavior, incentives, and hidden controls not described.

