

WHITE PAPER

Real-Time Risk Scoring in Cross-Border Payments

The next evolution of AML — why batch screening is dead, what real-time means in practice, and how the fastest-moving payment firms are rebuilding compliance for a world that settles in seconds.

Table of Contents

1. Executive Summary	3
2. The Speed Problem: Payments Got Faster, Compliance Didn't	4
3. From Batch Screening to Real-Time Scoring	5
4. Anatomy of a Real-Time Risk Score	6
5. The Machine Learning Advantage	7
6. Regulatory Expectations: What the Rules Now Require	8
7. Implementation: What It Takes to Go Real-Time	9
8. About KYCEER	10

1. Executive Summary

Cross-border payments now settle in seconds. Compliance screening still runs in batches. That gap is where financial crime lives — and where regulators are now pointing.

10s

Maximum settlement time for SEPA instant payments under EU regulation, effective October 2025

35%

Of cross-border retail payments credited within one hour — against a G20 target of 75%

~50%

Reduction in false positives achievable with ML-powered real-time scoring vs. rules-based batch screening

The shift from batch AML screening to real-time risk scoring is not a technology upgrade. It is a structural change in how compliance works. When a payment settles in 10 seconds, there is no time for a batch job to run overnight, no time for an analyst to review an alert before funds move, no time for a second chance.

This white paper explains what real-time risk scoring actually means in the context of cross-border payments, why it matters now, and what it takes to build it. Every number cited comes from a named source: the FSB, SWIFT, the European Central Bank, FATF, or published enforcement data.

Who This Is For

Compliance officers at payment institutions, money transfer operators, and fintechs that process cross-border transactions. Engineering teams building or upgrading payment infrastructure. Executives evaluating the gap between their current AML stack and where regulators expect them to be by 2026.

The core argument is simple: batch screening was designed for a world where payments settled in days. Instant payments require instant compliance. The firms that close this gap first gain a measurable advantage — in speed, cost, and regulatory standing. The firms that don't close it are running a risk that grows with every transaction.

2. The Speed Problem

Payments got faster. Compliance didn't. Here is what that mismatch looks like in numbers.

2.1 The Instant Payments Explosion

The global payment infrastructure is undergoing its most significant transformation in decades. Multiple forces are converging simultaneously:

SEPA Instant (EU)

Since January 2025, all EU Payment Service Providers must receive euro instant payments within 10 seconds. As of October 2025, they must also send them under the same standard. Batch screening is no longer fit for purpose — sanctions and fraud checks must run in real time.

ISO 20022 Migration

November 2025 marks the end of SWIFT's MT/ISO 20022 coexistence period. ISO 20022 becomes the exclusive standard for cross-border payments. As of late 2024, 28% of SWIFT cross-border traffic had migrated, with the top 175 banks (80% of volume) confirmed ready.

FedNow & CHIPS (US)

FedNow launched in July 2023, bringing real-time gross settlement to the US domestic market. CHIPS migrated to ISO 20022 in April 2024. In July 2025, Bank of America completed Fedwire ISO 20022 implementation, unifying all major US clearing infrastructure on a common standard.

G20 Roadmap Targets

The G20 targets 75% of cross-border payments credited within one hour by end-2027. The FSB's 2025 progress report shows we're at 35% for retail and 55% for wholesale. The gap is closing — and compliance must keep pace.

2.2 The Compliance Time Budget

Here is the fundamental arithmetic of real-time compliance:

Payment Type	Settlement Window	Available Screening Time	What Must Happen
Traditional SWIFT (MT103)	1-5 business days	Hours to days	Batch screening overnight, manual review during business hours
Same-day value (SWIFT gpi)	Same business day	Minutes to hours	Near-real-time screening, expedited manual review
SEPA Instant / FedNow	<10 seconds	<2 seconds	Fully automated real-time scoring, no manual intervention possible pre-settlement
Visa Direct / Mastercard Send	<30 minutes	<5 seconds	Real-time scoring at initiation, post-settlement monitoring for complex cases

The 2-Second Compliance Window

For a SEPA instant payment, the entire compliance check — sanctions screening, fraud scoring, AML risk assessment, and the accept/reject decision — must complete in under 2 seconds to leave enough time for the payment to process and settle within the 10-second regulatory limit. This is not a performance optimisation exercise. It is a hard constraint that reshapes the entire compliance architecture.

2.3 What Happens When Compliance Can't Keep Up

When screening latency exceeds the settlement window, payment firms face three options — all of them bad:

- ✓ **Block everything pending review** — destroys the customer experience and defeats the purpose of instant payments. Customers leave.
- ✓ **Let payments through unscreened** — creates regulatory exposure. This is what the \$3.3 billion in transaction monitoring fines in 2024 were about.
- ✓ **Opt out of instant payments entirely** — increasingly not viable as regulators mandate participation (SEPA instant) and customers expect it.

Real-time risk scoring is the fourth option: screen every transaction before settlement, automatically, in milliseconds.

3. From Batch Screening to Real-Time Scoring

Batch screening checks transactions against lists. Real-time scoring evaluates risk in context. The difference is not speed alone — it is a fundamentally different approach to compliance.

3.1 How Batch Screening Works (and Why It Breaks)

Traditional batch screening follows a simple pattern: collect transactions during the day, run them against sanctions lists and rule sets overnight, generate alerts the next morning, and assign analysts to review. This model worked when payments took days to settle. It breaks in three ways when payments are instant:

Timing Failure

If a sanctioned entity sends a payment at 9:01 AM and the batch runs at midnight, funds have been in the beneficiary's account for 15 hours before the alert fires. In an instant payment system, the funds are irrecoverable within seconds.

Context Failure

Batch screening evaluates each transaction in isolation. It cannot see that the same customer sent 47 payments of \$999 to 47 different beneficiaries in the same country over the past 3 hours — because those transactions are sitting in a queue waiting for the batch job.

3.2 What Real-Time Risk Scoring Changes

Real-time risk scoring replaces the batch model with a streaming architecture. Every transaction is scored at the moment of initiation, using all available context:

Dimension	Batch Screening	Real-Time Risk Scoring
When it runs	Scheduled (hourly, nightly)	At transaction initiation, before settlement
What it checks	Name against sanctions lists, amount against thresholds	Name, amount, counterparty, corridor, velocity, device, behavioural pattern
Context available	Single transaction in isolation	Full customer history, session activity, peer group behaviour
Decision speed	Hours to next business day	<200 milliseconds
Output	Binary alert (match / no match)	Continuous risk score (0–100) with explainability
False positive rate	90–98% typical	40–60% achievable with ML models
Human review	Required before decision on every alert	Only for scores above threshold; low-risk auto-cleared

<200ms

Target latency for real-time risk scoring (sanctions + AML + fraud combined)

36%

Detection accuracy improvement with ML-powered AML platforms vs. rules-only (2025 industry data)

13.4%

CAGR of cloud-native AML deployments, reflecting the shift to streaming analytics

The Score vs. The Alert

The most important conceptual shift is from binary alerts to continuous scores. A batch system says "match" or "no match." A real-time scoring system says "this transaction has a risk score of 73 out of 100, driven primarily by an unusual corridor for this customer (weight: 0.35), a counterparty with limited transaction history (weight: 0.25), and an amount 4.2x the customer's 90-day average (weight: 0.20)." The score enables automated decisioning. The explainability enables audit.

4. Anatomy of a Real-Time Risk Score

A risk score is not a single number. It is a composite of signals, each weighted by relevance, computed in milliseconds, and traceable for audit.

4.1 The Signal Categories

An effective real-time risk score for cross-border payments draws from five categories of signal, each evaluated simultaneously:

1. Entity Risk

The customer's inherent risk profile: KYC verification status, risk rating (low/medium/high), PEP or sanctions proximity, source of funds documentation, account age, and historical alert-to-SAR conversion rate. This is the baseline — computed at onboarding and updated continuously.

2. Counterparty Risk

Who the money is going to: known or first-time beneficiary, beneficiary country risk rating, beneficiary institution risk, and any adverse media or watchlist proximity on the receiving party. For cross-border payments, counterparty risk is often the highest-weighted signal.

3. Transaction Risk

The characteristics of this specific payment: amount relative to customer's typical behaviour, currency, payment purpose (if declared), whether it is structured just below reporting thresholds, and time of day. Anomaly detection models compare the transaction against the customer's established pattern.

4. Corridor Risk

The risk profile of the payment route: originating jurisdiction FATF rating, destination jurisdiction FATF rating (21 countries grey-listed as of 2025), specific corridor risk based on known typologies, and whether the corridor is flagged for elevated enforcement activity.

4.2 Velocity and Behavioural Signals

The fifth category — velocity and behavioural analysis — is what separates real-time scoring from batch screening entirely. These signals can only be computed in real time because they depend on what the customer is doing right now:

Signal	What It Measures	Why It Matters
Transaction velocity	Number and value of transactions in the last 1, 6, 24, 72 hours	Structuring and rapid movement of funds are invisible in batch; they are the strongest signal in real time
Beneficiary fan-out	Number of unique beneficiaries in recent time windows	A customer sending to 15 new beneficiaries in one day is a fundamentally different risk than one sending to a regular recipient
Round-tripping	Funds received and immediately sent out to a different party	Classic layering pattern; detectable only when inbound and outbound transactions are correlated in real time
Channel switching	Customer suddenly changes from app to agent, or shifts from one corridor to another	Behavioural shift may indicate account takeover or mule activity
Peer comparison	How this transaction compares to other customers with similar profiles	An amount that is normal for a business customer is abnormal for a student; peer-group context reduces false positives

The Structuring Problem

Structuring — breaking large amounts into smaller payments to avoid reporting thresholds — is the single most common money laundering typology flagged by regulators. In a batch system, detecting structuring requires analysing a full day's transactions after the fact. In a real-time system, the fifth payment of \$2,900 (just under a \$3,000 threshold) from the same customer in the same day triggers an elevated score on the fifth transaction, before it settles. The difference between detection and prevention.

5. The Machine Learning Advantage

Rules-based systems tell you what you already know to look for. Machine learning finds what you don't.

5.1 The Numbers

The case for ML in AML is no longer theoretical. 2025 industry data shows measurable, reproducible improvements:

36%

Improvement in detection accuracy with ML-powered AML platforms vs. rules-only systems

~50%

Reduction in false positives reported by institutions deploying ML for transaction monitoring

95%

Accuracy rate achieved by AI-driven fraud detection systems globally (2025)

5.2 What ML Does Differently

Rules-Based Approach

- Fixed thresholds: "flag any transaction over \$10,000"
- Static lists: check name against sanctions databases
- Binary output: match or no match
- Context-blind: same rule for a student and a business
- Easily evaded: criminals learn the thresholds
- High false positives: 90–98% typical

ML-Augmented Approach

- Dynamic thresholds: adjusts based on customer profile and behaviour
- Fuzzy matching: handles transliterations, aliases, partial names
- Continuous scoring: 0–100 risk score with feature attribution
- Context-aware: peer-group comparison, behavioural baselines
- Adaptive: model updates as new typologies emerge
- Lower false positives: 40–60% achievable

5.3 Explainability: The Regulatory Requirement

Regulators do not accept black-box decisions. FATF, the FCA, and the ECB all require that compliance decisions be explainable and auditable. This means ML models used in AML must provide:

- ✓ **Feature attribution** — which input signals contributed to the score and by how much
- ✓ **Decision rationale** — why a specific threshold triggered escalation or auto-clearance
- ✓ **Model documentation** — training data, validation methodology, performance metrics, and known limitations
- ✓ **Ongoing monitoring** — model performance tracked over time with alerts for drift or degradation
- ✓ **Human override capability** — analysts can override any automated decision with documented reasoning

"Technology-driven compliance is now mandatory. Regulators are encouraging AI-native transaction monitoring for real-time detection, automated KYC for improved efficiency, and predictive analytics identifying emerging risks before escalation."

— Moody's, AML in 2025 Report

The Feedback Loop

The most valuable property of ML-based scoring is the feedback loop. Every analyst decision — confirming a true positive, dismissing a false positive, filing a SAR — becomes training data that makes the model more accurate over time. Rules-based systems do not learn. They must be manually updated. An ML system that processes 100,000 transactions per day with analyst feedback on 500 escalated cases per day is continuously improving. After 12 months, the model has learned from 180,000+ analyst decisions. This is the compounding advantage that rules cannot replicate.

6. Regulatory Expectations

Real-time risk scoring has moved from best practice to regulatory expectation. FATF, the EBA, and national regulators are explicit: continuous risk assessment is no longer optional.

6.1 What the Rules Now Require

Regulator / Standard	Requirement	Effective
EU Instant Payments Regulation	PSPs must screen all instant payments against sanctions lists before settlement, within the 10-second processing window. Batch screening explicitly insufficient.	Jan 2025 (receive) / Oct 2025 (send)
FATF Recommendation 16	Originator and beneficiary information must be transmitted with wire transfers and made available for sanctions screening. Real-time screening implied for instant settlements.	Ongoing (updated guidance 2024)
EBA Guidelines on ML/TF risk factors	Institutions must apply a risk-based approach that includes ongoing monitoring of business relationships and transactions. Static, periodic reviews are insufficient for high-risk relationships.	2024 revision
SWIFT ISO 20022 mandate	All cross-border payments must use ISO 20022 messaging. The richer data structure enables and is intended to support more effective real-time screening.	November 2025
G20 / FSB Roadmap	75% of cross-border retail payments credited within 1 hour by end-2027. Compliance cannot be the bottleneck.	Target: end-2027

6.2 The Enforcement Signal

Regulators are backing these expectations with enforcement. The \$4.6 billion in global AML penalties in 2024 (Fenergo) tells a clear story about where the bar is being set:

\$3.3B

Penalties for transaction monitoring failures in 2024 — the single largest enforcement category

\$161M

AML penalties against money transmitters in 2025 alone, second only to crypto exchanges

95%

Of global penalty value concentrated in North America, reflecting heightened US enforcement

6.3 The ISO 20022 Compliance Opportunity

The migration to ISO 20022 is not just a messaging format change. It is a compliance upgrade. ISO 20022 messages carry significantly richer data than legacy MT messages:

What ISO 20022 Adds

- ✓ Structured originator and beneficiary name and address fields (vs. free-text in MT103)
- ✓ Purpose of payment codes (standardised, machine-readable)
- ✓ End-to-end transaction reference for full chain traceability
- ✓ LEI (Legal Entity Identifier) fields for corporate transactions

What This Enables for Risk Scoring

- ✓ More accurate sanctions screening (structured names reduce fuzzy match noise)
- ✓ Purpose-based risk rules (certain payment purposes are higher risk)
- ✓ Full payment chain visibility (layering detection across intermediaries)
- ✓ Automated entity resolution (LEI links to corporate registries)

The Data Quality Dividend

Institutions that invested early in ISO 20022 adoption are already seeing compliance benefits. Structured data reduces sanctions screening false positives by 20–30% compared to free-text matching, simply because the system can distinguish between a beneficiary named “Ali Hassan” and a street address containing “Hassan Ali Road.” This is not a technology improvement. It is a data quality improvement — and it compounds with every screening check.

7. Implementation: What It Takes to Go Real-Time

Moving from batch screening to real-time scoring is an architecture change, not a vendor swap. Here is what the transition looks like in practice.

7.1 Architecture Requirements

Component	Batch Architecture	Real-Time Architecture
Data ingestion	File-based ETL (daily/hourly)	Event-driven streaming (Kafka, Pulsar, or equivalent)
Screening engine	Database queries against list snapshots	In-memory scoring with sub-200ms SLA
Customer context	Looked up per batch run	Pre-computed feature store, updated in real time
Decision engine	Alert queue for next-day review	Automated accept/escalate/reject with configurable thresholds
List updates	Daily or weekly refresh	Continuous sync with <15 minute propagation
Scalability	Sized for peak batch load	Elastic, auto-scaling with transaction volume

7.2 The Build vs. Buy Decision

Building a real-time risk scoring engine in-house is possible but expensive. The typical trade-offs:

Build In-House

Timeline: 12–18 months to production. **Team:** 4–8 engineers plus data scientists. **Ongoing cost:** Infrastructure, list subscriptions, model maintenance. **Best for:** Large institutions with unique risk models and dedicated ML teams. **Risk:** Regulatory validation of proprietary models is complex and slow.

Buy (API Platform)

Timeline: 2–8 weeks to production. **Team:** 1–2 integration engineers. **Ongoing cost:** Per-transaction or subscription pricing. **Best for:** Payment firms that want to focus on their core product, not compliance infrastructure. **Advantage:** Pre-validated models, managed list updates, shared intelligence across the network.

7.3 Performance Benchmarks

These are the latency and throughput numbers that a production real-time scoring system should meet:

Operation	p50 Latency	p99 Latency	Pattern
Sanctions screening (single name)	15ms	80ms	Synchronous, pre-settlement
AML risk scoring (single transaction)	25ms	120ms	Synchronous, pre-settlement
Combined score (sanctions + AML + fraud)	50ms	200ms	Synchronous, parallel execution
Batch re-screening (list update, 100K entities)	—	<5 min total	Asynchronous background job
Sustained throughput	10,000+ transactions per second		

2–8 wks

Time to production with an API-based real-time scoring platform

<200ms

End-to-end scoring latency (sanctions + AML + fraud combined) at p99

10K+/s

Sustained transaction throughput for production-grade scoring infrastructure

The Transition Path

Most institutions do not switch from batch to real-time overnight. The proven transition path is: (1) Deploy real-time scoring in shadow mode alongside existing batch screening for 4–8 weeks. (2) Compare results — verify that real-time catches everything batch catches, plus more. (3) Gradually shift decision authority from batch to real-time, starting with low-risk transaction types. (4) Retire batch screening once real-time has been validated across all transaction types and risk levels. This phased approach satisfies both engineering rigour and regulatory expectations for model validation.

8. About KYCEER

KYCEER is an AI-native compliance platform built for the speed of modern payments. Every capability described in this white paper — real-time sanctions screening, ML-powered transaction scoring, velocity analytics, and explainable risk decisioning — is available through a single API.

Real-Time Sanctions Screening

Sub-200ms screening against OFAC, EU, UN, UK, and 50+ additional lists. Fuzzy matching with transliteration and alias support. Continuous re-screening on list updates with <15 minute propagation.

Transaction Risk Scoring

ML-powered risk scoring for every transaction. Entity risk, counterparty risk, corridor risk, and velocity signals combined into a single explainable score. Configurable thresholds for automated accept/escalate/reject decisions.

KYC & KYB Verification

Identity and business verification across 195+ countries. Sub-10-second verification for standard checks. Risk-adaptive flows that adjust depth based on customer risk profile and corridor.

Case Management & AI Agents

Full investigation workflow from alert to SAR filing. AI compliance agents that automate alert triage, case narrative generation, and regulatory report drafting. Every state transition accessible via API.

Built for Speed

- ✓ **API-first architecture** — RESTful APIs with sub-200ms SLAs, designed for inline payment processing
- ✓ **Streaming analytics** — velocity and behavioural signals computed in real time, not batch
- ✓ **ISO 20022 native** — structured data ingestion for richer screening and lower false positive rates
- ✓ **Explainable ML** — every score includes feature attribution for audit and regulatory review
- ✓ **Feedback loop** — analyst decisions continuously improve model accuracy
- ✓ **Multi-region deployment** — EU and UK data residency options with automated failover

Level39, One Canada Square, London

KYCEER is headquartered at Level39, Europe's largest technology accelerator for finance, cyber-security, and smart-city technology. Located in Canary Wharf, London — at the centre of the global financial services industry and the regulatory ecosystem that governs it.

Sources

FSB, G20 Roadmap for Cross-Border Payments: Consolidated Progress Report (October 2025) • European Commission, Instant Payments Regulation (EU) 2024/886 • SWIFT, ISO 20022 Migration Status Update (November 2024) • FATF, Updated Guidance on Recommendation 16 (2024) • European Banking Authority, Guidelines on ML/TF Risk Factors (2024 revision) • Fenargo, AML Enforcement Action Report (2024) • Moody's, AML in 2025 Report • Silent Eight, 2025 Trends in AML and Financial Crime Compliance • Bank of America, Fedwire ISO 20022 Implementation (July 2025) • BIS, Enhancing Cross-Border Payments: State of Play (2025)



Payments move in milliseconds. Does your compliance?

Talk to our team about real-time risk scoring for your cross-border payment infrastructure.

[Book a Demo](#)

kyceer.com

Level39, One Canada Square, London E14 5AB

contact@kyceer.com