resistant ai

Lucie Rehák Novotná

3/2025

About Resistant AI

- **17 years of AI experience** in fighting cyber and financial crime
- **9 PhDs in AI** with **100s of patents**
- Team of **100+** and growing, offices in **New York; London; Prague**
- Backed by the best

GV NOTION Index Ventures

Credo. Seedcamp

Who we work with

> When we think about geo-expansion we always talk with Resistant.
>
> **Head of Risk & Compliance, FINOM**

> Probably the best tool we have in our review flow.
>
> **Solutions Manager, Payoneer**

> The smoothest implementation of tech that we have ever experienced. No downtime, no interruption of business operations.
>
> **Head of Financial Crime, Holvi**

COMPLY ADVANTAGE®
Payoneer
FINOM
KBC
BOV Bank of Valletta
Habito
MONETA MONEY BANK
Deloitte.
ANNA
HOLVI
LUCINITY

100+ customers & partners

resistant ai

Some of the main developments impacting FinCrime compliance

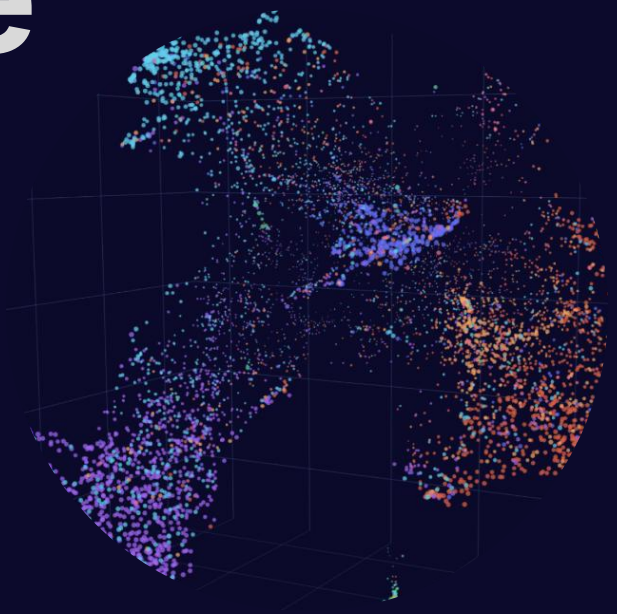
Fraud victim reimbursement

Information sharing

Instant cross-border payments

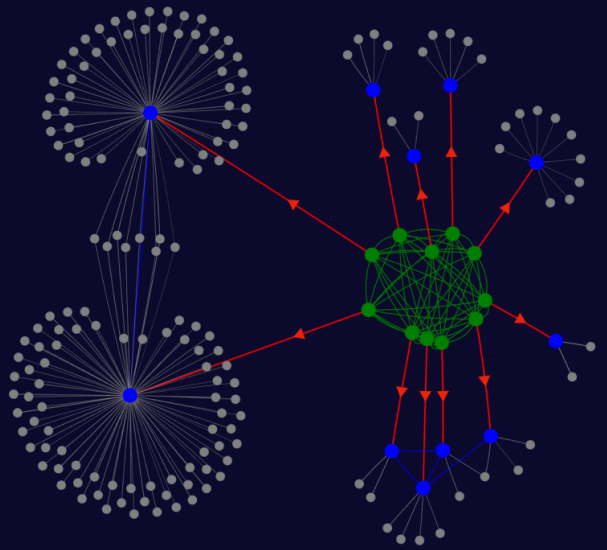
resistant ai

The fight against **financial crime** is due a revolution



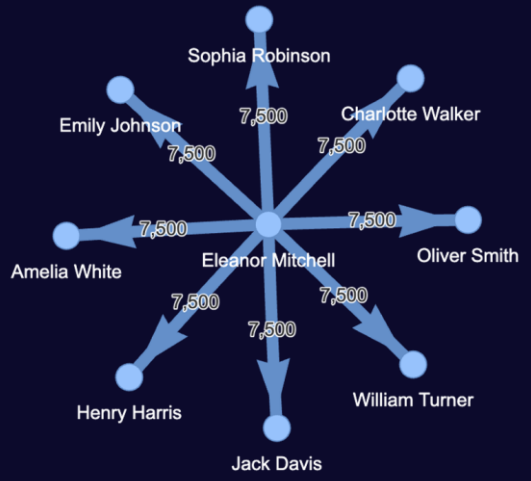
resistant ai

AI helps you to achieve (far) more with less



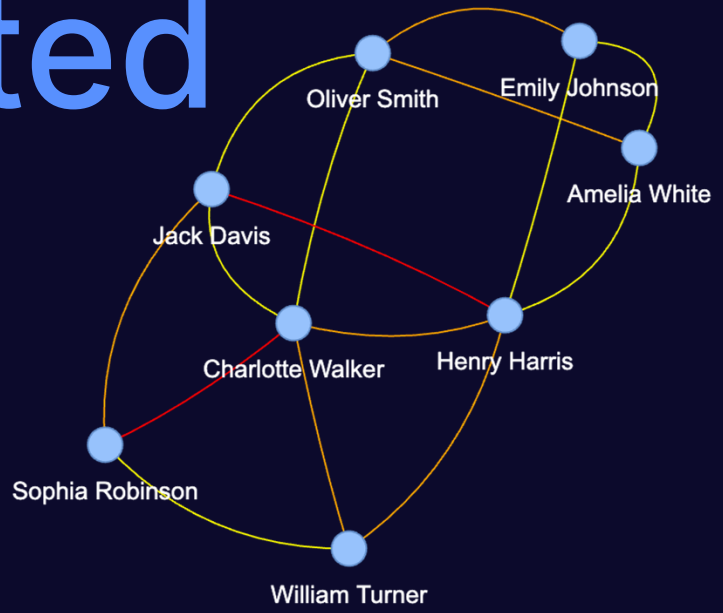
resistant ai

Achieving more is now expected by the regulators



resistant ai

Achieving more is now expected by the customers



resistant ai

Precision

The industry standard is that 95+% Anti-Money Laundering (AML), fraud, or sanction detections are false positives.

Recall

The vast majority of fincrime and fraud goes undetected. UNODC estimates that up to 5% of global GDP gets laundered annually.

Speed

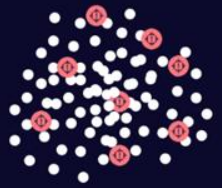
Monitoring of instant payments, both domestic and international, as customers expect immediate service as the norm.

resistant ai

The potential of AI solutions in Compliance

Reduce Fraud Losses

High precision fraud prevention



REAL-TIME CONTROLS

immediate verdict → PASS / HOLD

Reduce AML Risk

Broad capture FinCrime typologies

POST-TRANSACTION MONITORING

suspicious activity alert

Reduce Operational Costs

Operational efficiencies

ALERT QUEUE

manageable queue and more efficient investigations

resistant ai Advanced Detections

resistant ai Alert Prioritisation

resistant ai

Enhancing your transaction monitoring controls

Deploying advanced AI models and their combinations

resistant ai
Transaction
Forensics Engine

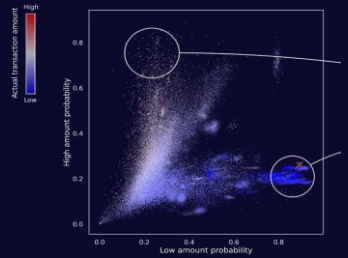
Large Language Models (LLMs)

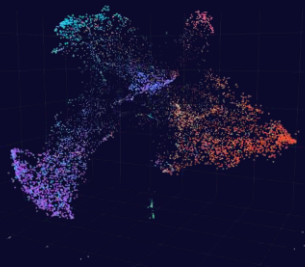
Threat Propagation

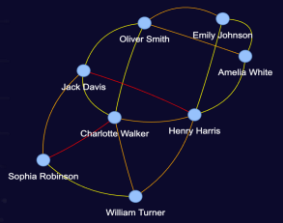
Clustering

Segmentation

Graph-based methods





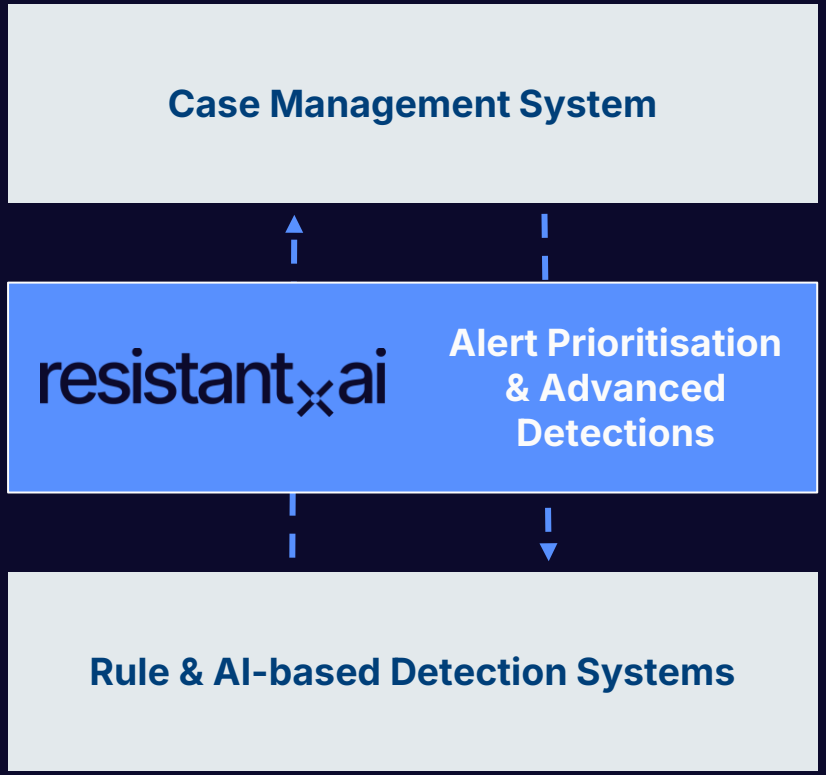


and more...

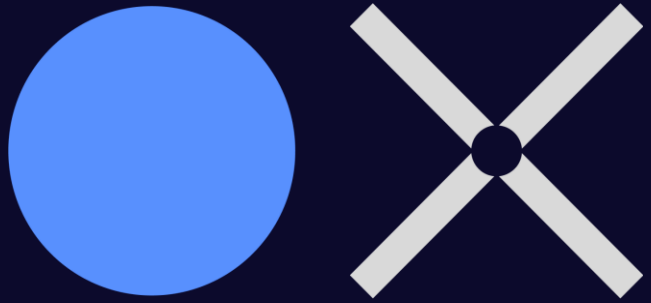
Transaction Forensics

Enhance, don't replace

- AI-native solution layers on top of existing transaction monitoring systems
- Models built to expose **new & complex** criminal behaviours
- **Adaptable & future proof** solution
- Time to value - **get the benefits of AI in weeks** - without having to go through a costly & painful migration
- Once a customer - **you'll never pay for new models release**!



resistant ai



Authorised-Push-Payment fraud

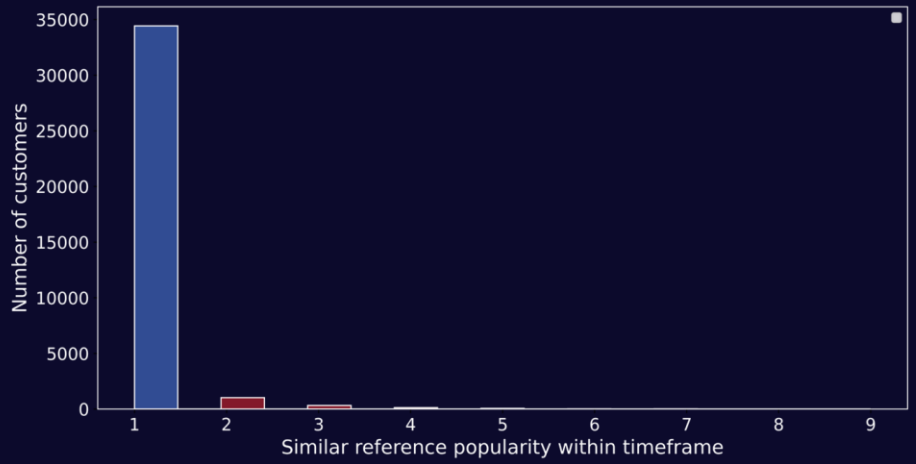
Aiming for high precision*

*All PII has been anonymised with the use of random name generator. Any similarity with existing persons or businesses is purely coincidental.

Preventing opportunistic fraud campaigns

Anomalous reference popularity

- Fraud campaigns during big events / product launches: Taylor Swift concerts, Euro tournament, new iPhone, etc.
- Semantically similar descriptions with suspicious frequency of use per given customer are alerted by RAI
- Within a short timespan, customer interacted with number of counterparties that was unexpected for similar kind of reference

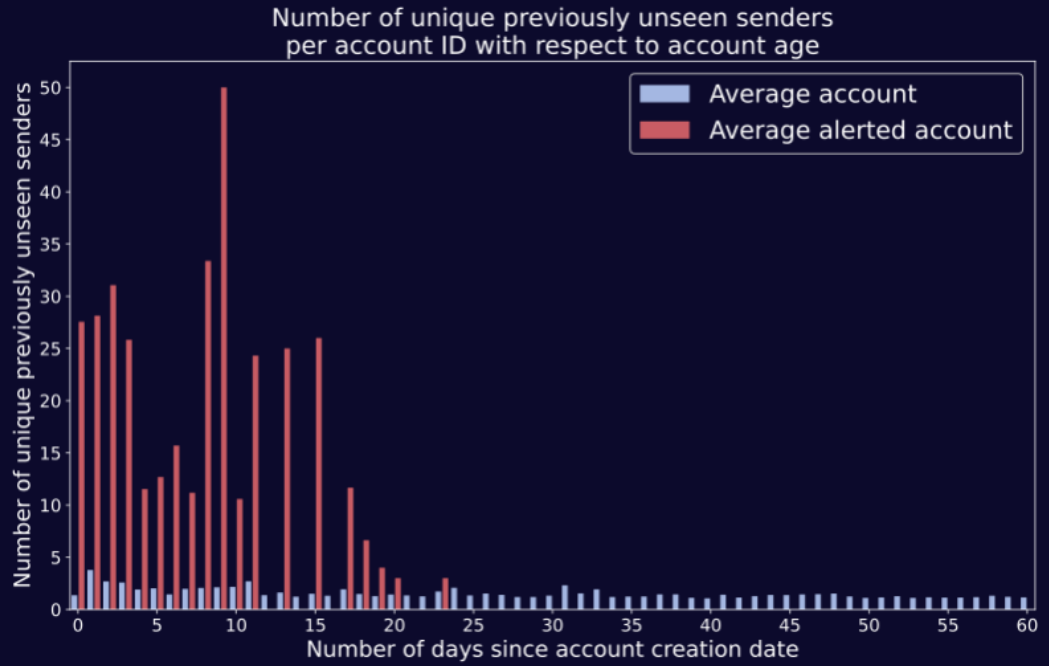


Counterparty FIRST SEEN	TRANSACTION DATETIME	REFERENCE	AMOUNT	CLIENT STATUS	FRAUD REPORTED	RAI VERDICT
17/10/2023	17/10/2023 17:49:36	taylor swift	£ 145	accepted	TRUE	HOLD
17/10/2023	17/10/2023 19:01:48	Taylor Swift	£ 120	accepted	TRUE	HOLD
17/10/2023	17/10/2023 19:32:05	Ticks	£ 240	accepted	TRUE	HOLD

Preventing opportunistic fraud campaigns

New account counterparty burst

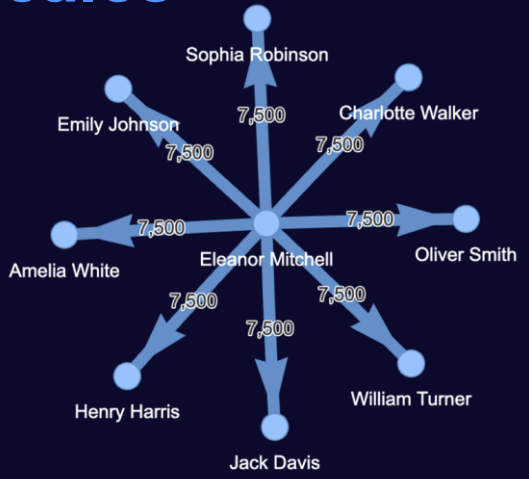
- Fraudsters create an account and try to defraud as many people as possible within a short period of time
- RAI detects bursts in the number of newly created counterparties with similar behaviour for every account
- Probabilistic thresholding to increase precision



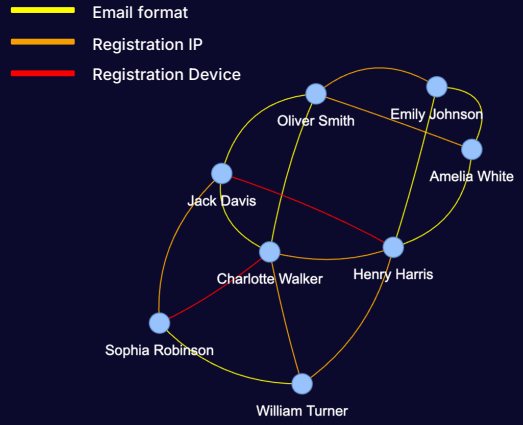
Identifying organised fraud before it scales

Detecting coordinated attacks

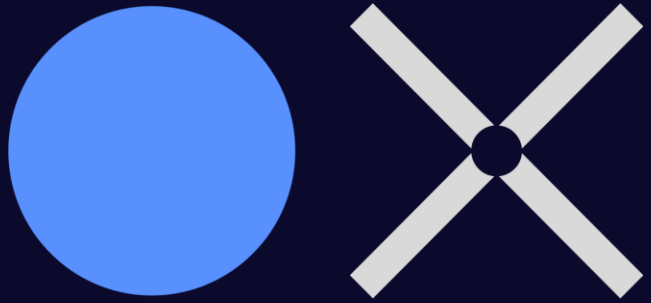
- Clustering based on the first transaction
- Accounts created in one day received the first transaction from the same counterparty
- Accounts share email format, registration IP, registration device, or other identity features



Account Creation Date	First Name	Last Name	Customer Email	Registration IP	Registration Device
2023-07-11	Oliver	Smith	smith8786@gmail.com	122.52.98.12	F5V24N9C6BM1Y7
2023-07-11	Emily	Johnson	8786johnson@gmail.com	122.52.98.87	X3S29N1V7L6R8M
2023-07-11	William	Turner	87turner86@gmail.com	122.77.45.175	H6I2L8K37UP1T9
2023-07-11	Charlotte	Walker	walker8786@gmail.com	122.77.45.132	W8ZE2I7T56R3O4
2023-07-11	Henry	Harris	8786harris@gmail.com	122.77.45.109	XRM3S29N1V7L68
2023-07-11	Sophia	Robinson	87robinson86@gmail.com	122.38.11.35	W8ZE2I7T56R3O4
2023-07-11	Jack	Davis	davis8786@gmail.com	122.38.11.61	XRM3S29N1V7L68
2023-07-11	Amelia	White	8786white@gmail.com	122.52.98.22	P5UQ7N0C9L11Y2



resistant ai



Complex money laundering

Aiming for high recall

Graph based detection of complex flows

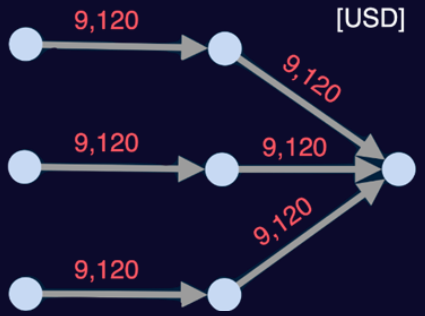
resistant ai

Dense cluster of nodes and/or edges with correlated weak indicators becomes a strong signal.

Limit avoidance

Dense group nodes connected by edges near the decision boundary of smurfing detector.

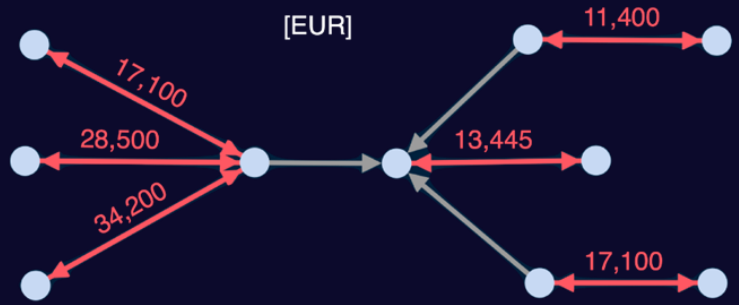
Example:



Mirror trading

Cluster of accounts with high percentage of mirror trading edges (same amounts exchanged both ways).

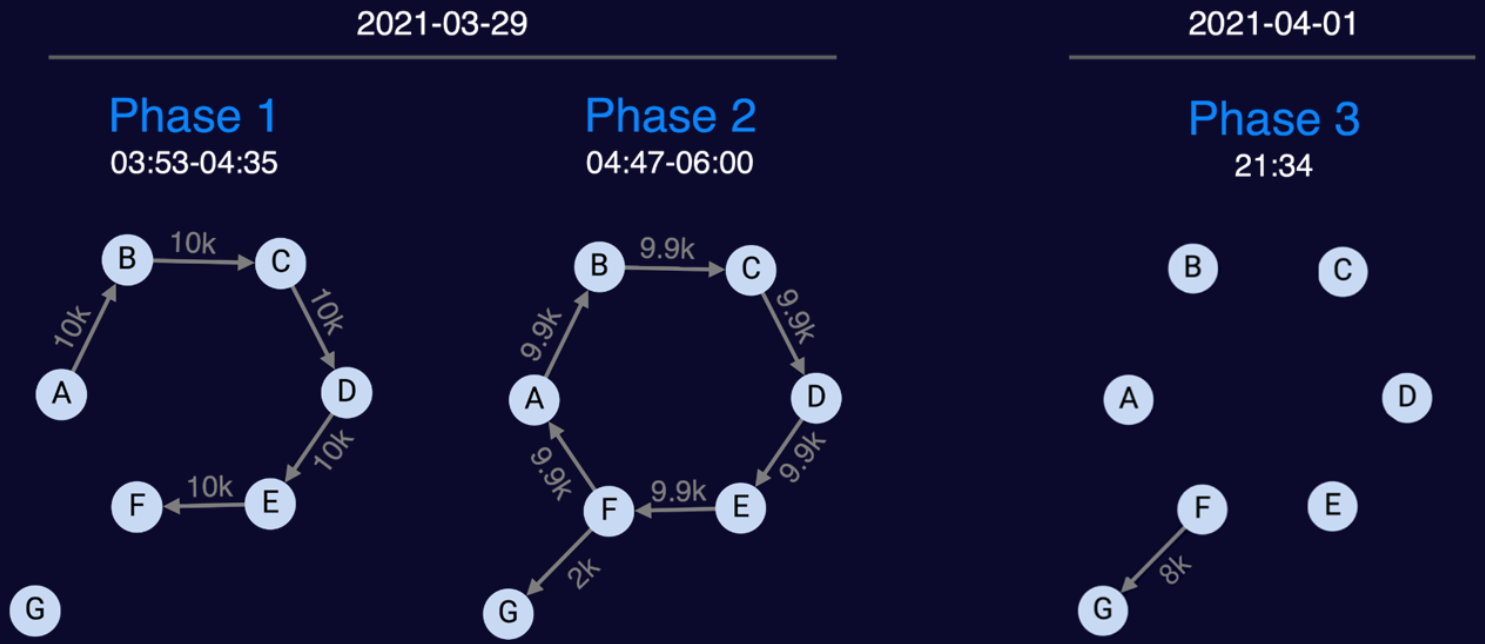
Example:



Graph based detection of complex flows

resistant ai

Detected by long chain and smurfing graph. All accounts with usual activity between 50 and 200 USD



Detection of behavioural anomalies

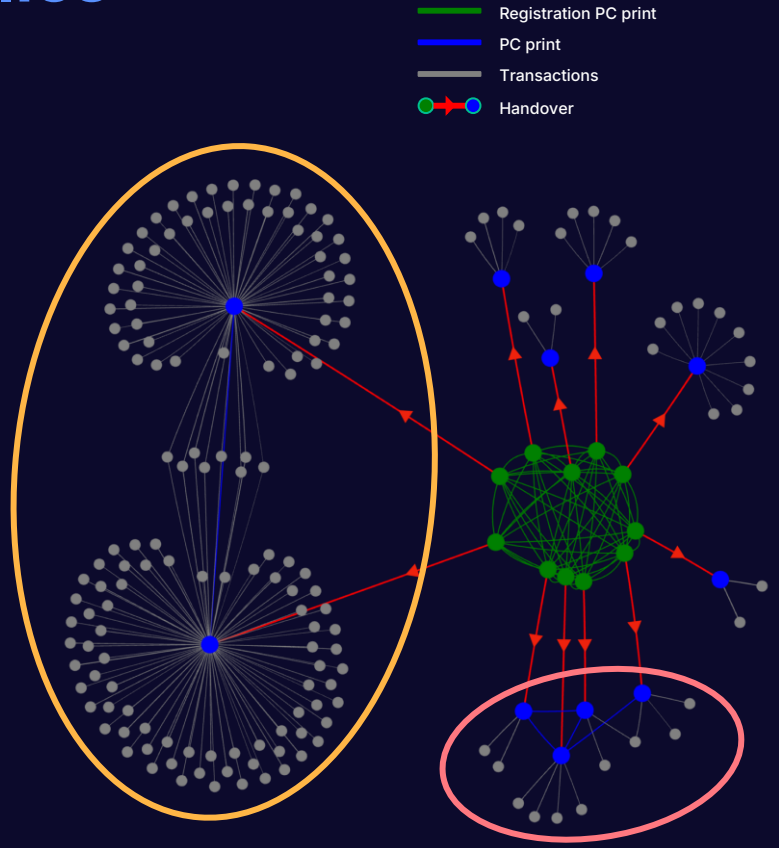
Account handover

Group of accounts created by single entity and then handed over to different entities for money-muling or other illicit activity

Accounts created by single entity share the same registration device and follow similar initial sequence of actions

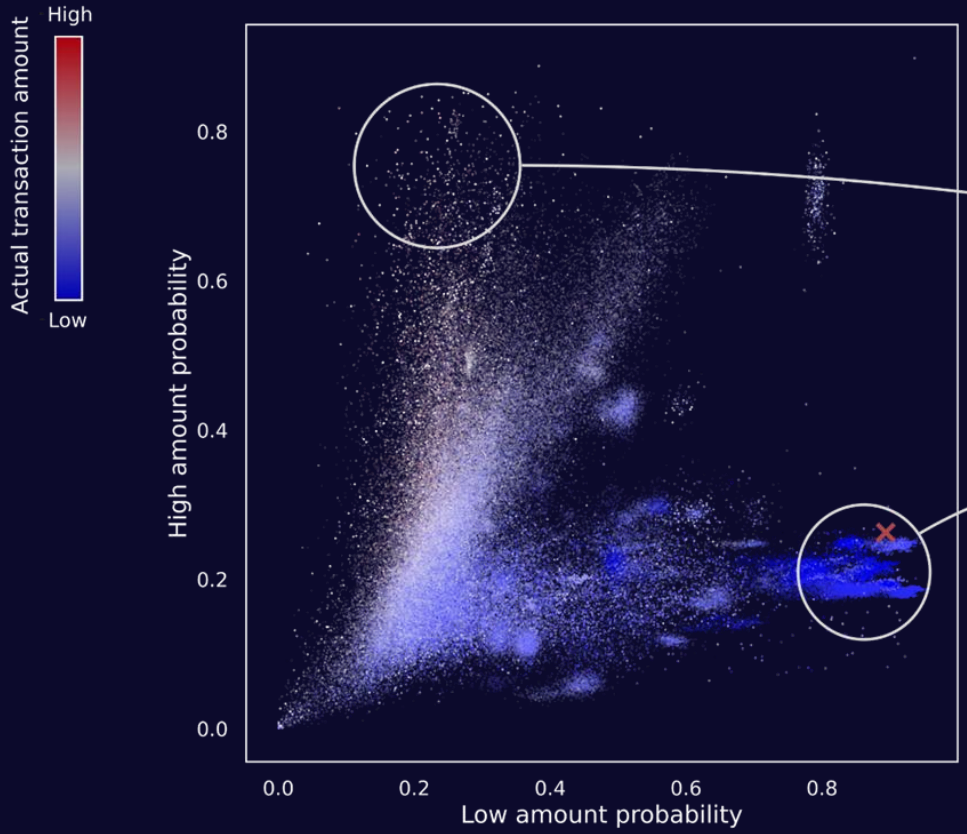
↓

Following the handover, the accounts are segregated into unconnected groups



Leveraging large language models

Reference Amount Discrepancy - leveraging the power of LLMs



RAI uses trained language models to predict reference-amount discrepancy.

References usually associated with high amount transactions:

- 7.1k USD: loan settlement no.120778/12
- 80k USD: Seed Investment S VC

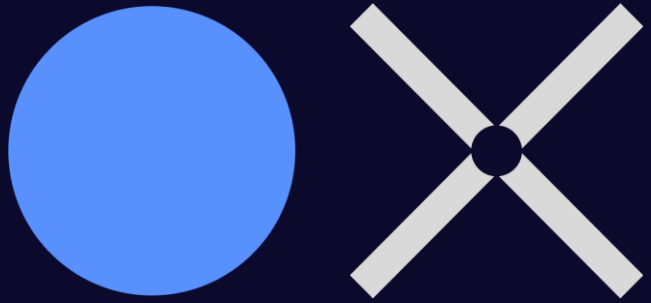
References usually associated with low amount transactions:

- 0.40 USD: Extra ATM withdrawal fee
- 8.49 USD: KAUFLAND MESTOBELIL GIEN

Detected anomalous transaction:

- × 9.5k USD: Uber ride London

resistant ai

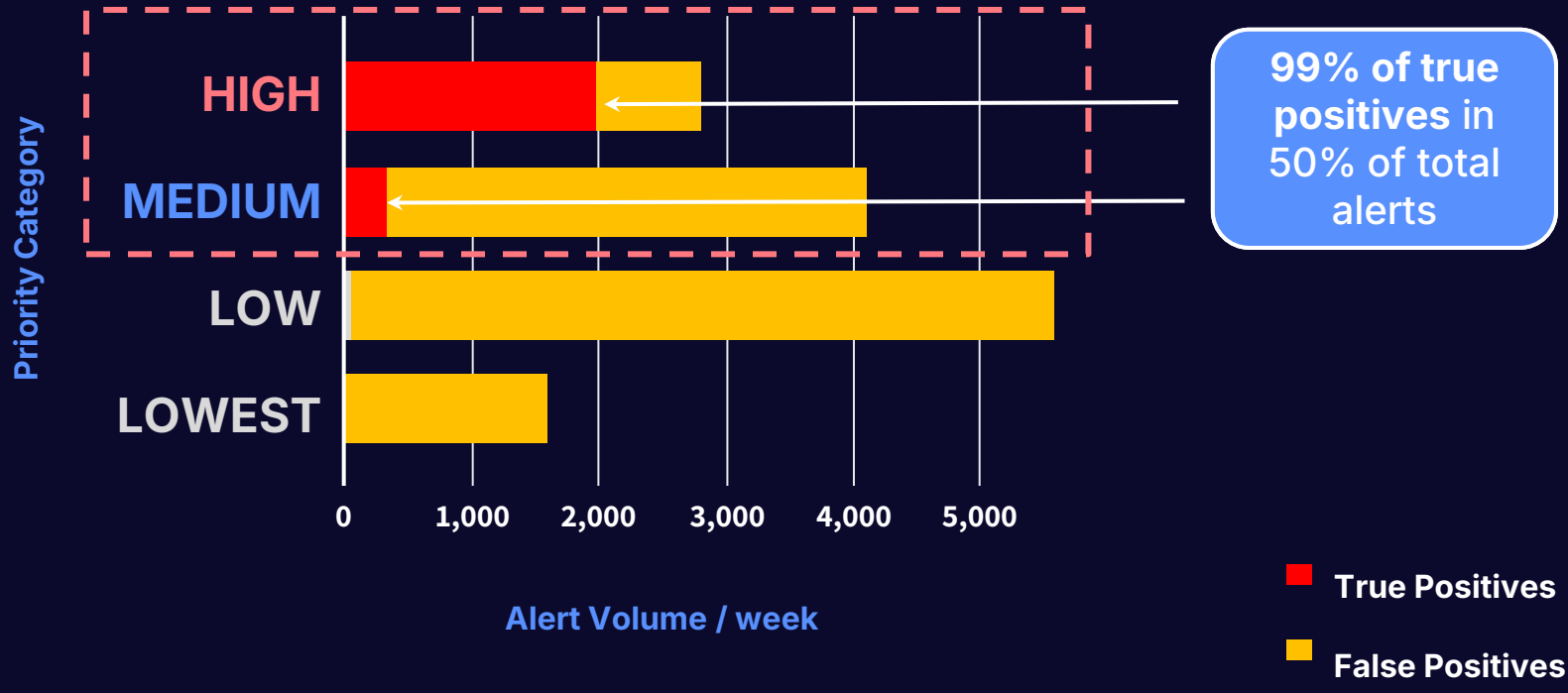


Alert prioritisation

Aiming for risk-based efficiency

Alert prioritisation

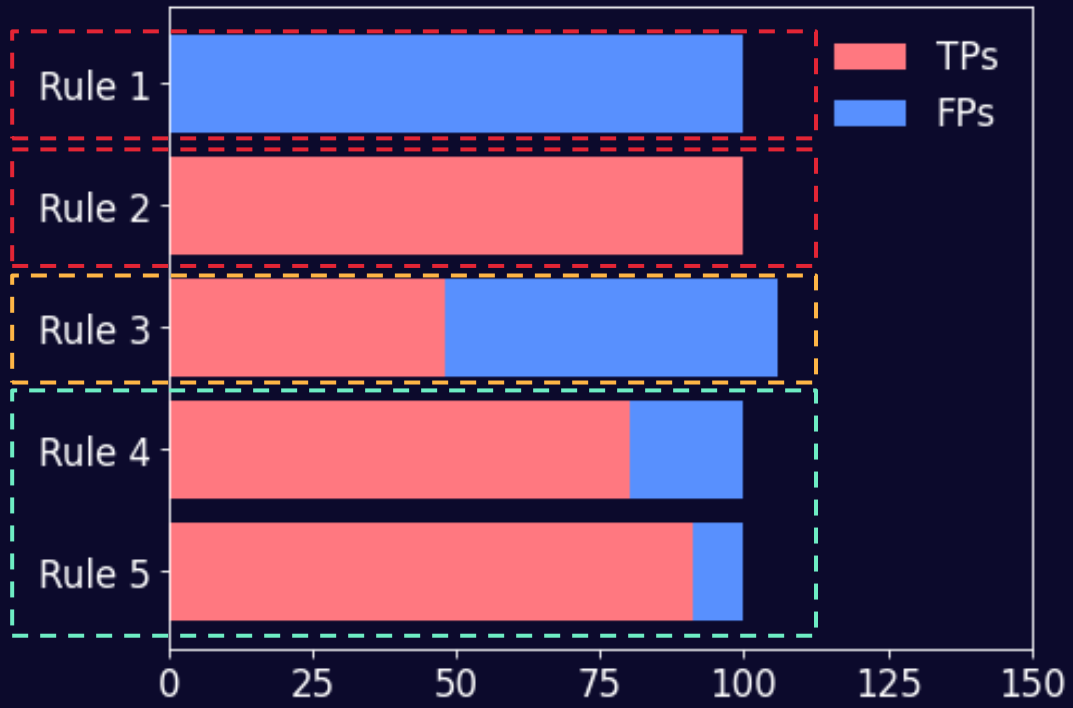
Take a true risk-based approach to alert management



resistant ai

Reduce false positives

Understand their root cause and take an informed approach



Identify underperforming rules

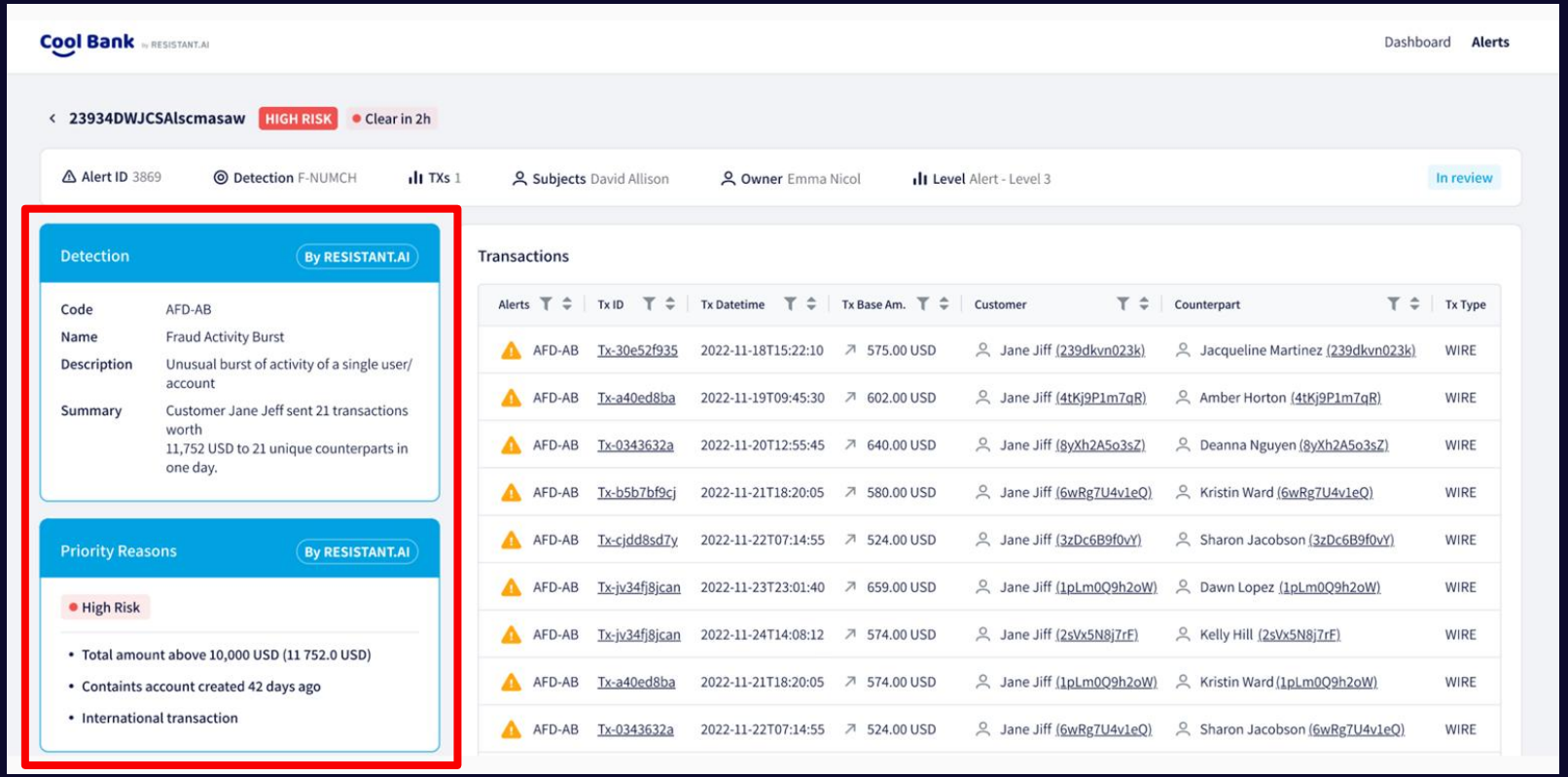
Take a safe approach to over performing ones

Optimise those with room for improvement

➡ Understand where your true hits come from

Alert prioritisation

Context and explainable outcomes for analysts



Understand your risk, consistently

Concrete examples of High and Low risk detections

Key Risk Factors AI

High Risk Factors

- mirror trading relative amount difference below 1%
- total amount above 1,000,000 USD (1,895,015 USD)
- multiple transactions (8) with the same counterparty within one day

High risk factors

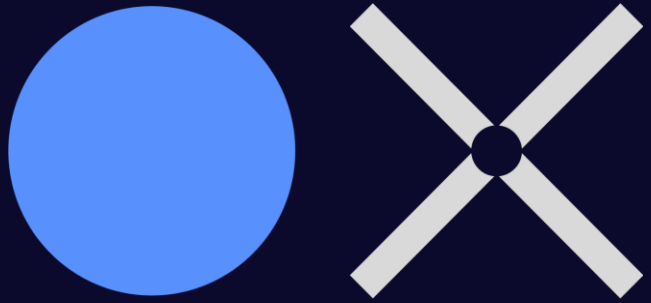
Low risk factors

Key Risk Factors AI

Lowest Risk Factors

- domestic transaction
- similar behaviour previously labeled as false positive (2 times)
- low risk customer

resistant ai



Risks, myths, and regulatory approaches

The main risks of relying on AI *and how to mitigate them...*

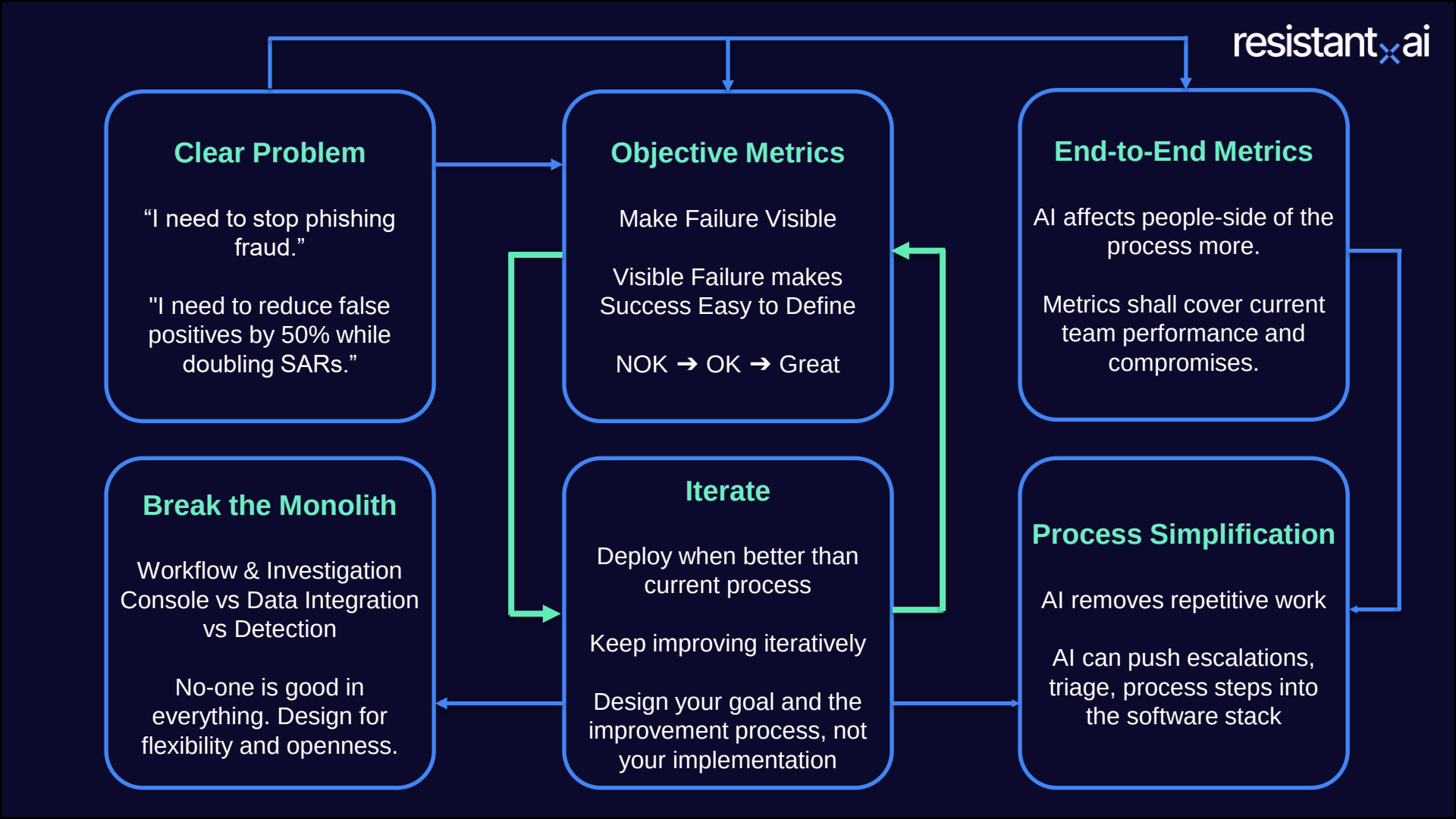
resistant ai

risk	*mitigant*
Unfair or discriminatory bias	Proactive model monitoring alerts to disproportionate importance of potentially sensitive features such as gender or nationality in the model decisions.
Building in analyst mistakes or malicious intent	Machine learning feedback loops are weighted to avoid a disproportionate impact of single occurrences. Data drift is proactively monitored. Below and above-the-line tests are conducted.
Relying on a single model approach	Ensemble modelling ensures that controls are balanced and comprehensive. Individual models provide verdict in the context of the ensemble, based on their relative confidence compared to other approaches.

Debunking the myths: AI in AML and Fraud

resistant ai

myth	*reality*	*tips*
We don't have enough data	Unsupervised approaches do not require millions of transactions to work well. Specialised vendors can provide valuable experience from your peers.	→ Conduct a Proof of Concept with a vendor(s) to assess feasibility → Benchmark results against your current controls
It's too expensive to rip and replace our systems	Enhancement solutions can sit on top of your current tech-stack, simply improving what you already use.	→ Identify your biggest risks and pain points (Fraud? AML? FPs?) → Scope the right solution for your problem and start there!
It's a blackbox	Explainability is crucial, specialised vendors design their solutions to be predictable, measurable, and explainable.	→ Require auditability and explainability → Combine approaches intentionally designed to target specific behaviours
We're not ready to use complex AI	The right solution will make your life easier, not more complicated. AI should be doing the boring job so your analysts can investigate true hits.	→ Adopt a side-by-side, phased approach to fully understand the solution before you rely on it solely → Ensure your teams receive sufficient and tailored training


resistant ai
Clear Problem
“I need to stop phishing fraud.”
"I need to reduce false positives by 50% while doubling SARs.”
Objective Metrics
Make Failure Visible
Visible Failure makes Success Easy to Define
NOK ➔ OK ➔ Great
End-to-End Metrics
AI affects people-side of the process more.
Metrics shall cover current team performance and compromises.
Break the Monolith
Workflow & Investigation Console vs Data Integration vs Detection
No-one is good in everything. Design for flexibility and openness.
Iterate
Deploy when better than current process
Keep improving iteratively
Design your goal and the improvement process, not your implementation
Process Simplification
AI removes repetitive work
AI can push escalations, triage, process steps into the software stack

resistant ai

The regulatory approaches we see across EMEA

1 Expecting innovation and encouraging full AI

Regulators like the **UK's FCA** and **PSR** expect firms to make **full use of available technology** and take all possible advantage of innovation. Regulatory sandboxes exist to test novel approaches safely.

Many clients rely solely on AI driven detections, with appropriate safeguards and documentation.

resistant ai

The regulatory approaches we see across EMEA

2 Encouraging innovation and supporting combined deployments

Many EU jurisdictions encourage **combined approaches**: hard rules based on values, volumes, and country risk combined with AI driven prioritisation as well as detection models.

Many clients combine old and new approaches to start with, and turn off legacy rules in a phased, data-driven approach.

resistant ai

The regulatory approaches we see across EMEA

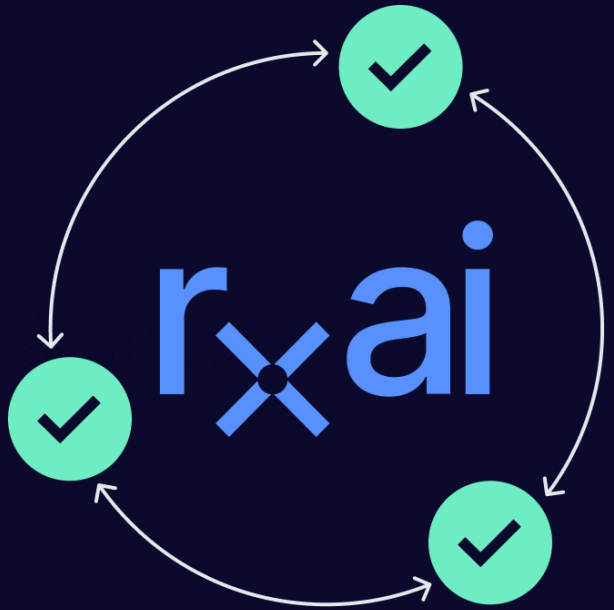
3 Agnostic, letting the private sector drive innovation and policy interpretation

Some regulators have not issued explicit policy or guidance on the use of AI, and are letting firms "figure it out" and **substantiate their innovation with data and own experience**.

Uncertainty in policy interpretation can slow down progress, but also presents an opportunity to innovative firms to gain a competitive advantage when it comes to customer fraud protection etc.

Waiting—the biggest mistake in AI for (FR)AML

resistant ai



- Don’t wait to replace your systems.
  Augment what you have
- Don’t wait for perfect data.
  There’s no such thing
- Don’t wait for a criminal event.
  Be ready before they scale

THANK YOU