

Increasing Compliance with Existing Investigator Capacity

Predictive Risk & Behavioural Profiling | Sarkar Islam, SAS Institute

IOTA 30th General Assembly | Budapest | 18 June 2026



The Compliance Challenge

5-20%

of anticipated revenues
lost to noncompliance globally

~€135B

in VAT & income tax
repayments processed annually

10B

data items interrogated
across multiple sources

Tax administrations face growing volumes, increasingly complex fraud schemes, and constrained investigator resources.

The question: How do we improve compliance outcomes without proportionally increasing headcount?

Why Risk-Based Analysis — Now Powered by AI

Traditional Approach

- Fixed rules and thresholds
- Random audit sampling
- Manual review of every flagged case
- One case can take up to 15 days to review
- High false positives drain investigator time
- Fraud networks invisible to single-case reviews

Risk-Based Approach

- ML models score every claim dynamically
- Entity resolution links related parties
- Same case can be investigated in significantly less time, not 15 days
- Network analytics reveals organised fraud rings

+ AI Assistant

- Summarises case findings in plain language
- Guides what to check next in the investigation
- Surfaces similar past cases and their outcomes
- Recommends the best course of action
- Human investigator always takes the final call

Moving from "find a needle in a haystack" to "the system shows you where the needles are"

From Months to Minutes: AI-Augmented Investigations

15 days → 1–2 days

case review time
traditional vs. risk-based
approach

~7x

more cases per investigator
per year, illustrative uplift

+€250M

indicative annual revenue
uplift
from recovered tax
(illustrative)

Same team, more cases, more revenue without adding headcount.

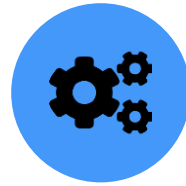
End-to-End Analytics Pipeline



01

Integrate

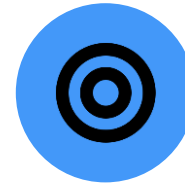
Multi-source data ingestion & entity resolution



02

Analyse

ML, AI & network analytics & behavioural profiling



03

Score

Dynamic risk scores with explainable factors



04

Act

AI-prioritised alerts, targeted interventions & audit selection

End-to-end analytics pipeline, from raw data to AI assisted investigator action

Data Foundation & Entity Resolution

Multi-Source Data Integration

- Tax returns & declarations (VAT, income tax, corporate)
- Banking & financial transaction records
- Import/export & customs data
- Company registration & directorship data
- Third-party data (merchant acquirers, platforms)
- Cross-border information exchange (DAC, CRS)



Entity Resolution

Match & merge: joins records across systems that refer to the same real-world entity

Resolve identities: works even with name variations, address changes, and missing data

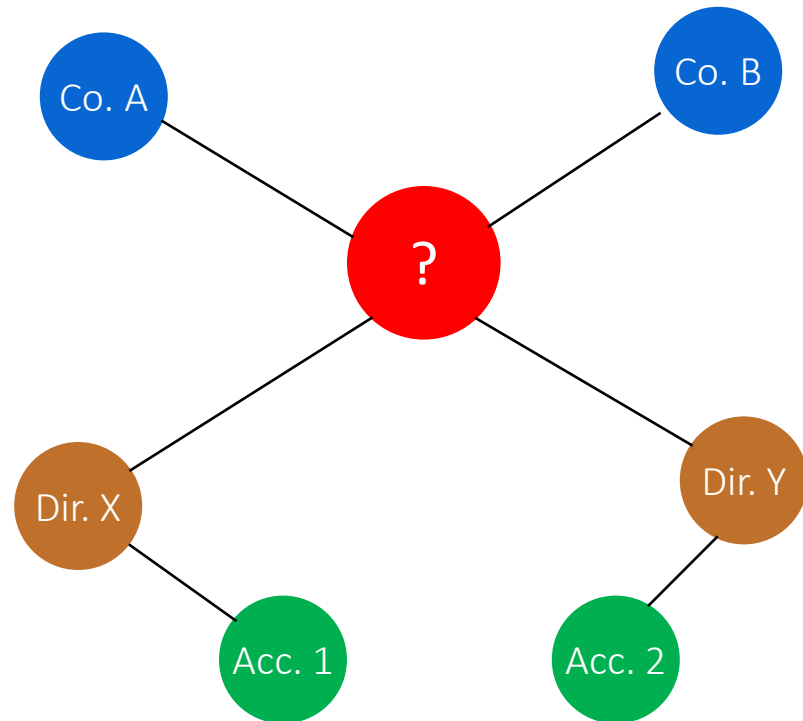
Link related parties: links through shared attributes (addresses, phone numbers, bank accounts, directors)

Build a 360° view: of each taxpayer and their connections

A large European tax authority processes 10 billion data items across these sources to build a comprehensive risk picture

Network Analytics

Revealing Hidden Connections in Tax Data



Carousel Fraud Detection

Network Analytics and Pattern matching on millions of transactions surfaces suspect VAT carousel chains across cross-border transactions

Shared Director Networks

Identify company clusters controlled by the same individuals using different identities

Financial and Transactions Flow Mapping

Track money or goods movements across transactions or accounts to reveal layering and hidden relationships

Cross-Border Linkages

Connect domestic entities to international networks using information exchange data

Why this matters for analyst capacity: Network analytics and pattern matching surface the organised VAT fraud rings that manual review misses, so the same team closes more cases.

Risk Scoring

How It Works

- Multiple analytical methods run concurrently on each claim
- ML models are trained on historical compliance outcomes
- Combine behavioural indicators, financial anomalies, entity scoring, network analytics to create one score
- Each case assigned a composite risk score (0–100)
- Scores include explainable contributing factors

Risk Tiers

High Risk (75-100)

Medium Risk (40-74)

Low Risk (0-39)

Key principle: Risk scores guide investigator attention, they are not a determination of guilt. Professional judgement always has the final say.

Matching Response to Risk



High Risk - Investigate

- Top-ranked high value cases routed straight to investigators
- Case insights, evidence, and network context ready at the start of every investigation
- AI assistant accelerates the case investigation by guiding investigators automatically



Medium Risk - Nudge

- Personalised reminder or clarification request
- Chance to self-correct before enforcement
- After each nudge taxpayers are re-analysed and re-scored
- No response triggers a stronger second nudge or escalation to high risk
- Resolves most issues without opening a full case



Low Risk - Fast-Track

- Straight-through processing for compliant taxpayers
- Faster refunds and less friction for honest filers
- Investigator capacity preserved for real risk

One risk score, three proportionate responses, turning analytics into action across the whole taxpayer base

Governance & Safeguards

Warning: *A flawed risk-scoring system wrongly accused thousands of families of fraud, forcing the entire government to resign, a stark reminder of why the below safeguards must hold.*

Model Governance

- Periodic validation & recalibration of risk models
- Documented model development and approval processes
- Independent review of model performance

Fairness & Proportionality

- Risk scores never treated as proof of noncompliance
- Audit selection should be evidence-based
- Guard against over-targeting specific groups or sectors

Transparency & Explainability

- Every risk score accompanied by contributing factors
- Clear escalation thresholds and rules
- Human investigator makes final decision on every case

Key Takeaways

- 01** Risk-based analysis lets tax administrations do more with the same investigator capacity, focus effort where it matters most
- 02** Entity resolution and network analytics uncover organised fraud that individual case reviews may miss entirely
- 03** Multi level risk scoring helps a graduated response, enforcement for the high-risk, support for the compliant
- 04** AI augments investigators by summarising cases, surfacing related entities, and recommending next steps
- 05** Governance, fairness, and explainability are non-negotiable, algorithms advise, humans decide
- 06** The technology exists today and is proven at scale, the challenge is organisational adoption and integration

Continue the conversation

Explore SAS insights on modern tax administration, ethical AI and compliance transformation.



Visit the SAS booth and scan the QR code to access the SAS Tax Knowledge Hub.



Scan to open the Tax Knowledge Hub



IOTA General Assembly | Budapest | 16–18 June

Copyright © SAS Institute Inc. All rights reserved.

