Open Source Intelligence and AI Handbook

blackdot
solutions

Contents

Why use AI for open source intelligence?

Interest in adopting solutions that leverage artificial intelligence (AI) for financial crime compliance is at an all-time high. AI promises to overcome inefficiencies and strengthen outcomes for financial institutions in different parts of their anti-financial crime workflow: Know Your Customer (KYC) and onboarding, customer due diligence, screening, and transaction monitoring. For these reasons, industry bodies increasingly advocate adopting AI as part of a risk-based model. The value of AI is most recognised in increasing efficiency by automating tasks, eliminating or minimising low-value work, and quickly identifying relevant information for human teams to investigate.

Because resources can be better and more strategically allocated, firms can scale and grow.

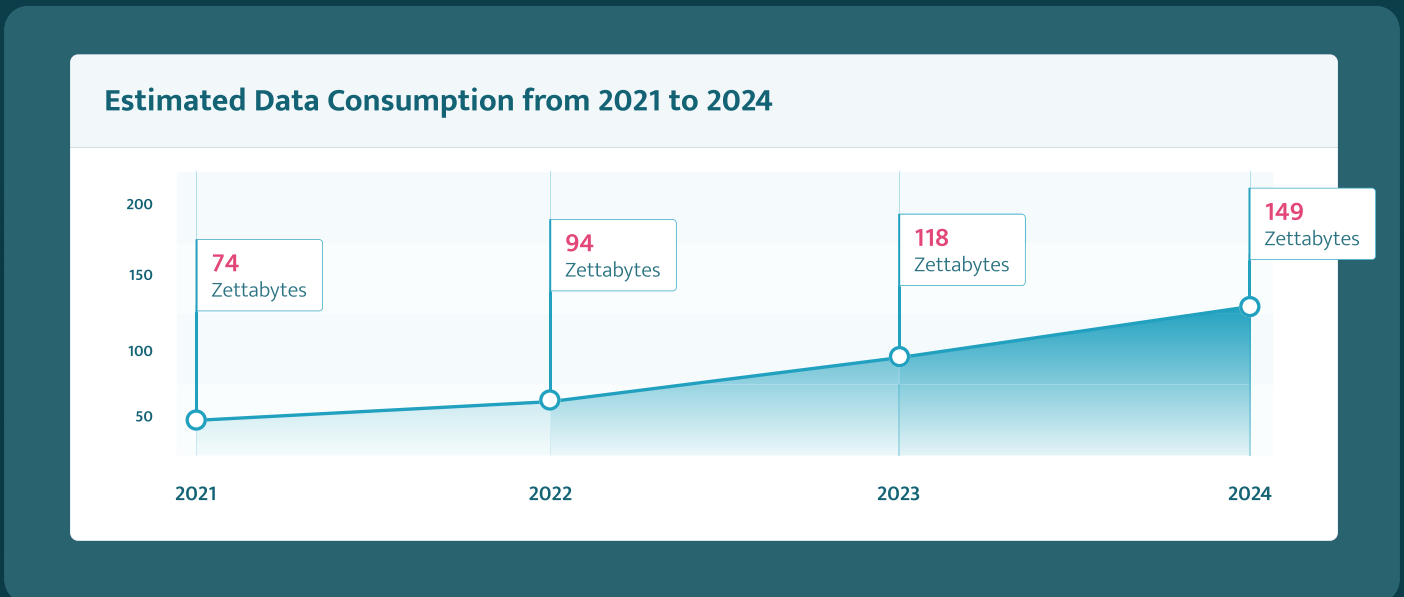
As part of a risk-based model where *knowing* your customers and the risks they pose is vital, firms should ideally leverage all available information, including online open sources, to gain an accurate view of their client's risk profile. The importance of leveraging new data sources has even been recognised by the UK government, with the development of INDEX, an information and data exchange initiative designed for the intelligence assessment community.[1]

Financial institutions cannot afford to ignore the growing amounts of unstructured online data, but assessing and managing it remains an overwhelming task without the right tools and technology. Through practical guidance, this handbook will highlight the advantages for financial institutions of using solutions that harness the power of AI in open source intelligence and investigations as part of their anti-financial crime programmes. It will also address and dispel some common misconceptions about using AI, through interviews with industry experts.

Five considerations when using AI for OSINT operations

1. Assessing a wealth of reliable sources

In screening and investigations in particular, gaining a complete picture of risk requires scanning a breadth of online open sources for information about your subject, including sources often omitted from curated databases. The real and current risk picture lies in the *live* snapshot of your subject, which includes less "official" sources like social media profiles, smaller media outlets, and online forums. However, an astounding quantity of online data is potentially available, with some estimates of around 1.134 trillion MB being created every day. How can investigators cope with assessing enormous volumes of hits?



Source: Finances Online

A critical component is understanding source reliability. When it comes to open sources, not all information is equally weighted. Within the open source intelligence world are many high-quality and reliable sources that investigators and researchers must consider first. For example, in certain jurisdictions like the United States, official records like court judgments and campaign contribution information are readily available through government websites. In some states, marital records, criminal records, and bankruptcy records can also be accessed. These are official sources managed by a government body and openly accessible to the public.

In addition to government records, quality media publications are also generally trustworthy sources of information. Journalists who report for respectable publications must abide by the ethical rules of journalism and meet high standards before publishing. Well-known titles that enjoy a favourable reputation due to a longstanding history of fact-based reporting, like the *New York Times*, can be considered reliable open sources. Well-researched stories can point to emerging scandals, violations, or reputational issues. For example, the *Financial Times* helped expose the now-insolvent payment processor Wirecard for fraud well before the German regulator took notice.[2] Quality and reliable investigative journalism can appear on other websites as well as those of the publications themselves, from Twitter posts to high-standard, well-sourced leaks such as those from the International Consortium of Investigative Journalists.

> Financial institutions are constrained by privacy restrictions in all the individual jurisdictions we work in. We work in partnership with law enforcement who are also restrained by borders in many of the actions and activities they can undertake. Journalism doesn't always suffer from these same restrictions. Generally, taking information across borders is not a challenge for investigative journalists as it is for the rest of us. We need to recognise that and recognise there is an important role investigative journalists play in the whole ecosystem of investigations.
>
> **Nick Lewis** | Managing Director, High Risk Client Unit - FCC, Standard Chartered Bank

Clearly, reliable forms of open source intelligence should have priority in an investigation, setting the tone and guiding the research process. To complement these strong, credible sources, some less reliable open sources, like social media profiles, may help corroborate suspicions or form an overall assessment.

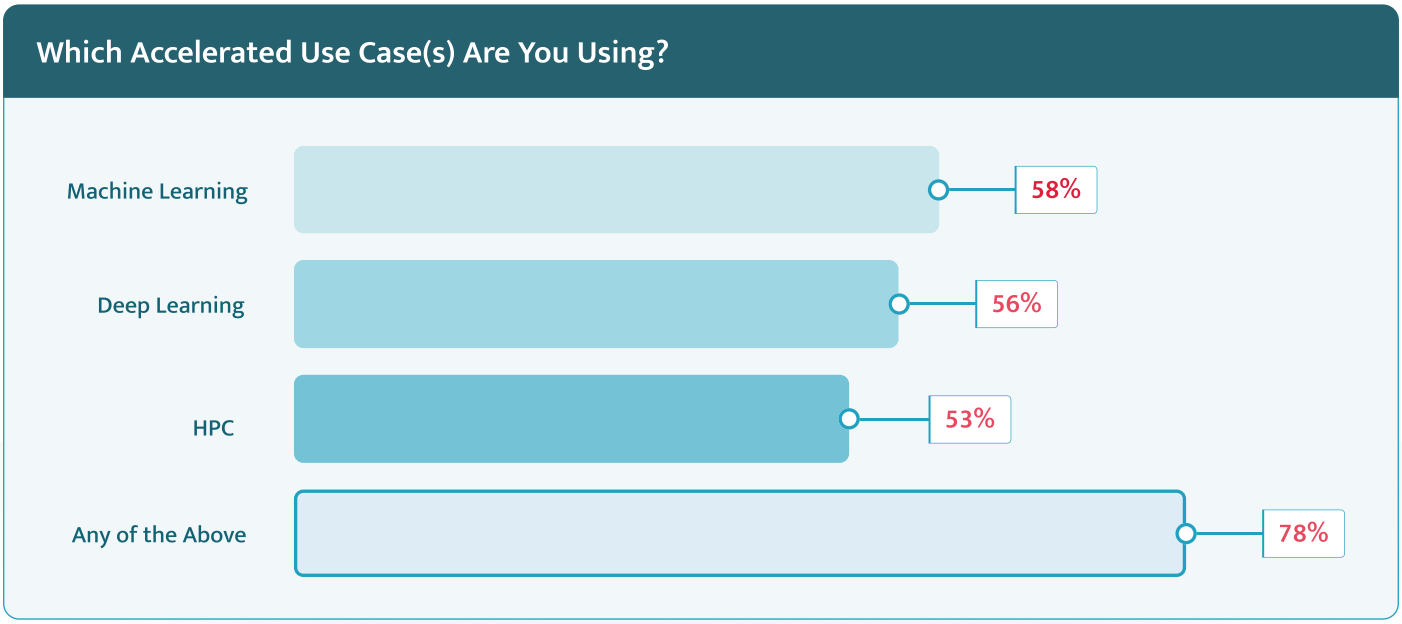
For example, a customer who has declared a certain salary figure on KYC files which contrasts with a lavish lifestyle touted on their social media profile may raise suspicion and lead to further investigation. Another example, as seen with the high-profile short squeeze of GameStop, is the potential to use online forums such as Reddit to provide early warning signs of compliance issues. While certainly a less reliable source of information, a financial institution conducting a background check on a company and its officers might find information in online forums useful, such as that pointing to abusive or manipulative trading activity.[3]

Another important factor for financial institutions to consider is that, unlike prosecutors, regulated entities are obliged to act on and report *suspicions* rather than prove things beyond a reasonable doubt. A range of open sources can give you enough information to form evidence-based judgments, whether that involves combining internal sources with government records, social media posts, or online publications.

Regulators expect regulated institutions to take immediate action when holding suspicions, not wait for indisputable proof. AI-powered tools can help retrieve relevant information from a range of online sources, present findings in an understandable and organised manner, and give the analyst a clear investigative path forward. By looking for information shared across several sources and automatically differentiating between reliable and less reliable, the right tool will speed up research processes and improve investigations.

Ideally, investigators should be able to immediately evaluate source credibility to conduct more efficient research *without* omitting sources that may be less reliable but potentially valuable. AI tools which are trained by its users to clearly differentiate between sources in terms of reliability help investigators perform due diligence through open sources more efficiently and easily. These features help give investigators a clearly marked pathway for their investigation, allowing them to prioritise their research, base conclusions on the most credible sources available, and leverage less-reliable sources to form or back up suspicions or guide them on where to look next.

In an ideal world, financial institutions would conduct thorough research on all of their customers to identify all potential areas of risk. In reality, this is clearly not possible - and is not expected by the regulators, who instead promote the risk-based approach of focusing resources on higher-risk customers. Tools which allow investigators to quickly and easily see which customers *are* higher risk, by reviewing multiple open sources and assigning a clear risk score, are game-changing in achieving this balance. They allow firms to review a greater portion of their customer base while allowing analysts to quickly triage their findings and hone in on relevant cases for further review.



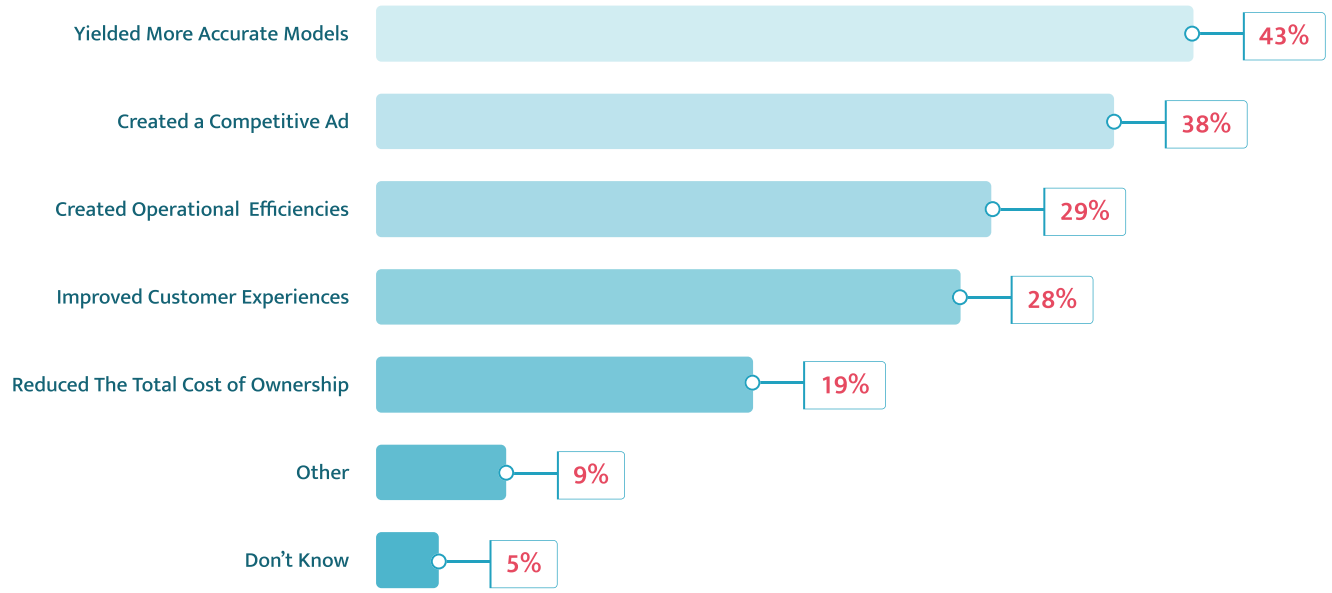
Source: State of AI in Financial Services, Nvidia

2. Cost effective

One of AI's most widely recognised benefits is its ability to increase operational efficiencies. When analysts conduct open source research, they are often met with an overwhelming amount of hits, giving them the time-consuming and onerous task of sifting through results that may not even apply to their subject. Even results that *do* relate to their subject may not be relevant to the context, such as negative news findings that don't entail financial crime or pose a reputational risk. In this context, the greatest asset to investigations is also its greatest challenge: large amounts of data.

For some financial institutions, the burden of *too much* information can lead to the elimination of open source intelligence in investigations entirely, making firms vulnerable to illicit actors. For those financial institutions considering using open source intelligence but rightly concerned with potential inefficiencies and extra costs due to high data volumes, technologies like user-driven AI and automation provide a viable solution.

What Benefits Are You Seeing From Your AI Investment?



Source: State of AI in Financial Services, Nvidia

In a world of ever-growing online data, AI can be given contextual information to identify true hits through entity resolution technology, which checks for other identifiers and data points like age, gender, or nationality to confirm identity.

Instead of an analyst manually confirming that a finding relates to their subject, tools can do it automatically and save time. Technology allows institutions to scale this process, systematically filtering through results in a way that humans cannot.

Additionally, a branch of AI known as natural language processing can assess if findings denote actual risk, instead of retrieving non-pertinent information about a subject. For example, if a subject is quoted in a media article criticising the levels of corruption in a particular sector, this does not represent a potential risk - it is vital to understand the connection between the subject and the word "corruption".

Finally, good AI tools categorise relevant hits in an optimal way that is easy for the researcher to use, allowing analysts to focus on actually *analysing* the findings in a methodological and orderly manner rather than determining their relevance. These factors combined allow for faster investigations, saving financial institutions both time and money.

By freeing up people to focus on more cognitive tasks rather than sifting through potential hits, AI can eliminate low-value work and achieve a reduction in unnecessary staffing costs. This increased efficiency allows financial institutions to scale and clear backlogs effectively without hiring expensive human support.

Clearing screening backlogs has multiple advantages. In addition to the obvious benefit of locating bad actors or suspicious activity, firms that clear backlogs can onboard more clients safely - taking on new business and boosting their revenue without contributing to the accumulation of unprocessed tasks and taking unwanted financial crime risk. In a similar vein, once a backlog is cleared AI-enabled tools allow firms to scale and take on new business *without* necessarily hiring more resources to manage the increased volumes, making it a growth enabler rather than a cost.

> Technology tools can demonstrate to the regulator, the customer, and banking partners that a firm has the ability to control their financial crime risk better than their competitors. Employing these tools help a financial institution show that they can go into a riskier market and offer more products. In this way the compliance function enables the business to expand. The compliance department can show that they have the tools, processes, and intelligence that can handle the risk better — making a strong business expansion case.
>
> **Haibo Zhang** | Data analytics expert and former Global Head of Analytics for Financial Crime and Threat Mitigation at HSBC

Ultimately, all firms must consider their return on investment (RoI). For smaller firms with smaller compliance teams and budgets, finding ways to protect RoI is arguably even more crucial. In these cases, using AI for open source intelligence and investigations is advantageous because firms can carry out their anti-financial crime programme *without* hiring new headcount or straining existing teams, by eliminating the mundane task of sorting through false alerts. Some firms may hesitate to adopt AI solutions because of the costs of implementing new systems. Replacing existing systems and processes wholesale is onerous and expensive, and in such a scenario the budget previously invested into a now obsolete system is squandered, and staff must train on a new tool, which drains resources. However, an ideal AI-powered open source intelligence tool works *with* existing systems, augmenting and optimising a firm's capabilities without creating unnecessary operational burdens. Such tools allow for smooth implementation and the ability to reap immediate benefits.

Features to look for in an AI tool to ensure a good RoI:

Ability to triage and prioritise alerts

Varied and suitable delivery mechanisms

(e.g. SaaS platform, cloud hosted, on premise, etc.)

Ability to sort and rank sources for reliability

Ability to differentiate and categorise between types of risk

(e.g. financial crime risk, serious and organised crime, political, extremism, offshore leaks, etc.)

Augment existing systems, rather than 'ripping and replacing'

Ability to integrate into case management systems

Fully explainable, configurable, and easy to use

3. Supporting (not replacing) human teams

In any economic climate, attracting and retaining talent is invaluable for a firm's long-term success. While AI tools can help protect a financial institution's bottom line from an efficiency standpoint, they can also help analysts carry out their work more effectively and rewardingly. Open source investigations as part of anti-financial crime compliance can be exciting, meaningful, and mentally stimulating work. Ideally, professionals should be able to use their research and analytical skills to triangulate relevant data sources, interpret their findings, and detect illicit actors. However, without the appropriate technology and tools, employees carrying out open source research can be inundated with irrelevant information, quickly making their work mundane and repetitive. Incessantly sifting through irrelevant hits is a low-value task and can lower morale over time. Not only is this repetitive work boring, it can also lead to 'false positive fatigue', where analysts feel burned out from mindlessly running through checklists. 'False positive fatigue' will ultimately lead to unsatisfied employees *and* higher turnover, and can also lead to a lack of mental alertness, meaning there is a higher likelihood of missing something of interest.

Another burden experienced by analysts is performing manual searches over multiple platforms. This is clunky, time-consuming, and can break an analyst's concentration when conducting an investigation. Instead, an optimised system has a clear interface and intuitive user experience, which allows employees to perform their job smoothly. Employing AI to speed up, optimise, and streamline tasks allows staff to engage with *real* findings related to financial crime activity, honing their skills and developing their subject matter expertise. As a result their jobs are more fulfilling, allow for career development and growth, leading to a healthy company culture. These factors translate to better staff retention, which positively impacts a firm's profitability. Since machines are excellent at

performing repetitive tasks consistently without becoming fatigued and making mistakes in the way humans do, they are best put on tasks that can be automated and codified for more consistent and accurate outcomes.

A key to unlocking these advantages is delegating processes and tasks in a way that plays to the strength of both AI and human capabilities. The human brain is good at contextualising, critical thinking, and drawing from experience to make decisions. Conversely, AI is good at identifying patterns, automating routine tasks, and analysing connections over large swathes of data. Accounting for this, powerful AI tools have been developed to showcase visualisations or categorise data in more comprehensible ways, allowing analysts to perform their work faster.

> Increasingly, financial crime investigations are about identifying networks. When we understand the network, we can understand the pollution an exited client has left behind: how many other clients were complicit or engaged in this activity? To do that, we need to gather a range of information and visualise it in a network chart.
>
> **Nick Lewis** | Managing Director, High Risk Client Unit - FCC, Standard Chartered Bank

4. Increasing regulatory expectations

The adoption of new technologies, including information processing abilities, is increasingly encouraged by industry bodies. In a 2021 report, the Financial Action Task Force (FATF), the global money laundering and terrorist financing watchdog, highlighted the ability of big data and new technologies to make anti-money laundering and counter-terrorism financing faster, cheaper, and more effective. Specifically, the report highlighted technological capabilities like natural language processing and fuzzy matching tools as more efficient at reducing false positives in adverse media screening.[4] Similar technology-positive sentiments were echoed by the European Banking Federation in a press release that warned against conventional methods and stated the future of anti-money laundering lies in using innovative technologies which add to human experts' judgement.[5]

More recently, the Wolfsberg Group has explicitly supported the use of advanced technologies for financial crime compliance programmes.[6] In negative news screening in particular, the Group noted the limitations of internet search engines like Google, even when using Boolean operators. Because these manual searches rely on algorithm-driven mechanisms that "filter out content to match individual browsing preferences", the results can be inconsistent and "vary greatly across different users".[7] Instead, search engines should be used as one part of an open source intelligence investigation, coupled with more sophisticated methodologies.

Data analytics expert Haibo Zhang posits that the movement away from Boolean operator searches has already happened - noting ChatGPT, essentially, already performs some of these same searches - and that increasingly sophisticated research engines stand to be the next trend. These newer tools "don't simply provide the results of a search, but also the context" and enable humans to draw their own conclusions. According to Zhang, they are set to "become more widespread and common practice" in the anti-financial crime compliance industry.

Regulators are placing a growing emphasis on effectiveness rather than tick-box compliance, with traditional tools becoming less and less effective in the face of complex criminal activities. Looking forward, firms will likely need to prove their systems are not only compliant but also effective. With the growing amount of online data available, using simple tools to conduct due diligence and investigations is no longer efficient or practical. Instead, financial institutions will be increasingly expected to harness the power of data fusion, or the combining of open source intelligence and internal data, to conduct investigations and assess risk. Nick Lewis, Managing Director of the High Risk Client Unit at Standard Chartered Bank, highlights the importance of data fusion given that politically exposed persons (PEPs) and kleptocrats:

> Often use family and extended family networks to hide their ill gotten gains, a network that is often not visible in CDD data. For example, if a PEP uses a family member as a lawyer this is often not apparent until all information is fused together. Then you can begin to see the connections and the network, determine that the network poses a risk, and take action against the entire network.

5. Explainable tools

While regulators are increasingly expecting financial institutions to use sophisticated technology and to be more effective in their anti-financial crime efforts, they are also clear that such technology must also be understandable. Just as all processes and controls within an anti-financial crime programme must be explainable, well-documented, and justifiable, so must any new technology. As the FATF states, "Regulated entities must be able to explain, and remain responsible for, the principles and technical details of the innovative solutions before deploying these new technologies."[8] Financial institutions must be prepared to explain tools to regulators, banking partners, and compliance staff members to different degrees of depth.

Many firms hesitate to adopt AI because of its prevailing stereotype as an opaque, confusing, and unexplainable black box, with complex algorithms that only a data scientist can interpret. However, AI solutions *can* be explainable by design, easy to understand and justify. As noted by data analytics expert Haibo Zhang, one vital component of a good AI tool is providing long lasting capability that is rooted in a clear methodology - "tools that sell you a pipeline with a methodology that clearly fits your MRM (model risk management) and operational framework are future-proofing your firm." Good AI tools allow you to customise and tune your system in a way that is understandable. Another vital element for explainability and optimisation is a tool that allows for collaboration between team members. Allowing users to work in real-time with their colleagues gives space for collaboration between different skills and perspectives - ultimately contributing to a stronger and more thorough investigation and result. This real time and immediate collaboration is increasingly important as criminal activity becomes more complex.

For firms looking to adopt a third party AI tool for open source intelligence investigations, a good tool can help them scale safety without needing to hire their own data science team. A good AI provider will also support a firm with integration and calibration rather than offering an off-the-shelf product without ongoing expert support.

Given the context of increasing AI use, it's only natural that supervisory bodies will require more regulation around its application. Recent moves for regulation, like the US's Blueprint for an AI Bill of Rights[9] and the Wolfsberg Group principles for responsible AI and machine learning[10] discuss the ethics surrounding AI implementation. Ultimately, these requirements will advance the development of AI tools, ensuring they become stronger and more transparent.

> When an investigator approaches leaks or open-sources searches, they must not think too limited. Consider them as a starting point and not an end point, because you can grow your intelligence out of it enormously. Technology tools are a fantastic medium to enrich data, coupled with an investigative mindset.
>
> **Graham Barrow** | Director, Dark Money Files

What's next?

If you're ready to identify risk faster and more effectively, here's some practical advice to help you select and implement an AI-driven solution:

1. Decide what external data is required and where

It's important to be clear about which part(s) of your team's workflows OSINT might be useful in, so that you can build the most effective process possible. It's common to find that different kinds of external data may fit into different parts of the process: many teams use a limited range of external data sources for screening checks but employ additional sources, such as social media and corporate records data, when a case becomes more complex. A good solution will be able to adapt to meet your team's needs.

2. Consider integration requirements

Will any potential solution need to integrate with an existing case management system? How will you fuse internal and external data? It's vital to be clear on your integration requirements from the start so that you can assess whether a vendor will meet your needs.

3. Find providers who offer flexible and explainable solutions

Following on from the above, identifying a provider with the flexibility to meet your needs is essential. A good technology vendor will have integrations with a range of data sources, as well as enable integration with your preferred providers and internal data. In the case of AI, it's essential that any vendor can explain how a decision has been reached and that any risk scoring models can be customised according to your risk appetite and terms.

4. Identify potential cost savings and understand time to value

Technology that works with external data is less time-consuming to implement than solutions for internal data, which may require extensive data normalisation. Look for vendors who will work with you to ensure fast time to value and overall process optimisation.

Book a free consultation

AI-driven open-source intelligence solutions like Blackdot's Videris help investigators streamline their research and investigation in a single interface. It allows investigators to use OSINT effectively in every investigation, from screening checks to complex financial crime investigations.

With AI capabilities that screen high data volumes and rank sources for relevance, built in visualisation capabilities and advanced analytics, investigators can speed up their processes and improve investigation outcomes.

Book a demo today

References

[1] Public

[2] The Financial Times

[3] Reuters

[4] FATF: Opportunities and Challenges of New Technologies for AML/CTF

[5] EBF: Demystifying AI for AML

[6] Wolfsberg Group: Wolfsberg Principles for Using Artificial Intelligence

[7] Wolfsberg Group: Negative News Screening FAQs

[8] FATF: Opportunities and Challenges of New Technologies for AML/CTF

[9] The White House

[10] Wolfsberg Group: Wolfsberg Principles for Using Artificial Intelligence

Diagram Sources

Finances Online

State of AI in Financial Services, Nvidia