BLUEVERGE ENTERPRISE SOLUTIONS
EXECUTIVE INTELLIGENCE PAPER

The 95% Problem

Why Modern Compliance Begins Where Sanctions Screening Ends

A data-driven examination of the false-positive crisis, the hidden cost of manual review, and the operating model compliance leaders are building to replace “screen-and-clear” with risk-based, AI-assisted, audit-ready decisioning.

Prepared for:
Chief Compliance Officers · MLROs · AML & Risk Officers · FinTech, Banking, Payments & iGaming Executives

Empowering Business and Compliance through Intelligence Systems

Executive Summary

Sanctions screening has become the most universally deployed — and most quietly inefficient — control in financial crime compliance. Nearly every regulated institution screens. Yet across banking, payments, FinTech and iGaming, the same uncomfortable pattern repeats: the overwhelming majority of alerts a screening engine produces are false positives, and the real cost, risk and regulatory exposure sit not in the match itself, but in everything that happens after it.

Industry accounts commonly place false-positive rates between 90% and 99%. In practical terms, for every hundred alerts an analyst reviews, a handful — sometimes none — represent a genuine concern. The screening engine has done its narrow job; the institution is left with the far larger job of assessing, investigating, documenting and deciding. That work is manual, repetitive, costly, and increasingly scrutinised by regulators who now expect not just detection, but demonstrable effectiveness, explainability and a defensible audit trail.

This paper argues that the compliance industry has been optimising the wrong half of the problem. The market competes on list coverage and match algorithms — the front of the process — while the expensive, risk-bearing back of the process (assessment, guidance, documentation, decision) remains under-served. The institutions pulling ahead are those rebuilding their operating model around a simple shift:

THE SHIFT IN ONE LINE

- FROM: Screening → Alert → Manual Review
- TO: Screening → Risk-Based Assessment → AI-Assisted Guidance → Documentation → Human Decision

The evidence assembled here — drawn from industry commentary, regulator enforcement data, vendor positioning and compliance surveys — points to four conclusions:

1. False positives are the primary driver of compliance cost and staffing strain, and the prevailing fix is no longer “more reviewers” but better data, risk-based prioritisation, and AI-assisted triage that keeps humans in control.
2. Screening engines alone are insufficient. Even market-leading vendors increasingly bolt on case management, audit trails and workflow because the engine, by itself, does not close the compliance loop.
3. Regulators are raising the bar from detection to effectiveness — demanding testing, tuning evidence, explainability and change-management governance that most manual, screenshot-based processes cannot produce.
4. AI adoption is accelerating, but its value depends on governance. The defensible model is human-in-the-loop: AI that explains, prioritises, and documents — never one that decides unaccountably.

For compliance leaders, the strategic question is no longer “Do we screen?” It is “What happens after the alert — and can we defend it?” This paper sets out the problem with data, examines why current approaches fall short, and describes the modern compliance operating model now emerging in response.

The Compliance Crisis Nobody Talks About

Every regulated institution talks about financial crime risk. Far fewer talk candidly about the day-to-day reality of the control most of them rely on most heavily: sanctions and watchlist screening. The uncomfortable truth, well known inside compliance teams but rarely stated plainly to the board, is that the process generates enormous volumes of noise, consumes disproportionate human effort, and produces remarkably few genuine hits.

This is not a failure of any single vendor or team. It is a structural feature of how screening works. Name-matching against large, frequently-updated watchlists — across multiple jurisdictions, with messy customer data, transliteration variance, common names and incomplete identifiers — inevitably produces large numbers of plausible-but-wrong matches. The system is designed to err on the side of flagging, because the regulatory cost of a missed true match (a false negative) is severe. The result is a deliberate, structural bias toward over-alerting.

The consequence is a quiet crisis of productivity and proportion. Analysts spend the majority of their time confirming that alerts are not what the system feared. Onboarding slows. Payments are held. Costs climb with volume. And because so much of the work is repetitive clearing of low-value alerts, the function struggles to direct its scarce expertise toward the genuinely elevated risks that warrant it.

THE CORE STATISTIC

90–99%

- of sanctions screening alerts are commonly reported as false positives across name-matching and transaction workflows.
- Independent technical reviews find false-positive rates that “exceed 90%.” Vendor and RegTech summaries cite figures as high as “95% of alerts are false positives.”
- Source: ACAMS commentary; academic/technical reviews; RegTech vendor analyses (2024–2025).

Framed differently: institutions have invested heavily in the ability to detect, and comparatively little in the ability to resolve efficiently and defensibly. The detection layer is mature and commoditised. The resolution layer — assessment, prioritisation, investigation, documentation — is where the cost lives, where the regulatory exposure concentrates, and where the least technology has historically been applied. That imbalance is the crisis this paper examines.

■ RECOMMENDED VISUAL — The Screening Funnel

A vertical funnel showing 100% alerts at the top narrowing through ‘false positives (90–99%)’ to a thin sliver of ‘true matches’ at the bottom — with a callout that the entire width of the funnel still requires human review effort. Reinforces that volume, not hit-rate, drives cost.

The False Positive Trap

The false positive is often treated as a tuning nuisance — a number to be optimised down by a few percentage points. That framing understates its strategic impact. False positives are not merely an efficiency problem; they are a trap that shapes the entire economics and risk posture of a compliance function.

Why the rate is so high

Screening quality, industry commentary consistently notes, depends less on raw access to sanctions lists than on data quality, matching logic, entity resolution and workflow design. When customer data is incomplete or inconsistent, when matching is broad to avoid missing true hits, and when entity structures and jurisdictions are hard to manage consistently, alert volumes inflate. The list is rarely the problem. The context around the match is.

This is why the same names, the same low-value alerts, recur. Without contextual matching — the ability to weigh identifiers, resolve entities, and apply risk rules — the engine cannot distinguish a genuine concern from a coincidental name overlap. It defers that judgment entirely to the human, every time.

The trap, defined

The trap has three jaws. First, the volume of false positives forces institutions to scale human review — adding cost that grows with business volume rather than with actual risk. Second, the monotony of clearing near-identical low-value alerts produces reviewer fatigue, which paradoxically raises the risk of the one error that matters most: missing a true match buried in the noise. Third, because each clearance is a judgment, every one of them must — in a well-run program — be documented to a standard a regulator would accept. At 90–99% false positives, that is an enormous documentation burden attached to outcomes that are, by definition, almost always negative.

THE TRADE-OFF REGULATORS ARE WATCHING

- Tuning screening less sensitively reduces false positives — but risks false negatives, which carry severe regulatory consequences.
- Supervisors increasingly expect institutions to justify this trade-off with evidence: testing, back-testing, and documented tuning rationale.
- The goal is not fewer alerts at any cost — it is defensible effectiveness.

Escaping the trap does not mean screening less. It means changing what happens to an alert the moment it is generated: scoring it by confidence and severity, prioritising by risk, and equipping the analyst to resolve and document it efficiently. That is a shift from a binary “match / no match” engine to a risk-based assessment model — a theme this paper returns to in depth.

The Hidden Cost of Manual Reviews

The cost of false positives is easy to underestimate because most of it is hidden in plain sight — distributed across thousands of short, individually-trivial review tasks. No single alert feels expensive. The cumulative burden, however, is substantial, and it compounds as transaction and customer volumes grow.

AN ILLUSTRATIVE EXAMPLE FROM THE FIELD

- A large bank reported that roughly one in three wires was flagged for review.

~2 min

- average human review time per cleared alert — fast, yet overwhelmed by volume.
- When a third of all wires alert and almost all are cleared manually, even rapid reviews accumulate into an unsustainable operational load. Speed per case cannot offset scale of cases.
- Source: industry case commentary cited in sanctions-screening analyses (2024–2025).

The hidden costs extend beyond reviewer salaries. They include the opportunity cost of expert compliance staff spending their time on low-value clearance rather than genuine risk; the business cost of delayed onboarding and held payments, which damages customer experience and revenue; the maintenance cost of legacy systems tuned to over-alert; and the risk cost of fatigue-driven error. Industry analyses of payments-heavy institutions emphasise that high transaction volumes, combined with PEP and adverse-media noise, inflate alert counts and increase both latency and cost.

Crucially, these costs scale with business growth rather than with risk. An institution that doubles its transaction volume does not double its financial-crime risk — but under a manual-review model, it may well double its review burden. This decoupling of cost from risk is precisely what a risk-based operating model is designed to correct: directing effort toward elevated-risk cases and resolving the routine majority efficiently.

■ RECOMMENDED VISUAL — Cost Scales With Volume, Not Risk

A line chart with two diverging lines as transaction volume increases: 'Review cost (manual model)' rising steeply and linearly, versus 'Actual risk exposure' rising gently. The widening gap between them is shaded and labelled 'avoidable cost.'

The Human Bottleneck

Behind every alert statistic is a person. The false-positive problem is, ultimately, experienced as a human bottleneck: too many alerts, too few qualified reviewers, and too little prioritisation to ensure the right cases reach the right people at the right time.

Compliance teams across the industry report material staffing shortages even as alert volumes climb. The talent is scarce and expensive, and the work — repetitive review of low-value alerts — is precisely the kind that erodes engagement and retention. Surveys of compliance functions consistently identify improving data quality and automation as top priorities specifically to relieve this staff pressure. The industry is not asking for more people to do the same work; it is asking for a way to do less of the low-value work at all.

Three constraints on analyst productivity

1. Repetition. Analysts spend disproportionate time on near-identical, low-value alerts, leaving less capacity for genuinely complex or elevated-risk cases.
2. Insufficient prioritisation. Without risk-based scoring, a coincidental name match on a low-risk customer can demand the same queue position as a credible match on a high-severity listing.
3. Fragmented data. Information needed to resolve an alert is often scattered across systems, forcing analysts to assemble context manually before they can even begin to assess.

The strategic implication is important: the bottleneck is not solved by hiring alone. Adding reviewers to an over-alerting system increases capacity linearly while costs rise and the underlying inefficiency remains. The leverage lies in reducing the volume of alerts that require human attention (through better data and risk-based prioritisation) and in making the human review that remains faster, more consistent and better-documented (through AI assistance that keeps the analyst in control).

WHAT COMPLIANCE TEAMS ACTUALLY WANT

- Fewer irrelevant alerts — without missing true matches.
- Clear prioritisation — so scarce expertise is spent on elevated risk, not coincidence.
- Faster, consistent resolution — with documentation produced as a by-product, not an afterthought.
- This is the design brief for the modern compliance operating model.

Why Existing Screening Solutions Still Fall Short

If false positives and manual-review burden were simply a matter of buying a better screening engine, the problem would have been solved long ago. The persistence of these pain points — even among institutions using market-leading vendors — tells a more revealing story. The limitations are not at the edges of the market; they run through its centre.

A scan of how the major vendors position themselves is instructive, because vendors market most heavily against the pain points their customers most want solved. Across the leading platforms, the same themes recur: minimising false positives, improving matching quality, integrating screening into workflow, and reducing analyst effort. The consistency of this messaging is itself evidence that these problems remain unsolved at scale.

The pattern across the market

Leading vendors emphasise machine learning and advanced matching to reduce false positives and prioritise alerts — an implicit acknowledgement that traditional screening creates extensive manual effort and delay. Others stress data-quality enhancement and superior matching logic to cut alert noise, or advertise case management, audit trails and workflow embedding — a tacit admission that a screening engine on its own is often not enough. Independent commentary reinforces the point: weak investigation and escalation processes create audit risk, especially when ownership is unclear or decisions are poorly documented.

Recurring market pain point	What vendor positioning reveals
High false positives	Heavily marketed ML and ‘advanced matching’ indicate this remains the dominant, unsolved complaint.
Manual-review burden	Automation and alert-prioritisation messaging signal persistent analyst overload.
Workflow fit	Emphasis on integration and workflow embedding shows screening rarely fits cleanly into existing processes.
Weak documentation / audit trail	Bolt-on case management and audit-trail features reveal that the engine alone does not produce defensible records.
List-update & data quality	Promises of ‘real-time updates’ and entity resolution point to ongoing data-quality and coverage gaps.

The conclusion is not that these are poor products — many are excellent at detection. It is that the market has largely optimised the front of the process while leaving the back of the process — assessment, guidance, documentation, decision — comparatively under-served. “Screening-engine-only” is, for a growing number of institutions, simply not enough. The next sections examine the environments where this gap bites hardest, and the operating model emerging to close it.

The iGaming Compliance Challenge

No sector illustrates the false-positive trap more sharply than iGaming. Online gaming and betting operators combine several characteristics that compound screening difficulty: very high volumes of transactions, predominantly low individual values, fast onboarding expectations, a global and mobile customer base, and intense regulatory attention to anti-money-laundering and responsible-gaming obligations.

The structural profile — high-volume, low-value flows — is precisely the environment in which traditional screening generates the most noise. A betting platform onboarding thousands of customers and processing large numbers of small deposits and withdrawals will, under broad name-matching, generate enormous alert volumes. Yet the operator faces the same regulatory expectation as a bank: screen effectively, resolve defensibly, and document everything.

This creates an acute version of the central dilemma. iGaming operators cannot afford the onboarding friction of slow, manual review — their commercial model depends on fast, frictionless customer experience. But they equally cannot afford the regulatory consequences of weak controls in a sector under sustained supervisory scrutiny. They are squeezed between speed and rigour more tightly than almost any other industry.

WHY iGAMING FEELS THE TRAP HARDEST

- High transaction volume → very large alert volumes under broad matching.
- Low value per transaction → manual review cost is disproportionate to the value screened.
- Fast-onboarding business model → review latency directly damages revenue and experience.
- Heightened regulatory scrutiny → weak documentation is a serious, sector-specific exposure.

The practical path forward for high-volume sectors, supported across industry guidance, is tiered, risk-based screening: pre-filtering and risk scoring to suppress noise, real-time list management, centralised case handling, and targeted human review reserved for genuinely elevated risk. In other words, the answer for iGaming is not to screen less, but to move decisively from “screen-and-clear-everything-manually” to a risk-based model that concentrates scarce human judgment where it matters. This is the same operating-model shift the whole industry is undergoing — iGaming simply feels its urgency first.

Regulators are no longer accepting “resource constraints” as a defence

Recent enforcement makes the stakes concrete. Sweden’s gambling regulator fined Videoslots for serious AML shortcomings, including inadequate customer-knowledge work and insufficient measures to assess money-laundering and terrorist-financing risk, and issued further AML penalties against other operators through 2025. In the United Kingdom, the Gambling Commission fined Corbett Bookmakers for AML and social-responsibility failures, and penalised ProgressPlay £1 million for weaknesses in risk assessment and source-of-funds controls. The pattern is unmistakable: passive monitoring, generic warnings and checkbox compliance are not enough, and supervisors are notably unsympathetic to “we lacked the resources” when controls prove ineffective.

What regulators expect of iGaming operators now is explicit: risk-based customer due diligence, timely escalation, source-of-funds checks where warranted, and documented decision-making — with intervention when behaviour or transaction patterns indicate risk, not after-the-fact review. Compliance must be operationalised inside the product and payments flow, not bolted on afterward. That expectation maps directly onto the operating model this paper describes: score by risk, assist the analyst, document as you go, and keep a human accountable for the decision.

Lessons from Recent Regulatory Enforcement

The cost of getting compliance wrong is not theoretical. Sanctions enforcement has intensified, and the pattern of recent actions carries clear lessons about where supervisory expectations are heading — and why a screen-and-clear model is increasingly difficult to defend.

The U.S. Office of Foreign Assets Control (OFAC) published numerous civil penalty actions through 2025, maintaining a steady stream of settlements and fines against corporates and individuals. In the United Kingdom, the Office of Financial Sanctions Implementation (OFSI) operates active enforcement processes and publishes decisions that underline its compliance expectations. Industry enforcement round-ups summarise significant AML, CFT and sanctions penalties across banking, payments, crypto and broker-dealer sectors in recent years — evidence of focused, cross-sector supervisory attention.

What enforcement patterns reveal

Three lessons stand out for compliance leaders. First, supervisors increasingly assess not just whether an institution screened, but whether its program was effective — properly tuned, tested, and capable of catching what it should. Detection alone is no longer a defence. Second, documentation and governance are central: regulators expect institutions to demonstrate how decisions were made, by whom, on what basis, and with what evidence. Weak audit trails are themselves a finding. Third, change management matters: programs are expected to evolve with risk, with formal governance over tuning and configuration changes.

THE ENFORCEMENT MESSAGE, DISTILLED

- Regulators are shifting the standard from detection to demonstrable effectiveness.
- They expect explainability — a clear, evidenced rationale for every material decision.
- They expect documented testing, tuning, and change-management governance.
- A process that cannot produce a defensible record is a regulatory exposure, regardless of intent.

This is the strategic significance of the enforcement trend: it directly penalises the weakest part of the traditional model. A manual, screenshot-and-spreadsheet review process may technically “screen,” but it struggles to produce the consistent, explainable, well-governed evidence that modern supervision demands. The enforcement environment is, in effect, pricing in the cost of the documentation and effectiveness gaps this paper has described — and rewarding institutions that close them.

The Rise of Risk-Based Screening

If the binary “match / no match” engine is the root of the false-positive trap, the response taking hold across the industry is risk-based screening: an approach that does not merely flag a possible match, but assesses and prioritises it. The shift is conceptual before it is technical. It treats an alert not as a yes/no event, but as a question with a measurable degree of confidence and a measurable degree of consequence.

Two scores, not one flag

Risk-based screening separates two questions that the traditional model collapses into one. The first is identity confidence: how likely is it that the screened subject genuinely is the listed party? This is driven by how many identifying attributes align — name, date of birth, nationality, identifiers — and how many contradict. The second is risk severity: how serious is the listing that was matched? A match against a terrorism or proliferation program carries different weight than a match on a lower-tier list.

By scoring both — a Match Confidence Score and a Risk Severity Score — the model enables genuine prioritisation. A high-confidence match against a high-severity listing escalates immediately. A low-confidence coincidental name overlap on a low-risk customer is triaged accordingly. Scarce analyst attention flows to where confidence and consequence are both high. This is the mechanism that breaks the decoupling of cost from risk described earlier.

Traditional Model	Risk-Based Model
Screening → Alert → Manual Review	Screening → Risk-Based Assessment → AI-Assisted Guidance → Documentation → Human Decision
Binary match / no match	Confidence score + severity score
Every alert treated equally	Alerts prioritised by risk
Volume drives cost	Risk drives effort
Documentation is a manual afterthought	Documentation produced through the workflow

Expressed in operation, risk-based screening converts an undifferentiated flood of alerts into a prioritised queue. Rather than returning a flat list of possible matches, it produces, for each potential match, a measure of identity confidence and a measure of risk severity, and applies configurable thresholds — review, potential-match, escalation — so each institution can tune the model to its own risk appetite. Screening stops being a binary gate and becomes the first step of a risk-based assessment, focusing analyst productivity on the matches that genuinely warrant it.

THE SHIFT THIS ENABLES

- From: every alert competes equally for analyst time.
- To: confidence × severity determines priority — effort follows risk.
- Configurable thresholds let each institution encode its own risk appetite, with the rationale documented

for examiners.

The Future of AI-Assisted Compliance

Risk-based scoring addresses prioritisation. The second half of the modern model addresses the resolution itself — and this is where AI is reshaping compliance work. The momentum is unambiguous: compliance surveys show many institutions already deploying AI and machine learning in screening, with a substantial majority planning adoption over the next one to three years, specifically to reduce false positives, prioritise alerts and improve explainability.

AI ADOPTION MOMENTUM

80%+

- of surveyed institutions are either already deploying AI in compliance or planning to — with reported figures such as 41% already using AI and 38% planning adoption.
- The stated goals are consistent: reduce false positives, prioritise alerts, and improve explainability.
- Source: compliance and RegTech industry surveys (2024–2025).

But the research carries an equally consistent warning: the value of AI in compliance depends entirely on governance. Regulators expect model explainability, data quality, testing and human-in-the-loop design. An AI that produces unexplained conclusions is not an asset in a regulated environment — it is a liability. The defensible model is not automation that replaces the analyst, but augmentation that makes the analyst faster, more consistent, and better-documented while keeping the human firmly in control of the decision.

Where AI genuinely helps

Industry analysis identifies the strongest, safest AI opportunities clearly: alert clustering, entity resolution, risk-based prioritisation, case summarisation, and suggested dispositioning — always with human review retained for edge cases. The common thread is that AI handles the assembly, explanation and documentation work that consumes analyst time, while the judgment remains human. The best AI-assisted tools, the analysis concludes, will be those that reduce false positives while preserving a clean audit trail for regulators.

Expressed in operation, an AI compliance assistant does precisely what the research identifies as valuable and defensible: it explains why an alert was generated, assesses whether it is more likely a false positive or a genuine match, identifies the specific information missing to resolve it, suggests appropriate next steps and enhanced-due-diligence actions, and drafts the investigation note — so documentation is produced as a by-product of the work rather than a chore after it. Critically, the defensible design rests on one principle: the assistant assists, explains and documents, but does not decide. The analyst makes and owns the final determination.

THE GUARDRAIL THAT MAKES AI DEFENSIBLE

- An AI compliance assistant should explain and document — never decide unaccountably.

- The analyst owns the decision; the assistant owns the reasoning support and the record.
- Result: faster resolution and a stronger, explainable audit trail at the same time.

The Importance of Documentation and Audit Readiness

Throughout this paper, one theme recurs across every section: documentation. The false-positive burden is, in large part, a documentation burden. The enforcement trend is, in large part, a documentation expectation. The AI opportunity is, in large part, a documentation opportunity. If there is a single thread connecting the cost problem, the regulatory problem and the technology solution, it is the quality and defensibility of the record.

Documentation and audit-trail weaknesses remain among the most common industry gaps. Public commentary repeatedly notes that weak investigation and escalation processes create audit risk — particularly when ownership of decisions is unclear, or when the rationale for a decision is not captured in a form a regulator would accept. In a manual model, documentation is typically an afterthought: an analyst resolves an alert, then separately writes up a note, often inconsistently, sometimes long after the fact, frequently in a format that does not survive examination.

The modern operating model inverts this. Documentation is produced through the workflow, not bolted on after it. When an AI assistant drafts the investigation note as part of resolving the alert, and a case-management system captures the decision, the evidence and the rationale in a centralised, examinable record, audit readiness becomes a property of the process rather than a periodic scramble. This is not merely tidier — it is the difference between a defensible program and an exposed one.

Expressed in operation, a governance and case-management layer centralises the work: it maintains investigations, evidence and decisions in one place; preserves a consistent, examinable audit trail; and supports the testing, tuning records and change-management governance that regulators increasingly demand. Where risk-based screening prioritises and an AI assistant accelerates and documents, the case-management layer is the connective tissue that turns individual well-handled alerts into a program that can be demonstrated, examined and defended as a whole.

AUDIT READINESS AS A PROPERTY OF THE PROCESS

- Centralised investigations and decisions — not scattered notes and screenshots.
- A consistent, examinable audit trail produced as work happens.
- Support for testing, tuning evidence and change-management governance.
- Examination readiness becomes continuous, not a periodic fire-drill.

Building the Modern Compliance Operating Model

The threads of this paper converge on a single, coherent operating model — one that several leading institutions are already assembling, and that the evidence suggests will become the industry standard. It is best understood not as a product, but as a redesigned flow, with technology and people each doing what they do best.

THE MODERN COMPLIANCE OPERATING MODEL

- 1. SCREENING — cast the net (detection, the commoditised layer).
- 2. RISK-BASED ASSESSMENT — score confidence and severity; prioritise by risk.
- 3. AI-ASSISTED GUIDANCE — explain, assess, suggest next steps, draft documentation.
- 4. DOCUMENTATION — capture evidence, rationale and decision in a centralised, examinable record.
- 5. HUMAN DECISION — the analyst makes and owns the final determination.

The philosophy underlying this model is straightforward: people, systems and AI, each in their proper role. Technology casts the net and removes the noise. AI assembles context, explains, and documents. The human exercises judgment where judgment is required. Nothing in this model replaces the compliance professional; everything in it is designed to make that professional more effective, more consistent, and better-supported — and to make the institution's program demonstrably effective and defensible.

The components map cleanly onto capabilities. Risk-based screening performs steps one and two. The AI compliance assistant performs step three and the drafting within step four. The case-management and governance layer consolidates step four into an examinable whole. And because the model's effectiveness depends as much on people as on systems, capability-building — training that raises analyst effectiveness and overall compliance maturity — ensures the human at the centre of the model is equipped to use it well.

■ RECOMMENDED VISUAL — The Operating Model Flow

A horizontal five-stage flow diagram: Screening → Risk-Based Assessment → AI-Assisted Guidance → Documentation → Human Decision. Under each stage, a small label showing the responsible layer (engine / scoring / assistant / case management / analyst). Use BlueVerge blue for system stages and orange for the final human-decision stage to emphasise human control.

Strategic Recommendations for Compliance Leaders

For compliance leaders weighing how to respond, the evidence supports a clear, practical agenda. The following recommendations are ordered to deliver the fastest relief from cost and risk first, building toward a fully modern operating model.

1. **Measure your false-positive rate honestly.** You cannot manage what you do not measure. Establish your true false-positive rate, average review time, and alert volume — these three numbers quantify the hidden cost and build the business case for change.
2. **Move from binary alerts to risk-based scoring.** Prioritise by match confidence and risk severity so that scarce analyst attention follows actual risk, not alert volume. This single shift addresses the decoupling of cost from risk.
3. **Adopt AI assistance — with human-in-the-loop governance.** Deploy AI to explain, assess, prioritise and document, while keeping the analyst in control of every decision. Insist on explainability; reject black-box dispositioning.
4. **Make documentation a by-product, not an afterthought.** Re-engineer the workflow so that the investigation note, evidence and rationale are captured as the work happens, in a centralised, examinable record.
5. **Treat audit readiness as continuous.** Centralise investigations and maintain testing, tuning and change-management evidence so that examination readiness is a standing property of the program, not a periodic scramble.
6. **Invest in people alongside systems.** Technology raises the ceiling; trained analysts reach it. Build compliance capability and maturity so the human at the centre of the model uses it to full effect.
7. **Tune to your own risk appetite — and document why.** Use configurable thresholds to reflect your institution's risk appetite, and record the rationale. Regulators expect you to justify the trade-off between sensitivity and operational feasibility.

These steps are mutually reinforcing. Risk-based scoring reduces the volume reaching analysts; AI assistance accelerates the review that remains; workflow-native documentation makes each resolution defensible; case management makes the whole program examinable; and capability-building ensures people use the model well. Adopted together, they convert compliance from a cost centre fighting alert volume into a controlled, demonstrable, risk-based function.

Conclusion

Sanctions screening is necessary, universal, and — on its own — no longer sufficient. The industry has spent years optimising the front of the process: bigger lists, cleverer matching, faster updates. Yet the persistent reality of 90–99% false positives shows that the hardest, most expensive, most scrutinised work was never the match itself. It is everything that happens after: the assessment, the investigation, the documentation, the decision.

That is the 95% problem. Not that screening produces false positives — it always will — but that the traditional operating model leaves an institution to resolve that flood manually, inconsistently, and often indefensibly. The cost scales with volume rather than risk. The documentation strains to meet a rising regulatory bar. The scarce expertise of compliance professionals is consumed by clearing coincidences rather than confronting genuine threats.

The institutions moving ahead have recognised that the answer is not to screen more, but to rebuild what happens after the alert. They are adopting risk-based assessment to prioritise by confidence and severity; AI assistance to explain, accelerate and document while keeping humans in control; case management to make the program centralised and examinable; and capability-building to ensure their people use these tools well. People, systems and AI — each in its proper role.

Modern compliance, in short, begins where sanctions screening ends. The leaders who internalise that shift will spend less, defend more, and direct their best people at the risks that genuinely matter. Those who do not will continue to pay the rising, hidden cost of the 95% problem — one two-minute review at a time.

About this paper

This paper sets out the problem, the data and the strategic insight. The companion Executive Briefing — The Solution Blueprint goes further: it details the operating model in practice, compares the options available to compliance leaders, and maps a practical implementation path, including how BlueVerge's capabilities (AMRAC RBS, the BELA Compliance Agent, ACQUA-MS, and BELA Training) fit within it. To request the Executive Briefing or discuss your environment, contact the BlueVerge team.

This paper is provided for informational and educational purposes. Statistics are drawn from publicly available industry commentary, regulator publications, vendor materials and compliance surveys (2024–2025) as summarised in the supporting research; figures are frequently reported as ranges and should be validated against primary sources before citation. Nothing herein constitutes legal or regulatory advice.