

AML VENDOR CHECKLIST

A PRACTICAL, NO-FLUFF FRAMEWORK TO PICK A SCREENING PARTNER THAT WON'T FAIL UNDER REAL RISK.

1

VENDOR FIT AND SCOPE

Start with alignment the vendor should match your risk model and not force you into theirs.

- What AML screening use cases are supported: sanctions, PEP, watchlists, adverse media, KYB, and ongoing re-screening?
- Is AML screening their core product or a secondary add-on?
- Proven in your industry, geography, and regulator class?
- Can they show references in your exact risk tier (size, exposure, regions)?
- What is proprietary vs third-party (lists, media sources, enrichment)?
- 12–18 month roadmap: compliance-driven or sales-driven?

COVERAGE DEPTH ACROSS SANCTIONS, PEPs & WATCHLISTS

Coverage doesn't necessarily mean how many lists you cover, but about how well those lists reflect reality.

SANCTIONS COVERAGE

- Covers key global regimes (OFAC, UN, EU, UK OFSI, etc.) plus domestic/local lists where you operate?
- Handles domestic data issues with duplicates, missing DOBs, and inconsistent scripts?
- Shows source authority and listing timeline per record?
- Does it unify overlaps into one resolved identity profile?
- Interprets primary vs secondary sanctions per jurisdiction?

PEP COVERAGE

- Includes municipal/regional/local PEPs, not just top leadership?
- Handles disputed/contested territories transparently?
- Maps, Relatives and Close Associates (RCAs) structurally?
- Continuous monitoring of role changes/exit/re-entry?

WATCHLIST/ENFORCEMENT INTELLIGENCE

- Clear definition of watchlists vs sanctions/PEP?
- Includes enforcement categories of debarments, fugitives, regulatory actions, fraud/scam bulletins, and professional disciplinary lists?
- Let's you filter decisions by watchlist type?

ADVERSE MEDIA (IF CLAIMED)

- Coverage of global, local and official/court sources?
- Multi-language support?
- Severity/relevance scoring, not noisy keyword spam?
- Linked to a resolved identity?

3

MATCHING ENGINE QUALITY

The engine is the difference between operational chaos and true risk control to minimise false positives.

Hard requirement: run a live test using your worst-case real names.

- Matching method: fuzzy, phonetic, and AI/NLP hybrid?
- Handles hard realities:
 - multilingual transliterations
 - script-to-script matching (Arabic ↔ Latin, Cyrillic ↔ Latin, etc.)
 - honorifics/initials/hyphens/surname-first formats
- Can distinguish:
 - same spelling / different people
 - different spelling / same person
- Provides match explainability (why this score)?
- Thresholds tunable by list type (sanctions vs PEP vs watchlist)?
- Shares benchmark performance on hard/local names?

RISK SCORING AND DECISION LOGIC

Your policy should drive decisions, not vendor defaults.

- Scoring model declared rules / ML/hybrid?
- Tunable by jurisdiction, list type, customer/KYB type, and product risk?
- Decision states supported auto-clear, auto-reject, escalate, and monitor?
- Feedback loop from analyst outcomes with audit trail?
- Evidence of false-positive reduction after tuning?

QUICK “KILL SWITCH” RED FLAGS

Any of these will **compromise** compliance outcomes

- No domestic sanctions depth.
- PEP lists are limited to top-tier officials only.
- Watchlists were wrongly merged into sanctions/PEP.
- Black-box scoring.
- No live stress test on your own name.