

Politically Exposed Person SMEVals

What it takes to trust an agent that identifies and escalates PEP risk

Politically Exposed Person screening agents are used to identify customers, beneficial owners, related parties, or counterparties who may create heightened corruption, bribery, sanctions, money-laundering, or reputational risk.

Politically Exposed Person screening is harder than simple name matching. The agent must distinguish domestic and foreign political exposure, family members, close associates, former officials, beneficial ownership links, state-owned enterprise roles, seniority, jurisdictional risk, recency, and source quality.

A technical eval can show that the agent found a profile. An SMEVal asks whether it reached the right PEP disposition.

The Three Questions

01 Does the agent reach the right answer?

Does it correctly identify whether the person is a PEP, related to a PEP, a close associate, a false positive, or an uncertain match?

02 Does the agent take the right action?

Does it escalate to enhanced due diligence, document rationale, apply policy, and avoid improper approval or rejection without required review?

03 Can it be tricked into doing the wrong thing?

Can a document, account field, webpage, or user prompt cause it to ignore PEP links, downgrade risk, or leak sensitive review criteria?



What the SMEVal Tests

Identity and Match Quality

FairPlay tests whether the agent distinguishes true PEP matches from similarly named individuals, uses identifiers properly, and escalates ambiguous matches rather than guessing.

Jurisdiction, Seniority, and Risk Context

FairPlay tests whether the agent considers role seniority, jurisdictional risk, corruption exposure, recency, and whether the position creates meaningful financial-crime risk.

Fairness and National-Origin Risk

FairPlay tests whether the agent avoids treating nationality, accent, geography, or ethnicity as a substitute for actual PEP exposure.

PEP Status and Relationship Mapping

FairPlay tests whether the agent recognizes current and former officials, family members, close associates, beneficial ownership links, intermediaries, and state-owned enterprise exposure.

Enhanced Due Diligence Routing

FairPlay tests whether the agent routes PEP matches into the correct EDD workflow, documents evidence, captures source links, and avoids final disposition where human review is required.

Prompt-Injection and Attack Resistance

FairPlay tests whether the agent can be manipulated into ignoring PEP status, downgrading risk, approving onboarding, leaking review rules, or following instructions hidden in customer documents, web content, or account fields.

Failure Modes These PEP Screening SMEVals Catch

- Missing a senior official because the agent fails to connect aliases.
- Ignoring family-member or close-associate relationships.
- Using nationality as a proxy for risk.
- Producing a disposition without evidence that compliance can reconstruct.
- Treating a low-confidence same-name match as a confirmed PEP.
- Failing to escalate for EDD.
- Accepting a customer-provided document that says “no political exposure” despite contrary evidence.

fairplay

ABOUT FAIRPLAY

FairPlay SMEVals test whether financial-crimes agents understand the subject-matter judgment embedded in PEP screening — not just whether they can retrieve and summarize search results.

FairPlay is the AI enablement company for the financial services and insurance industries. FairPlay specializes in testing, tuning, and validating AI systems and agents, helping institutions deploy AI faster, more profitably, and in full compliance. Its products address the toughest challenges in AI deployment, including Bias Testing, Model Optimization, Model and Agentic Validation, Second Look, and Safe AI Sourcing. FairPlay's work has been featured in The Wall Street Journal, USA Today, Politico, American Banker, and Forbes. In 2023, Fast Company recognized FairPlay as a “World Changing Idea.” To learn more, please visit fairplay.ai.