KYC2020

WHITEPAPER

TRANSACTION SCREENING · DECISION INTEGRITY

RDS-HIVE

A Decision-Integrity Framework for Transaction Screening

ABSTRACT

Despite decades of investment in rules-based systems, canonical data models, and artificial intelligence, transaction screening continues to generate high false-positive rates and rely heavily on manual review. This paper argues that the primary contributor is not inadequate detection technology, but the way decision-making is governed across the screening lifecycle.

AUTHOR	TITLE	ORGANIZATION	VERSION	YEAR
Rajeev Bahri	CEO & Head of Product	KYC2020 LLC	Final Draft	2025

KYC 2020 SEE FAR

1. Executive Summary

Transaction screening is commonly treated as a detection problem. In practice, the dominant operational challenge is not missed detection, but excessive screening scope driven by uncertainty in raw transaction data.

To avoid compliance and audit failure, institutions routinely screen more than necessary. This behavior is intentional, policy-driven, and difficult to unwind through tuning alone.

This paper reframes transaction screening as three distinct decisions:

> *WHAT to screen, HOW to interpret screening outcomes, and WHEN and HOW to validate automated outcome decisions. Conflating these decisions obscures uncertainty, drives defensive over-screening, and limits safe automation.*

01	**WHAT to Screen** Governing screening scope before execution begins
02	**HOW to Interpret** Separating screening outcomes from transaction payment decisions
03	**WHEN / HOW to Validate** Proportionate audit, validation and decision finality

RDS-HIVE governs these decisions explicitly by preserving multiple representations of transaction data, applying independent deterministic and probabilistic decision-makers, and authorizing outcomes through deterministic, policy-aligned controls. Probabilistic models inform decisions but do not authorize them.

By separating detection, interpretation, and validation, the framework constructs confidence before execution, reduces unnecessary escalation, and enables proportionate automation without sacrificing auditability.

> **A Note on the Role of Screening**
> This framework does not diminish the role of screening. Deterministic, policy-driven screening remains a critical and non-negotiable component of transaction compliance — the execution engine that produces the signal set from which decisions are drawn.
>
> What RDS-HIVE addresses is the governance architecture around that engine: how scope is defined before screening runs, how results are interpreted after it completes, and how outcomes are validated and finalized. Screening is essential to the framework; it does not, however, define or envelop it.

The design principles and capabilities of the KYC2020 screening engine — including its deterministic gateway and name screening architecture, matching logic, and proof-of-work auditability — are covered in detail in the **DecisionIQ Whitepaper**.

What Does "HIVE" Mean?

The name HIVE reflects the structural logic of the framework. A hive does not operate through a single point of control. It functions through independent, specialized contributors whose outputs are coordinated by shared rules to produce reliable, collective outcomes.

RDS-HIVE applies this same principle to transaction screening decisions. Multiple independent makers — deterministic and probabilistic — each contribute a representation-specific proposal. No single maker is authoritative. A deterministic checker, aligned to AML policy, evaluates and reconciles those proposals into a single, governed outcome.

HIVE: Heterogeneous Independent Verification Engine — a decision architecture in which independent contributors inform outcomes that only policy-aligned, deterministic controls can authorize.

2. The Core Problem: Uncertainty Before Screening Results in Over-Screening

Raw transaction messages are inherently ambiguous. They contain partial identifiers, free-text narratives, abbreviations, and contextual references that are often insufficient to determine screening relevance with confidence. Faced with this uncertainty, institutions rationally adopt a defensive posture: screening everything that could plausibly matter in order to avoid compliance and audit failure. The result is excessive screening scope and downstream alert volume.

False positives are therefore not primarily the result of inadequate matching logic or poor tuning. They are the consequence of uncertainty being pushed forward rather than resolved early.

Canonical data models restore structure and consistency and are essential for auditability, but canonicalization alone often collapses ambiguity by forcing a single interpretation where multiple plausible interpretations exist.

Financial institutions have been appropriately cautious in adopting probabilistic and AI-based systems within transaction screening. Decisions that cannot be reproduced deterministically or re-executed under audit scrutiny are routinely challenged, even when narrative explanations are provided. This caution is well-founded and reflects legitimate model risk and governance concerns.

At the same time, the most consequential screening errors occur before detection begins, when determining what information within a transaction is relevant to screen. Free-text fields and narratives require inference and contextual recognition that deterministic methods alone continue to struggle with.

RDS-HIVE addresses this tension through a representation-diverse decision model. Probabilistic models contribute contextual inference as one of several independent

representations of the same transaction, while deterministic representations preserve structural certainty. No single representation is treated as authoritative.

This approach enables human-like inference to inform relevance early, without allowing probabilistic systems to define outcomes, execute screening, or authorize decisions.

3. Decision 1 — WHAT to Screen

3.1 Why 'WHAT to Screen' Is the Primary Failure Point

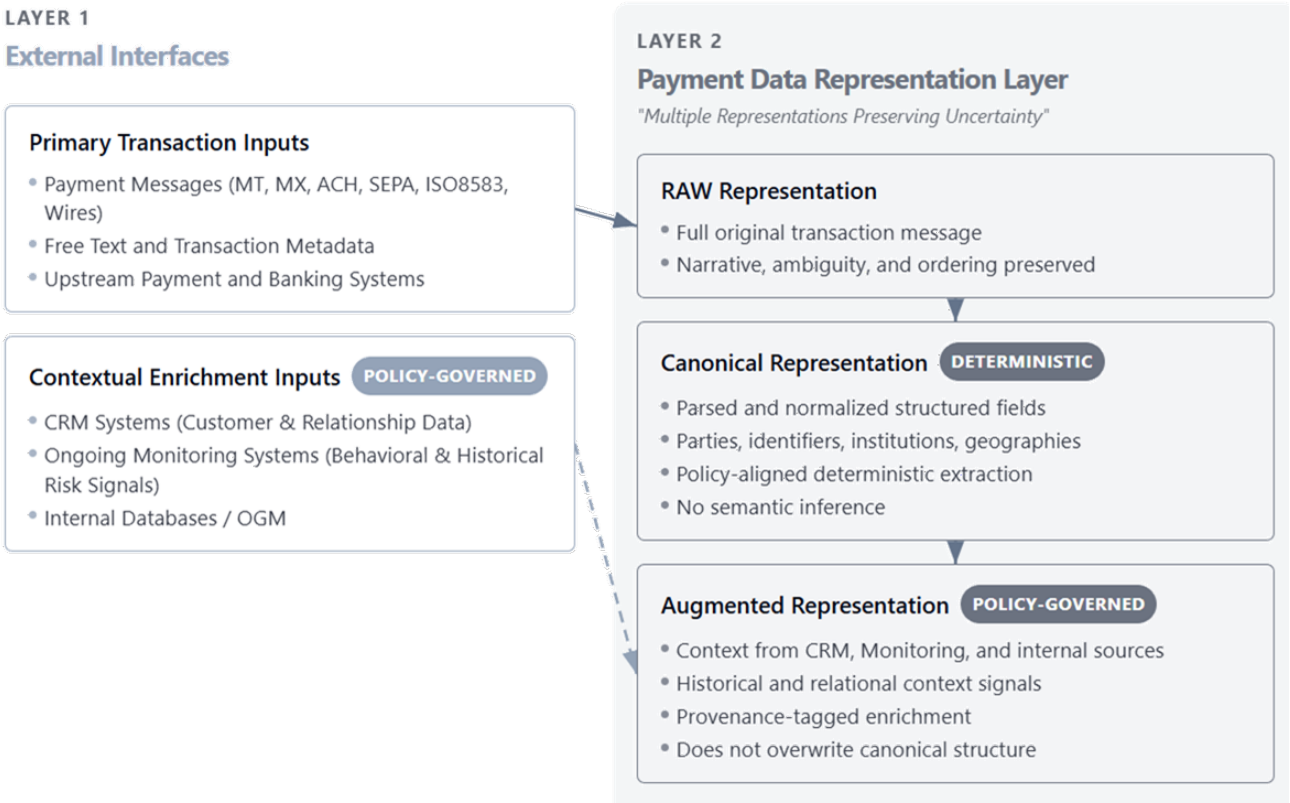
The largest source of false positives in transaction screening does not arise from poor matching logic or inadequate lists. It arises earlier, at the decision of what elements of a transaction should be screened at all.

Most institutions operate under low confidence at this stage, defaulting to defensive over-screening to avoid regulatory failure.

3.2 Multiple Representations, Not a Single 'Correct' View

No single representation of a transaction is sufficient to determine screening scope with confidence. Each transformation applied to raw transaction data preserves some information while discarding or resolving other aspects.

RDS-HIVE therefore derives and preserves three parallel representations from the same RAW transaction message, rather than forcing early convergence:



RAW Representation

The original transaction message retained in full, including narrative fields, ordering, and free-text content. No assumptions are resolved. This representation preserves linguistic ambiguity and contextual nuance and is well suited for probabilistic reasoning, while remaining non-authoritative.

Canonical Representation

A deterministic parsing process extracts and normalizes structured elements such as party roles, identifiers, and policy-relevant fields. Canonicalization restores structure required for policy enforcement, while deliberately avoiding semantic inference.

Augmented Representation

The canonical representation is enriched with additional context intended to reduce uncertainty without collapsing alternatives. Augmentation may include inferred meaning from free text using probabilistic models (LLM), customer history, transaction patterns stored via vector databases, and selectively approved external context via CRM and banking system integrations. It is strictly policy-governed and source-constrained.

3.3 Independent Makers Propose Screening Scope

Determining what should be screened requires multiple independent decision proposals, not reliance on a single model or method. In this paper, maker refers to a decision role, not a specific technology. Both deterministic systems and probabilistic models may act as makers.

- A probabilistic maker operating on the RAW representation may propose a broader screening scope due to unresolved ambiguity.
- A deterministic maker operating on the canonical representation may propose a narrower scope aligned strictly to structured policy triggers.
- A probabilistic maker operating on the augmented representation may refine scope using additional contextual signals.

Each maker independently proposes what should be screened. No maker is privileged by default, and disagreement across makers is treated as a signal of uncertainty, not as an error to be eliminated.

3.4 Deterministic Checker and Policy-Safe Consensus

Maker proposals are evaluated by a deterministic checker aligned to AML policy and regulatory regime. The checker does not reinterpret the transaction or introduce new inference. Its role is to:

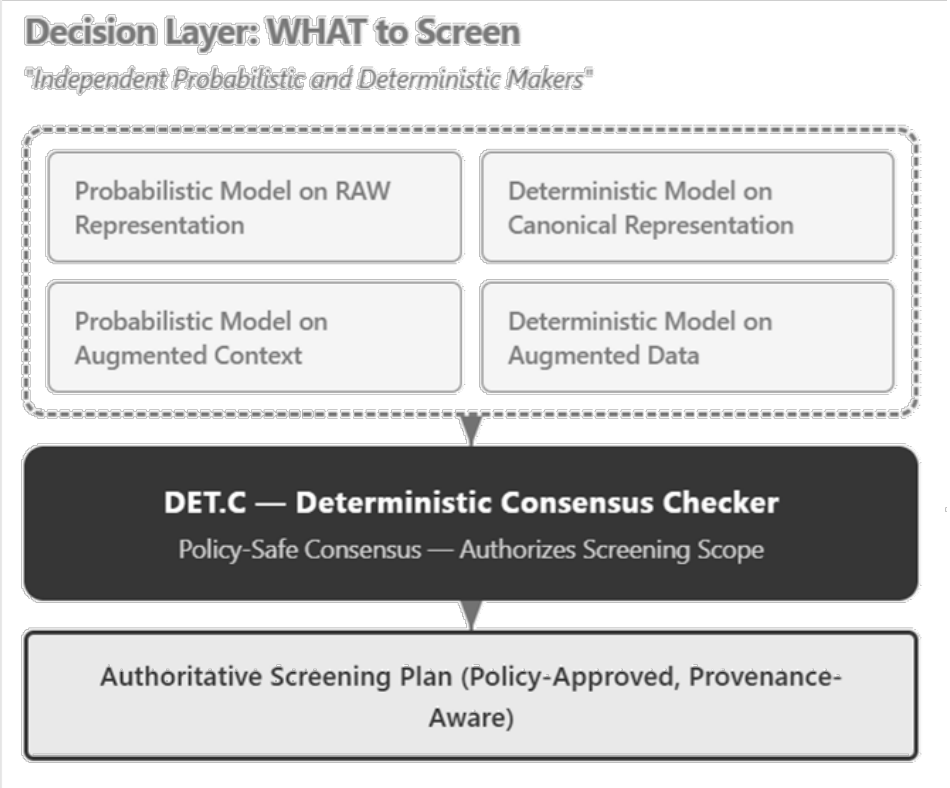
- ensure mandatory policy coverage is met
- prevent under-screening
- resolve disagreement using explicit, documented rules

> *Where proposals converge, confidence is established. Where they diverge, policy — not model dominance — determines the outcome. The result is a single, policy-authorized decision: this is what must be screened.*

4. Deterministic Screening Execution

Once the decision of WHAT to screen has been authorized, screening execution is no longer a decision. It is a deterministic process governed by applicable policy and regulatory regime. At this stage, screening scope is fixed: lists, thresholds, and matching logic are predefined, and execution produces a reproducible outcome set without introducing inference or contextual judgment.

RDS-HIVE does not alter how screening engines operate. It governs when and how they are invoked, ensuring that deterministic screening is applied only after scope has been established with confidence and authorized by policy.



4.1 KYC2020 Screening Service: Design Approach and Trade-offs

KYC2020 provides an integrated screening service designed to operate within this framework and expose policy-driven trade-offs to customers. KYC2020 separates screening into two logically distinct domains:

Gateway Screening	Name Screening
Deterministic evaluation of non-negotiable indicators such as jurisdictions, vessels, bank identifiers, accounts, and other prohibited or high-risk attributes.	Screening of parties using fuzzy matching with name variant intelligence, and jurisdiction-specific logic. Relational attributes inform name matching and not as independent triggers.

Parallel versus Gated Execution

KYC2020 supports configurable orchestration of gateway and name screening, driven by customer policy and risk appetite:

- **Parallel:** Parallel execution, where gateway and name screening run concurrently, is typically optimal in high-throughput environments.
- **Gated:** Gated execution, where gateway screening precedes name screening and may defer name matching upon a policy-mandated HOLD, can reduce latency and cost without bypassing controls.

Sequencing is a configuration choice governed by policy, not a change in control rigor.

Policy-Bounded Scope Adjustment

Gateway outcomes may inform the name screening plan within policy-defined limits (e.g., list selection or matching thresholds). Such adjustments improve efficiency while remaining fully bounded by customer policy and regulatory obligations.

4.2 Screening Service Requirements and Boundaries

For RDS-HIVE to function as intended, the screening service must operate strictly as a deterministic execution component, not as a decision authority. At a minimum, a screening service must satisfy the following requirements:

Requirement	Description
Deterministic Execution	Screening results must be reproducible for a given input and configuration.
Complete Proof of Work & Auditability	Screening outcomes must expose inputs evaluated, lists and rules applied, match logic, thresholds, and dispositions.
No Policy Authority	The screening service must not determine screening scope, interpret results, or authorize outcomes.
Logical Separation of Gateway and Name Screening	Non-negotiable indicators must be screened independently of entity name matching; relational attributes must not trigger alerts in the absence of a high-confidence name match.
Tolerance for False Positives	Screening may favor recall over precision where necessary, provided results are transparent, reproducible, and subject to downstream governance.

> *Screening services that fail to meet these criteria risk embedding implicit policy judgment, generating uncontrolled alert storms, and undermining decision confidence across the framework.*

5. Decision 2 — HOW to Interpret and Evaluate Screening Outcomes

5.1 Interpretation Is a Separate Decision

Screening outcomes are often treated as decisions by default. This practice conflates detection with interpretation and is a major contributor to unnecessary escalation and manual review.

> *Screening execution produces signals, not outcomes. The decision of how to interpret those signals and whether to authorize payment, hold, or escalate is a separate decision with its own uncertainty and controls.*

At this stage, the input is singular: a single screening outcome set governed under one policy regime, even if it contains multiple hits. Ambiguity now arises from relevance, context, and policy interpretation rather than data selection.

5.2 Deterministic and Probabilistic Interpretation

Effective interpretation requires two complementary forms of reasoning:

Deterministic Interpretation	Probabilistic Interpretation
Enforces non-negotiable policy constraints: – List type and regime obligations – Country risk rules – Mandatory escalation thresholds – Prohibited party conditions *These rules define the outer bounds of permissible outcomes.*	Evaluates: – Name similarity relevance – Contextual linkage between hit and transaction – Narrative explanations for apparent matches *Probabilistic methods, including LLMs, add semantic clarity. They inform interpretation; they do not authorize outcomes.*

5.3 Multiple Independent Makers

As with Decision 1, RDS-HIVE applies multiple independent makers at the interpretation stage. Makers may include:

- Deterministic risk engines proposing policy-allowed dispositions
- Probabilistic models proposing relevance-based assessments

Each maker proposes a disposition, such as PAY, HOLD, or ESCALATE, along with supporting rationale. No single proposal is authoritative.

5.4 Deterministic Checker and Outcome Authorization

A deterministic checker, aligned to AML policy, evaluates maker proposals by asking:

- Are any mandatory policy constraints violated?
- Does probabilistic reasoning contradict deterministic rules?
- Is confidence sufficient to permit automation?

> *The checker does not reinterpret evidence. It authorizes or blocks proposed outcomes based on explicit policy logic.*

6. Decision 3 — WHEN / HOW to Validate and Finalize

6.1 Validation Is a Policy Decision

After an outcome has been authorized in Decision 2, the system must determine whether validation is required, what form it should take, and when the decision is final.

Validation is not universal. Applying the same level of review to every transaction is inefficient and unnecessary. Effective governance requires validation effort to be proportionate to risk, confidence, and regulatory obligation.

Decision 3 governs this proportionality and establishes decision finality.

6.2 Validation Options

RDS-HIVE supports multiple validation paths, selected dynamically by policy:

Path	When Applied
No Validation	Confidence is high and no regulatory or policy trigger requires review.
AI-Based Validation	An independent probabilistic checker confirms consistency, rationale, and absence of new adverse context.
Mandatory Human Validation	Required for defined high-risk scenarios, jurisdictions, or thresholds.
Random / Sampling-Based	Applied for quality assurance, audit coverage, and control effectiveness testing.

These options are contextual, not hierarchical.

6.3 Separation of Roles

Validators (whether AI or human) do not reinterpret the transaction from scratch. Their role is to:

- confirm that prior decisions were made within policy
- challenge inconsistencies or gaps in rationale
- approve or block finalization

Authority flows from policy, not from the validator's identity.

6.4 Finality as a First-Class Control

Decision finality is explicit and auditable. For each transaction, the system records:

- whether validation was required
- which validation path was applied
- why additional review was not necessary

> *This prevents silent assumptions and supports defensible outcomes under audit.*

7. Governing Decision Integrity Across the Screening Lifecycle

RDS-HIVE is not a screening model or detection technique. It is a governing architecture that defines how decision authority is constructed, constrained, and validated across the transaction screening lifecycle.

The framework preserves multiple representations of the same transaction, applies independent deterministic and probabilistic decision-makers, and authorizes outcomes through deterministic, policy-aligned checkers. Disagreement between makers is treated as a signal of uncertainty and resolved through explicit policy logic rather than forced automation.

> *By separating inference from authority and validation from execution, RDS-HIVE enables automation where confidence is earned, while preserving auditability and control.*

8. Regulatory Alignment and Auditability

Supervisory guidance such as the U.S. Federal Reserve's SR 11-7 emphasizes that models that influence risk decisions, directly or indirectly, must be governed, validated, and used within appropriate controls. While RDS-HIVE is an architectural framework rather than a model or compliance assertion, its design aligns closely with these expectations.

RDS-HIVE supports regulator-defensible decisioning by ensuring that:

- probabilistic models inform decisions but do not authorize them
- independence is enforced by design rather than process
- decision lineage is explicit and preserved

- validation and finality are policy-driven and auditable

By separating inference, authorization, and validation, the framework constrains model risk rather than assuming it away. This approach is applicable across institutions, regimes, and technologies without prescribing specific algorithms or controls.

9. Real-Time Payments as a Governance Stress Test

Real-time payment systems do not introduce new financial crime risks; they expose weaknesses in decision governance. Latency constraints reduce tolerance for manual intervention, while regulatory expectations for control and accountability remain unchanged. In real-time environments, this approach is no longer viable. Decisions must be governed and authorized before execution.

RDS-HIVE supports real-time operation by resolving screening scope early, separating interpretation from validation, and applying audit and review selectively based on policy. Decisions may be finalized immediately where confidence is sufficient, while validation is applied conditionally without blocking payment flow.

> *Real-time viability is achieved not by relaxing controls, but by placing them where they are effective.*

10. Conclusion — From Screening to Decision Integrity

Transaction screening has long been treated primarily as a detection and matching problem. While successive generations of systems have improved specific capabilities, the core challenges of over-screening, excessive false positives, and sustained reliance on human review have persisted.

This paper has argued that a primary contributing factor is not insufficient intelligence, but insufficient decision governance. By collapsing fundamentally different decisions into a single screening pipeline, traditional architectures obscure uncertainty and incentivize defensive behavior.

RDS-HIVE addresses this by explicitly separating transaction screening into three governed decisions: WHAT to screen, HOW to interpret screening outcomes, and WHEN and HOW to validate and finalize decisions. The framework preserves multiple representations of transaction data, applies independent deterministic and probabilistic decision-makers, and authorizes outcomes through deterministic, policy-aligned controls.

Crucially, probabilistic models inform decisions but do not authorize them. Disagreement is surfaced as a control signal, validation is applied proportionately by policy, and decision finality is explicit and auditable.

> *The result is not fewer controls, but better ones. Over-screening is reduced at its source, automation is enabled where confidence is earned, and human review is reserved for cases where uncertainty remains irreducible.*

As payment systems accelerate and regulatory expectations continue to rise, transaction screening architectures must move beyond optimizing matches toward preserving decision integrity. RDS-HIVE provides a practical, flexible, and regulator-defensible foundation for that shift in transaction screening governance.

A final note: nothing in this framework diminishes the role of screening. Deterministic, policy-driven screening remains an indispensable component of transaction compliance and a core part of the RDS-HIVE architecture. RDS-HIVE governs the decisions that surround it — not in place of it. Readers seeking a detailed treatment of the KYC2020 screening engine, its gateway and name screening design, and its proof-of-work auditability should refer to the DecisionIQ Whitepaper available at www.kyc2020.com or contact us via email at sales@kyc2020.com

Appendix A – Worked Example: MT103 Free-Text Storm and Controlled Resolution

Transaction Screening 'ALERT STORMS'

RAW MESSAGE (AS RECEIVED)

```
:20:TRX-20240518-998771
:23B:CRED
:32A:240518USD1254830,00
:50K:/789456123
AL NOOR TRADING L.L.C
P O BOX 44567
DUBAI UAE
:52A:XYZBAEAD
:56A:BKTRUS33
:57A:SABRSARI
:59:/00987654321
GLOBAL INFRASTRUCTURE SOLUTIONS LTD
12 INDUSTRIAL ZONE
BAGHDAD SITE OFFICE PHASE II
REF: IRAQ POWER PLANT UPGRADE
CONTRACT NO IP-7781
:70:INVOICE 7781
PAYMENT FOR CONSULTING AND
TECHNICAL SERVICES RELATED TO
GAS TURBINE INSTALLATION
:71A:SHA
```

WHY THIS CREATES A FIRCO STORM

A traditional transaction screening system will trigger multiple independent alerts:

NAME-BASED ALERTS

AL NOOR TRADING L.L.C
Common Middle East business name. Fuzzy matches against multiple OFAC / UN entities.

GLOBAL INFRASTRUCTURE SOLUTIONS LTD
Generic corporate name. Partial overlaps with sanctioned entities.

GEOGRAPHIC / KEYWORD ALERTS

BAGHDAD, IRAQ, GAS, POWER PLANT, TURBINE

BANK / ROUTING ALERTS

- Correspondent routing through multiple jurisdictions
- Intermediary bank paths interpreted as exposure

Result: Dozens of alerts, mixed severity, no clear prioritization.

Appendix A – Worked Example: MT103 Free-Text Storm and Controlled Resolution

RDS-HIVE Framework in Action

STEP 1: PARALLEL REPRESENTATIONS

RAW Representation
- Free-text fields contain high-risk keywords
- Geographic references appear in unstructured text only
- Multiple plausible interpretations exist

Canonical Representation (DET)
- Ordering customer and beneficiary normalized
- No sanctioned country in destination or routing fields
- Banks resolved to valid BICs
- No mandatory jurisdiction trigger identified

Augmented Representation (LLM)
- "Baghdad Site Office" interpreted as **project location**, not payment destination
- Customer history shows repeated cleared payments for same contract
- External context confirms civilian infrastructure program
- No evidence of prohibited sector or entity

STEPS 2–7: PROCESSING PIPELINE

STEP 2: MAKER PROPOSALS

LLM Maker (RAW)
Proposes broad screening due to unresolved free-text ambiguity

DET Maker (Canonical)
- Beneficiary name screening
- Ordering party name screening
- **No geographic gateway screening**

LLM Maker (Augmented)
- Name screening only
- Downgrade free-text geographic relevance

Result: Disagreement on whether geography should be screened.

STEP 3: POLICY AUTHORITY

Deterministic Checker
- Geographic screening required only when country appears in structured routing or destination fields
- Narrative references alone are insufficient

Authorized Screening Scope
Screen ordering and beneficiary names only. Decision rationale recorded with policy citation.

STEP 4: GATEWAY SCREENING
- Country lists → CLEAR
- Vessel / asset lists → CLEAR
- Bank identifiers → CLEAR

No gateway indicators triggered.

STEP 5: NAME SCREENING
- Fuzzy matches generated for both parties
- All matches below escalation threshold after:
  - Name disambiguation
  - Jurisdiction weighting
  - Customer history correlation

Screening Outcome Package
- Full match trace · Thresholds · Scoring rationale · Proof of Work

STEP 6: INTERPRETATION
- Absence of gateway risk
- Historical consistency
- Contract continuity
- Civilian infrastructure context

Auto-Decision: PAY
Risk Score: Medium-Low · Proof of Decision generated

STEP 7: CONDITIONAL VALIDATION
- PAY decisions with Iraq-related free text fall into **audit band**
- **LLM Checker** reviews rationale → ACCEPT
- No human escalation required

OUTCOME
- ✓ **No name screening explosion**
- ✓ Full audit trail preserved:
  - Why keywords did not trigger geography screening
  - Why matches were cleared
  - Why decision was authorized

WHY THIS EXAMPLE IS IMPORTANT

This example demonstrates that:

- FIRCO storms are caused by **unresolved ambiguity**, not weak screening
- Free-text keywords are signals, not decisions
- Deterministic policy authority prevents probabilistic overreach
- Screening scope control matters more than match tuning
- Auditability is preserved without **universal escalation**