



NOCTULA RISK INTELLIGENCE

RESEARCH & INTELLIGENCE WEEKLY REPORT

The Structural Deficits in Financial Crime Investigations

The Cost of Fragmented Data and Inconsistent Risk Thresholds

Doğancan İpek

October 2026

KEY TAKEAWAYS

- **Spending is large; outcomes are hard to measure.** Financial institutions spent an estimated \$206.1 billion on financial crime compliance (LexisNexis, 2023), while Europol estimated that 1.1% of criminal profits in the EU were confiscated in 2010-2014. These measure different things, but together they point to an effectiveness gap.
- **Volume is not value.** At 19 large US banks, about 16 million alerts in 2017 produced roughly 640,000 SARs, and a median of 4% of SARs drew law-enforcement follow-up (BPI, 2018). BPI treats that follow-up figure as a likely under-count.
- **Two structural deficits.** Fragmented data hides multi-hop networks, and there is no shared quantitative reference for how strong a relationship must be to count as risk.
- **Relational Intelligence is a promising remedy with early evidence.** Entity resolution, relationship graphs, calibrated and explainable scoring and outcome feedback address both deficits, but bank-level results are mostly vendor-reported, academic benchmarks rely largely on public cryptocurrency data, and the approach carries governance, privacy and data-quality risks.
- **Next step.** Measure outcomes rather than volumes, resolve entities first, and pilot network analytics on one bounded typology against existing rules.

SUMMARY

Anti-money laundering (AML) controls consume very large resources, while measurable recovery of criminal proceeds remains low. This report is a narrative evidence review, not an original empirical study. Drawing on regulator, standard-setter, industry-association and vendor sources available through January 2026, it examines the pipeline from transaction monitoring through alert investigation to suspicious activity reporting, and identifies two structural deficits. The first is a data-mapping deficit: customer and transaction data held in separate systems obscures multi-hop networks. The second is a calibration deficit: there is no shared quantitative reference for how strong a relationship must be to count as risk, so thresholds and alert volumes vary widely between institutions. We define Relational Intelligence as an approach that combines entity resolution, relationship graphs, calibrated and explainable scoring, and outcome feedback; we review the evidence for it and its limits; and we outline practical steps for institutions and for the sector.

Keywords: Anti-Money Laundering, Graph Analytics, Entity Resolution, False Positives, Relational Intelligence

Disclosure. Noctula Risk develops AI-powered financial crime prevention technology for banks and FinTechs, and the recommendations in this report align with that business. Sources are cited so readers can verify the evidence independently, and figures that originate with vendors are identified as such.

1. INTRODUCTION: THE EFFECTIVENESS GAP

The United Nations Office on Drugs and Crime (UNODC) estimates that the amount of money laundered globally in a year is between 2 and 5 percent of global GDP, or roughly \$800 billion to \$2 trillion (UNODC, 2024). Such figures are uncertain, and Europol cautions that many published estimates of criminal profits are unreliable (Europol, 2016). They are best read as an order of magnitude rather than a measurement.

Recovery appears small by comparison. Europol estimated that between 2010 and 2014, 2.2 percent of estimated criminal proceeds in the EU were provisionally seized or frozen and 1.1 percent were finally confiscated (Europol, 2016). That figure covers asset recovery across organised crime generally, not the performance of banks' detection systems, and it is now a decade old, but it remains the most cited EU benchmark.

The cost side is large and persistent. A LexisNexis study estimated global financial crime compliance costs at \$206.1 billion (LexisNexis, 2023), a vendor-run survey that covers sanctions and KYC controls as well

as AML. Later LexisNexis regional studies reported \$61 billion for the US and Canada and \$85 billion for EMEA (LexisNexis, 2024a; 2024b); we found no newer global total. Penalties are volatile: Fenergo, itself a compliance-technology vendor, tallied AML, KYC, sanctions and due-diligence penalties of \$6.6 billion in 2023, \$4.6 billion in 2024 and \$3.8 billion in 2025, with the 2023 total including a single \$4.3 billion penalty on a crypto exchange (Fenergo, 2024; 2026).

These figures differ in geography, year and unit, so we do not divide one by another (Exhibit 1). Together they support a narrower claim: resources are large, measurable yield is limited, and the industry itself is questioning the trade-off. In 2024 the Wolfsberg Group, whose members are major international banks, stated that the value derived from ever-rising volumes of suspicious activity reports is not proportionate to effective outcomes (Wolfsberg Group, 2024).

Exhibit 1. Key figures, with period, geography and caveats

Measure	Figure	Period / geography	Source	Caveat
Money laundered globally (estimate)	2-5% of global GDP (\$800 billion to \$2 trillion)	Annual, global	UNODC (2024)	Estimate; methodology is dated and contested
Criminal profits confiscated	1.1% (2.2% frozen or seized)	2010-2014, EU	Europol (2016)	All organised crime, not bank detection; Europol calls its figures conservative
Financial crime compliance cost	\$206.1 billion	2023 study, global	LexisNexis (2023)	Vendor-run survey; covers sanctions, KYC and related controls. Latest global total found
AML, KYC, sanctions and CDD penalties	\$6.6bn, \$4.6bn, \$3.8bn	2023, 2024, 2025, global	Fenergo (2024; 2026); FT (2023)	Vendor tally; a few large cases dominate, and totals for 2022 differ between reports (about \$5bn vs \$4.2bn)
SARs filed	4.7 million (about 12,870 a day)	FY2024, US	FinCEN (2025)	All SAR types, including fraud
SARs drawing law-enforcement follow-up	Median 4%	2017, 19 large US banks	BPI (2018)	Proxy that BPI considers likely under-inclusive
Alerts converting to SARs	About 4% (16 million alerts; 640,000+ SARs)	2017, 19 large US banks	BPI (2018); author's calculation	A SAR is a report, not confirmed crime

Sources as listed; full references at the end of the report. Figures are not directly comparable and are not combined in this report.

2. SCOPE, METHOD AND LIMITATIONS OF THE EVIDENCE

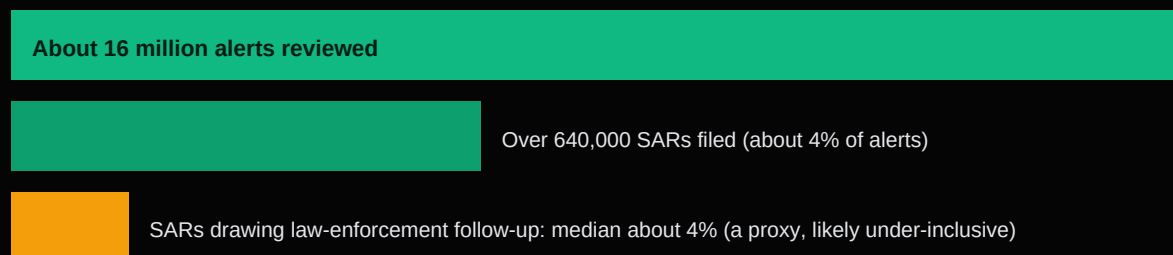
This report is a narrative review, not a systematic review, and it reports no original data. Sources were selected for relevance to monitoring, alert handling and relationship-based analytics, and drawn from international bodies (UNODC, FATF, the Wolfsberg Group), law-enforcement and financial-intelligence publications (Europol, FinCEN), industry research (the Bank Policy Institute), vendor-published studies (LexisNexis, Fenergo, Google Cloud) and selected academic work. Online sources were retrieved in October 2026.

Three limits apply. First, many cost and false-positive figures originate with vendors that have commercial interests, and we flag them. Second, figures differ in geography, year and definition. Third, there is no agreed metric of AML effectiveness: BPI notes that no established measure exists of whether reports are useful to law enforcement, so its study used law-enforcement follow-up as a proxy that is probably under-inclusive, because authorities may use filings without telling the bank (BPI, 2018).

3. EVIDENCE: HIGH VOLUME, LOW MEASURABLE YIELD

BPI's survey of 19 large US banks gives the most detailed public view of the pipeline (Exhibit 2). In 2017 the banks reviewed about 16 million alerts and filed over 640,000 SARs, an alert-to-SAR ratio of about 4 percent by our calculation, and a median of 4 percent of SARs drew follow-up inquiries from law enforcement (BPI, 2018). Nearly half of AML personnel at those banks were engaged in administrative tasks, such as data and process quality assurance, rather than work directly tied to reporting. For scale, US institutions filed 4.7 million SARs of all types in fiscal year 2024, about 12,870 a day (FinCEN, 2025); the BPI figures cover AML SARs at 19 banks only, so the two are not directly comparable.

Exhibit 2. From alerts to law-enforcement follow-up



Schematic: bar widths are illustrative, not to scale. Source: BPI (2018), 19 large US banks, 2017. The alert-to-SAR ratio is the author's calculation. Follow-up is a proxy that BPI considers likely under-inclusive.

This does not mean reporting is without value. The FBI linked between 16 and 40 percent of its investigations in fiscal year 2024 to SARs or currency transaction reports, depending on the crime type (ABA Banking Journal, 2025). The more careful conclusion is that the value per report is hard to measure and unevenly distributed. Industry commentary also describes a defensive-filing dynamic, in which institutions file broadly to limit regulatory risk, but we are not aware of a direct measurement of its extent.

False-positive rates are usually quoted from vendor sources. Google Cloud, for example, states that more than 95 percent of system-generated alerts prove to be false positives at first review (Google Cloud, 2023). The BPI-derived ratio above points the same way, since roughly 96 percent of alerts did not become SARs, but neither figure is audited regulator data. Pol (2020) goes further, arguing that the AML regime has not demonstrated policy effectiveness and estimating that it affects less than 0.1 percent of criminal finances; that is a critical position, and a contested one.

4. DEFICIT ONE: THE DATA-MAPPING GAP

Customer information typically sits across retail systems, KYC repositories, CRM tools and payment platforms. BPI found that banks used, in some cases, more than 20 different IT systems for BSA/AML compliance (BPI, 2018). The Wolfsberg Group likewise notes that customer behaviour and attributes, combined with transactions, give broader insight into suspicious activity than transaction monitoring alone (Wolfsberg Group, 2024). When systems do not share a common view of who is who, relationships between accounts go unseen.

Two standard typologies illustrate the problem. **Structuring**, or smurfing, splits a large sum into many small transfers across accounts; each account looks ordinary, and only linking the accounts to a common controller or beneficiary reveals the pattern. In BPI's survey, 18 percent of alerts at banks that recorded alerts by activity type related to structuring (BPI, 2018). **Money-mule hubs** receive funds from many unrelated senders and pass them on quickly; the signal lies in the pattern of connections, not in any single row (Exhibit 3, an illustrative example, not real data).

Exhibit 3. The same activity in a tabular and a network view

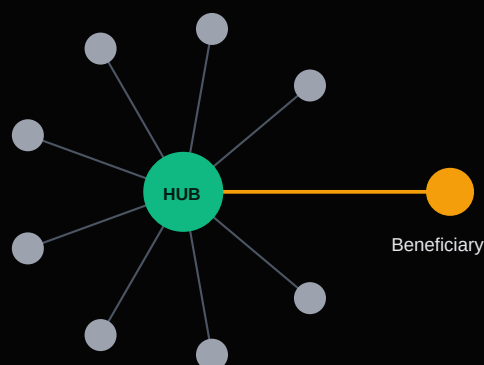
What a per-account rule sees

Account	Credits, 24h	Alert
A-01	Small, below threshold	None
A-02	Small, below threshold	None
A-03	Small, below threshold	None
A-04	Small, below threshold	None
A-05	Small, below threshold	None

Each account looks ordinary on its own.

Illustrative only; not real data.

What the same data looks like as a network



Eight unrelated senders, one hub account, one beneficiary.

Seeing such patterns depends on two capabilities. The first is **entity resolution**, which links records that refer to the same real-world person or business across systems; without it, a network is built on duplicated or mismatched nodes. The second is storing and querying relationships as first-class data. Relational databases can represent graphs, and the SQL:2023 standard added property graph queries (ISO/IEC, 2023), while native graph engines are designed to traverse relationships outward from a starting entity. The practical question is therefore less "SQL versus graph" than whether an institution resolves entities and can run multi-hop queries at investigation time. Published performance comparisons are often vendor-produced and workload-specific, so institutions should benchmark on their own data.

5. DEFICIT TWO: NO SHARED REFERENCE FOR RELATIONAL RISK

FATF Recommendation 1 requires a risk-based approach, under which institutions calibrate controls to their own risks (FATF, updated periodically). That is deliberate and appropriate, because risk appetite and customer bases differ. The gap lies within it: there is no commonly accepted quantitative reference for how strong a link, such as a shared device, address, counterparty or beneficial owner, must be to indicate elevated risk, or how to weigh several weak links together.

Consider a shared device identifier. One institution may alert on any match, another only above a value threshold. Neither is wrong in itself, but their alert volumes and SAR yields cannot be compared or benchmarked, and thresholds tend to be set by judgement and adjusted to analyst capacity rather than to measured precision. It is plausible that such differences are exploitable by sophisticated actors, but we are aware of limited empirical evidence on how often this happens, so we treat it as a hypothesis.

A shared reference would complement the risk-based approach, not replace it. It could include common definitions of relationship types and evidence strength; confidence scores calibrated against outcome data (alert disposition, SAR decisions and law-enforcement feedback); documented rationale and back-testing for thresholds; and sector-level benchmark metrics. The Wolfsberg Group's 2025 statement points in this direction, urging that new monitoring approaches be introduced through a responsible framework with proper validation (Wolfsberg Group, 2025).

6. RELATIONAL INTELLIGENCE: A WORKING DEFINITION

We use Relational Intelligence to mean the combination of four components:

1. **Entity resolution.** Records are unified into one profile per real-world entity. Deterministic rules apply where verified identifiers exist, such as national ID or company registration numbers; probabilistic matching with an explicit confidence value applies where they do not. Every match is auditable and

reversible.

2. **Relationship graph.** Customers, accounts, devices, addresses, counterparties and beneficial owners are nodes, and relationships are typed, time-stamped edges, supporting multi-hop queries.
3. **Calibrated, explainable scoring.** Scores estimate likelihood and are calibrated against outcomes, and each alert shows which relationships and evidence drove it. "Deterministic" here means reproducible: the same inputs give the same output under documented logic. It does not mean the risk is certain. A score is an estimate, not proof.
4. **Outcome feedback.** Investigator dispositions, SAR decisions and any law-enforcement feedback flow back to recalibrate models, under validation and governance.

7. WHAT THE EVIDENCE SUPPORTS, AND WHAT IT DOES NOT

What it supports. HSBC, working with Google Cloud, reported that its machine-learning monitoring system identified two to four times as much suspicious activity while cutting alert volumes by about 60 percent (Google Cloud, 2023). This is a vendor-and-bank-reported result for ML risk scoring that includes network behaviour, not a controlled study and not specific to graph databases. In the academic literature, Colladon and Remondi (2017) analysed more than 33,000 operations at an Italian factoring company and found that network metrics could predict client risk profiles. Weber et al. (2019) released a labelled graph of Bitcoin transactions and found that a random forest outperformed graph convolutional networks, while inviting work to combine the two approaches. A 2024 preprint reviewing 97 papers and benchmarking methods on the same dataset found that network analytics added predictive power, with graph neural networks performing best, but noted that most advanced network methods have not reached mainstream AML adoption, that data availability is the main obstacle, and that the best-performing models are hard to interpret (Deprez et al., 2024). Cryptocurrency data differs from bank data, so the transfer is not automatic.

What it does not support. We are not aware of independent, peer-reviewed, production-scale evidence that graph-based methods reduce false positives by a specific amount at a bank. Several risks also need management:

- **Governance.** Explainability, validation and model-risk oversight are prerequisites; FATF (2021) discusses both the opportunities and the challenges of new technologies, and the Wolfsberg Group (2025) stresses validation.
- **Entity-resolution errors.** A wrong merge creates false links; a wrong split hides real networks.
- **Privacy and legal limits.** Data use and sharing are constrained by laws such as the EU GDPR (Regulation 2016/679) and Türkiye's Law No. 6698 on the Protection of Personal Data, and by rules against tipping off.
- **Legitimate hubs.** Landlords, marketplaces and payroll providers look like hubs; network methods can generate their own false positives.
- **A single-institution view.** One bank's graph sees only its own slice of a network that may span many institutions.
- **Cost and migration.** Replacing or integrating legacy systems takes time and investment.

8. RECOMMENDATIONS

For institutions

1. **Measure outcomes, not volume.** Track alert-to-SAR conversion, SAR usefulness feedback, handling time per alert and the share of effort spent on administrative tasks.
2. **Fix entity resolution first.** Document match rules, confidence thresholds and measured error rates.

3. **Pilot on one bounded typology.** For example, mule detection, run as a champion-challenger test against existing rules with success metrics set in advance.
4. **Calibrate relationship thresholds against outcomes.** Record the rationale and back-test periodically.
5. **Build governance in from the start.** Include explainability, independent validation, model-risk oversight and privacy review.

For the sector

6. **Develop shared definitions** of relationship types and evidence strength, and publish benchmark metrics that institutions can compare against.
7. **Strengthen feedback from authorities to institutions.** The Wolfsberg Group has argued that such feedback is the best way to judge whether reports are useful (Wolfsberg Group, 2024).

9. CONCLUSION

Compliance spending is large and persistent, while measured recovery of criminal proceeds remains low: Europol estimated that 1.1 percent of criminal profits were confiscated in the EU in 2010-2014. These are different measures and should not be reduced to a single ratio, but together they indicate an effectiveness gap that better data and better-calibrated detection could narrow.

We have argued that two structural deficits sit behind part of that gap: fragmented data that hides networks, and the absence of a shared quantitative reference for relational risk. Relational Intelligence, as defined here, is a promising response, and its goal is not mathematical certainty but risk assessment that is consistent, explainable, calibrated against outcomes and open to audit. The evidence so far is early: bank-level results are largely vendor-reported, and academic benchmarks rely mostly on public cryptocurrency data. The sensible next step is controlled, well-governed pilots measured on outcomes rather than on alert volumes.

REFERENCES

- ABA Banking Journal. (2025, June). *FinCEN releases figures on BSA filings*.
<https://bankingjournal.aba.com/2025/06/fincen-releases-figures-on-bsa-filings/>
- Bank Policy Institute. (2018). *Getting to effectiveness: Report on U.S. financial institution resources devoted to BSA/AML & sanctions compliance*. Washington, DC: BPI. Summary:
https://bpi.com/wp-content/uploads/2019/05/BPI-Reviewing-AML_CFT-Effectiveness.pdf
- Colladon, A. F., & Remondi, E. (2017). Using social network analysis to prevent money laundering. *Expert Systems with Applications*, 67, 49-58. <https://doi.org/10.1016/j.eswa.2016.09.029>
- Deprez, B., Vanderschueren, T., Baesens, B., Verdonck, T., & Verbeke, W. (2024). *Network analytics for anti-money laundering: A systematic literature review and experimental evaluation* (preprint). arXiv:2405.19383.
- Europol. (2016). *Does crime still pay? Criminal asset recovery in the EU: Survey of statistical information 2010-2014*. The Hague: Europol.
https://www.europol.europa.eu/cms/sites/default/files/documents/criminal_asset_recovery_in_the_eu_web_version.pdf
- European Union. (2016). Regulation (EU) 2016/679 (General Data Protection Regulation).
- FATF. (updated periodically). *International standards on combating money laundering and the financing of terrorism & proliferation: The FATF Recommendations*. Paris: Financial Action Task Force.
- FATF. (2021). *Opportunities and challenges of new technologies for AML/CFT*. Paris: Financial Action Task Force.
- Fenergo. (2024). *Global AML fines research report 2023*. <https://www.fenergo.com/aml-report>
- Fenergo. (2026, January 13). *Global financial regulatory penalties fall by 18% in 2025 as enforcement shifts from US to EMEA and APAC*. <https://resources.fenergo.com/newsroom/global-financial-regulatory-penalties-fall-by-18-in-2025-as-enforcement-shifts-from-us-to-emea-and-apac>
- Financial Times. (2023, January 19). *Global anti-money laundering fines surge 50%* [reproduced by The Irish Times].
<https://www.irishtimes.com/business/2023/01/19/global-anti-money-laundering-fines-surge-50/>

- FinCEN. (2025). *FinCEN Year in Review for FY 2024*.
<https://www.fincen.gov/sites/default/files/shared/FinCEN-Infographic-Public-2025-508.pdf>
- Google Cloud. (2023, June 21). *Google Cloud launches AI-powered anti money laundering product for financial institutions*.
<https://www.googlecloudpresscorner.com/2023-06-21-Google-Cloud-Launches-AI-Powered-Anti-Money-Laundering-Product-for-Financial-Institutions>
- ISO/IEC. (2023). *ISO/IEC 9075-16:2023, Information technology: Database languages SQL, Part 16: Property graph queries (SQL/PGQ)*.
- LexisNexis Risk Solutions. (2023, September 26). *Study reveals global financial crime compliance costs for financial institutions total more than U.S.\$206 billion* (True Cost of Financial Crime Compliance Report, 2023). <https://risk.lexisnexis.com/global/en/about-us/press-room/press-release/20230926-global-financial-crime-compliance-costs>
- LexisNexis Risk Solutions. (2024a, February 21). *True cost of financial crime compliance study: U.S. and Canada* (\$61 billion). <https://risk.lexisnexis.com/global/en/about-us/press-room/press-release/20240221-true-cost-of-compliance-us-ca>
- LexisNexis Risk Solutions. (2024b, March 6). *True cost of financial crime compliance study: Europe, the Middle East and Africa* (\$85 billion).
<https://risk.lexisnexis.com/global/en/about-us/press-room/press-release/20240306-true-cost-of-compliance-emea>
- Pol, R. F. (2020). Anti-money laundering: The world's least effective policy experiment? Together, we can fix it. *Policy Design and Practice*, 3(1), 73-94.
- Republic of Türkiye. (2016). *Law No. 6698 on the Protection of Personal Data*.
- UNODC. (2024, November). *Improving regional investigations on money laundering and asset recovery*. https://unodc.org/roca/en/NEWS/news_2024/november/improving-regional-investigations-on-money-laundering-and-asset-recovery.html
- Weber, M., Domeniconi, G., Chen, J., Weidele, D. K. I., Bellei, C., Robinson, T., & Leiserson, C. E. (2019). *Anti-money laundering in Bitcoin: Experimenting with graph convolutional networks for financial forensics*. KDD '19 Workshop on Anomaly Detection in Finance, Anchorage, AK. arXiv:1908.02591.
- Wolfsberg Group. (2024, July 1). *Statement on effective monitoring for suspicious activity*.
<https://wolfsberg-group.org/news/69>
- Wolfsberg Group. (2025, August 27). *Statement on effective monitoring for suspicious activity, Part II: Transitioning to innovation*. <https://wolfsberg-group.org/resources/202>