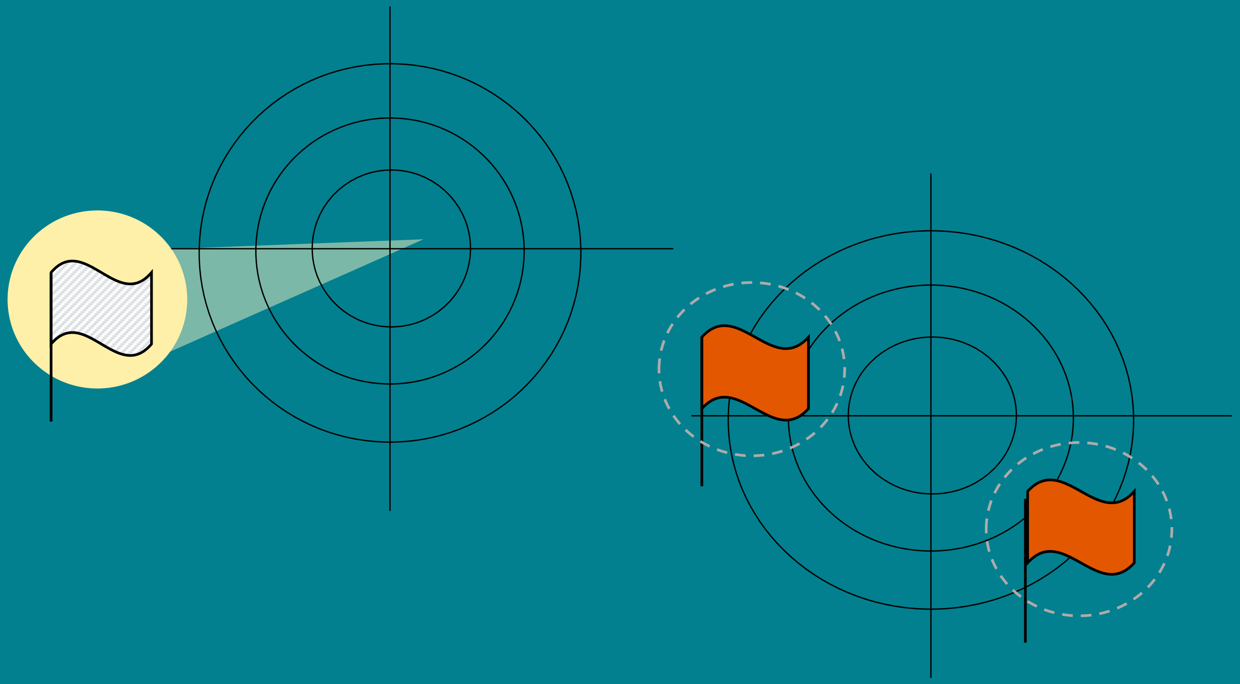
EBOOK

False Positives vs. False Negatives: Which One Actually Puts You at Greater Risk?

Understanding the difference to reframe your risk management approach



False Positives vs. False Negatives: Which one actually puts you at greater risk?

Overview

It's a common complaint of compliance leaders in the financial services industry: too many false positives—alerts flagged by legacy, rules-based, or keyword-based monitoring solutions that turn out to be irrelevant, noisy, or low risk. Industry-wide, the rate of false positives may be as high as 95%, causing firms to expend significant resources on the investigation of each alert.

But there is a quieter and potentially costlier problem that often receives far less attention: false negatives.

False negatives occur when **genuine risk goes undetected** and **no alert is triggered at all,** typically as a direct consequence of the same, legacy technology. Unlike false positives, false negatives don't appear in queues, metrics, or complaints. Instead, they surface later as regulatory failures, reputational damage, or enforcement actions—when it's already too late.

A U.S. bank was **fined $3 billion and sanctioned** for failure to "conduct adequate risk management and oversight" in its anti-money laundering (AML) program."

Source: *CNBC, October 2024*

What's inside this guide

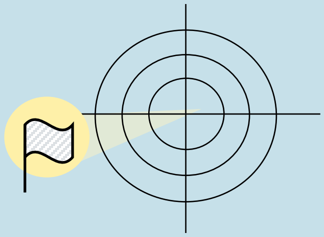
This guide explores:

- The differences between false positives and false negatives.
- Why false negatives may be the greater hidden threat.
- How leading organizations are reframing adverse media monitoring to deliver more precise outcomes.
- How AI can help reduce false positives and surface false negatives.

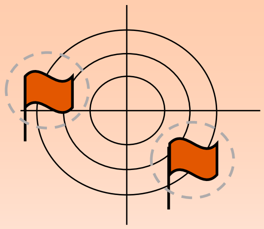
Understanding the difference between false positives and false negatives

In adverse media monitoring, two error types determine system performance:

What are they?



False positives
Alerts triggered through automated monitoring that appear risky but prove irrelevant after investigation.



False negatives
Risk events that should trigger alerts but did not.

While both errors matter, they do *not* carry equal consequences.

Side-by-side comparison

	False Positives	False Negatives
Visibility	Highly visible; appear in dashboards and queues	Invisible until damage occurs
Operational Impacts	Wasted analyst time; inefficiency	Undetected misconduct, missed criminal or reputational risk
Business Outcomes	Inefficient processes, but limited long-term damage	Regulatory findings, enforcement actions, lasting reputational harm
Costs to Organization (example)	~$100 in analyst review time	Potential millions of dollars in legal, financial, and reputational impact

The bottom line: False positives are noisy; false negatives are silent. But the silence of not knowing could be the greatest risk to your firm.

Why false negatives may be a bigger threat to your organization

1 Missed fraud during onboarding could lead to financial loss

If a regulated business (bank, fintech, insurance, investment platform) accidentally lets a prohibited customer through due to a false negative, consequences could include heavy fines, audits, remediation requirements, increased regulatory scrutiny and/or mandatory reporting. Any organization, regulated or not, also risks reputational damage from not identifying potential bad actors.

2 Bad actors may be harder to detect *after* onboarding

Once inside the onboarding “front door,” bad actors can blend in. Transactions may look normal at first and risk signals may only appear after the damage is done. One bad actor can inflict losses that are orders of magnitude higher than the operational cost of reviewing a few false positives.

3 High false negatives rates may reflect widespread deficiencies in risk and fraud monitoring

More than one instance of a false negative could suggest gaps and inefficiencies in customer and vendor monitoring processes, such as: investigation fatigue from chasing down too many initial false positives; lack of a continual monitoring process; or lack of capabilities to monitor a broader selection of information available on the web (i.e., unstructured data).

Risk management is about managing uncertainty, not eliminating it. But by focusing too much on operational efficiency and false positives—and monitoring only a few structured watchlists or sanctions sources—organizations can expose themselves to significant, unmanaged risk.

Modernizing your risk management approach: Reframing adverse media monitoring to help reduce false negatives

With the flow of data in multiple formats reaching unprecedented levels, organizations that rely primarily on traditional adverse media monitoring for customer onboarding may be falling behind.

Some leading organizations are reframing their approach to adverse media monitoring by asking a different question:

Is it worth reviewing more alerts if it meaningfully reduces the chance of missing serious false negative risks?

False positive

$100
in time and resources to investigate

False negative

Potential millions
in legal penalties, plus potential business disruption, regulatory restrictions, and reputational damage

When teams compare the marginal cost of an extra review to the potential cost of a missed risk event, the answer often becomes clear: Accepting a trade-off between investigating false positives and broadening monitoring capabilities to surface more false negatives may be a rational and necessary decision when the alternative can be regulatory fines or failure.

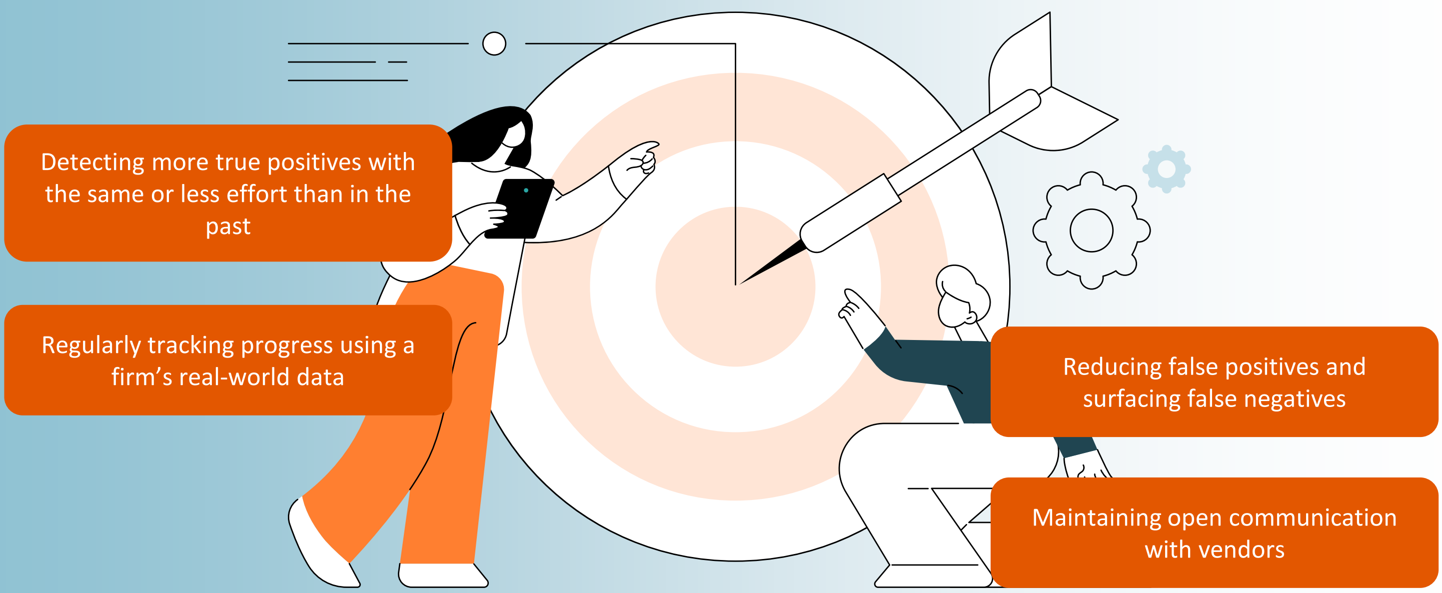
Using advanced AI and continuous monitoring to help detect more true positives and surface false negatives

With **AI and large language models (LLMs),** it's now possible to continuously monitor vast volumes of data in multiple formats quickly. The AI-based models help extract meaningful context so organizations can disposition alert cases more efficiently and reduce the window of time during which potential risk goes undetected**.**

The result: Broader, deeper coverage with less manual burden.

The benefit: Greater efficiency that many firms strive for, plus the ability to more easily surface real risks to investigate, which may have been overlooked in the past.

Adopting an advanced, LLM-reliant AI system can help navigate the delicate balance between potential risk exposure and operational effort generating measurable outcomes.



With the right AI-based adverse media monitoring system in place, compliance teams and risk managers can **make smarter decisions, act faster, and reduce exposure to hidden threats.**

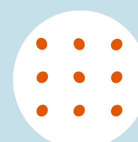
How SaifrScreen℠ helps firms balance risk

SaifrScreen enables firms to more accurately and efficiently identify potential risks in full customer and vendor populations for further investigation. The solution uses advanced AI, including machine learning (ML) technology, natural language processing (NLP), and large language models (LLMs), as well as behavioral science to understand context. Additionally, SaifrScreen crawls and indexes internet data 24/7 to provide ongoing review and monitoring that can surface more precise early warning notifications faster.

How SaifrScreen uses AI-based models to help surface false negatives:

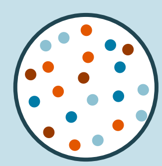
Most traditional adverse media screening and monitoring methods focus solely on structured data (e.g., sanctions, wanted lists, and watch lists), which only represent ~20% of internet data and can be slow to be updated.

Structured



SaifrScreen continuously reviews large populations against publicly available information on the web (i.e., unstructured data), such as news media, government sources, arrest and court record aggregators, and more, to help **identify more potential indications of financial or reputational risk sooner for further investigation.**

Unstructured



SaifrScreen extends its reach to unstructured data, including:

230K	190	160	23B	millions
Online sources	countries	languages	webpages	of webpages added daily

Sources subject to change

Searching this remaining ~80% of internet data can reveal valuable details and enables firms to pinpoint and further investigate potential threats such as fraud as soon as they become known.

saifr

About Saifr

Saifr redefines how compliance operates with advanced AI technology, the right data, and deep industry expertise. Built within Fidelity Investments' innovation incubator, Fidelity Labs, Saifr harnesses the power of AI agents to help address the limitations and inefficiencies within traditional compliance frameworks, helping safeguard organizations from regulatory and reputational risks. Saifr helps clients save time, reduce costs, and improve accuracy while protecting their firms. Our AI-powered risk prevention and management solutions include capabilities for marketing compliance review, adverse media monitoring, and electronic communications surveillance. Learn more at [https://saifr.ai.](https://saifr.ai)

1252691.1.0