Industry Perspectives on Best Practices - Leveraging on Data Analytics and Machine Learning Methods for AML/CFT

March 2024

Contents

1 Introduction

1.1 Background

Overview

Automation and digitalisation advancements have improved our ability to detect and therefore fight financial crime. Conversely, they have also introduced new opportunities and more sophisticated methods for committing fraud and financial crime.

As a global financial hub, Singapore needs to be ever vigilant to combat a spectrum of financial crimes. Financial Institutions ("FIs") have spent increasing amounts to fight financial crime over the years, from US$ 3.1 billion in 2019, to US$ 3.8 billion in 2020, and more recently hitting US$5.7 billion (S$7.8 billion) in the past year[1]. These amounts were invested to enhance existing controls, or adopt new controls facilitated by data analytics to improve effectiveness and efficiency. One of the focus areas is on scam prevention arising from the significant increase in the number of scams and cybercrime cases from 7,493 cases recorded in 2018 to 33,669 cases in 2022[2].

This paper will discuss how FIs based in Singapore have adopted data analytics to improve their control effectiveness and operational efficiency in detecting and preventing financial crime. It reviews enhancements since the publication of our first paper in 2018, titled, Industry Perspectives – Adopting Data Analytics Methods for AML/CFT[3] and discusses emerging techniques and capabilities that are being developed.

Through sharing insights and experiences of the member banks, the paper is hoped to serve as an information paper for FIs to gain broader perspective of the future directions in fighting financial crime. It is not seeking to set recommended minimum standards rather to provide awareness and possible directions for FIs to consider in the adoption of data analytics.

1.2 Developments since 2018

In our 2018 paper, "Industry Perspectives – Adopting Data Analytics Methods AML/CFT", we have covered an introductory overview with use cases of data analytics adopted by FIs in facilitating anti-money laundering and countering the financing of terrorism ("AML/CFT"), highlighting the levels of adoption as well as key considerations and future focus areas for industry and private-public collaboration.

Since the 2018 publication, the adoption of data analytics to address challenges within the AML/CFT space has gained significant maturity across the industry. FIs have also ramped up their systems and processes towards a more proactive and automated approach to improve their efficiency and effectiveness in financial crime surveillance.

Many data analytics methods, which were previously identified to be partially in deployment or exploration, have been successfully implemented and deployed in production.

Some examples would be the adoption of link analysis for detection of money laundering networks and deployment of machine learning models to prioritise alerts. Some banks have also adopted analytics to enhance customer risk assessment.

1.3 Objective

This paper provides a progress update on the use of data analytics in combating financial crime since the 2018 paper. Hopefully, topics discussed will inspire and encourage FIs who have yet to engage with or expand on data analytics to adopt and broaden their use of the same, to keep up with the fast pace of financial sector development and emerging crime typologies.

The paper will cover the current landscape of financial crime and share insights on the core foundational pillars supporting the successful implementation of such solutions. This will be followed by a section on how FIs have leveraged data analytics at a macro and network level, converging into the outcomes of customer and transaction multi-level risk surveillance.

The paper then delves into the use of technology and data analytics to combat six different prioritised financial crime risk types identified for Singapore, namely:

- Sanctions evasion;
- Misuse of legal persons and legal arrangements;
- Trade-based money laundering;
- Tax evasion;
- Fraud and scams; and
- Money mules.

FIs can apply these selected use cases as benchmarks to drive continual improvements in their technology and data analytics capabilities to keep pace with the ever-growing size and complexity of financial crime risks.

Finally, the paper will conclude with upcoming trends and detail the considerations for next generation's data infrastructure, exploring the advanced analytics deployed by member banks in different areas to stay ahead of financial crime.

[1] Source: Lexis Nexis Risk Solutions Study: LexisNexis Risk Solutions Study Reveals Global Financial Crime Compliance Costs for Financial Institutions Totals More Than U.S.$206 Billion | LexisNexis Risk Solutions
[2] Source: Singapore Police Force report: "Annual Scams and Cybercrime Brief 2022"
[3] Source: ACIP 2018 paper: acip-working-group-paper---data-analytics-for-aml.pdf (abs.org.sg)

2 Current landscape

2.1 Overview

The effective use of technology and data analytics can play a pivotal role in combatting money laundering /terrorism financing ("TF") risks within FIs. The successful and effective application of data analytics in the AML/CFT space requires three foundational pillars:

- Organisational Structure and Skillset;
- Governance and Model Risk Management;
- Data Architecture and Solution Design.

This section sets out some key considerations under each pillar for FIs to consider as they progress in their data analytics journeys.

The section then discusses two key conceptual frameworks that have evolved since the 2018 paper – a multi-layer risk surveillance framework leveraging data analytics; and a multi-pronged approach to managing financial crime risks.

Finally, this section shares use cases across the six different risk types, as defined in Section 1.2 above.

2.2 Foundational pillars

2.2.1 Organisational structure and skillset

To harness the power of data with analytics, FIs could begin establishing a robust structure and cultivate the necessary skillsets within their AML/CFT teams.

Various FIs have adopted different structural setups, mainly characterised as decentralised, centralised or hybrid.

In a decentralised structure, the AML/CFT analytics teams are set up within individual business units or departments (i.e., Commercial, Corporate, Retail). This allows teams to implement analytics solutions relevant to the business through deep understanding of unit operations, tailor risk assessments and provide accountability at the business level. However, challenges presented by such segregated teams include inconsistent methodologies, duplication of efforts, and difficulty in data aggregation across the FI.

Meanwhile, a centralised structure comprises of a single team being responsible for all data analytics in the FI, including AML/CFT analytics. This structure promotes consistency in methodologies, tools, and processes across the FI, and facilitates efficient resource utilisation and holistic risk assessment. Additionally, centralisation makes aggregation and analysis of data easier which enables the identification of cross-business unit patterns and emerging risks. The challenges faced by a centralised structure are limitations in understanding or catering for business-specific requirements and delays in providing necessary updates.

A hybrid structure combines elements of both the centralised and decentralised approaches by having a central analytics team supported by specialist teams residing in the various business units. The central team is responsible for designing and implementing AML/CFT analytics solutions for the FI, ensuring consistent methodologies and processes. The specialist teams collaborate with the central team during the initial development stage. They are also responsible for the operation and maintenance of analytics solutions at the business unit level, leveraging on their unit-specific expertise to work with the central team to make timely adjustments for business requirements.

The hybrid structure also offers flexibility in resource allocation and decision-making processes, capitalising on the benefits of both the centralised and decentralised structures. However, the structure is highly reliant on collaboration between teams, requiring close coordination and communication to be effective.

The following table summarises the characteristics, pros and cons generally associated with each analytics organisational structure.

Key Focus Areas	Decentralised	Hybrid	Centralised
Characteristics	Each business unit or department within the FI has its own dedicated analytics team responsible for AML/CFT	FI maintains both centralised and decentralised analytics capabilities	A single analytics team is responsible for handling AML/CFT activities across entire FI
Pros	Closer alignment with specific business units to have deep understanding of their operations, unique AML/CFT risks and value proposition	Standardisation of methodologies, tools, and processes while maintaining business unit specific expertise	Consistent application of methodologies, tools, and processes across the bank
	Rapid response to department-specific issues or emerging risks	Enhanced collaboration and communication between the centralised and decentralised teams	Efficient use of resources and expertise
	Promotes accountability within business units for AML/CFT compliance	Flexibility to allocate resources based on specific needs or risks	Easier aggregation and analysis of data across the bank, enabling holistic view of AML/CFT risks
Cons	Lack of consistency in methodologies, tools, and processes across the FI	Requires careful coordination and communication between teams	Limited understanding of specific business unit operations and nuances
	Potential duplication of efforts and inefficiencies	Potential complexity in decision-making processes and governance	Potential delays in responding to department-specific issues or emerging risks
	Challenges in aggregating data and obtaining a holistic view of the FI's AML/CFT risk	Increased resource requirements	Challenges in maintaining effective communication and collaboration with various business units

It is important for an FI to assess its needs and objectives in establishing a data analytics team to establish an organisational structure that fits the specific requirements. The following table summarises the considerations in determining the appropriate organisational structure for the data analytics teams.

Key Focus Areas	Decentralised	Hybrid	Centralised
Responsibility	Specific business units can take on such responsibilities	A centralised team is responsible for firm-wide risks while business specific risks are addressed by specialised team	A centralised team responsible for the entire FI is more efficient
Alignment	Close alignment with business units' needs is required	Alignment varies according to business units' needs	Alignment with business units' needs can be managed
Expertise	Deep understanding of business units is important	Mix of broad and deep understanding of business units is desired	Broad expertise can still meet the needs of the organisation
Consistency	Differences in methodologies is manageable	Consistent methodologies adapted by specialist teams and approved by central teams	Consistent methodologies are important
Efficiency	Potential duplication is manageable	Efficient resource utilisation is needed	Efficient resource utilisation is needed
Data Aggregation	Lower need for aggregating data	Aggregation of data is required	Aggregation of data is required
Communication	Collaboration between units can be managed	Centralised collaboration is critical for implementation	Centralised collaboration is important for implementation
Decision-Making	Business unit-centric decision-making is required	Business unit-centric decision-making is required	Centralised decision-making is suited for the organisation
Governance	Unit-specific governance is manageable	Centralised governance with business specific considerations is important	Centralised governance is important

Case examples of different organisation structures

Purely decentralised or centralised structures have not been commonly observed as the disadvantages associated with each option could have undesirable impact on FIs in the long run. Instead, most FIs choose to sit within the spectrum by adopting a varying combination of both structures, i.e., hybrid.

Centralised

One member bank has a model that includes elements of both a centralised and hybrid set up, which gives the flexibility and autonomy to experiment (fail fast) and focus on executing approved initiatives.

The centralised data science team, comprising data scientists and model developers, handles all analytics and machine learning activities covering areas such as employee compliance, disclosure of interest and AML/CFT. This specialised team collaborates with strategy and AML technology teams to experiment and develop machine learning models tailored for transaction monitoring. This gives the FI the ability to experiment rapidly and in an agile manner.

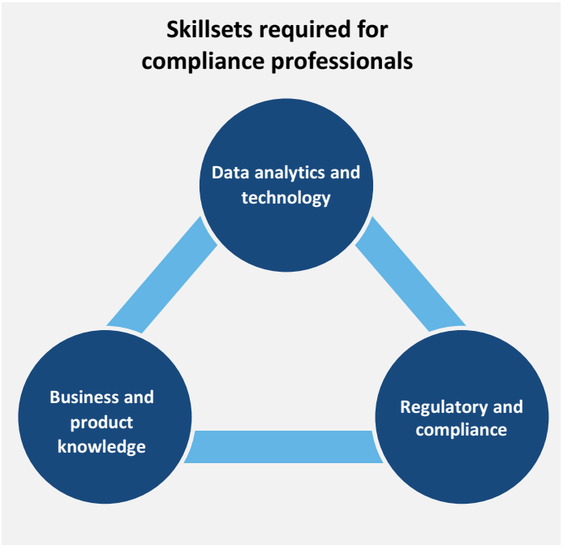
Once an initiative is formally approved, a project working group is formed involving members from various specialised teams, ranging from strategy, AML investigations, program execution, data science, business and AML technology. The initiative is then coordinated and managed by a program execution team. This gives the FI the ability to have dedicated resources for an approved initiative.

Hybrid

Another member bank has a hybrid model, where a pool of centralised data scientists serves as an extension to the specialised team and provides additional resources and technical Artificial Intelligence/Machine Learning ("AI/ML") expertise when needed. The bank shared that the structure has helped deliver solutions efficiently through the collaboration between centralised and specialised data science teams, complemented with domain analytics expertise holding in-depth financial crime and data knowledge. The outcome has been faster development of experiments and roll-out of innovative solutions.

Developing data analytics professionals with financial crime compliance knowledge

Apart from organisational structure, the complexity of financial crime necessitates a combination of business and technical expertise to effectively build advanced technological solutions for combatting financial crime. The skillsets required of financial crime compliance professionals have shifted to demand a suite of skills across "data analytics and technology", "regulatory and compliance", and "business and product knowledge" (see right illustration). Within a team, it will be optimal to have a mixture of individuals who have common base knowledge in AML but also specialise in different and complementary skills.



Ideally, a data analytics professional specialising in AML/CFT will have a blend of business acumen and technical proficiency of the three areas – capable of understanding the intricacies of financial transactions and effectively utilising data-driven methodologies to detect suspicious patterns and anomalies. Additionally, a deep understanding of the AML/CFT regulatory landscape, risk assessment frameworks and industry best practices allows teams to develop advanced analytics solutions which align with industry standards.

In preparation for the future AML/CFT landscape, the Institute of Banking and Finance ("IBF") has designed a framework tool[4] which details the required skills for current and emerging roles in the industry. Some industry-wide capabilities identified by IBF as essential by professionals for the rapidly changing industry are digital awareness, agile and entrepreneurial thinking, data analysis, future communication, human centred design, and risk and governance in the digital space.

Building a future-ready AML/CFT team

FIs could consider a targeted hiring approach when finding data analytics professionals to join the team. This means ensuring they have the necessary skillsets (i.e., Data Science, Business Analytics) fit for the role.

As data analytics professionals hired would have strong fundamentals in analytics and technology, FIs would need to help them grow the aspects of "Business and product knowledge" and "Regulatory and Compliance" to prepare such individuals for roles in financial crime compliance. This could be via onboarding or formal compliance training, providing the needed financial crime compliance domain knowledge and business training to effectively contribute or problem-solve for emerging risks and patterns.

FIs may also consider using cross functional teams comprised of professionals specialised in data analytics and those in AML/CFT for daily operations. Such arrangements provide additional avenues for team members to exchange skillsets and gain invaluable domain knowledge in areas outside of their professions.

As talents progress in their career, they can be equipped or upskilled by taking on AML/CFT qualifications to elevate their professional expertise in the domain. While not impossible for professionals in AML/CFT compliance to upskill in analytics and technology, the learning curve for such technical skills may likely be steeper.

It is worthwhile to note that compliance-related knowledge and skillsets are often built over time through career experience, beyond the classroom.

As a final option, FIs can partner with external companies, learning institutions or industry bodies to provide valuable training opportunities to their talent. Such partnerships could also forge strong industry working groups for innovation on AML/CFT solutions.

Key takeaways

This section discusses the first of three fundamental pillars to apply data analytics successfully and effectively. FIs must have an environment capable of using productive analytical solutions, requiring the following:

Organisation structure: As different FIs have varying objectives, sizes and resources, there is no one-size-fits-all approach for FIs to adopt. Based on these factors, FIs can identify where in the spectrum of possible structures they would fit best, to give themselves the ability to innovate and focus on core initiatives effectively.

Skillset: Future Financial Crime Compliance ("FCC") professionals will require skillsets across technology and analytics, regulatory policy and compliance, and an appreciation of customer journeys and business products. FIs could define learning curricula to support their teams' career growth and learning in this regard. Leveraging on external certification courses could also be beneficial.

[4] Source: IBF Framework Tool: IBF Skills Framework for Financial Services

2.2.2 Data infrastructure and solution design

To facilitate the adoption of data analytics, one area for FIs to consider is the appropriateness of data infrastructure/platform (storage and associated compute) infrastructure for the analytics and solution design. This might require FIs to build or enhance existing platform to meet the objectives of financial crime use cases and broader AML/CFT. In doing so, the FIs may consider adopting existing FI-specific data landscape to handle both structured and unstructured data along with market-specific regulatory compliance and expectations on data movement across borders, residency, retention, and security. There are a number of key considerations to keep in mind, including:

Rapid Experimentation and Deployment

The proposed data infrastructure would include seamless access to the data platform, ideally a separate ring-fenced sandbox compute environment that allows access to production data besides bringing in necessary custom datasets to facilitate rapid experimentations, build and subsequently maintain AI/ML models, allow back-testing and forward-testing (simulation) of analytics models, and finally, allow deployment and execution of end-to-end solutions so that the results could be thoroughly verified by various users such as data analysts, machine learning engineers, and data scientists before deploying them in production environment.

Scalability and Reusability

Financial crime detection and prevention programs often require the analysis of millions of transactions and associated linkages. Consequently, sufficient capacity and computational resources to handle the data load with acceptable performance is required from data storage and infrastructure. The infrastructure must be able to scale up, without significant performance degradation and latency, to meet the increasing workloads and demands of various typologies and use cases, detection techniques, connecting the dots with external data (from public sources, vendors, regulatory and law enforcement agencies) through entity resolution capabilities, and importantly, the evolving needs of the organisation from growth, risk appetite, control effectiveness, and operational efficiency/productivity standpoints.

Reusability of solution or solution components (code, API, data, etc.) helps to address faster time to value and requires the technical design of the solutions to be modular for ease of integration with other components and systems on a modern technology stack, either on premises or on cloud, and storing key intermediate results such that any new requirement can leverage these data faster and need not process from step 1. Therefore, it is important to apply design patterns, architecture principles, and time to insights trade-offs to consciously build reusable assets during operationalisation of any solution.

Reusability also provides additional benefits by reducing development costs, improving maintainability, and enhancing consistency and accuracy of data consumed and the results achieved.

Flexibility and Ease of Use

The data infrastructure would ideally support a diverse set of products, technologies, and tools that might be required to store, retrieve, compute, visualise, or process AI/ML pipelines for a range of analytics use cases, while remaining easy to use. These could be supported by persona-based training, knowledge sharing, and use of playbooks and guides to help the business and technology users onboard quickly and use them effectively.

Data Analytics Solution Development Life Cycle

The development of a robust model benefits from a defined foundation and a process of continual improvement. The following steps outline a typical model development life cycle, which can be easily adapted for data analytics solutions as well.

1. Model design	Detail the purpose of the model, how it will be used, changes to existing business processes or if a new business process will be created.
2. Data preparation and validation	Document data sources and validate if the data used to generate features for training the model is fit for purpose. Provide justification for the use of any personal data.
3. Features engineering	Document motivations and assumptions made when creating features, features selection process and any transformations applied to the source data.
4. Model Development	Develop the model through an iterative process to ensure the final proposed model shows consistent performance on desired criteria for both the development and validation dataset.
5. Model validation	Validate the model using a holdout data set which is previously not used to train the model.
6. Back testing	Test the model using historical data, to understand how the model performs when the underlying data patterns and distribution change.
[Optional] **7. Forward testing**	Test the model using new data, which is previously not available during development phase, to ensure that the model performance does not deteriorate drastically.
[Optional] **8. Parallel testing**	Conduct parallel testing if the model is meant to replace an existing model, system or process, to compare the differences between the two.
9. Model monitoring	Set up processes to produce dashboards or reports to allow for the monitoring of model performance on a regular basis. Perform model re-training when performance drops below a pre-defined threshold.

There is no ideal or standardised data architecture and solution design that should be followed. The above is used as a reference, to highlight key areas that can be considered in building a data platform (or an area within an existing data platform) that is suited for financial crime analytics.

Effective Data Platform, Data Architecture, and Data Management

It may be of interest for FIs to pay careful attention to this foundational capability and building block to enable highly accurate models and sustain advanced financial crime analytics. The data infrastructure should also take into consideration the resources needed for the data analytics development life cycle and be able to support it.

Data infrastructure typically consists of one or more Data Platforms, which is a collection of best-of-breed tool and technologies to collect, store, process, analyse, and govern data, either on-premises or on cloud, and is typically developed at an enterprise level to support various business functions and use cases. Data Platforms are implemented either as centralised aka *Data Fabric* or decentralised by geographies or functions aka *Data Mesh*, or in combination, depending on the needs, operational factors, and constraints.

The key enablers to the success of a Data Platform comes from two inter-related disciplines – Data Architecture and Data Management and are essential to get the best value out of data. Data Architecture is the blueprint covering the design and structure of FI's data assets/data landscape and defines how data is organised, stored, accessed, used, and disposed in alignment with business objectives through policies and standards. It is strategic in nature and forms the foundation for Data Management. FIs may consider defining a modular data architecture that uses best-of-breed and perhaps, open-source components that can be replaced with new technologies as needed, without affecting other parts of the data architecture.

Data Management focuses on ensuring that data is managed effectively and efficiently throughout its lifecycle, in line with the Data Architecture, through set of policies, processes, procedures, and tools. Essentially, it is the ongoing upkeep through day-to-day implementation of and execution to the data architecture.

There are few key specialised areas within Data Management that FIs may consider paying special attention to and address gaps, if any, through a separate book of work as enablers as listed ahead. These pointers are kept brief to remain within the scope of this paper.

a. Data Integration and Lineage
Data ingestion and transformation pipelines, with data lineages captured automatically. Capturing data lineage allows changes to datasets to be communicated to downstream users.

b. Data Model and Storage
Special attention should be paid to the type of data storage used as performance and cost needs to be balanced, and the design of the data models and accompanying design principles.

c. Data Discovery and Metadata Management
Users should be able to discover datasets by being able to search the metadata captured for each dataset.

d. Data Access, Security, and Privacy
Security controls and processes should be in place, such as tokenisation and access control to name a few.

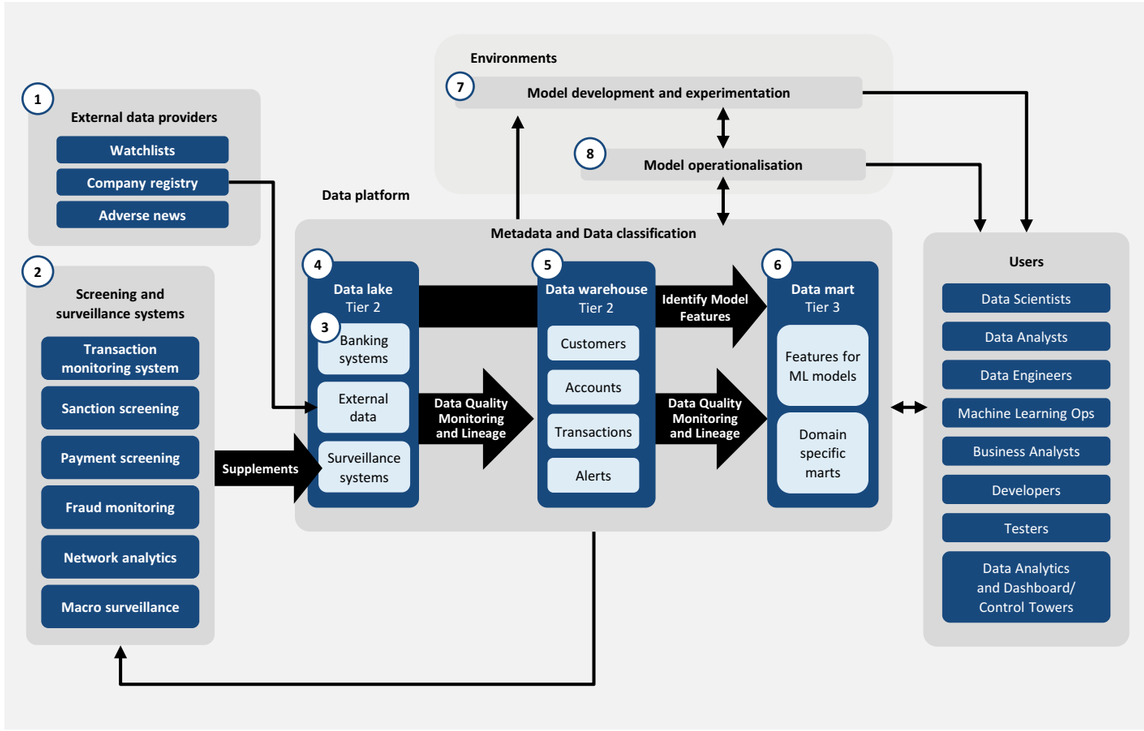
e. Data Quality consists of core dimensions such as Accuracy, Completeness, Consistency, Timeliness, Uniqueness, Validity/Relevancy would encompass any on-going data reconciliation processes between source and target systems. Value based and count based reconciliation between source and target datasets should be in place and performed regularly to ensure data quality.

f. Data Ownership
Clear data ownership should be established to ensure that the data is treated as an enterprise asset.

g. Data Governance including Responsible and Ethical Data usage
Use of the data should be governed and detailed in the previous chapter on Governance.

One advantage in moving to a centralised data platform construct is that data within the organisation can be easily leveraged upon for various enterprise needs, including financial crime compliance objectives. The following diagram provides an overview of a potential data platform architecture.

Data Platform Architecture and Technological Infrastructure for reference



The following section further elaborates the key components of the data:

1. External data providers
Data required for background checks, verify company records, and watchlists/Politically Exposed Person ("PEP") for sanctions screening are often sourced from external data providers, such as company registries and specialised vendor offerings. The extracted data should be evaluated, tested, and managed in accordance with the governance structure.

2. Screening and Surveillance systems
Screening and surveillance systems to conduct monitoring and screening of customers, accounts and transactions stored in the data platform as single validated source as opposed to sourcing from various operational systems in the FI. There may be some performance considerations in using data platform to integrate for some real-time use cases such as Fraud detection and blocking vs. integrating directly. Alerts generated, along with, supplementary data, are then ingested into the data platform, effectively forming a feedback loop and enable other downstream surveillance and analytics systems to use them as a data source.

3. Banking systems
Banking systems refer to the internal, operational product and transactional systems that process and contain customers and relationships, accounts and transactions data of various products, services, and geographies that are of interest.

4. Data Lake (Tier 1)
The data lake (Tier 1) of the data platform is the centralised repository of raw source data for the entire organisation. Raw mirrored data from banking systems, external data providers and surveillance systems are found and stored here, with light transformation for security and discoverability purposes. FIs should conduct regular periodic checks and value-based reconciliation between the original data sources and the organisation's data lake to ensure data completeness and accuracy, while rectifying any data gaps or errors.

5. **Data Warehouse for Financial Crime (Tier 2)**
The data warehouse (Tier 2) of the data platform stores all the transformed data processed from raw data found in the data lake (Tier 1). The Data Warehouse consolidates all customers, accounts, and transactions data, including financial crime alerts generated from applied screening and surveillance systems, for performing monitoring, data diagnosis and analysis of the results. Tier 2 data will be consistently monitored and checked against Tier 1 data to relevant data quality dimensions and metrics are within defined limits and undertake corrective actions to resolve exceptions in a timely manner.

6. **Data Marts (Tier 3)**
The data mart (Tier 3) of the data platform stores relevant data in domain specific marts that are built off the data previously processed in the data warehouse (Tier 2). The data will be tailored to address specified use cases like transaction monitoring and machine learning model features earlier identified with Tier 1 data.
Similarly, between Tier 2 and Tier 1, Tier 3 data will be consistently reconciled against Tier 2 data to ensure relevant data quality dimensions and metrics are within acceptable limits.

7. **Model Development and Experimentation Environment**
Experimentation environment (or Pre-production) is typically required for users to build and test different rule-based and AI/ML models and other analytics solutions to validate results, adjust algorithms where necessary. The experimentation environment is secure with dedicated computing resources so as to not impact production and leverage seamless access to data stored in the data platform layers (T1 to T3) as well as bring necessary custom datasets as required to allow rapid development. Once the models are satisfactory and finalised, they are implemented and operationalised in the production environment (point 8). This is also leveraged for ongoing validation and re-training of the models (refer to table on previous page).

8. **Model Operationalisation Environment**
Production environments is where the approved models from the experimentation environment (point 7) are operationalised and executed at scheduled frequencies, using standard controlled pipelines enforcing governance and tighter security standards. Regular model monitoring and maintenance according to the governance structure should be included in the implementation process.

Key takeaways

A well-defined data infrastructure allows for the continual development and improvement of data analytics solutions for the detection and prevention of financial crime. The ideas discussed in this section serve as guiding principles that businesses may adopt in their development of data analytics solutions.

Experimentation and deployment
Businesses can reference and adopt a development cycle for their data analytics that incorporates testing and retraining. Testing with production data in an experimentation environment may provide greater improvements and the cycle of testing and retraining can help the data analytics solution stay current or ahead of financial crime.

Scalability and reusability
As businesses grow and financial crime evolves, the need for advanced data analytics calls for the corresponding need to upscale data infrastructure in order to support the increasing resources and demands from the implemented solutions. A good data infrastructure allows for this upscaling with minimal performance latency. Businesses may also keep solution components reusable for faster time to value and greater ease of integration with other systems in their technology stack. This may see reduction in development costs while improving maintainability and consistency in results achieved.

Effective data platform
A sound data infrastructure is dependent on the data platform it builds off. The key to a successful data platform is an appropriate data architecture that adequately supports business objectives, and which serves as the foundation for data management, including the effective and efficient utilisation of data through the lifecycle of data analytics development.

2.2.3 Governance and model risk management

As FIs adopt more technology and data analytics to combat financial crime, they are collecting and processing a larger volume and variety of data, ranging from sensitive and confidential personal data to financial transactions. This increasing use and reliance on data has in turn generated greater regulatory focus on the need to address data and model related risks.

The changing landscape has also resulted in the emergence of new risk areas like protection of intellectual property, data privacy, data sovereignty, proof of identity (using biometric information such as voice, fingerprint, or facial biometrics), and data provenance. To effectively address new risk areas and mitigate potential losses, regulators and FIs have been proactively refining existing data and model governance guidelines and frameworks. The increasing adoption of data analytics and machine learning models have also urged FIs to focus on model governance which formalises model risk management activities for implementation.

In Singapore, the key focus areas relating to the adoption of data analytics include concerns over fairness, governance, accountability, transparency, and data protection. With the above driving forces, it is necessary for FIs to re-examine existing governance and model risk management structures.

Governance framework

FIs need to have a basic governance framework in place.

A basic governance framework is essential to ensure that analytics tools or machine learning models used by FIs generate reliable insights.

This is particularly important in the areas of data management, analytics, and model design.

- Data governance is a fundamental component consisting of policies, procedures, and standards of data management, that are applied throughout the data lifecycle, covering collection, processing, storage, and disposal (refer to Section 2.2.3 Data Infrastructure and Solution Design).
- Model governance is responsible for controlling access to models, and monitoring model related activities, including model changes, model outputs, and compliance with regulatory requirements.

In establishing the basic governance required for an analytics solution, essential considerations include the:

- Purpose of data analytics solutions
- Business process and impact analytics
- Roles and responsibilities
- Documentation of data sources, analytics solutions, controls, and monitoring

With a basic governance framework, FIs could then customise each component based on the purpose, complexity and risk of solution to achieve optimality.

As data analytics solutions vary, there is no single standardised governance framework that is suitable for all data analytics solutions. Similarly, the level of governance applied to each solution can be tailored based on the complexity and risk of the data analytics solutions or models. These two components, complexity of the solution and risk associated, also forms the basis for considering the types of model risk management activities to be implemented.

To determine the level of governance required, it is critical for FIs to conduct comprehensive assessment of the solution prior to any development. A possible approach is to first classify the solution based on its nature and complexity, subsequently evaluating the materiality of the solution by assessing potential associated risks. Depending on the materiality, FIs could then decide on the level of governance required, which in turn impacts the risk management components and level of approvals required.

With evolving risks and changing regulatory landscape, FIs may need to conduct periodic reviews and model validation to assess the adequacy and risks of the proposed solution or model, customising the level of governance to address concerns brought up during the reviews.

Data analytics solutions can be classified into three distinct types: visualisation, rule-based, and machine learning models.

Data visualisation

- Display of datasets and statistics from transformation or aggregation in the form of charts, networks, colours to highlight patterns and trends.

Rule-based analytics

- Use of logical formula around known thresholds and controls to provide a binary result.
- Often expressed as "if this then that".
- Created and tuned by Domain Experts to generate risk insights.

Machine learning models

- Use of supervised, unsupervised, semi-supervised, or reinforcement learning models trained with historic data.
- Can be trained with existing AML/CFT typologies to identify suspicious activity or generate new findings from datasets directly.
- Subsequently applied to new data to provide forecasting or predictive analysis.

FIs may note that more complex analytics tools require more effort for calibration and review and may not always be ideal depending on the use case. In the implementation of complex solutions, FIs will find assurance in applying the appropriate level of governance to assure reliability of insights generated.

Risk classification

To determine the risk of a data analytics solution, the types and likelihood of impacts of the solution, and the level of autonomy it has must be considered.

The impact of a solution refers to the consequences from decisions made using the solution. The purpose and use of an analytics solution would determine the types, severity, and probability of impacts faced. For example, an improperly managed or unethically used data analytics solution for customer segmentation could potentially have a greater negative impact than a solution intended to enhance internal efficiency.

The autonomy of the solution refers to the extent of human intervention in the model's output. While extensive human intervention increases the risk of human error or reduces the cost benefit to implementing an analytics solution, too little human intervention might give rise to a "black box" situation where the reviewer cannot understand the model's output.

Assessing the impact type and likelihood, as well as the autonomy of a solution helps FIs to identify and prioritise associated risks and develop targeted mitigation plans. Consequently, an appropriate level of governance can be applied to address these risks and ensure explainability of the model.

The table below details the two considerations that make up the risk of a data analytics solution:

Impact and likelihood	Accessibility to a particular product or service by customer
	Inequality towards individual or groups of customers/employees
	Potential economic loss
	Potential risk of litigation, regulatory warning, material dispute or reputation loss
Autonomy	Human in the loop: Humans can influence or determine the outcome
	Human out of the loop: Humans are not involved in the entire process and do not influence the outcome

Increasing transparency of "black box" machine learning models

Machine learning models are often regarded as "black box" due to complex mathematical concepts behind them and non-transparent internal mechanisms, resulting in unexplainable outputs. Additionally, models are often layered upon each other or used in unison, which complicates the analytics, rendering generated outputs harder to explain and justify. As such, it is essential for FIs to establish a comprehensive understanding of the models in the decisioning process and generation of outcomes, to achieve confidence in using the model and provide assurance to regulators.

Clarity of the model's output and decisions made, otherwise known as "model explainability", is critical for governance and model risk management as it defines how much of the model's decision-making capabilities can be explained and assessed by FIs.

For example, a good understanding of how the analytics models make decisions help FIs explain why model outputs meet business requirements and cannot be misused. FIs would also be aware of solution limitations, allowing them to identify and understand risks associated with the model and enact target mitigation plans to ensure output quality.

Ethical use of AI

As machine learning models are increasingly being embedded into decision making processes, prudence must be applied to ensure that such models are used ethically. When left unchecked, potential unethical use of data could exist, leading models to inadvertently start discriminating or being biased towards certain groups of people.

In response to the potential risks, MAS issued in November 2018 a set of principles covering fairness, ethics, accountability, and transparency ("FEAT"), to promote responsible use of artificial intelligence and data analytics in finance[5]. While these guiding principles are not regulatory obligations, it is vital for FIs in Singapore to adhere to them to foster trust with the public and regulators in the use of advanced analytics in AML/CFT solutions.

These principles can supplement the existing governance established by FIs in the form of assessment and peer review. In current and future technology implementations, assessments in accordance with FEAT principles can help ensure sufficient coverage and assurance in the use of the models. Further, a peer review process assures the quality and integrity of the adoption of data analytics. This layer of check helps FIs reap the benefits of utilising data analytics while maintaining appropriate oversight.

Key takeaways

The governance strategies and structures discussed in this section form the industry's best practices which FIs can begin employing to enhance their AML/CFT risk surveillance. Dependent on the firm's objectives and/or compliance standards, FIs might benefit from a governance plan that aligns to the risk and complexity of the analytics solutions used to ensure effectiveness and suitability to operational needs and regulatory expectations (e.g., business scale, supporting resource availability, etc.). The following summarises the attributes of a well-defined framework:

Governance
FIs could consider establishing a well-defined governance structure with clear roles and responsibilities, reporting paths and documentation of the solutions. FIs may be keen to tailor the level of governance depending on the identified risks.

Ethical use of AI
Apart from the model performance of AI/ML models, FIs might want to consider the ethics of the model to ensure it results in no unintended harm. This can be achieved by incorporating FEAT principles and peer assessments into the process of developing an AI/ML model, thereby boosting the confidence of not only the end users but also regulators.

[5] Source: MAS information paper - Principles to Promote FEAT in the Use of AI and Data Analytics in Singapore's Financial Sector

2.3 Applying data analytics to combat financial crime

Overview

The application of AML/CFT analytics throughout the industry remains varied in degrees of adoption, largely focused on enhancing traditional AML monitoring practices like Know-Your-Client ("KYC"), customer risk rating, and account and transactions monitoring.

According to Financial Action Task Force ("FATF") in 2021[6], one of the main challenges in effective implementation of AML/CFT measures is the inadequacy in identifying, assessing, and managing risk at a large scale, with majority of FIs basing their efforts on a combination of automated and static analyses on pre-selected risk factors, supplemented by human review. With greater use of analytics, advancements made to previously exploratory technology solutions[7] and development of FI-to-FI information sharing platform – Collaborative Sharing of Money Laundering / Terrorism Financing (ML/TF) Information and Cases ("COSMIC"), the industry is ready for newer AML/CFT risk surveillance methods to detect and facilitate investigations into fund flows and monitor customer activity at any scale efficiently and effectively.

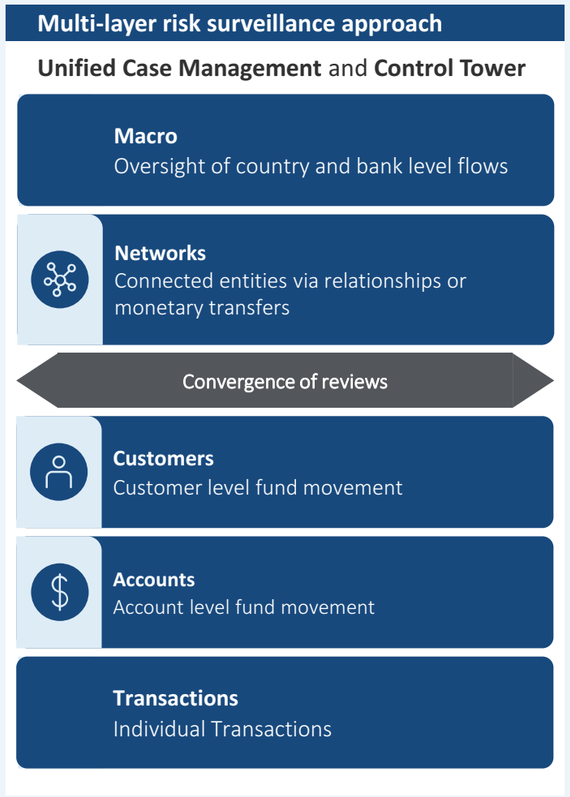
This section will cover:

1. The emergence of a multi-layer surveillance framework, thereby enabling transactions monitoring to be more dynamic and effective; and
2. A multi-pronged approach to better manage financial crime risks, with the example demonstrating how this has been leveraged to manage indirect sanctions risks.

2.3.1 Multi-layer risk surveillance

The Multi-layer risk surveillance is a conceptual framework that leverages on data analytics to perform monitoring and detection across multiple levels of aggregated activity.

Rule-based or analytics models at the transaction, account or customer levels enable FIs to analyse and detect behaviour and transaction patterns via the bottom-up view, focusing on customers with unusual fund movement or behaviour. Adding capabilities at the network and macro level then combines overall money flows of groups and individuals and provides insights at the broader level which are not apparent when entities are being reviewed in silo. The convergence of these multiple levels of surveillance can allow rapid detection of scaled activity by triangulating the top-down and bottoms-up views. The following diagram provides an overview of such a framework.



[6] Source: FATF paper: Opportunities and Challenges of New Technologies for AML/CFT (fatf-gafi.org)
[7] Source: ACIP 2018 paper: Industry Perspectives - Adopting Data Analytics Methods for AML/CFT (mas.gov.sg)

2.3.1.1 Customer, accounts and transactions

The foundation layer towards achieving a holistic surveillance framework involves the continuous assessment and monitoring of customers, accounts and transactions. These include conducting KYC and risk assessments as part of customer due diligence ("CDD") prior and during customer onboarding, transaction monitoring, as well as name and payment screening post customer onboarding. Acknowledging the importance, most FIs have already established a basic governance framework for the systems and processes involving KYC/CDD, TM and screenings. As these systems and processes are performing in an optimal level for effectiveness, FIs are focusing on the next step to build upon the basic governance framework and further enhance using AI/ML to drive efficiencies.

One example is the application of AI/ML on TM alerts to reduce false alerts and prioritise alerts for investigation. The approach reduces time and effort on potential false alerts without significant manpower increase by redirecting resources to focus on higher priority alerts which are more likely to be true while lower priority alerts may be treated with slower review processes or hibernation. Where hibernation is applied, FIs should consider the potential of new risk concerns and if they have sufficient controls or safeguards against such an occurrence.

Additionally, with the efficiency gained from the application of AI/ML, FIs can better reinvest their resources and efforts into monitoring on a wider scope in the network and macro levels, forming a more holistic surveillance.

2.3.1.2 Network level surveillance

The network level primarily leverages techniques such as network analytics. Network analytics is amongst the more advanced technologies adopted by FIs in their fight against financial crime. Previously identified to be in early stages of adoption in the 2018 AML/CFT Industry Partnership ("ACIP") data analytics paper, this technology is now a mature and well-established tool leveraged by most members of the Working Group ("WG"). In the region, the role of this technology in network analytics was also explored by the Hong Kong Monetary Authority together with Deloitte with the paper "AML Regtech: Network Analytics" in May 2023. This paper details how network analytics have been applied within AML/CFT and support the banking sector[8].

Network analytics has been critical in strengthening FI compliance with regulatory requirements like MAS Notice 626 Para 6.22 (b), which states that Banks need to put in place adequate systems to detect complex and unusual patterns of transactions.

The application of this technique has been cited most frequently in the investigations process. In the past, the process for reviewing customer counterparties and connected parties was often siloed, rendering a concurrent review of multiple customers challenging.

With network analytics, FIs can combine transactions of multiple customers, and overlay connected party relationships and adversity indicators on both counterparties and connected parties regardless of the number of parties involved. In other instances, it has helped FIs to detect networks and interconnections among bad actors, changing or new relationships, and rapidly evolving criminal behaviour.

To facilitate network analysis, network visualisation tools are developed to address two key challenges of manual analysis of networks for AML/CFT investigations:

1. Repetitive collation and assembly of data from multiple data sources
2. Establishing interconnections amongst data (e.g., common counterparties, common addresses) in the network.

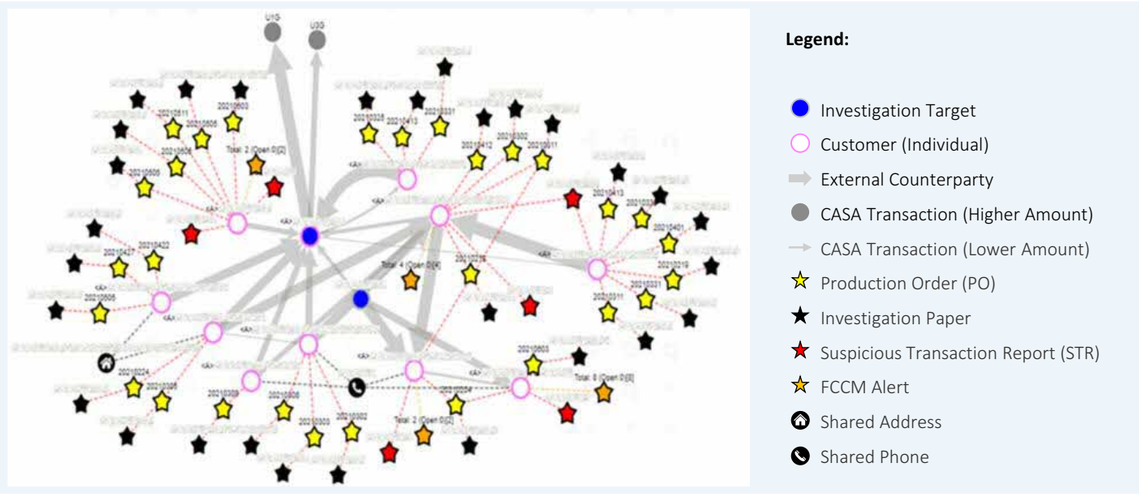
With the development of such tools, analysts can better understand typologies, thereby encouraging FIs to develop more complex data analytics approaches to detect those typologies at scale. For these reasons, many banks have developed in-house network analytics to facilitate investigations.

One member bank built a network analytics tool with the following data points:

1. Latest customer profile information
2. Transaction data with a rolling lookback window of 5 years
3. Suspicious transaction reporting ("STR") data, production orders ("PO") data and transaction alerts monitoring data ingested from a data pipeline.

The tool allows investigators to concurrently query for more than one subject of interest along with corresponding data of counterparties and or connected parties up to two hops away from the initial subjects of interest. The visualisation generated elucidates networks surrounding the subject(s) of interest based on customer-to-customer relationships and transactional relationships, and reveals interconnections and adversities (e.g., STRs, POs) among the subject(s) of interest and their relevant counterparties.

[8] Source: Hong Kong Monetary Authority: AML Regtech: Network Analytics (hkma.gov.hk)



The above diagram provides a sample view of the platform.

Further, the platform allows investigators flexibility to customise the network visualisation query to the investigation's needs. For instance, implementing relevant transaction amount filters to de-noise the visualisation generated, or filtering the network to identify nodes that have had past STRs filed on them.

Additionally, the bank also integrated adversity-relevant red flags into the visualisation to aid investigators' in capturing such insights. For example, a red flag indictor for misuse of legal persons typology was incorporated into the platform, namely displaying common addresses shared by more than one node (entity) in the network.

Network analytics is used alongside traditional reviews, as an additional layer of investigation used to detect interconnections among customers and external counterparties that may have otherwise been missed through siloed review of customers.

With network analytics, the bank observed greater efficiency - almost 60% reduction in time spent on a sample complex investigation case. In traditional manual reviews, investigators required long hours to query several systems for relevant customer and transaction information such as information related to same name matches on non-bank counterparties. Collated information is then overlaid to identify common connections among multiple subjects of interest and adversity indicators in the network, which is also time consuming.

These manual and tedious work steps have been streamlined to a few clicks on the network analytics tool, improving efficiency in conducting investigations.

More importantly, the bank saw greater effectiveness in the discovery of interconnected customers beyond Level 2 (or Level 1 subject's immediate connected parties / counterparties) in a complex network. Such interconnections would not have been easily detected in silo reviews of each customer, particularly as the number of subjects of interest increases, cross-referencing counterparties and connected parties in the network becomes highly complicated, resulting in the detection of nexus between parties to become exponentially more difficult.

In addition to greater efficiency and effectiveness, a key benefit of leveraging such technology is in the filing of higher quality STRs. The visualisation of linkages between parties of interest and the ability to query on a party of interest enables the FI to identify complex networks and/or hidden linkages to support in the STR filing process.

This is just one example in which network analytics has proven beneficial in helping FIs and reviewers combat financial crime in terms of the speed, scale, and materiality, allowing FIs to detect and identify any unusual transaction patterns at a faster rate compared to traditional manual reviews especially when complex networks are involved, and to learn and respond to any new typologies quicker.

2.3.1.3 Macro level surveillance

The macro level leverages aggregation of customer payments to monitor flows at a macro-level, for example between countries and/or banks, to identify significant deviations from regular flow patterns. For instance, the quantum (example by volume, value or frequency) of transactions with countries or institutions of concern can be determined and alerted should they exceed pre-determined thresholds. Suspicious patterns can then be investigated by expanding on key players or networks, periods, or specific transactions for granularity.

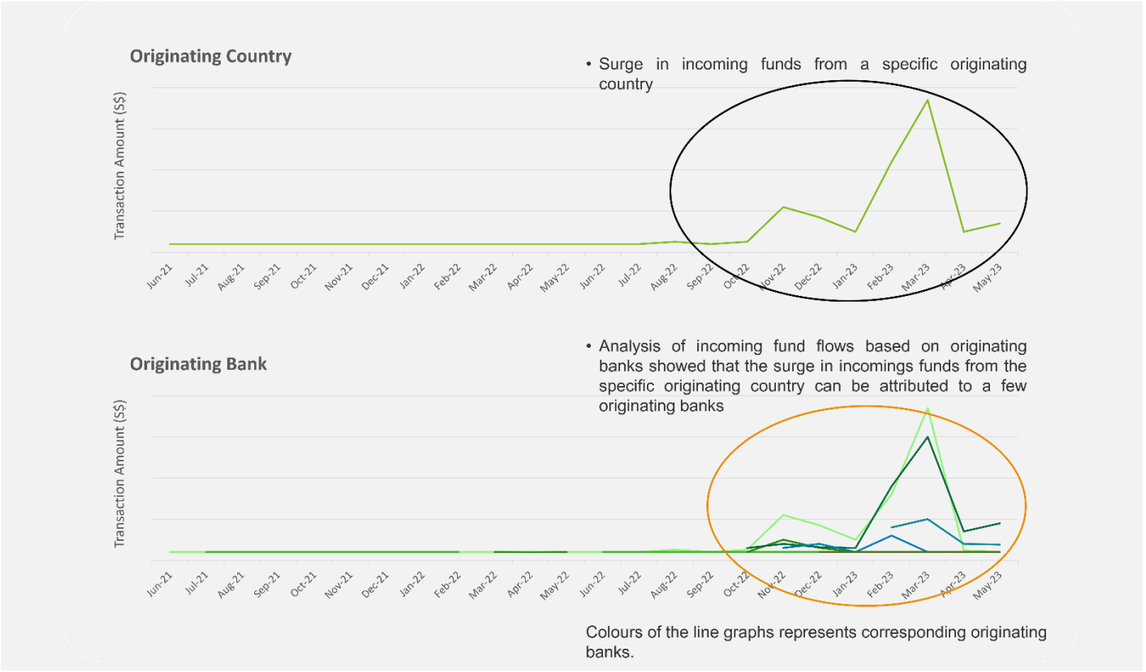
This surveillance can manage risks relating to:

- Sanctions Evasion
- Tax Evasion
- Money Mules

The data elements that make up key features of a macro framework or dashboard example below are:

- Inbound vs outbound flows
- By country and by bank
- By transaction currency
- Historical trend view and period-on-period views
- Pre-established thresholds that will trigger highlighted flows for review and investigation. For example, one bank adopted the use of Inter-Quartile Range (IQR) to flag out anomalies for monthly review
- Ability to drill down to top customers contributing to the highlighted flows

To see how macro-level surveillance can be applied operationally, please refer to use case 2 of section 2.4.2, “Misuse of Legal Persons”.



2.3.1.4 Unified case management

Today, transaction monitoring, name screening, client due diligence and fraud detection each often have their own case management system that do not communicate with each other. Enabling a shared data platform, both across these risk domains and across business units, would allow for consolidation of data across different banking systems within the same bank using a common unified case management. A unified case management provides banks with enhancement opportunities across key functionalities within the AML ecosystem.

Moving to a common unified case management system has a few advantages:

1. Increase in effectiveness by having a holistic view of a customer’s risk profile (for example when reviewing a TM alert, the reviewer would be able to know if the customer was flagged in any fraud alerts or CDD review).
2. Increase in efficiency by being able to leverage on the reviews and background work performed in other risk surveillance processes.
3. Data enrichment, where data from the case reviews and conclusions can be used to enhance, train, build and validate analytical models.

2.3.1.5 Control tower

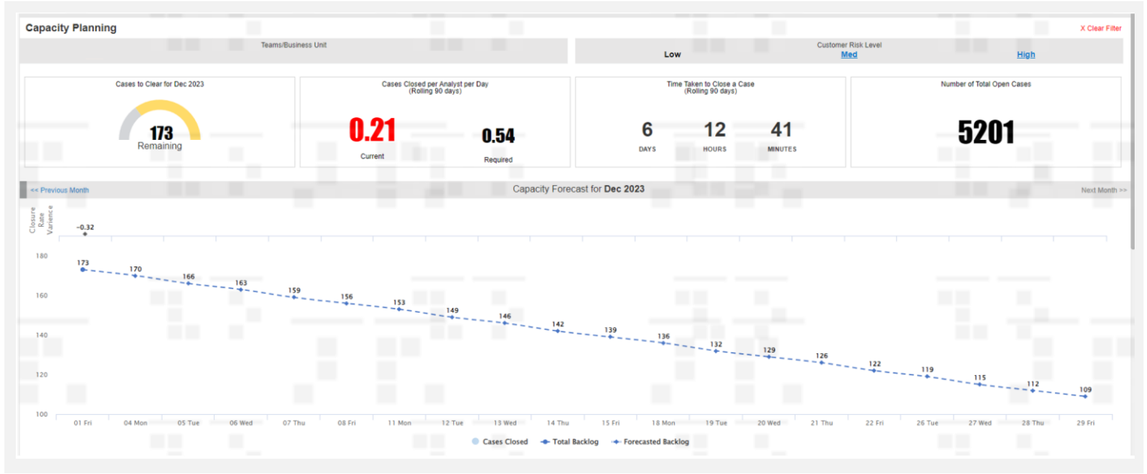
Taking a page from how control towers are used to manage air traffic at an airport, a control tower is about near to real time monitoring of financial crime risks using leading, lagging, and operational indicators. These indicators measure clear effectiveness metrics and are used to drive decision making and intervention when necessary.

A control tower enables an organisation to take control of its complex business processes by making timely decisions and actions to achieve the desired business outcome.

It is a visual presentation of:

- Outcomes to driver linkages.
- Alerts to breaches of threshold or changes to trends etc that require attention, decision or intervention.
- Allows the user to trigger or undertake corrective actions.
- Provides a robust feedback loop to take in responses to changes introduced, planned or unplanned.

The example below shows operational indicators in the control tower dashboard which help team leads make necessary interventions or adjustments to resources, ensuring sufficient capacity for the forecasted alert volume.



2.4 Selected use cases – Solutions adopted to address risk areas identified

Based on the collective experience of implementing technology solutions to combat financial crimes, the working group has selected six risk areas where the implementation of such solutions has been most effective, and which have a higher level of maturity in terms of adoption across the member banks.

These use cases highlight examples of solutions adopted to address each risk area, detailing measures of effectiveness and efficiency, and challenges that each member bank overcame to successfully introduce each model.

2.4.1 Sanctions evasion

Sanctions are a widely used tool employed by governments and international organisations to address various issues, including terrorism, nuclear proliferation, and destabilising actions by some nations. Sanctions evasion is the intentional act of attempting to remove or conceal the involvement of sanctioned places, entities or individuals in a transaction or series of transactions. FIs should consider identifying and assessing the risks of sanctions evasion when dealing with their customers and take appropriate mitigating measures commensurate with the level of risks identified.

Given the complexity of transactions, it is important for an FI to understand that their exposure to sanctions evasion risks may not be only via direct clients but also their counterparties and owners or controllers. Case study one explores the use of technology to combat this problem. Meanwhile case study two considers the complexity of sanctions considering not solely the customer but also their location in determining whether certain sanctions may apply.

Use case 1

One member bank has built and leveraged a network analytics tool to facilitate investigations linked to sanctioned entities.

Overview and the problem statement

Whilst FIs perform due diligence on their customers, it may not always be sufficient to rely solely on first layer sanctions screening of customers without considering the business activities and risks posed by their counterparties in assessing potential sanctions evasion risks. For a customer with a known sanctions nexus, it might be useful in certain circumstances to perform a lookback review of its network of counterparties up till two hops away from the subject of interest, holistically considering red flags and interconnections in the network.

One member bank had identified a customer, "Company X", through hits in the bank's IP address monitoring system. Company X was detected to have performed multiple login attempts via an online banking from a sanctioned country, which suggested business associations or dealings in or with the sanctioned country. During the review, Company X was further noted to have received payments from third party entities. As a next step, the bank was required to further investigate the third-party payors.

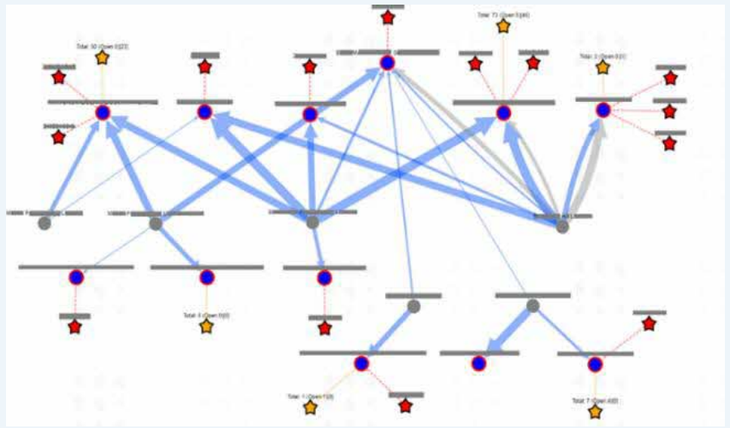
The bank noted from the review that one of the third-party payors, Company A is a Sanctioned Person by an applicable sanctions authority for knowingly engaging in a significant transaction for the purchase, acquisition, sale, transport, or marketing of petrochemical products from a sanctioned country. The transactions between Company X and Company A took place prior to the designation.

Two other counterparties, "Company B" and "Company C", were observed to have made significant transactions to Company X and shared the same address as Company A. All three counterparties shared common traits such as having a sole director and shareholder of the same nationality, no online presence, and similar name structure, although they are not known to share any common ultimate beneficial owners or connected parties.

With the help of network analytics for the investigations, the bank was able to further identify counterparties of these third-party payors, which were also triggered for review. As the three counterparties of concern were non-bank customers, the use of the network analytics tool has helped the bank to visualise their Level 2 network (i.e., the Level 3 network relative to the very first subject of interest).

What was the solution selected?

The member bank leveraged an internally developed network analytics tool, to further investigate the third-party payors. The diagram below shows a brief overview of the tool.



Legends:

- Investigation Target
- ECustomer (Corporate)
- External Counterparty
- CASA Transaction
- SWIFT Transaction
- C2C Relationship
- Suspicious Transaction Report (STR)
- FCCM Alert

Measures of effectiveness

The tool is effective in helping the bank facilitate investigations relating to sanctioned entities. The bank was able to proactively identify suspicious connected parties who were non-bank customers. Further, it is worthwhile to note that the two non-sanctioned counterparties of concern, Company B and Company C, who were identified and added to the bank's internal watchlist, were subsequently featured in public sanctions-related adverse news published more than a year later, where they were reported to be front/shell companies controlled by a sanctioned bank as part of its sanction evasion network.

In addition, the bank was also able to identify other suspicious counterparties of non-customer third party payors, resulting in a total of eight STRs filed on Company X and seven other customers that were observed to have transacted with the three counterparties of concern.

Measures of efficiency

The use of network analytics brings greater efficiency to the member bank by reducing two days of manual trawls (including data preparation, trawl time and review of results) to only one hour, which translates to an approximate 93% reduction in turnaround time.

What were the challenges and considerations whilst adopting the solution?

The use of network analytics brings greater efficiency to the member bank by reducing two days of manual trawls (including data preparation, trawl time and review of results) to only one hour, which translates to an approximate 93% reduction in turnaround time.

Use case 2

Another member bank applied data analytics to flag mismatch in domicile and digital footprints linked to sanctioned countries.

Overview and the problem statement

As internet banking is becoming the norm, customers no longer must be physically present in bank branches to perform transactions or raise requests regarding accounts. It becomes increasingly difficult but important for FIs to take proactive approach to detect any discrepancies in the customers identity.

One proactive approach is to detect potentially suspicious deviations between client declared domicile and domicile access data captured in the electronic banking platform based on the resolved geolocation of the clients' IP address, specifically with respect to potential accesses from-sanctioned countries for the purposes of detecting sanctions evasion.

What was the solution selected?

The member bank collected IP address data from the electronic banking platform and resolved customer information, including details of their declared domicile/tax domicile. The IP address data for each user account was resolved to country-level geolocation data, using an accurate external vendor reference data set. The external reference data set also enabled the bank to identify if IP addresses were using Virtual Private network ("VPN") or other IP address anonymisation services by matching the virtual addresses, allowing for such accesses to be discounted for the purpose of the review.

The data was then compared between the declared country domicile and the matching IP address country domicile data. Additionally, rules were also set to identify any platform accesses from any sanctioned countries.

Measures of effectiveness

For cases with any such accesses, the bank performed reviews on the customer profile to identify if there was any declared indirect (i.e., through connected parties) connection to sanctioned countries. In cases where a previous connection was declared and assessed, data was reported for information purposes only. Any new information on sanctioned countries accesses where a previous connection was not assessed were escalated on an urgent basis. A report was produced to set out the details of all hits, included client details, electronic banking platform access details, associated transaction details, and control assessment findings.

No 'true positive' cases (i.e., any sanctioned countries access was by a party without any known nexus to any sanctioned countries), were detected in the member bank's Hong Kong or Singapore locations. The control frequency was progressively reduced to perform only on a six-monthly basis as continued results were cleared without having any such risk, and as the sanctions risk landscape stabilised. Further, the bank has also expanded the solution for use in other locations with electronic banking platforms where some true positives were detected and promptly managed.

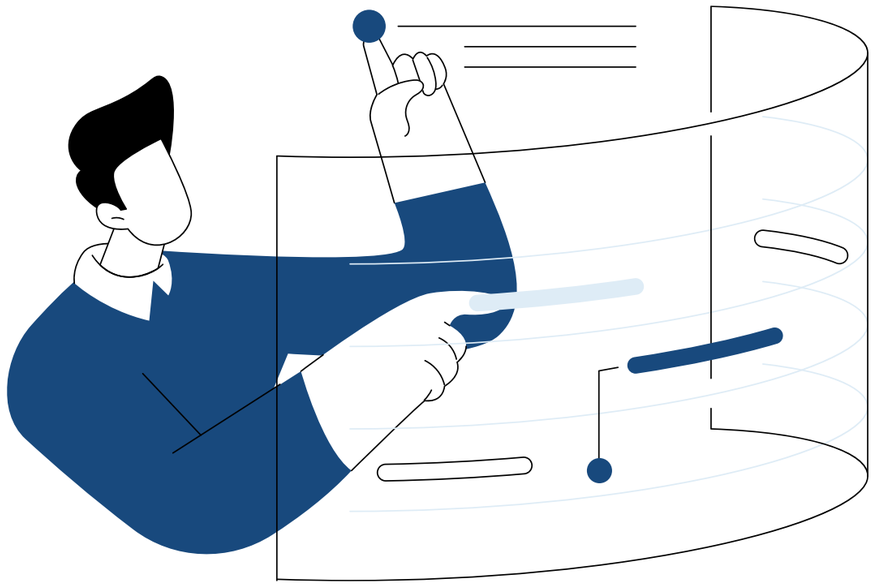
Overall, this newly devised control helped to provide additional comfort regarding potential sanctions exposure in response to the rapidly developing risk environment and proved to be a useful compliment to expand on the existing framework of sanctions controls.

Measures of efficiency

Results were relatively straightforward to identify – an automation workflow was established to perform the analysis on a recurring basis. The additional source of information regarding client domicile also enabled more rapid identification and review of key client sanctions risk scenarios.

What were the challenges and considerations whilst adopting the solution?

No challenges identified. This control development was a relatively straightforward pivot from the existing capabilities established to review IP address geolocation from the electronic banking platform, hence the bank was able to seamlessly deploy the solution. Due to this capability foundation, the control was able to be rapidly deployed in the context of an emerging sanctions circumvention risk landscape. For example, after the commencement of the armed conflict in Ukraine.



2.4.2 Misuse of legal persons and legal arrangements

Shell companies or shell corporations are companies that do not have significant assets, active operations, or employees. They are primarily used to make transactions acting in a pass-through capacity or facilitating cross border activity. Where there are complex structures and or secrecy laws in place, owners of such companies could be obfuscated and difficult to establish.

Shell companies can be misused for illegitimate purposes such as money laundering, terrorist financing, fraud, sanctions, and tax evasion. Modern technologies such as machine learning and big data analytics can assist FIs to identify shell companies and prevent illegal transactions, as evidenced in the use cases below.

In June 2023, the MAS published an information paper on the Effective Use of Data Analytics to Detect and Mitigate Money Laundering/Terrorism Financing Risks from the Misuse of Legal Persons[9], which further emphasised the importance of combating this risk by deploying data analytics.

Use case 1

One member bank has applied data analytics to identify misuse of legal persons and legal arrangements.

Overview and the problem statement

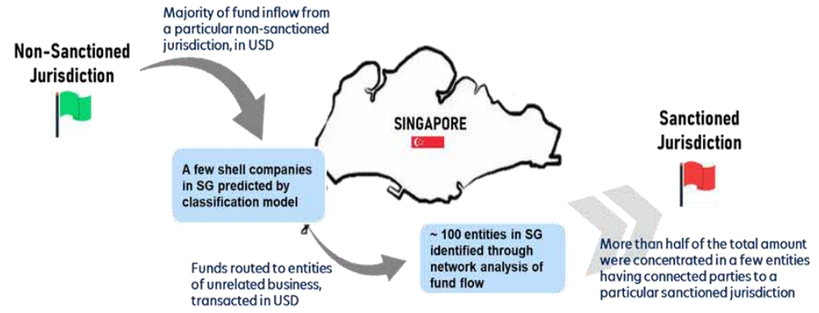
The misuse of legal persons to facilitate third-party payments to sanctioned countries is an illicit practice of concern to member banks. One method is to establish legal persons in non-sanctioned jurisdictions, commonly known as "shell companies." These shell companies facilitate transactions that would otherwise be subject to scrutiny or outright rejection. By exploiting complex ownership structures, nominee directors and false documentation, perpetrators conceal their involvement in illicit activities, making it difficult for authorities to trace and prosecute them effectively. The consequences are far-reaching as it enables sanctioned countries to bypass restrictions, gain access to funds, and potentially finance illegal activities. Additionally, the misuse of legal persons undermines the integrity of financial systems.

What was the solution selected?

The solution involved the application of classification models and network analysis.

1. Classification models to predict the likelihood of shell companies are developed using various machine learning techniques (such as random forest and eXtreme Gradient Boosting ("XGB") by analysing data on financial transactions (domestic and cross-borders), company information and ownership structure, demographic profiles of related individuals and connections to confirmed shell companies. The models produce a risk score for each customer and those with the highest risk scores are prioritised for investigations.
2. Network analysis of fund flows can serve as a second layer of analysis to detect common networks of transaction counterparties by entities with high-risk scores. Unusual transaction patterns and abnormally high value of flow-through funds with counterparties of different or unrelated business natures, raise further red flags for illicit shell company activities. Where the funds are largely transacted to entities with connected parties to sanctioned jurisdictions or having notable dealings with other entities from sanctioned jurisdictions, the suspicion of circumventing sanctions is further increased.

In one specific case shared by the member bank, the combined use of risk scoring models and network analysis resulted in the detection of a few high-risk shell companies (as predicted by the classification model) having received funds from a common non-sanctioned jurisdiction over a period of several months. The funds were subsequently transacted to close to 100 entities in Singapore of unrelated business nature. More than half of the total amount transacted was concentrated in a few entities who had connected parties from or notable dealings with a particular sanctioned jurisdiction. Several other red flags were uncovered during investigation, leading to STR filing on the high-risk shell companies and several of their transaction counterparties.



Measures of effectiveness

The findings led the member bank to file multiple STRs for shell companies predicted by the classification model as well as several of the transaction counterparties that were uncovered through the network analysis and connected to a sanctioned jurisdiction.

Measures of efficiency

The network analysis enabled investigators to quickly identify entities having significant transactions with the high-risk shell companies predicted by the classification model, some of which transacted with more than one high-risk shell company. Such insights helped prioritize investigation efforts, for the large number (~100) of entities identified to have undertaken such transactions.

What were the challenges and considerations whilst adopting the solution?

While the concept of suspicious transactions between entities of unrelated business nature is intuitive and easy to understand, it is often difficult to determine based on transaction data alone. Consequently, the bank still needed to incorporate human expertise and judgment in the review process.

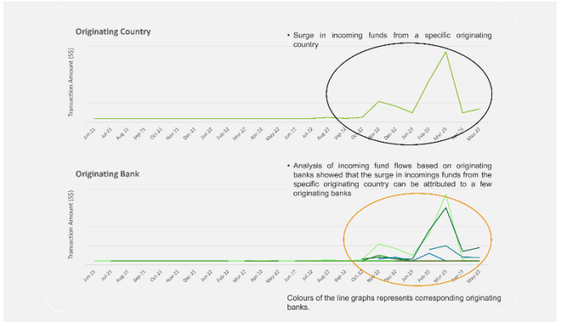
[9] Source: MAS June 2023 Paper – Effective Use of Data Analytics to Detect and Mitigate ML/TF Risks from the Misuse of Legal Persons – effective-use-of-data-analytics-to-detect-and-mitigate-mltf-risks-from-the-misuse-of-legal-persons.pdf (mas.gov.sg)

Use case 2

One member bank utilised macro-level surveillance to detect anomalous flows in payment corridors.

Overview and the problem statement

The misuse of legal persons presents a significant risk concern for potential money laundering and other illicit financial activities. The detection of anomalous flows at an individual account level represents a challenge with the bank's day to day monitoring system as the flows would not be apparent.

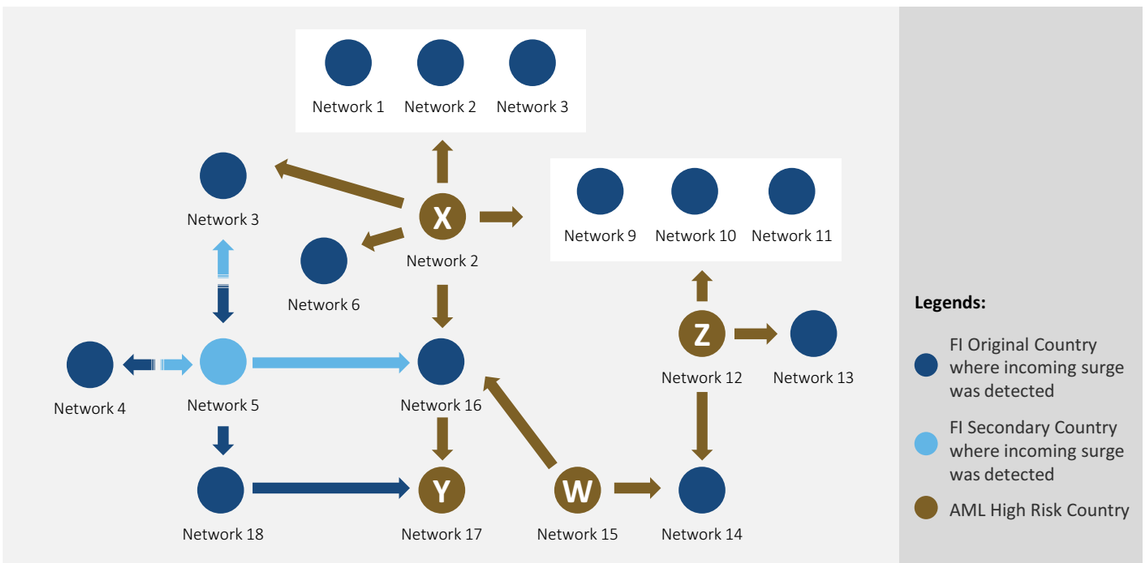


Macro-level Surveillance

What was the solution selected?

The solution involved the use of macro level surveillance and network analytics.

1. The use of macro-level surveillance enabled the bank to conduct macro/country-level monitoring of payment flows to identify unusual and large fund spike flows from network of bank accounts in a particular country ("originator") to Singapore, resulting in a disproportionately large value of inflows and outflow through Singapore over a short span of time.
2. Through the use of network analytics, the bank also noted that many of the entities appear in different networks involving different bank accounts from the originator country.



Measures of effectiveness

The member bank successfully detected a surge and persistent uptrend in incoming transfers from two countries with high AML risk, in the span of a quarter, which was unusual as typical flows with these countries are very low.

Further review at the bank level revealed that transfers originated from a select few banks prompting a comprehensive review of the bank customers involved in these transfers. Leveraging network analytics, fund movements and relationships were further traced to a large network of 50 customers and extended fund flows to another two countries with high AML risk. In addition, some entities of concern who were customers of the bank in another location were identified to be linked with the original network and have also been exited as a result.

Measures of efficiency

Through macro-level surveillance and network analytics, the member bank was able to speed up intervention and detect the large network and spike flows within months of its existence.

What were the challenges and considerations whilst adopting the solution?

While performing this solution, the member bank had to consider data privacy and sharing considerations across geographies.

Hence, the initial assessment leveraging key indicators was performed at Group level where reviewers can access dashboard information across locations. In close partnership across geographies, deep-dive reviews were conducted in the respective countries and outcomes are surfaced back centrally to allow for a holistic view across the region.

2.4.3 Trade fraud and trade-based money laundering

The international trade system is subject to a wide range of risks and vulnerabilities that can be exploited by criminals, for example, the volume, complexity, and reliance on physical documentation. Trade-based money laundering is the process of exploiting these vulnerabilities to disguise illegitimate proceeds and move them through the international trade system. This can be achieved in various ways including misrepresentation of quantity, quality, or price of the goods and the use of existing criminal or trafficking networks. Use case one explores how analytical platforms can be utilised to detect fraudulent behaviour at the client level. Use case two explores how network analysis can be utilised to identify hidden relationships and criminal networks.

Use case 1

One member bank has developed scenarios to detect fraudulent behaviour using analytical tools and network analysis.

Overview and the problem statement

Trade Fraud Risk Management in the trade and lending portfolio requires close surveillance to detect potential fraud risks for post-transaction monitoring.

Due to the tightening market environment and significant fraud impairments, the member bank deployed trade fraud surveillance analytical tools to review and dispose the false positives that are generated by the tools, while also holistically engaging the business, credit, and fraud risk teams on cases of interest requiring risk mitigation actions. The key goal, is to prevent trade fraud, reduce losses and improve profitability.

What was the solution selected?

Trade fraud analysis is done at client-level, triggering alerts against thresholds of various scenarios which were injected into the surveillance tool platform.

Platform's capabilities are constantly updated and strengthened to reflect evolving material trade fraud events relevant to the member bank and threats from fraud events that have happened outside of the member bank.

Against these risk scenarios, a network analysis is then derived at client-level which links internal data (such as trade and cash transactions) and external data from various available sources to provide a robust framework for identifying fraud risk relevant to the member bank. Thus, using data networks to overlay scenarios which investigate relationship patterns and fund flows that are symptomatic of fraudulent behaviour in trade transactions.

Additionally, work is ongoing to improve detection of financial statement misrepresentation that has also been on the rise, and the bank has worked to highlight financial red flags that complement trade and transactional red flags.

Examples of key typologies identified as part of the review of cases include:

1. Trade within Social Network Group:
    - Complex group structure and non-disclosure of trades with related parties adding another layer to the supply chain
    - Co-mingling of funds i.e., trade flows between related parties and cash pooling funds making it difficult to ascertain the end use of funds
    - Inability to provide transportation documentation
2. Elevated risk due to:
    - Relationships with new, small and/or unknown counterparties
    - Loose Facility structuring
    - When the bank does not have end to end visibility of trade and cash flow
    - Round value and repeated invoice amounts and cash deposits
    - Weak financial health or anomaly found in financial statements such as mismatch of financials between group entities.

Measures of effectiveness

Fraud impairment losses have significantly reduced over time. The use of analytics has enabled the member bank to better identify material risk events and proactively de-risk to avoid future potential fraud losses.

Measures of efficiency

With the constant enhancement of the analytical tools combined with expertise of our fraud investigations, the member bank was able to focus its investigators' time on material trade fraud cases to investigate the anomalies identified from the network analysis (for example: blacklisted entities, undisclosed related party transactions, inflow of funds from entities not aligned with its business model).

What were the challenges and considerations whilst adopting the solution?

Potential fraudsters are taking advantage of the document intensive nature of the trade finance business using fake invoices and transport documents which are hard to predict using data models. Investigators are required to review hundreds of documents before concluding a potential case, which means the cost of reviewing such alerts would remain high.

Furthermore, there are difficulties in adopting advanced AI/ML models due to lack of use cases in trade and lending fraud. Higher unstructured data and reliance on documentary trades also requires data labelling efforts.

Additionally, accelerating digitalisation and evolving trade corridors requires the member bank to adopt newer scenarios and typologies to detect anomalies

Use case 2

Another member bank has explored the use of network analysis to identify hidden relationships and potential criminal networks.

Overview and the problem statement

Trade financing remains a relevant payment conduit for financial crime perpetrators to facilitate movement of funds across jurisdictions, predicated on seemingly innocent trade financing transactions. The inherent financial crime risk typologies include but are not limited to fictitious trades for purposes of obtaining loans, over-and-under invoicing to transfer funds, tax evasion, collusion, etc.

While conventional transactional surveillance and monitoring tools flag potential suspicious transactions based on specific transactional indicators, the rules do not show the potential social connections and/or hidden relationships.

At this juncture, the bank is exploring the basic form of network analysis, focusing on:

a. Counterparty / Counterparties concentration
b. Related trading entities that exist in the bank's trade financing network

What was the solution selected?

The starting point of the analytical review is predicated on (i) specific customer types, for instance, higher risk customers by nature of business activities and/or by (ii) the type(s) of trade financing facility / facilities accorded i.e., higher risk products, for instance, invoice financing.

The following dataset and features are broadly categorised, as follows:

- Customers' demographics and risk profile.
- Counterparty / counterparties concentration if any within a defined time period.
- Changes in the counterparty-composition over time e.g., any observed pattern(s) in the suppliers and buyers.
- The profile(s) of the counterparty / counterparties.
- Presence of the same counterparty / counterparties in the bank's trade financing network within a defined time period.
- High risk jurisdictions and potential incongruent locations of the trading parties e.g., Supplier, Buyer, Consignee, etc.

Measures of effectiveness

Current approach is generally useful in identifying potential hidden social relationships, which would be further validated by (a) corporate registry searches, (b) checks with the International Maritime Bureau, etc. Further analysis may indicate involvement of common Non-Vessel Operating Common Carriers.

Measures of efficiency

Current approach provides a good feedback loop to the Credit team.

What were the challenges and considerations whilst adopting the solution?

While performing this solution, challenges identified include the need to:

1. Re-balance the staffing composition mix for simple as well as complex analysis and
2. Expand the relevant detection scenarios and features.

2.4.4 Tax evasion

Tax evasion is an illegal act in which a person or entity deliberately avoids or underpays a true tax liability. To determine whether the act of failure to pay is intentional, certain factors will be considered such as whether the party committed fraud or deliberately hid their income offshore.

The below use case explores how geolocation data can be utilised to identify whether a customer is fraudulently declaring an incorrect tax domicile to pay lower taxes.

Use case 1

One member bank had applied data analytics for the detection of potential tax evasion.

Overview and the problem statement

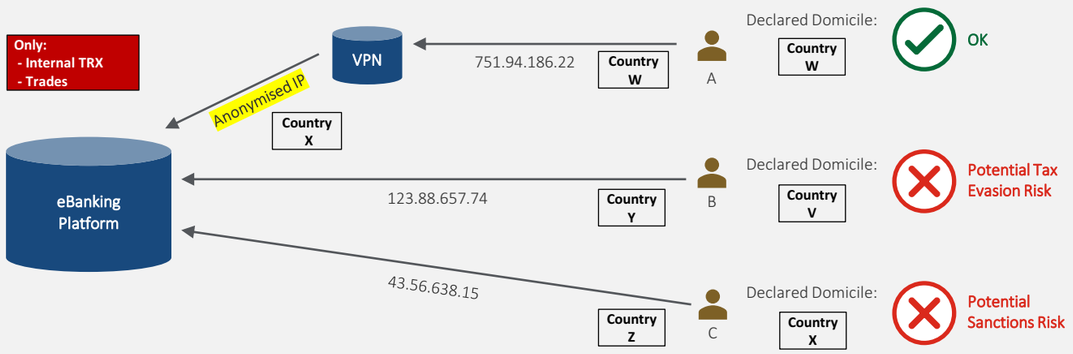
Over the years, authorities and firms across the globe have been trying to integrate AI tools and machine learning to prevent and detect tax evasion. Numerous options have been explored including taxpayer profiling and segmentation, risk scoring of tax operations, analysis of the supply chain, and identification of abnormal transactions.

In this case, the member bank implemented a solution to detect potentially suspicious deviations of client declared domicile, compared to the trends in domicile access into the eBanking platform based on the resolved geolocation of the clients' IP address to detect potential tax evasion.

What was the solution selected?

IP address data was collected from the eBanking platform and resolved to customer identities, including details of their declared domicile/tax domicile. The IP address data for each user account was resolved to country-level geolocation data, using an accurate external vendor reference data set. The external reference data set also enabled the identification of IP addresses that matched the use of VPN or other IP address anonymisation services, allowing for such accesses to be discounted for the purpose of the review.

The data was then compared between the declared country domicile against the matching IP address country domicile data. Rules were set to identify sustained log-in behaviour – specifically a requirement that there be a minimum of accesses across 6 contiguous months, from the mis-matched domicile – e.g., to exclude clients going on holiday or business trips – and hits were then flagged.



Measures of effectiveness

The designed solution was able to effectively detect the targeted typology, including the flagging of physical vs virtual accesses from specific countries (e.g., through a use of a VPN). This capability will enhance other related client investigations, allowing for the creation of a 'storyline' of client log-ins and the domiciles matching those log-ins. This additional intelligence source will further enrich the overall client profile and lead to more effective investigations and more effective client risk management. Historically, information of client domicile was generally limited to the information declared by the client – therefore, this capability adds an additional form of independent client location check. Some additional effectiveness considerations are described further in the following challenges section.

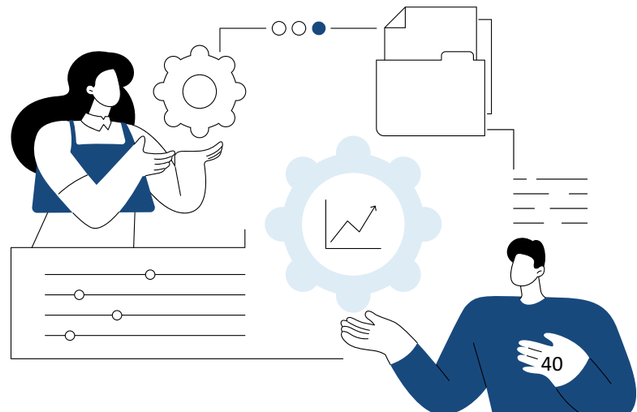
Measures of efficiency

Regarding the specific review, results were relatively straightforward to identify – one team member was able to perform the analysis in less than 1-month total time. As the data will be used for other purposes – either through periodic 'wrong domicile' controls, or in relation to specific client investigations – the additional domicile intelligence information will allow for more efficient identification of key client risk considerations.

What were the challenges and considerations whilst adopting the solution?

Hits were escalated to the member bank's Financial Intelligence Unit ("FIU") team. FIU's standard process, upon receipt of escalations, is to perform detailed review of the client scenario – including general profile review, name/media searches, and transaction review. For the IP address/eBanking review, the cases escalated to the FIU team were able to be explained with non-concerning explanations (e.g., was client's child logging in from foreign location).

It was assessed therefore that the control should be considered for incorporation into controls sitting closer to the business teams, where such explanations could be provided, without requiring the extensive review performed by the FIU team. Any potentially suspicious cases detected by the business-aligned control would then be escalated for FIU review.



Use case 2

One member bank had applied clustering for the detection of potential tax evasion.

Overview and the problem statement

Tax evasion is a key risk identified in the member bank and a risk of concern for regulators. Traditional AML techniques in detecting tax evasion risk have not been as effective as needed, given the sophisticated profile of ultra-high net worth customers who have complex structures for the purposes of their investment strategies, as well as hold multiple citizenships/residences worldwide. Therefore, the bank needs to use advanced techniques to allow the bank to better detect tax evasion risk as well as emerging and new tax evasion risk typologies.

What was the solution selected?

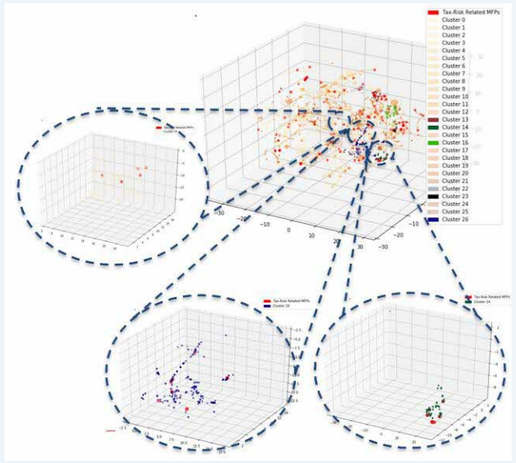
Tax Evasion Clustering Analysis ("TECA") adopts an internally developed semi-supervised machine learning model that analyses tax evasion related clients' characteristics, to then identify similar clients with Tax Evasion risks for review. Some examples of client clusters resulting from the model were:

Clients with complex structures and multiple accounts set up with no reasonable economic justifications.

Clients with high volume of unusual multi-directional transaction flows internally or externally, potentially for layering or co-mingling of personal wealth and business-related funds.

Clients with multiple citizenships/residences and having Golden Visa passport, potentially to circumvent tax filing obligations.

Clients with mismatches in the Tax Residency declaration against other demographic details such as nationality, residency, source of wealth countries, contact details and mailing addresses, as well as differential tax rates among the mismatched jurisdictions that can potentially aid in tax evasion.



The TECA model allows for refresh and retraining of the model on a monthly basis. Features considered in the model include full client account structures and their associated entity relationships, tax residency and other client demographic related characteristics, tax rates of all jurisdictions, and transactional related characteristics (such as daily moving window to track rapid transactional flow within each account structure over the past year) etc.

TECA employed the following key design features that enables the analysis to be targeted to tax evasion risks to reduce false positive alerts, minimising unnecessary manual reviews and therefore maximising review efficiency:

Features of the model were designed by financial crime domain experts with several decades of financial crime investigation experience. The identified key tax evasion-related client characteristics and typologies enabled the machine learning process to be more targeted towards the risk-related typology.

- The semi-supervised machine learning algorithm employs clustering before overlaying with insights on Tax Evasion risks (using STRs and Trigger Event Reviews) to identify high density tax evasion risks clusters. The learning process focuses on the concept of similarity, and the "closeness" to tax evasion risks so as to be more targeted towards higher risk clients.

TECA helped to uncover high-density clusters for deeper analysis that enabled the bank to self-trigger additional tax evasion cases for holistic data analytics reviews. These eventually led to the filing of STRs as well as enhancements in the bank's financial crime compliance policy and procedures.

Measures of effectiveness

Through TECA, the bank has filed tax evasion STRs from the cases identified by the model, which involved hundreds of millions (in USD) of transactional values and Assets Under Management ("AUM") at point of case escalation.

STRs filed were on additional risks identified that were otherwise not uncovered by business as usual ("BAU") control processes. The STRs filed through the escalations from TECA reflected the added effectiveness of this data analytics capability that helps to improve the bank's anti-financial crime control efforts.

The additional insights from TECA and the STRs filed were also shared with the teams managing the BAU control process, so as to enlighten and enhance their awareness of the tax evasion risks typologies and to enhance the policy and procedures at the BAU control gateways. The sharing included the BAU anti-financial crime processes with the bank's front-line staff (i.e., Relationship Managers, Business Managers and Business Compliance Managers), as well as Compliance Line 2 anti-financial crime control teams such as onboarding and periodic review, to raise their tax evasion risk awareness and be vigilant to unusual customer behaviour or transaction patterns.

Measures of efficiency

Based on the statistics of the Level 2 analytics review process, the bank has achieved a STR yield of 68%, which involved hundreds of millions (in USD) of transactional value and AUM at point of case escalation. The analytics technique employed in TECA clearly demonstrated its ability to reduce false positive alerts, minimising unnecessary manual reviews and therefore maximising review efficiency.

On top of that, the transparency of the TECA model provided in-depth insights into the key features and red flag typologies that resulted in the escalation of the self-triggered cases.

The granular data utilised in features engineering as well as the insights from the key features are provided to the bank's financial crime compliance DA investigators to focus on targeted risks upon escalation. The immediate availability of key information reduced the need for additional retrieval of client's information and reduced the time spent by investigator by 50% for each case. Therefore, review efficiency is enhanced as investigators can spend their time on review tasks and less on other manual tasks.

What were the challenges and considerations whilst adopting the solution?

While modelling the solution, the bank had identified two key challenges and overcame the challenges. Firstly, the identified tax evasion related characteristics of clients are temporal in nature, and it is critical to identify the right period in history to model the tax evasion related characteristics, especially for clients identified as having tax evasion risks. The bank paid special attention to accurately identify the specific periods for modelling of every client with tax evasion risks.

Secondly, the initial set of identified tax evasion related characteristics is not exhaustive, and new sets of characteristics and red flags might be discovered during investigations. It is highly recommended for the bank to instil a continuous learning process, including the development of new features to improve the machine learning model.

2.4.5 Fraud / Scams

Fraud is an intentionally deceptive action designed to enrich the perpetrator by stealing from a victim. A scam is designed to trick a victim into giving away money, personal details, or data. Types of fraud include credit card fraud, wire fraud, securities fraud, and tax fraud. Types of scams include impersonation scam, inheritance scam, internet love scam, job scam and investment scam.

With technological advancements, many jurisdictions around the globe have reported consistent growth in the global magnitude and scale of online scams[10]. Illicit proceeds from these scams are often transferred to foreign jurisdictions and may be further laundered through the financial systems of other third-party jurisdictions. There is a need for FIs to continue strengthening their capabilities to detect online scams and related money laundering and enhance suspicious transaction reporting, as scam syndicates continuously evolve their modus operandi.

Overview and the problem statement

Scams have been on the rise and have become more and more sophisticated in recent years. Not only are they evolving quickly technologically, but scammers have also managed to refine their social engineering to new heights. Vulnerable individuals are targeted and manipulated into transferring funds to the scammers. In 2022, scam victims in Singapore lost $660.7M to scams[11].

Due to the speed at which the scammers evolve to circumvent detection and prevention efforts, existing rules-based methods though robust, need to be bolstered to form a holistic method to effectively fight fraud. Moreover, traditional approach using rules rely on hard thresholds and specific conditions to block transactions, which inevitably lead to high volume of false positives. When genuine transactions (i.e., false positives) are held, it negatively impacts customer experience as customers will need to wait for the transaction to be cleared. These false positives also increase the workload for the operations team responsible for contacting the customers to verify the transactions.

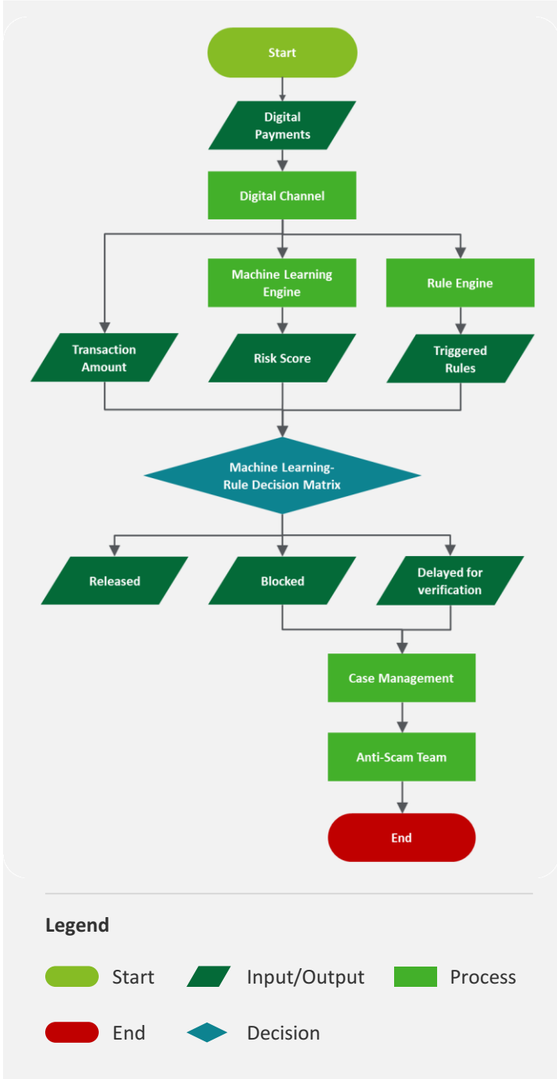
What was the solution selected?

As machine learning models can handle more complex scenarios, the member bank built a model to potentially reduce the false positive rate, and at the same time identify outliers and behaviour changes to surface potential new scam cases. In addition, machine learning models can be continuously auto trained to identify new fraudulent patterns even before new modi operandi are established, thereby giving banks an edge in the never-ending battle against scammers.

The member bank trained a machine learning model using XGB on a set of labelled historical data with over 300 features engineered. The features fall into six main categories: (1) financial events, (2) non-financial events, (3) beneficiary account behaviour, (4) customer outflow behaviour, (5) customer money temperature, and (6) customer demographics.

The machine learning model was implemented alongside the existing rule engine. Based on a combined decision matrix, comprising of the risk score provided by the model, transaction amount and whether any rules were triggered, the transaction will be either blocked, delayed for verification, or released.

The following diagram provides an overview of the process.



As transactions are intercepted in flight, the model is expected to produce a risk score within 10 milliseconds. This requirement also influenced the bank on the selection of XGB as the machine learning model for fraud detection. With the use of machine learning models, it is imperative that the team verifying the blocked or delayed transactions is able to understand why the decision was made. Hence, the bank enhanced its case management system to highlight model features that contributed most to the decision.

Measures of effectiveness

With the model, the bank can identify an additional 6% to 10% of very high-risk alerts that would have been missed by the rule-based engine.

Measures of efficiency

The bank also observed an overall increase in efficiency by freeing up investigator capacity as the model auto-releases 12% to 15% of rule-based alerts which it ranked as very low risk.

What were the challenges and considerations whilst adopting the solution?

As the bank receives a significant volume of transactions every second, it is important that the model can accurately process, and risk score each transaction in real time. As such, the bank had to balance the trade-off between model performance and model runtime to meet the requirements.

Since transactions take place in real time, the bank also considered appropriate data architecture to serve the model with the required features. Features were categorised into buckets, based on computation needs, so that minimal computation is needed in real time.

To maximise the optimal benefits from a machine learning model, the model is auto tuned on a daily or even near real-time basis, considering patterns of latest scam cases. In this way, new typologies can be picked up and detection can be put in effect very quickly.

[10] Source: FATF CEF Paper - Illicit Financial Flows from Cyber-enabled Fraud (fatf-gafi.org)
[11] Source: The Straits Times - https://str.sg/wt9B

2.4.6 Money mules

Money mules are people who help criminal syndicates launder proceeds of crime, either knowingly or unknowingly. Criminal syndicates usually recruit money mules to move illegally acquired money by providing their own accounts to help receive and transfer these funds to legitimise it. Recruitment of money mules usually includes several methods – they were promised a financial gain or commission, solicited via online scams, or are being manipulated by a friend or romantic partner.

Traditional transaction monitoring tools often prove ineffective in identifying money mules accounts, given that there are no specific rules in transaction monitoring systems that identify money mules. Banks must look at holistic machine learning methods to achieve a higher detection rate of money mules. Use case one explores the use of data analytics to identify suspicious activity, whilst use case two highlights the benefits of machine learning to detect mule activity.

Use case 1

One member bank has shared the use of data analytics to detect money mules.

Overview and the problem statement

The increasing trend of potential scams, money mules, take-over and/or misuse of personal accounts at the national level raises concerns. The current approach using static rules to mitigate such cases has known limitations which impact FIs ability to detect and proactively prevent these types of issues.

What was the solution selected?

The bank explored alternative detection methods to complement the existing transaction monitoring tool. The solution selected provided a holistic view of the connected/related counterparties as well as hidden relationships in the customers' portfolio to better identify connected illicit network(s).

The bank built and deployed a solution using Structured Query Language ("SQL"), which was utilised alongside network analysis. The data set and features used to build the solution are categorised into the following:

- **Customers' demographics data**
- **Transactional profiling data**: This includes customer inflow-outflow behaviour and beneficiaries' patterns. Considerations were made on the data set to include review prioritisation predicated on the top 10th percentile of the aggregated transaction amounts and transaction counts for the calendar month, analysis of month-on-month variance on aggregated transaction amounts, and "noises" where effects of deposit campaigns and promotions were taken into consideration to mitigate the false positives associated with "spikes" in the account movements.
- **Risk analysis data**: This considers the risk associated with various customers, e.g., single customer to multiple external parties and vice versa, multiple external parties to single customer, single external party to multiple customers and conversely.
- **Digital footprints**: This includes IP address, which is currently located in a separate system and yet to be integrated into the current monitoring.

Measures of effectiveness

The bank was able to effectively identify 30 STRs during the review, of which 25 were money-mule related and the remaining 5 were attributed to other suspicious activities. Further, post-mortem validation of the current transaction monitoring alert generation as well as the fraud system also indicated a high trigger correlation.

Measures of efficiency

As the solution remains under exploratory stage, the immediate efficiency is the better quality of alerts which ensures that the TM investigators time is spent evaluating true alerts.

What were the challenges and considerations whilst adopting the solution?

While developing the solution, the bank had made considerations to manage and distribute capabilities such that both simple and complex analysis can be handled appropriately. Further considerations were made to decide applicable scenario(s) for the network analysis tool.

Use case 2

Another member bank has applied machine learning model in the detection of money mules.

Overview and the problem statement

With scams on the rise in recent years, it has been reported that there is an emerging trend of Singpass users selling their accounts, allowing criminals to open bank accounts to launder such crime proceeds, totalling almost $1.3 billion in 2021 and 2022.

The increase in money mule activities – 19,000 cases were investigated in the past three years – prompted amendments to anti-money laundering laws to make it easier for authorities to prosecute money mules[12].

Financial institutions have a critical role to play in combating money laundering through accounts of money mules, by implementing data analytics solutions to detect suspicious transactions and account activities.

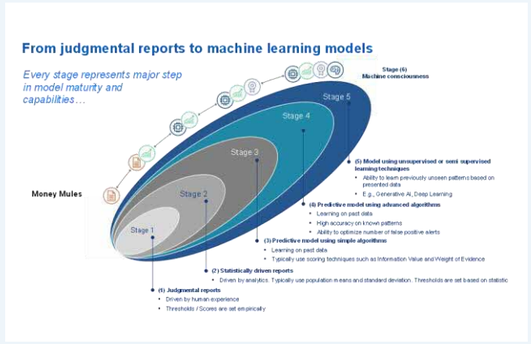
Overview and the problem statement

The bank implemented a predictive model to detect money mule accounts and put in place processes to review alerted cases promptly. The bank decided to use a scoring-based technique for its initial model, by analysing information value and weight of evidence of a broad range of features, including transaction data, customer demographic, account information, customer activities, geolocation, and other relevant external information.

The model is applied on the full customer base daily as it allows the bank to take prompt actions against suspected money mules and those with the highest scores are further investigated.

A dashboard was also created to facilitate investigations, showing historical transaction trends, connected parties and details of observed typologies for each customer being investigated. Upon investigation, highly suspicious customers may be exited or suspicious transaction reports (STR) may be filed with the authorities.

The member bank also shared that with the labelled data collected from the cases reviewed, it is looking into developing ensemble models using more advanced machine learning algorithms.



Measures of effectiveness

The member bank shared that the initial scoring-based model has been effective in detecting money mule accounts. Of the alerts generated by the model, about a third received production order or seizure order subsequently on the same day or after the alert date.

Measures of efficiency

Timely detection and action by banks to freeze accounts used for money mule activities are critical. The member bank achieved this through the daily refresh of the model/alerts and the use of relevant dashboard to facilitate investigation.

The member bank shared a case where a newly onboarded customer started exhibiting money mule typologies and transacting on his account two weeks after it was opened. The model raised an alert on the following day and the case was promptly investigated, leading to STR filing within a few days of the alert. The account holder allegedly handed over his internet banking login credentials to an unknown person and has been charged in court for helping scammers launder criminal proceeds.

What were the challenges and considerations whilst adopting the solution?

As banks and businesses move towards public-private partnerships, steps will be needed to ensure that banks are able to make the best use real-time intelligence to enhance future detection models.

One consideration that the bank had was the evolving typologies and changing regulatory requirements. Machine learning models learn from historical cases and may require frequent retraining to keep up with techniques used by money launderers. Even so, the models may not pick out new typologies fast enough. Proposed changes to the law may also impact the performance of existing models, as it may result in changes to money mule demography, the way money mules open their accounts, as well as their transaction patterns. To ensure that the model remains relevant, the bank had to keep up with the evolving typologies and changing regulatory requirements.

[12] Source: The Straits Time - https://www.straitstimes.com/singapore/update-new-anti-money-laundering-laws-regularly-to-deter-scammers-money-mules-experts

Key takeaways

The use cases highlighted in this section detail the rationale, effectiveness, measures, and challenges of implementing the analytics solutions of member banks in the six chosen priority risk areas. Their experiences can be summarised into the following:

Harvest efficiency and re-invest in effectiveness

Financial institutions (FIs) have been leveraging data analytics and machine learning techniques to dynamically assess customer risk and prioritize alerts. As the maturity and reliability of these solutions increase, FIs have successfully achieved efficiency gains while maintaining detection levels.

These savings are subsequently reinvested in more advanced data analytics, such as network and macro-level surveillance. These enhanced techniques excel in detecting large sums of money and interconnected networks with greater speed and on a larger scale. Consequently, the effectiveness of risk surveillance and coverage of review by analysts are significantly increased.

Experiment first, fail fast or scale

To emulate successful startups, FIs should adopt an experimental mindset and embrace the "fail fast" approach. This involves actively seeking out opportunities to test new ideas, rapidly identifying and addressing challenges, and swiftly pivoting data analytics led risk surveillance solutions as needed. By doing so, FIs can reap benefits more promptly, rather than investing excessive time and resources in unproductive endeavours.

Furthermore, fostering a culture that views failures as valuable learning opportunities encourages innovation and continuous improvement, ultimately driving FIs toward achieving their objectives more efficiently and effectively. Upon achieving success, the proven solution can then be scaled across the entire organization, potentially spanning multiple geographical locations.

Massive adoption of data analytics is largely in areas where you have good label data and good reference information

Good label data and reference information was essential to the success of these use cases. As FIs look to adopt data analytics for its significant and scalable benefits, they should identify areas that are most suitable for analytics and be aware of the efforts that go behind the development of an analytics model. For analytics to work, member banks ensure that there is widely available data that can be collected and processed in a consistent format. Often, this would mean that data labels and references used by the analytics model must be input in a standard format.

Continuous enhancements to adapt to shifting typologies

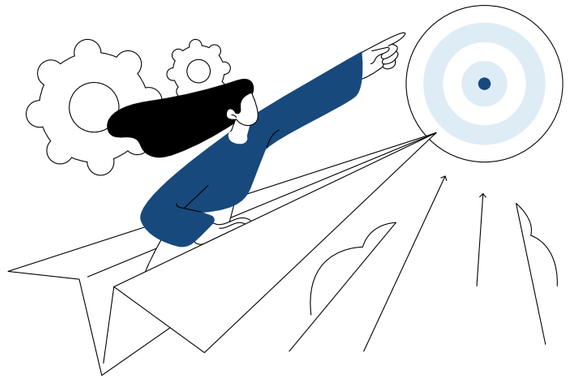
Commonly observed across use cases was the dynamism of the financial crime landscape, and the essential trait of always enhancing and learning new money laundering typologies. As member banks take on more proactive roles in the fight against fraud and money laundering, analytics solutions can and have been at the forefront of detecting new behaviours and potential typologies. In many cases, member banks have cited efforts to continuously enhance their monitoring solutions to ensure that the solutions remain relevant and beneficial by means of incorporating investigation outcomes or new regulatory requirements.

3 Looking ahead

As we look forward from where we are today, the fast pace of transactions and evolving financial crime typologies in the industry will demand improvements in effectiveness from existing detection methods using data analytics and advanced technological solutions.

To be successful in detecting and mitigating financial crimes (including money laundering, terrorism financing, fraud/scams, and sanctions), FIs are encouraged to commit to efficiently and effectively adopt advanced technology solutions across the end-to-end customer lifecycle.

Below are some upcoming themes that could influence the future of data analytics in AML/CFT.



1. **Convergence of AML, Sanctions and Fraud surveillance**
   As criminal organisations grow in complexity, size and sophistication, their activities are likely to span multiple financial crime areas concurrently. For example, laundering proceeds from fraudulent activities to fund terrorist groups in sanctioned countries.
   Therefore, there is a need to converge the existing surveillances to form one holistic surveillance model, correlating related but different criminal activities to reveal the complete picture.

2. **Integrated monitoring of digital assets**
   With the proliferation of crypto currencies and digital assets, screening and surveillance will need to be able to track assets as they transition on-chain and off-chain seamlessly. It is critical to seamlessly integrate the insights from such surveillance into the existing monitoring processes (covering Fiat). This is a greenfield area which will require considerable experimentation and investment; however, it is crucial to do so to prevent criminals from exploiting and abusing digital assets as it becomes increasingly common and easily accessible.

3. **Reducing time to detect**
   With criminals becoming more sophisticated and agile, the frequency of batch-based surveillance will likely need to continuously increase. Evident in efforts to combat transaction-based crimes like fraud and scams, surveillance is more effective when done more frequently, or real-time. To do this, considerable investment in data infrastructure and technology stack is required.

4. **Solutions and infrastructures that enables public and private information sharing**
   Public and private partnerships are essential to combating financial crime. Governments and law enforcement agencies need to work together with FIs and other businesses to share information and coordinate their efforts. Some examples of public and private partnerships to combat financial crime include:

   - Government agencies like FIUs that collect and analyse financial data to identify suspicious activity will often share information with financial institutions and other businesses to help them detect and prevent financial crime.
   - Public-private partnerships for information sharing ("PPPIS") are forums where governments and businesses can share information about financial crime threats and trends. This will help to improve the effectiveness of both public and private sector efforts to combat financial crime. For example, Joint Money Laundering Taskforce ("JIMLIT") in UK and ACIP in Singapore facilitate such partnerships.

5. **Use of AI/ML models to replace rule-based surveillance systems**
   Traditional rule-based surveillance systems are no longer able to keep up with the sophistication of financial criminals. As a result, FIs are increasingly turning to AI/ML models to detect and prevent financial crime. AI/ML models are increasingly used to analyse large amounts of data to identify patterns and anomalies that would be difficult or impossible to define and detect with rule-based systems and they can do so without sacrificing control effectiveness.
   Thus, they have begun to make a stronger case to replace legacy rule-based systems as opposed to just being complementary in nature. For example, AI/ML models can be used to detect suspicious transactions, identify money laundering networks, and detect fraud.

6. **Rise of Generative AI**
   Generative AI ("GenAI") is a rapidly developing field with the potential to revolutionise the way financial crime is detected and prevented. GenAI is still in its early stages of development, but it has the potential to revolutionise financial crime analytics efforts. GenAI can be used to generate synthetic data that can be used to train machine learning models to identify new and emerging financial crime patterns.

It is worthwhile to note that the considerations and trends in financial crime identified in the paper have similarly been highlighted by FATF[13] as areas of priority, showing the alignment of the concerns and future directions.

As data, analytics, and AI become more entrenched in the day-to-day functioning at most FIs, it is critical that data platform and data architecture for financial crime evolve to stay ahead of the bad actors and position the FIs as efficient in manging risks, effective in detecting key cases, and minimise regulatory and reputational risks that would arise when they are deemed deficient.

[13] Source: FATF CEF paper - Illicit Financial Flows from Cyber-enabled Fraud (fatf-gafi.org)

3.1 Advanced analytics solutions

The industry notes that technology implementations are progressing beyond intelligence-led models to holistic systems which provide parties (FIs and regulators) with a comprehensive oversight of all FCC aspects in an organisation. This will involve integrating intelligence-led models across all FCC areas such as transaction monitoring, name screening, and sanctions to provide a holistic vantage point of risk exposure.

To this end, below are some of the technologies which member banks have been experimenting with or which have been implemented by a minority of member banks at the point of drafting this paper, with an aim of increasing the effectiveness of identifying bad actors.

3.1.1 Behavioural biometrics based fraud and mules detection

Rethinking Onboarding

Cybercriminals exploit vulnerabilities in digital platforms, especially during the account opening process, using stolen, synthetic, or compromised identities. Such malicious endeavours not only lead to substantial financial losses but also tarnish the reputation of banks. Traditional fraud detection mechanisms are increasingly proving inadequate in this evolving landscape. As a result, behavioural biometrics has emerged as a significant advancement. This technology conducts a thorough analysis into a user's physical and cognitive digital behaviour, offering insights that are instrumental in distinguishing authentic individuals from potential fraudulent entities. With the rise in account takeover attacks, social engineering scams, and the proliferation of mule accounts, the need for advanced, continuous, and proactive fraud detection has never been more paramount.

Behavioural biometrics offers an additional solution to the challenges posed by digital banking fraud. Unlike traditional authentication methods that rely on static information (like passwords or PINs), behavioural biometrics focuses on the unique ways in which individuals interact with digital devices and platforms, involving:

- **Continuous authentication**: Instead of a one-time authentication at login, behavioural biometrics continuously monitors users' activities throughout their session. This ensures that even if a fraudster gains access initially, any anomalous behaviour during the session can trigger alerts.
- **Analysing physical and cognitive behaviour**: Behavioural biometrics evaluates a range of factors, from the way a user types or swipes on a touchscreen to their navigation patterns and response times. These behaviours are difficult, if not impossible, for fraudulent entities to mimic.
- **Detecting anomalies**: By building a profile of a user's typical behaviour, the system can quickly identify deviations from the norm, whether it's a result of account takeover, a social engineering scam, or a cybercriminal using a mule account.
- **Reducing false positives**: Traditional fraud detection systems can often flag legitimate activities as suspicious, leading to unnecessary friction for authentic individuals. Behavioural biometrics, with its nuanced understanding of user behaviour, can significantly reduce these false positives.
- **Enhancing user experience**: By relying on passive authentication methods, users are not burdened with additional authentication steps unless truly necessary, leading to a smoother digital banking experience.

Behavioural biometrics brings about the following benefits:

1. Proactive fraud detection
   Traditional systems react to fraud after it occurs. In contrast, behavioural biometrics identifies potential threats in real-time, allowing banks to take preventive measures before any damage is done.
2. Enhanced accuracy
   By analysing thousands of behavioural parameters, from mouse movements to touchscreen interactions, behavioural biometrics offers a higher degree of accuracy in distinguishing authentic individuals from fraudulent entities.

3.1.2 Perpetual KYC

Rethinking Customer Due Diligence

Currently KYC is performed at a specific time in the customer lifecycle – during onboarding and periodically post onboarding, where the interval of review is determined by the KYC profile and risk level aggregated from the customer profile.

Review frequency is dependent on AML risk ratings, e.g., in Retail or Commercial Banking, High and Medium Risk customers are subject to periodic CDD reviews every 1 and 3 years respectively, while Low Risk customers are only reviewed upon the occurrence of a trigger event. Quite often, the outcome of periodic review is simply an update in customer profile, if any changes, with no behavioural anomalies observed.

The dynamic Perpetual KYC ("pKYC") approach seeks to enable FI's to achieve higher effectiveness in AML risk management. In essence, it allows for the identification and assessment of more predictive risk attributes exhibited by customers, by leveraging data analytics to consider various attributes like changes in customer profile, customer's associated network relationships and account activity behavioural red flags.

"Perpetual KYC" is named as such because it involves regular checks in fixed frequency – monthly, weekly, or even daily – on the customer using a set of pKYC pre-defined criteria. Once hit, a CDD review of the customer will be triggered regardless of his AML risk rating.

The criteria applied in pKYC would typically consider a combination of profile, network links and account activity attributes. It is important to note that they should be differentiated against criteria used purely for transaction monitoring purposes, to avoid duplicate triggers. To elaborate, we may use significant change in account behaviour as a straight rule in transaction monitoring, while in pKYC, the criteria may be designed as a change of mandate followed by significant change in account activity.

Other than rules-based approach, a machine learning model-based approach for pKYC may also be considered. Regardless the detection solution used, robust back-testing is necessary to ensure that bad actors previously found during CDD periodic reviews will continue to be captured with pKYC, and in addition, other bad actors with suspicious profiles and patterns are raised by pKYC.

In moving towards pKYC, expected benefits may include:

1. More timely detection of customers with unusual behavioural red flags
2. Improved efficiency in allocation of KYC resources to manage AML risk
3. Targeted reviews to also address specific red flags, rather than a broad-based CDD review
4. All customers, regardless of AML risk ratings, are subject to pKYC, thus widening coverage

3.1.3 Machine learning based transaction monitoring

Rethinking Rule-based Transactions Monitoring

Limitations exist in the current scenario-based transactions monitoring approaches that are vulnerable to seasonal increases in alerts or customer segments that may be in line with expected activity, yet flagged as suspicious on other parameters, hence, requiring investigation. The industry has observed a move towards more Intelligence-led approaches to TM. Such a model uses more precise analytics to reflect a probability assessment for each customer, as opposed to a binary result from the scenario-based monitoring systems. The unique behaviours of every customer can be assessed and can contribute to the overall probability assessment and indicate a likelihood of financial crime risk. Intelligence-led approaches to TM and machine learning therefore are a potential strategic solution to structured scenario-based TM system limitations. To that end, certain banks have begun adopting machine learning models as the primary intelligence led detection tools for TM, phasing out scenario-based approaches.

The machine learning led approach to TM is a data driven strategic solution for consideration given it seeks the following outcomes:

- Identify financial crime risks in bank portfolios within a shorter timeframe;
- Eliminate unnecessary customer friction;
- Eliminate unnecessary frontline distraction;
- Deliver a simpler financial crime framework; and
- Report more useful information to law enforcement.

One member bank deployed a machine learning model for transaction monitoring and phased out scenarios led monitoring for two of its customer segments.

The model was developed to predict a customer's propensity for money laundering based on comparing activity against a known bad target set of previously reported and/or exited customers. The solution therefore does not use rules to trigger investigations in the same way rule-based scenarios do, instead the predicative model uses 'features' designed for risk assessment. To build the model, various types of data, i.e., transactional data, entity data, risk data, digital data, financial crime data and, open-source data etc., were used.

This leverages a 'Master Model' with pertinent features within the feature set, assesses their interplay and learns to recognise the behavioural patterns of "known bad actors". It ultimately derives a risk assessment probability for each one of our customers, supplemented with a supporting rationale indicating the activities/behaviours seen. The master model can then be supplemented with subject matter expertise – using knowledge of typologies and suspicious activity data sets; anomaly detection – leveraging time series analysis to detect abnormal behaviour against historical behaviour through signal processing; and networking analysis through transactional networks.

The machine learning led approach outperforms scenario-based approaches with the following:

- **Identification of risk:** One FI that adopted a machine learning led approach to TM has experienced initial positive signs on the identification of financial crime risk with STRs increasing, when compared to the scenario-based approach. This is particularly evident on the corporate side of the business, where inherent risks exist in the form of shell company abuse. Performance was equally sustained on the retail side of the business and deemed effective. Targets identified then are a feedback loop into the model post investigation to dynamically retrain for future detection of risk.
- **Reduction in false positives and investigation:** The same bank experienced a 60%-80% reduction levels in transaction monitoring alerts, increasing STR conversion rates more than 2-3 times. This allows more time for resources to focus on investigations rather than false positives, with the process open to streamlining into a holistic risk-based review conducted by one investigator, removing duplication and inefficiency created by hand-offs under a two-tier system, and managing quality under accreditation, manager review and quality assurance under one common set of procedures.
- **Dynamic feedback loops:** Modelling oversight and optimisation governance are equally in place to train the model via a feedback loop, supporting agility and education to counteract the ever-changing behaviour of illicit actors in the financial system. These feedback loops must be frequent, to increasingly critique a model - particularly within the embedment phase of adopting a machine learning led approach.

Migration from a scenario-based approach to a machine learning led approach to TM is an overhaul to the TM framework, with dependencies cutting across Information Security and Technology, Analytics and Modelling, and Compliance Investigations. Given the structural and process dependencies that are both internal and external in nature, there is a significant runway required for the end-to-end implementation of the approach.

Also, machine learning is potentially subject to unintended bias from the data sources used for feature creation and historical customer exit information used for model training. There are several controls in place throughout the model development process to ensure the risk of bias is mitigated. Model and data governance remain key pillars in upholding ethical data standards. Various controls and standards are further required within the process of adopting a machine learning led approach, given its reliance on model training compared with stagnant scenarios. A number of minimum standards can be applied to provide confidence and control within machine learning TM models, within model risk governance framework. From a TM testing capability perspective, the following items are critical (yet not limited to) for migrating to a machine learning led approach to TM:

- Minimum performance operating standards;
- Empirical proving;
- Design validation;
- Operational Readiness Testing; and
- Evidence of the capability.

Constant feedback loops, procedural controls and validation are critical to oversee the above, supported by the prevailing model risk governance standards are key control frameworks, as effective oversight with adopting the solution.

3.1.4 Dynamic review monitoring model

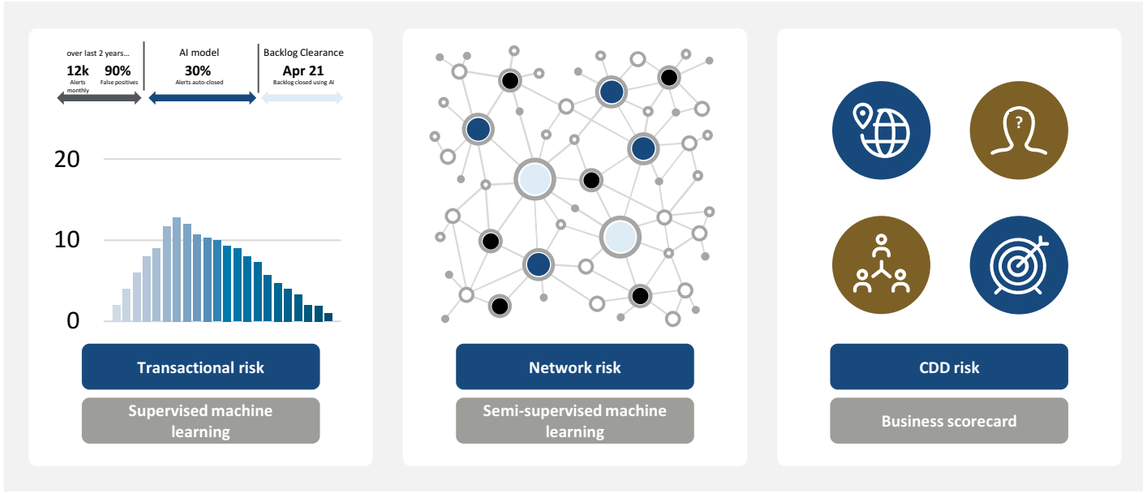
Rethinking Customer Risk Assessment

One member bank has implemented a Dynamic Review Monitoring Model ("DRM") to address the problem of outdated customer risk to predict the likelihood of future STRs on a customer. In their deployment of a DRM, risk scores are first allocated to customers via a model based on their proximity to known adversities such as STRs or POs. The scores allocated on each type of adversity together with other risk scores are later used as input features to the next model, Logistic Regression. Through Logistic Regression, each input feature is weighted according to its influence towards the likelihood of a STR filing, subsequently generating a final risk score on the customer.

Features contributing to the final risk score can be categorised into three main pillars:

1. Transaction risk pillar
   This pillar consists of features relating to customer transactions. A Light Gradient Boosting Machine model is employed to evaluate customer's transaction behaviours. The model assigns risk scores to every transaction monitoring alert closed in the previous month based on presence or absence of features found to be correlated with STR filing. The risk score indicates the level of AML risk and likelihood of STR filing required on the alert.
2. Networking risk pillar
   This pillar focuses on customer's transaction networks. A customised label propagation algorithm is used where the risk generated from the adversities in the first model, i.e., transaction risk pillar, is propagated through transactional and connected party links across the bank's customer base to other counterparties and customers. Adversity sources used include STRs, seizure orders and production orders. Customers with adversity indicators are first tagged with a high-risk score as a starting point, the scores are subsequently propagated along transactional and connected party linkages with an attenuation factor at each hop, to other customers in the bank.
3. Customer profile risk pillar:
   This pillar consists of features relating to customer profile. Rule-based approach is employed to quantify customer's profile risk based on various risk contributors such as whether the customer is a PEP, adverse information on the customer, customers with complex structure, and other factors. Each risk contributor is assigned a score by a subject matter expert, a customer risk score is then calculated by as summation of the risk contributors.

In summary, DRM learns from past STRs filed, selects the major features that correlates with these past STRs the most, and derives an aggregated score from all three pillars on each customer. A customer with a high DRM score would be investigated, acting as the final and overarching layer of the banks' AML/CFT controls to detect incremental high-risk actors not caught by existing controls. The following diagram summaries the three pillars of the DRM.



The bank observed achieved effectiveness in employing DRM. From testing done on DRM, the bank observed 35% of customers subject to Level 1 reviews were escalated to Level 2 reviews, and 29% of these escalations resulted in STR filings from testing done on the Emerging Business segment. Additionally, the DRM when integrated into the bank's network visualisation investigation tool allows for filtering of customer networks to focus on customers with the highest DRM scores, i.e. customers most likely to co-occur with reasons for suspicion.

In addition, DRM improves efficiency in allowing the bank to redistribute resources to focus on higher risk customers, as only customers with DRM score exceeding a pre-determined level are reviewed in each DRM monthly run.

The following table outlines potential challenges in the implementation of dynamic customer risk assessment and potential solutions to address the challenges.

Potential pitfall	Possible solution
Model tuning to factor for differences between customer segments As different customer segments within the bank have different customer profiles and prevalent typologies, optimum model weights can be expected to be different across customer segments.	This can be addressed by running the model on various customer segments separately.
Highly connected nodes As label propagation is used as the approach to network scoring, the existence of highly connected benign nodes such as government bodies (for example, CPF, IRAS) or merchants can cause high scores to be propagated to nodes in absence of truly suspicious connections. Further, some highly connected nodes can occasionally have adversity linked to them (for example, a PO tagged to CPF).	This can be addressed by making modifications to the model to whitelist the adversity tagged to benign highly connected nodes, and an attenuation factor was added to the model to reduce the degree of label propagation through highly connected nodes.
Learning from past STRs Current version of the DRM (which has Network Risk as one of the pillars), assigns a holistic risk score on a customer based on typologies exhibited by historical STR customers. This means the model is not tuned to immediately identify new, emerging risks which may not be apparent in past STRs filed.	The visual investigation platform can be equipped with filters to enable the network to be filtered down to key nodes (for example, common counterparties, nodes that transacted above a certain threshold amount).

3.1.5 Network discovery and scoring

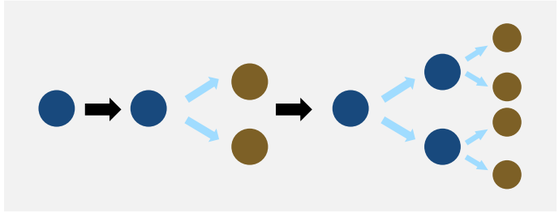
Rethinking Network Detection

As we have seen in earlier parts of this paper, Network Analytics is a key and important tool that many banks have deployed, albeit primarily as a visualisation tool. One member bank has embarked on the journey of “Network Discovery and Scoring”. There are broadly two parts to the solution:

- Detection of suspicious networks – ability to surface networks based on the collective profiles and patterns of parties in the network.
- Risk scoring of networks – ability to assign scores to defined networks for prioritisation.

The bank’s current solution is largely logic-based and was developed in close collaboration with financial crime compliance domain experts, leveraging true cases of suspicious networks to define. The three Network Discovery algorithms created are "Heuristic Expansion", "Transflow" and "Ringfenced Community" where further details are discussed below.

Beyond the three algorithms, the bank has been continuously identifying and experimenting new Network Discovery methods to broaden coverage on network typologies. It has also been experimenting with machine learning algorithms to perform risk scoring.



1. “Heuristic expansion”

This algorithm starts from a suspicious seed node, then develops into a network as it is iteratively expanded along connected party edges to search for other parties. This continues for ‘n’ rounds for as long as additional parties increase the overall score of the network.

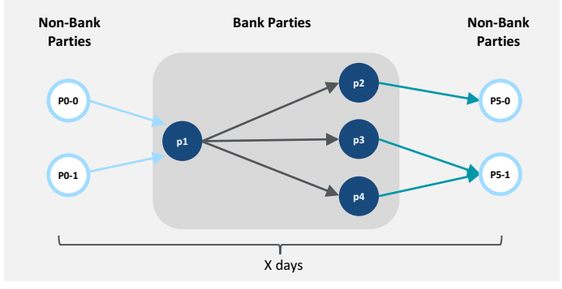
A comprehensive set of heuristic-based features at both node and network level was developed, which the algorithm uses to prune the network at each round of expansion and to decide the best iteration of the network to keep. This best iteration is based on the computed network score, which is rolled up using a combination of scoring assigned to each node and network level feature that is met.

To elaborate on the features used and computation of scoring,

- Examples of node level features could be excessive number of counterparties, or simply high risk CDD attributes
- An example of a network level feature could be the total amount involved in cyclic transactions in network, as percent of total transaction amount in network
- Each feature at both node and network level is assessed during each round of expansion, and their respective scores if hit are summed to give total network score

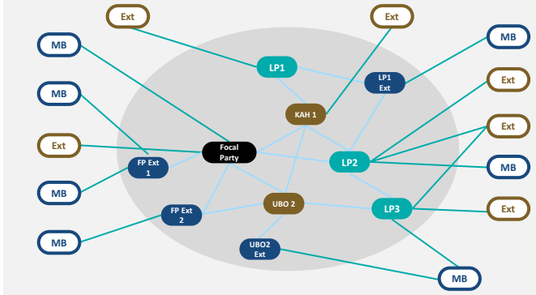
2. “Transflow”

During investigations, cases were identified where the money originates from outside the bank, passes through ‘x’ number of bank internal parties, then exits the bank’s network. The amount is large, and this chain of activities happens within a short timeframe. The Transflow algorithm was developed to detect such transaction patterns and identify the parties involved.



3. “Ringfenced Community”

This typology focuses on ringfencing a network of parties connected by relationships and monitoring the volume of flows between these connected parties vis-à-vis external parties. It assumes that legitimate companies will have a significant proportion of external flows during the course of day-to-day business operations. The diagram below depicts how the Ringfenced Community is being defined.



Legend -

- UBO: Ultimate Beneficial Owner
- KAH: Key Appointment Holder
- FP Ext: Focal Party External (Focal Party based in another Bank)
- LP: Linked Party (Party sharing the same KAH/UBO as the Focal Party)
- MB/Ext: Parties not linked to the Focal Party nor to parties in the ringfenced network

The bank is also working on a few experiments or future improvements, including:

- Machine learning methods to perform alternative ways of Network Scoring, to capture information not only more effectively at the node and edge levels, but also at the broader graph context.
- One of the hurdles faced when developing machine learning based scoring was the lack of labelled networks to learn from. A possible way to overcome this is to develop an algorithm to auto-generate labelled networks which we can then be used to train supervised machine learning models effectively.
- Leveraging graph databases which may help to traverse and explore graph edges faster.
- Finally, integration of the new data derived from Network Discovery and Scoring with the Network Link Analysis tool for front-end user consumption.

3.1.6 Generative AI

As we look further in the future, GenAI, short for "generative artificial intelligence," represents a cutting-edge advancement in the field of artificial intelligence. Its strengths lie in its ability to generate highly realistic and contextually relevant content, making it an invaluable tool for various industries. GenAI can create compelling marketing materials, generate realistic images, and even compose music and literature with remarkable proficiency. However, it also has notable weaknesses, such as the potential for bias in its outputs and a lack of true understanding or consciousness. In the finance sector, GenAI is being harnessed by companies to automate data analysis, streamline customer service through chatbots, and predict market trends. While GenAI offers incredible potential, businesses should consider navigating ethical and regulatory challenges (including data privacy requirements) to ensure responsible and unbiased usage in the finance sector and beyond.

Below we share some possible use cases of incorporating GenAI to aide in managing financial crime risks.

1. Adverse news screening

Adverse news screening requires performing searches via a search engine and reading through the linked results manually which is very time consuming. This job is perfect for GenAI, which can analyse large amounts of information easily and provide valuable summaries for the analyst.

2. Suspicious transaction report filing

When filing a STR, analysts put together a detailed commentary based on the findings observed during the review. GenAI can help to draft the commentary for the analyst to review before submission.

3.2 COSMIC – Industry level data sharing and FI user experience

3.2.1 Background

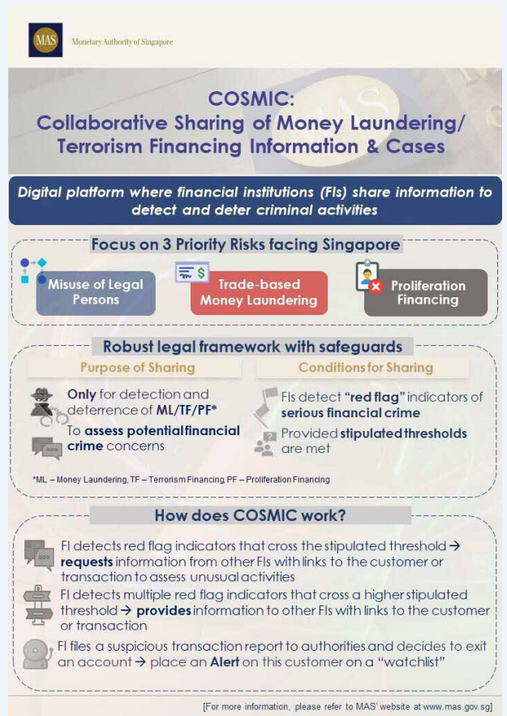
MAS is developing a digital information-sharing platform for AML/CFT, called COSMIC[14]. Through COSMIC, participant FIs will share with one another information on a confidential basis on customers whose profile or behaviour exhibits potential financial crime concerns. This will make it easier for FIs to detect and thereby deter criminal activity, by preventing criminals from exploiting information gaps across FIs.

COSMIC's initial focus will be on addressing three key financial crime risks facing Singapore, namely, misuse of legal persons, trade-based money laundering, and proliferation financing. The initial participants on COSMIC will be six major commercial banks in Singapore which have co-developed the platform alongside MAS, namely, DBS, OCBC, UOB, HSBC, Citibank and Standard Chartered Bank. Sharing will initially be voluntary, to allow participant FIs adequate time to familiarise themselves with this new paradigm. MAS plans to progressively expand COSMIC's coverage to include more FIs and risk areas and make sharing mandatory in higher-risk circumstances.

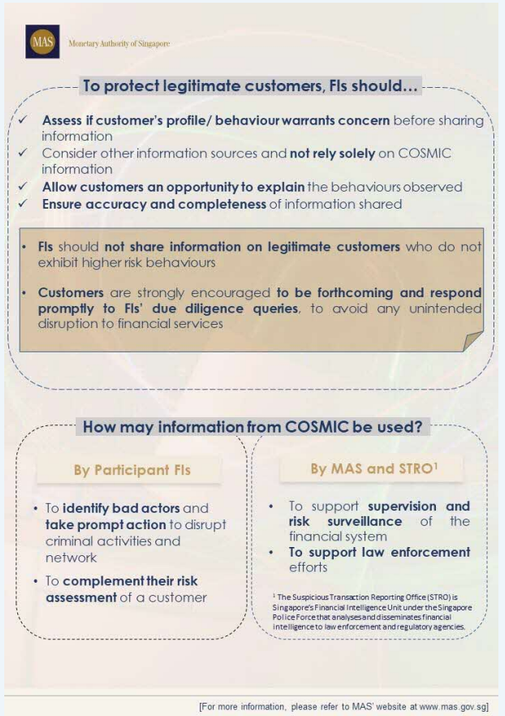
Modes of Sharing

When a customer meets the pre-determined materiality threshold, participant FIs may share information using one of three modes.

- Firstly, where a customer has exhibited some red flags and a participant FI requires more risk information to assess whether the customer is potentially involved in illicit activity, the participant FI may **request** information from another participant FI to facilitate this assessment.
- Next, where the customer's unusual activities cross a higher threshold, indicating a greater risk of the customer being involved in illicit activity, a participant FI can **provide** information to counterparts with nexus to the customer.
- Lastly, where a customer's activities exhibit the higher threshold of red flags, and a participant FI has filed an STR on the customer and decided to terminate the relationship, the FI may **alert** other participants by placing this customer on a "watchlist" within COSMIC, which all participants can check against current or prospective customers.



COSMIC Infographics[15]



COSMIC Infographics[16]

Data Analytics Integration

MAS has worked with the banks on the structured data format to enable FIs to seamlessly apply their data analytics capabilities on COSMIC data alongside their existing internal systems and databases, via APIs. FIs' financial crime analysts will have COSMIC embedded within their own user interfaces, minimising their operational burden.

COSMIC will identify linkages (e.g., via common parties or linked transactions), amongst seemingly unrelated networks and proactively warn the relevant FIs that their customers may be part of a larger illicit network. This will provide additional context which may trigger further reviews that pick out additional suspicious actors.

Legislation and Governance

The Financial Services Markets Act (FSMA) was amended in May 2023 to govern the sharing of information through COSMIC. In preparation for COSMIC's launch, MAS and the participant banks are working closely together to ensure the robustness and security of the COSMIC platform as well as participants' internal governance and controls to safeguard the use of and access to information from COSMIC. This includes setting clear baseline principles to ensure proper use of COSMIC information and to prevent undue de-risking of customers. Lastly, through its regular surveillance, MAS will look out for potential migration of risks to FIs that are not on COSMIC and engage these FIs to strengthen their defences.

[14]Source: MAS paper - Consultation Paper FI-FI- information sharing platform
[15]Source: MAS COSMIC Infographics - COSMIC Infographic page 1
[16]Source: MAS COSMIC Infographics - COSMIC Infographic page 2

3.2.2 Preparing for COSMIC

In preparation for COSMIC, participant FIs can consider the following aspects to reap the benefits of risk effectiveness while ensuring operational efficiency:

Automation of red flags

COSMIC red flags can be systematically identified where feasible (e.g. pass-through nature of transactions) to augment qualitative risk assessment.

Where automation is not feasible, relevant data points can also be aggregated for ease of review.

Enhancing case management system and integrating it with COSMIC

The existing case management system may need to be enhanced to handle COSMIC processes efficiently. Key features could include, but not limited to:

- Approval workflows
- Linkage between alerts and COSMIC cases
- Data aggregation (including automated red flags)
- Data access controls
- Control tower/dashboards

Customer name screening uplift

System capabilities could be uplifted to enable name screening of entities placed on COSMIC Alert list.

These may include the tuning of system to manage alert volumes and management of COSMIC alert list in conjunction with existing name screening watchlists.

Data infrastructure and security

Given the confidentiality of COSMIC data, the development of the technical infrastructure to support COSMIC should consider incorporating security by design, covering controls for data at rest, in transit and in use.

Appropriate access controls may need to be put in place to ensure COSMIC data is only accessible on a need-to-know basis, which should also apply to technical personnel such as database administrators and production support.

Connectivity

As an option to enable systematic end-to-end management of COSMIC data, an FI's internal systems could integrate with the COSMIC platform via APIs.

The development of the APIs should be based on the latest documented technical and security specifications for both FIs and MAS.

Control tower

A control tower is an up-to-date visualization of the key metrics driving service level agreements, operational efficiency, and risk effectiveness. This will serve as feedback to drive continuous improvement of COSMIC related processes.

Examples of metrics include turnaround time for a COSMIC request, number of “provide” cases reviewed per day, and conversion rate of “provide” to “alert” case respectively.

3.2.3 Leveraging COSMIC information for data analytics

Information shared over COSMIC could be leveraged to enhance existing data analytics solutions, for example:

Customer and alert risk scoring

Most banks perform risk scoring on customers or alerts to ensure that resources are focused on higher risk customers and activities. With COSMIC, information shared can be incorporated into risk scoring algorithms in FIs so that appropriate attention and focus is given where necessary.

Entity resolution with additional supplementary information

Being able to resolve entities for sanction, adverse news and watchlist screening has always been difficult, due to limited information about the entities. Information shared over COSMIC could supplement entity data to improve the performance of entity resolution, translating to greater efficiency savings with less false alerts.

Network link analytics

Networks built are currently limited to entities that have transacted with the bank. With COSMIC, these networks can now be enhanced and expanded with non-customer entities, allowing the discovery and identification of bigger financial crime networks.

Also, network link analytics will benefit from improved entity resolution in the previous point, as the networks will be more precise and accurate.

Over time, with greater experience and by working closely with the MAS, the data could also enable us to build on advanced data analytics models and frameworks shared previously, including perpetual KYC, dynamic risk monitoring and machine learning based transactions monitoring.

4 Conclusion

The 2018 publication aimed to provide insights and perspectives on leveraging data analytics for AML/CFT purposes. By initiating or advancing discussions on adopting and implementing such solutions, the publication sought to impact individual FIs at various stages of their analytics journey and across the industry as a whole.

The objective was to encourage broader adoption of data analytics in AML/CFT functions and decision-making processes where appropriate, enhance the utilization of existing tools, and promote effective collaboration between analytics tools and essential human input and judgment. This collaborative approach aimed to strengthen AML/CFT frameworks.

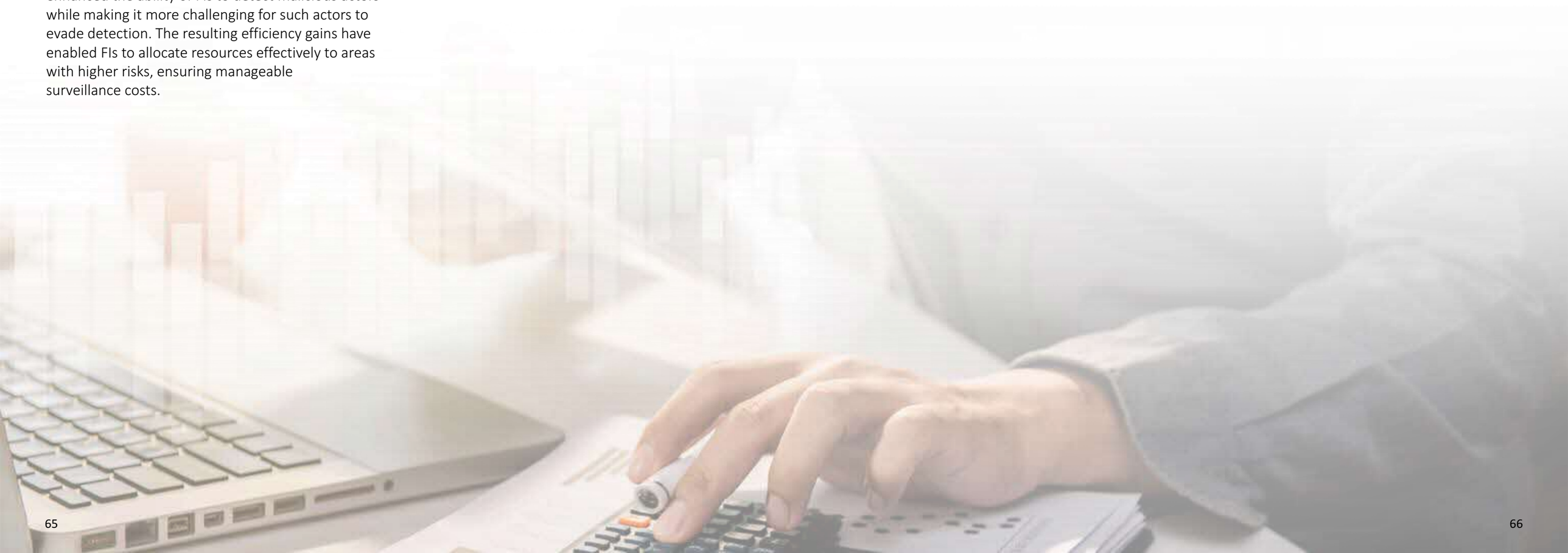
Industry's progress

Since the 2018 publication, FIs have successfully implemented numerous data analytics solutions that incorporate machine learning techniques in various risk areas. These solutions have significantly enhanced the ability of FIs to detect malicious actors while making it more challenging for such actors to evade detection. The resulting efficiency gains have enabled FIs to allocate resources effectively to areas with higher risks, ensuring manageable surveillance costs.

Looking ahead

As technology and analytics continue to advance, enabling the effective analysis of large data volumes, public-private partnerships will reach new heights. With MAS rolling out industry-wide initiatives such as COSMIC to facilitate information sharing across the industry, FIs are enhancing their technical infrastructure to enable seamless transfer of data, allowing for prompt investigations supported by comprehensive information. Intelligence shared could be used to further enhance FIs specific data analytics solutions in the future.

This strong collaborative approach effectively discourages potential threats and safeguards the integrity of the Singapore financial ecosystem.



5 Appendix

5.1 Glossary

Terms	Descriptions
ACIP	AML/CFT Industry Partnership
AI	Artificial Intelligence
AI/ML	Artificial Intelligence/Machine Learning
AIS	Automatic Identification System
AML	Anti-Money Laundering
AUM	Assets Under Management
BAU	Business As Usual
BO	Beneficial Owner
Black box	A system where the internal mechanics are invisible to users.
BU	Business Units
CDD	Customer Due Diligence
CFT	Countering the Financing of Terrorism
COSMIC	Collaborative Sharing of Money Laundering/Terrorism Financing (ML/TF) Information & Cases
DRM	Dynamic Review Monitoring Model
FCC	Financial Crime Compliance
FEAT	Fairness, Ethics, Accountability and Transparency
FIs	Financial Institutions
FIU	Financial Intelligence Unit
FATF	Financial Action Task Force
GenAI	Generative Artificial Intelligence
IBF	Institute of Banking and Finance
IP	Internet Protocol
IRAS	Inland Revenue Authority of Singapore
JIMLIT	Joint Money Laundering Taskforce
KYC	Know-Your-Client
MAS	Monetary Authority of Singapore
Model Explainability	The ability to understand and explain the results generated by the model in clear and simple terms that are easily comprehend by non-technical individuals.
Network Analytics	A study of structures, relationships, behaviours in a network.
Network Link Analysis	A technique used to identify trends and patterns of relationship between individuals/entities in a network.
PEP	Politically Exposed Person
pKYC	Perpetual KYC
PO	Production Order
RFI	Request For Information
SQL	Structured Query Language
STR	Suspicious Transaction Reporting
TECA	Tax Evasion Clustering Analysis
TF	Terrorism Financing
TM	Transaction Monitoring
UBO	Ultimate Beneficial Owner
VPN	Virtual Private Network
WG	Working Group – representatives from commercial banks operating in Singapore. Refer to Section 5.2.
XGB	eXtreme Gradient Boosting – a machine learning algorithm.

5.2 Data analytics WG members and other contributors

WG Members

Firm	Representative(s)
DBS Bank Ltd	Harsh Narula (Co-Chair) Ivena Seow Elisa Ang
Standard Chartered Bank	Tanty Muliani (Co-Chair) Grace Ho Toh Ying Hui Emmanuel Goh
Citibank N.A	Amit Kumar Alvin Ng Rajeev Radhakrishnan Nair Raja Chandrahasan Dan Ting Lip
HSBC Bank (Singapore) Ltd	Hazel Tan Kyle Austin
JPMorgan Chase Bank	Santanu Biswas Jerry Leong
Maybank Singapore Ltd	Jon Yeo Tung Kheng Chen Jee Meng
Oversea-Chinese Banking Corporation Ltd	Terence Gue Tricia Lee
UBS AG	Ravi Kakad Andrew Barker
United Overseas Bank Ltd	Denny Irawan Huah Cheng Jiann

Invitees

Firm	Representative(s)
Deloitte & Touche LLP	Kalyani Vasan Nai Seng Wong Claire Franklin

ACIP Secretariat Representatives

Monetary Authority of Singapore

Commercial Affairs Department, Singapore Police Force