Financial Stability
Institute

FSI Insights
on policy implementation
No 18

Suptech applications for anti-money laundering

by Rodrigo Coelho, Marco De Simoni and Jermy Prenio

August 2019

JEL classification: C45, G21, G38, O32
Keywords: anti-money laundering, suptech, innovation, data analytics

BANK FOR INTERNATIONAL SETTLEMENTS

FSI Insights are written by members of the Financial Stability Institute (FSI) of the Bank for International Settlements (BIS), often in collaboration with staff from supervisory agencies and central banks. The papers aim to contribute to international discussions on a range of contemporary regulatory and supervisory policy issues and implementation challenges faced by financial sector authorities. The views expressed in them are solely those of the authors and do not necessarily reflect those of the BIS or the Basel-based committees.

Authorised by the Chairman of the FSI, Fernando Restoy.

This publication is available on the BIS website (www.bis.org). To contact the BIS Media and Public Relations team, please email press@bis.org. You can sign up for email alerts at www.bis.org/emailalerts.htm.

ISSN 2522-2481 (print)
ISBN 978-92-9259-287-5 (print)

ISSN 2522-249X (online)
ISBN 978-92-9259-288-2 (online)

Contents

Suptech applications for anti-money laundering[1]

Executive summary

Use by financial authorities of advanced data collection and analytics tools enabled by new technologies is collectively called suptech. In the area of data analytics, development of such tools has been facilitated by advances in artificial intelligence and its practical application in machine learning, natural language processing and other advanced analytics capabilities. These tools have provided opportunities to enhance financial authorities' capacity.

Detecting potential anti-money laundering (AML) and combating the financing of terrorism (CFT) violations is one field where data analytics tools seem more advanced. This paper therefore dives deeper into these tools. In particular, it aims to explore the various data analytics tools used by authorities tasked with AML/CFT responsibilities, as well as their practical experiences in using such tools. Nine AML/CFT authorities are covered in this paper. Such authorities have either supervision or financial intelligence functions, or both.

AML/CFT supervision and financial intelligence functions have different mandates. Authorities with AML supervision functions are expected to ensure compliance by financial institutions with requirements to combat money laundering (ML) and terrorist financing (TF). Authorities with financial intelligence functions, ie financial intelligence units (FIUs), meanwhile, are expected to serve as national centres for the receipt and analysis of suspicious transaction reports and other information relevant to money laundering, and to disseminate the results of that analysis. FIUs sometimes also have AML supervision functions.

Both AML/CFT supervisors and FIUs need advanced data analytics tools to analyse the large volumes of information at their disposal. AML/CFT authorities typically receive substantial amounts of transactional and non-transactional data. On top of these traditional sources of data, some AML/CFT authorities are now actively collaborating with other government agencies and private entities to expand the scope of data available to them. Some authorities are also exploring the use of non-traditional sources of information (eg newspaper articles, social media) and integrating them with traditional information to come up with richer analyses.

The difference in mandates does not seem to affect the types of advanced data analytics tools the AML/CFT authorities are pursuing. AML/CFT authorities covered in the paper are in general pursuing similar advanced data analytics tools, such as network analysis, natural language processing, text mining and machine learning. These tools increase their ability to detect networks of related transactions, to identify unusual behaviours and in general to transform significant amounts of structured and unstructured data into useful information that contributes to their respective processes.

Authorities have used different strategies to develop these tools. AML/CFT authorities that are within the central bank, or prudential or conduct authority, generally benefit from the institutional

[1] Rodrigo Coelho (rodrigo.coelho@bis.org) and Jermy Prenio (jermy.prenio@bis.org), Bank for International Settlements, and Marco De Simoni (marco.desimoni@bancaditalia.it), Bank of Italy.

The authors are grateful to the representatives from the authorities interviewed; to the participants of the second FSI meeting on the use of innovative technology in financial supervision held in Basel on 5-6 June 2019 for the insightful discussions; and to Kuntay Celik, Matei Dohotaru, Tom Neylan and Greg Sutton for helpful comments and insights. We are also grateful to Esther Künzi and Cissy Mak for valuable support with this paper.

strategy to utilise innovative technology to help in supervision work and can develop these solutions in-house. For some AML/CFT authorities, taking advantage of ready solutions in the market may be more efficient. Others are actively collaborating with the academic community and promoting research in this field. Many of the authorities use a combination of these approaches. The optimal solution for a specific authority will depend on several factors such as the profile of the authority, the characteristics of the financial system that the authority oversees, and the legal framework in which the authority operates.

Efficiency gains seem to be the number one benefit of advanced data analytics tools, which could help capacity-constrained AML/CFT authorities. AML/CFT authorities have highlighted the gains in terms of time savings they achieved in using these advanced data analytics tools. This could translate to reallocation of resources or capacity from more manual work to more judgment-based work. Assessing effectiveness particularly of tools used by FIUs, however, is not that straightforward.

The benefits that these tools bring are particularly important for jurisdictions that have been heavily impacted by the unintended consequences of AML/CFT international standards, particularly de-risking. Jurisdictions most frequently exited by global correspondent banks seem to be those with weak AML/CFT supervisory and regulatory frameworks. That being so, the development of advanced data analytics tools as well as development of necessary skills could help strengthen these jurisdictions' frameworks and potentially reverse this trend.

However, the use of these innovative technologies gives rise to a number of challenges. First, computational capacity may be an issue, since these tools deal with large volumes of data. Second, data privacy and confidentiality requirements provide safeguards that AML/CFT authorities must consider in using certain data and external resources in developing data analytics tools. Third, assessing the effectiveness of these tools might be challenging, in particular for FIUs given the necessary time to prove the occurrence of a money laundering activity. Finally, tools based on supervised machine learning could lose their effectiveness over time, especially if not regularly updated with new training data, given the capacity of criminal organisations to change their behaviour in order to avoid detection.

There is scope for information-sharing among AML/CFT authorities on the data analytics tools they are developing or using in order to promote peer learning. Although the data analytics tools used by AML/CFT authorities are tweaked to reflect their mandates, the underlying methodologies of these tools are quite similar. There are therefore opportunities for peer learning through regular exchange of information and sharing of experiences on the development and use of these tools.

AML/CFT authorities that are just starting to develop their data infrastructure have a "late mover" advantage and may find it easier to integrate advanced data analytics tools. These authorities have the advantage of developing their data infrastructure from scratch without the burden of legacy systems. They can design it in a way that makes the data collection, validation and management processes seamless, while more easily enabling the integration of newly developed analytical tools.

ML/TF risks have international reach, so development of data analytics tools that are international in scope should be considered. The tools discussed in this paper are all national in scope. Money laundering, however, is an international issue, and criminal organisations tend to exploit loopholes anywhere in the world. Therefore, a strong argument could be made for international cooperation and collaboration in terms of developing data analytics tools with an international coverage.

Section 1 – Introduction

1. **Use by financial authorities of advanced data collection and analytics tools enabled by new technologies is collectively called suptech.**[2] In the area of data analytics, development of such tools has been facilitated by advances in artificial intelligence (AI) and its practical application in machine learning, natural language processing (NLP) and other advanced analytics capabilities. These tools have provided opportunities to enhance financial authorities' capacity.

2. **These suptech data analytics tools can be used by authorities for anti-money laundering (AML) and combating the financing of terrorism (CFT) purposes.** Broeders and Prenio (2018) provides a comprehensive overview of the suptech tools used by different authorities. In the area of data analytics, suptech tools can be found in market surveillance, misconduct analysis, microprudential supervision and macroprudential supervision. Detecting potential AML/CFT violations is one field where data analytics tools for misconduct analysis are used and seems more advanced.

3. **The use of these tools has the potential to significantly contribute to reducing money laundering (ML) and terrorist financing (TF) risks by strengthening jurisdictions' AML/CFT frameworks.** One estimate of AML/CFT compliance cost by financial firms put it at USD 25.3 billion per year for US financial services firms alone.[3] AML/CFT authorities are also devoting substantial resources to fight ML/TF to the extent possible. However, many of them, particularly those in emerging market and developing economies, have severe capacity constraints. Despite these efforts, estimates of money laundered worldwide are still staggering.[4] In addition, there seems to be slow progress globally in reducing ML/TF risks.[5] This has wide-ranging economic and social repercussions. The Financial Stability Board (FSB), for example, noted that jurisdictions most frequently exited by global correspondent banks seem to be those with weak AML/CFT supervisory and regulatory frameworks.[6] This has implications for international trade and remittance flows to these jurisdictions.

4. **In particular, these tools can help AML/CFT authorities[7] analyse the large volumes of information at their disposal.** AML/CFT authorities typically receive substantial amounts of transactional data from reporting institutions. These are then supplemented by non-transactional data (eg data from tax, customs and property registration authorities) to provide context. Some AML/CFT authorities are now actively collaborating with other government agencies and private entities to expand the scope of data available to them. Some authorities are also exploring the use of non-traditional sources of information (eg newspaper articles, social media) and integrating them with traditional information to come up with richer analyses. The emergence of big data analytics enables the integration of all this information from different sources to tell a coherent story.

5. **Advanced data analytics tools could also potentially improve the effectiveness of AML/CFT authorities.** Financial institutions traditionally relied on rule-based AML/CFT measures, which are shown to generate false positive alerts of around 90–95% leading to substantial resource implications.[8] So they

[2] Broeders and Prenio (2018) defines suptech as the use of innovative technology by supervisory agencies to support supervision. For this paper, we broadened the definition of suptech to include advanced data collection and analytics tools used by financial authorities, ie not only supervisors or regulators but other authorities as well, such as financial intelligence units (FIUs).

[3] LexisNexis Risk Solutions (2018).

[4] See for example MONEYVAL (2017).

[5] See eg Basel Institute of Governance (2018).

[6] FSB (2015).

[7] In this paper, the expression "AML/CFT authorities" is meant to encompass both AML/CFT supervisors and FIUs.

[8] Saaradeey et al (2019).

are now increasingly using machine learning applications to reduce the number of false positives significantly. The same improvements could be expected from the use by AML/CFT authorities of these advanced data analytics tools.

6. **While existing information remains scarce, there are already a number of advanced data analytics tools that are being pursued by different AML/CFT authorities.** A few papers (eg Kraft (2018), di Castri et al (2018)) as well as authorities' publications (eg AUSTRAC (2018)) and newspaper articles individually provide glimpses of the different initiatives undertaken by individual authorities in this field. In addition, the overview paper on suptech tools by Broeders and Prenio (2018) reveals that a number of authorities are already actively exploring such tools for AML/CFT purposes.

7. **International dialogue on the use of advanced data analytics by AML/CFT authorities, however, seems limited.** This has been pointed out in Kraft (2018), which sets out a framework for looking at the use of data analytics to develop intelligence on money laundering risks. It notes that the limited international dialogue inhibits opportunities for peer learning despite some concrete work already being done in this field. It should be noted, however, that the Egmont Group has an Information Exchange Working Group with a workstream on enhancing data analytics and effective communications. The work stream aims at "embracing innovation in technology to enable FIUs to be agile, responsive and effective when exchanging and exploiting information".[9]

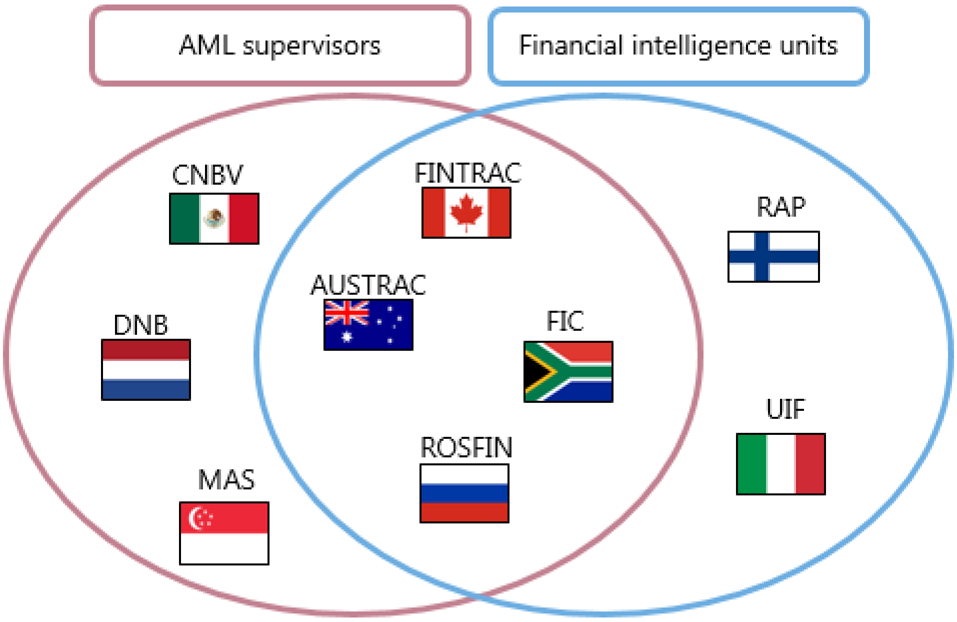
8. **This paper therefore aims to contribute to this international dialogue by focusing on and providing some specific examples and insights on the development of these tools.** Nine AML/CFT authorities are covered in this paper. These authorities were selected based on information available about their data analytics work or from recommendations from other AML/CFT authorities. Most of them were interviewed, while publicly available information was used for a few of them.

9. **AML/CFT authorities covered in this paper have either supervision or financial intelligence functions, or both (Figure 1).** Countries organise their AML/CFT authorities in different ways. AML/CFT supervision, for example, may be within the central bank or prudential or conduct authority,[10] or it may be within the financial intelligence unit (FIU). The financial intelligence function, on the other hand, may be a standalone entity, an agency attached to but independent from the central bank, or part of the law enforcement function.

[9] Egmont Group (2018).

[10] Prudential and conduct functions are also organised differently. Depending on the supervision model used, they may be under one body or separate bodies within or outside the central bank.

Figure 1 – AML/CFT authorities covered in the paper



AUSTRAC - Australian Transaction Reports and Analysis Centre
CNBV – National Banking and Securities Commission of Mexico
DNB – Netherlands Bank
FIC - Financial Intelligence Centre of South Africa
FINTRAC - Financial Transactions and Reports Analysis Centre of Canada
MAS - Monetary Authority of Singapore
RAP - Financial Intelligence Unit of Finland
ROSFIN - Financial Intelligence Unit of Russia
UIF - Financial Intelligence Unit of Italy

10. **AML/CFT supervision and financial intelligence functions have different mandates.** The Financial Action Task Force (FATF) Recommendations (FATF (2012)) expect AML/CFT supervisors "to supervise or monitor, and ensure compliance by, financial institutions with requirements to combat money laundering and terrorist financing". FIUs, meanwhile, are expected to serve as national centres "for the receipt and analysis of: (a) suspicious transaction reports; and (b) other information relevant to money laundering, associated predicate offences and terrorist financing, and for the dissemination of the results of that analysis".[11]

11. **The paper describes advanced analytics tools being used or developed by the covered AML/CFT authorities, and their practical experience in using/developing these tools.** Specifically, section 2 discusses the data-related challenges faced by AML/CFT authorities in pursuing their mandates. Section 3 dives into the specific data analytics tools that AML/CFT authorities use to address these challenges. It also discusses how authorities develop these tools and the associated issues they face. Section 4 provides some concluding thoughts.

[11] FATF Recommendations 27 and 29, respectively.

Section 2 – Data-related challenges faced by AML/CFT authorities

12. **AML/CFT authorities conduct data analysis in order to identify, assess and understand the money laundering and terrorist financing risks in their jurisdictions.** FATF recommendation 1 states that countries should identify, assess and understand their money laundering and terrorist financing risks. AML/CFT authorities are responsible for undertaking these assessments on an ongoing basis. These assessments are aimed to inform potential changes to AML/CFT laws, rules and regulations, help in the risk-based prioritisation of work and effective allocation of AML/CFT authorities' resources, and serve as input to the AML/CFT assessments conducted by financial institutions and designated non-financial entities.

13. **In conducting data analyses, AML/CFT authorities rely on both transactional and non-transactional information.** AML/CFT authorities typically look at transactional information in conducting their analyses. FIUs, for example, receive suspicious transaction reports (STRs) that consist of a combination of structured information (eg amount and entities involved in the transactions) and unstructured information (eg the reasons why the transactions are classified as suspicious). In general, AML/CFT supervisors do not have access to STRs, so they tend to look at transactional information that may exceed a certain threshold (eg data on wire transfers above a certain amount). AML/CFT authorities may also look at aggregated transactional data, such as all transactions involving certain counterparties or jurisdictions. AML/CFT authorities need to make sense of all this transactional information, so they also look at non-transactional information to provide context to the transactions – for example, information relating to politically exposed persons, countries in sanctions lists, property or business ownership, and business relationships.

14. **A number of authorities have gone beyond the traditional data sources in order to increase their AML/CFT surveillance capacity.** CNBV and UIF, for example, use information obtained systematically from press articles to enhance some of their monitoring processes. CNBV also makes use of information coming from social media and a customer database consisting of information about financial institutions' clients, including data from know-your-customer questionnaires and the client risk rating. Furthermore, FIC has created an open source analysis capability and uses information obtained from open sources to enhance its monitoring processes and intelligence products.

15. **In addition, a few countries have created collaboration platforms through which data are shared.** In 2017, the Fintel Alliance was created in Australia to combat money laundering, terrorism financing and other serious crime. It consists of a number of member institutions from the government and the private sector – both domestic and international – which work together, sharing and analysing financial intelligence. Its main objectives are to build up the resilience of the financial sector against criminal exploitation and to support law enforcement investigations into serious crime and national security matters. In the Netherlands, a joint venture among a number of public entities (DNB, police, prosecutors, tax authority, customs, and data providers such as the Chamber of Commerce and the Land Registry) has been established to share data from different sources and provide access to this large data set to each participant.[12]

16. **The combination of all this information results in large volumes of data.** For example, UIF receives about 100,000 STRs per year and about 100 million records of the SARA database (a database comprising anonymous and aggregated transactions above EUR 15,000). FINTRAC, meanwhile, receives almost 180,000 STRs, over 10 million large cash transactions and over 14 million electronic funds transfer reports per year. Moreover, FIC receives more than 330,000 STRs and over 4.8 million cash threshold and cash threshold aggregation reports per year (FIC (2018)).

[12] https://icov.nl.

17. **The need to analyse large volumes of data results in certain challenges, particularly in terms of data collection and management.** The existing data collection process at CNBV is still a time-consuming and inefficient manual operation. As a consequence, CNBV is now undertaking a project, with the help of Regtech for Regulators Accelerator (R^2A), which will allow financial institutions to submit reports digitally and automatically. The project includes central data storage, as well as automatic visualisation tools. The use of smart contracts is also being explored to address the data collection challenges. This is the case, for example, at AUSTRAC, which is now undertaking a project that involves using smart contracts for reporting international funds transfer instructions.

18. **With large volumes of data, data quality could also be a challenge.** This is particularly the case since most of the data that AML/CFT authorities analyse are generated from supervised entities' reports. Supervised entities range from big global banks to very small entities, with widely varying reporting capabilities. Authorities have established several automatic ways of checking the quality of data and sending them back to the supervised entity if necessary. Validations can be embedded in the automatic data collection process, which is what CNBV and AUSTRAC are developing. DNB, on the other hand, is working on the standardisation of data received, taking into account the characteristics of supervised entities. The data received will be automatically compared with data sent by peers to control for consistency in data volume and other dimensions.

19. **Matching data generated from different databases could also cause a problem.** This is the case, for example, in Australia, where there is no national ID system and so each government agency maintains its own database. In order to address this, data matching projects are being undertaken between AUSTRAC and other government agencies as well as private entities within the Fintel Alliance. In Europe, FIU.NET, a decentralised computer network supporting FIUs in the European Union, has developed a project that matches data from different FIUs when they exchange information.

20. **While automation of data collection and validation processes addresses some of the challenges, use of advanced data analytics tools is still necessary to enable the analysis of large volumes of information from disparate sources.** AML/CFT authorities are already exploring such data analytics tools in order to more efficiently analyse the information at their disposal and come up with a complete and coherent picture of AML/CFT risk.

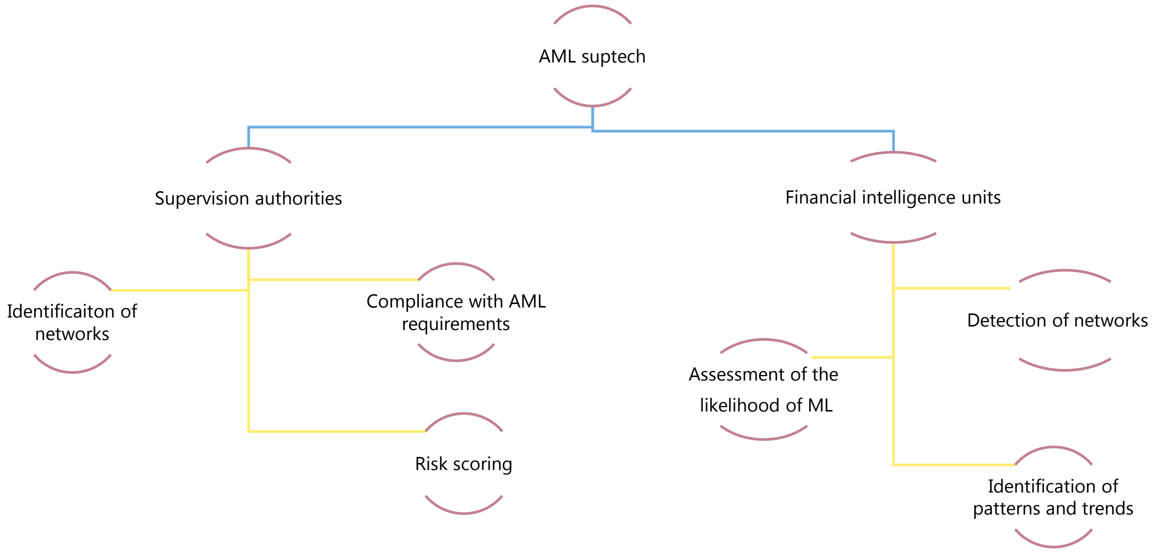
Section 3 – Data analytics tools

21. **AML/CFT supervisors and FIUs use similar innovative data analytical tools to achieve their objectives.** This includes network analysis, machine learning, natural language processing and text mining. Albeit with different objectives, both types of authorities use these tools to increase their ability to detect networks of related transactions, to identify unusual behaviours, and in general to transform significant amounts of structured and unstructured data into useful information that contributes to their respective processes.

22. **Advanced analytics employed by AML/CFT supervisors are mainly targeted at improving the compliance assessment of individual financial institutions.** The usual mandate of AML/CFT supervisory authorities is to supervise or monitor financial institutions and ensure their compliance with requirements to combat ML/TF. In order to do so, such authorities must assess the governance, risk management, internal controls, processes and systems in place in individual financial institutions to make sure they have sufficient capacity to deter the flow of illicit funds. This integrity risk assessment typically results from a combination of onsite and offsite activities. The advanced analytical tools developed and/or used by AML/CFT supervisory authorities seek to enhance the offsite assessment of the risk profile of individual institutions, which also contributes to the risk-based prioritisation of onsite inspections.

23. **FIUs' focus is primarily on identifying suspicious activity at both the micro and the macro level.** At the micro level, FIUs seek to identify suspicious activity in individual or networks of transactions to support their intelligence activities and also to contribute to the work of law enforcement authorities. Visualisation and automation tools which seek to find suspicious flows with a user-friendly interface belong to this category. At the macro level, FIUs' objective is to monitor trends and patterns in suspicious activities in order to improve their analysis and to raise market participants' awareness of the most relevant typologies. Data analytics tools help FIUs mainly in two ways. First, these tools allow FIUs to assess the likelihood of occurrence of money laundering activity and therefore help in the prioritisation of their work. Second, since these tools detect trends and patterns in potential ML/TF activities, they could decrease the reaction time of FIUs to a changing environment. This is important because money laundering is constantly evolving as criminals try to find new ways to launder illicit proceeds.

Figure 2 – Objectives of the tools employed by AML supervisors and FIUs


AML suptech
Supervision authorities
Financial intelligence units
Identificaiton of networks
Compliance with AML requirements
Risk scoring
Detection of networks
Assessment of the likelihood of ML
Identification of patterns and trends

Tools developed or used by AML/CFT supervisors

Risk scoring of supervised entities

24. **Advanced data analytics may be used to assess the overall risk of financial institutions.** In such cases, tools typically assign a rating to each financial institution depending on the likelihood of non-compliance with AML/CFT requirements. FINTRAC, for example, has developed a heuristic model that uses several risk factors related to an institution's profile, compliance history, reporting behaviour and other intelligence. These risk factors are determined by using different methods such as principal component analysis and geospatial analysis (eg closeness to borders) and are based on a large amount of data (eg STRs, wire transfer reports, large cash transactions, casino disbursements) received from reporting entities as well as from various other sources. Based on these risk factors, the model then ranks the institutions according to their likelihood of non-compliance with AML/CFT regulations. A challenger model is also developed each year using a combination of machine learning and heuristic methods to compare the performance of the model in production and update it, if required.

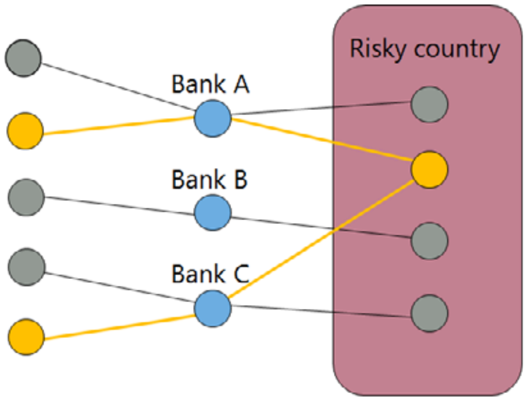
Identification of networks to which financial institutions are exposed

25. **A number of supervisory authorities use network analysis to assess the exposure of supervised entities to money laundering risk.** The construction of networks is not a goal in itself but the means to assess the compliance of financial institutions with AML/CFT requirements. Results of these analyses are normally used to assist in the prioritisation of onsite examinations and are sometimes shared with FIUs and law enforcement to support their work. DNB, for example, has set up a tool to assess the exposure of financial institutions to networks of suspicious transactions (Box 1).

Box 1

DNB application of network analysis using transactional data

DNB has set up a tool that uses transactional data to detect networks of related entities sending funds to the same counterparties in high-risk countries via different financial institutions (money transfer offices). Based on the outcomes of this analysis and the existing links with these networks, the risk profile of financial institutions is assessed. This assessment also serves as an input to define how supervisory resources are to be allocated. The tool is operational, and its results have already been used in a number of police operations and even led to the suspension of the licence of one financial institution.



26. **Networks can also be built using unstructured data.** MAS applies network analysis to STRs. Currently, network analysis is performed on the structured fields of the STRs. MAS is exploring the possibility of augmenting its current network analysis solution to include information extracted from the unstructured text of the STRs. This can be done by applying NLP to extract information such as the relationships of parties mentioned in the unstructured text of a single STR. Similarly, CNBV is developing a prototype application that looks at news referring to selected people and companies involved in money laundering, and then searches for these entities in several other data sources (eg social media, transactional data) in order to create networks of people or companies that may be engaging in money laundering activities.

Evaluation of financial institutions' compliance with AML/CFT requirements

27. **Innovative technologies increase the authorities' capacity to process substantial amounts of unstructured data.** At DNB, one of these tools, which is in the development stage, is used to increase the efficiency in the analysis of numerous, varied and sometimes lengthy documents called systemic integrity risk analysis (SIRA) reports, which are self-assessments that banks submit about their integrity risks. The tool allows supervisors to pose standardised questions and then searches for pieces of information across a large number of reports (eg number of high-risk customers) in order to answer the questions. The objective of the tool is to assess whether the bank has sufficient understanding of the risks

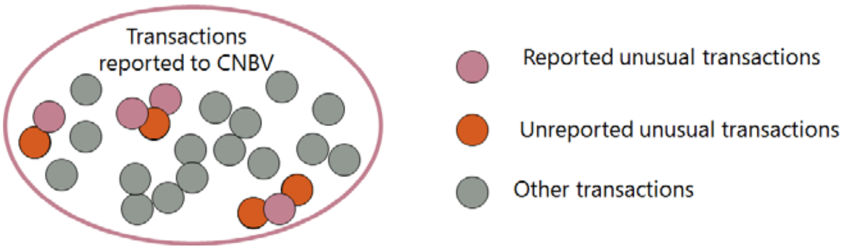
it is exposed to and also if internal controls are adequate to deal with these risks. The tool will use a combination of text mining and supervised machine learning to extract the responses. CNBV, meanwhile, is developing a tool that would identify unreported unusual transactions (Box 2).

Box 2

CNBV application of machine learning to transactional data

CNBV is testing a machine learning solution developed in collaboration with R^{2}A for the detection of unusual transactions which have not been reported by financial institutions. In this context, “manually found” as well as reported unusual transactions are used to train the machine, which then scans the whole data set to uncover other transactions with similar patterns. One of the underlying goals of this tool is to assess the risk profile of financial institutions based on the number of unreported unusual transactions. In other words, financial institutions with a large number of unreported unusual transactions are deemed more risky and therefore would be subject to more intensive supervision.

This prototype mainly relies on two data sources: a transactional database consisting of all transactions made by the clients of a financial institution in a given period, which includes those marked as risky and reported to the FIU; and a customer database containing personal information about the financial institution’s clients including data from know-your-customer questionnaires and the client risk rating. This prototype is based on cloud technology as a cost-effective storage solution for increasing volumes of surveillance data.


Transactions reported to CNBV
Reported unusual transactions
Unreported unusual transactions
Other transactions

Tools developed or used by financial intelligence units

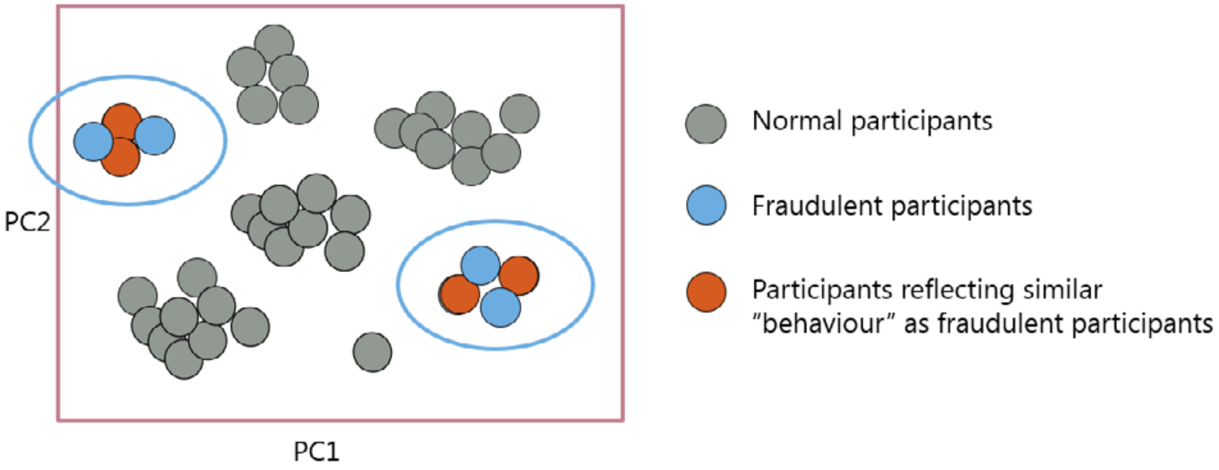
Detection of networks of entities involved in suspicious activities

28. **A number of FIUs are using network analysis to help uncover criminal activities.** In order to do so, FIUs mainly rely on the STR database, but they can also take advantage of other data sources at their disposal. Using these data, FIUs can combine information provided by various reporting entities in different periods of time to uncover hidden relations and determine the network. UIF, for example, has developed a tool that builds networks using STRs. UIF is also exploring a tool that would apply a combination of network analysis and self-organising map techniques in its gold declaration database (Box 3). Similarly, ROSFIN is currently developing a tool that will use STRs and other available information (eg public contracts, customs declarations, data on directors/founders of companies) to build the networks. FINTRAC uses STRs, wire transfers and large cash transactions to construct potential criminal networks in Canada. AUSTRAC, on the other hand, has been applying network analysis on its remittances database to detect suspicious cross-border activities.

Box 3

UIF application of network analysis and self-organising maps to the gold database

In Italy, UIF is testing a tool that combines network analysis and self-organising maps techniques to search the gold declaration database for fraudulent schemes. This process involves several steps. The first one is to build the network of all participants in the gold market using network analysis. The second step is to determine the behaviours of the participants in the network. These behaviours are defined by the characteristics of each participant in the network given by vector of descriptive and centrality network metrics. The third step is to reduce the number of dimensions of these characteristics by applying principal component analysis. On this data set, a self-organising map is trained in order to identify gold market operators' typical behaviour. Finally, using data from behaviours of fraudulent schemes, the previously trained self-organising map is applied to look for participants with similar characteristics.


PC2
PC1
Normal participants
Fraudulent participants
Participants reflecting similar "behaviour" as fraudulent participants

Assessment of the likelihood of money laundering activity

29. **Innovative tools may be used to assess the likelihood of money laundering activity in individual transactions.** In Finland, RAP has developed a tool that assigns a score to each STR based on several criteria such as the number of times the party has been reported in the last six months or whether the party has a criminal background or is subject to an ongoing investigation at the FIU. ROSFIN, on the other hand, has developed a tool that uses machine learning to detect potentially fictitious companies based on a sample of known shell companies. UIF is using big data to monitor wire transfers to and from selected countries. By combining structured and, to a lesser extent, unstructured data (eg press articles), the tool is able to calculate indicators that help in measuring the degree of anomaly of each flow.

Identification of patterns and trends in criminal activities

30. **FIUs also employ advanced data analytics tools to identify new trends and patterns in money laundering.** Money laundering methods evolve over time and authorities need tools to be able to react quickly in this fast-changing environment. The early identification of new typologies allows FIUs to disseminate this information to market participants and other authorities and also supports their own investigative efforts.

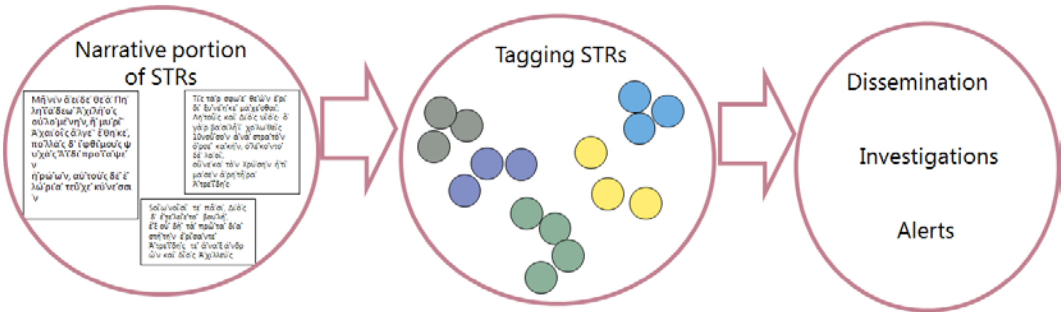
31. **Automatic classification tools help authorities identify new trends or assess how ML/TF typologies evolve over time as well as in specific geographical areas.** UIF, for example, is currently developing a tool to classify STRs according to the type of money laundering scheme perpetrated. First, information is extracted from the unstructured part of the STR using NLP. Then a supervised machine

learning technique is used to classify the STRs according to the identified typology. Similarly, FINTRAC is developing a text mining tool to classify STRs (Box 4).

Box 4

FINTRAC application of text mining to STRs

FINTRAC is developing a text mining tool to assess trends and patterns in ML/TF. The tool analyses narrative portions of STRs and discovers key concepts related to ML/TF, such as types of offences, criminal groups, jurisdictions of interest, and types of products and services exploited. With these key concepts as a basis, all STRs will be tagged, which allows FINTRAC to have a multidimensional view of the trends observed within the STRs and to provide actionable insights for its strategic intelligence group. The results will also help FINTRAC in developing strategic intelligence products such as operational alerts more efficiently and disseminating them to reporting entities to raise awareness about new typologies and trends.



How do AML/CFT authorities develop these tools?

32. **AML/CFT authorities explore or develop advanced data analytics tools in different ways.** Some are building in-house capabilities, while others are taking advantage of ready solutions in the market. Some are also actively collaborating with the academic community and promoting research in this field.

33. **AML/CFT authorities that are within the central bank or prudential or conduct authority generally benefit from the institutional strategy to utilise innovative technology to help in supervision work.** For example, UIF benefits from the Bank of Italy's multidisciplinary team on big data, which built a hardware and software infrastructure to deal with various types of big data analytics needs of the bank. DNB, meanwhile, has established a Supervision Innovation Office that coordinates the implementation of the digital strategy of the bank. It also has an informal Data Science Hub that serves as a venue for sharing ideas among people within the bank who are interested in data science. At MAS, the data analytics needs of AML/CFT supervisors are supported by its Supervisory Technology Office, which partners with MAS supervision departments in exploring and developing analytical applications.

34. **The mandate of FIUs is also leading them to explore advanced analytics tools.** The international community recognised the need on the part of FIUs for tools to help collect and analyse huge amounts of data. The United Nations Office on Drugs and Crime (UNODC), for example, developed the goAML application. It is an integrated database and intelligent analysis system with three primary solutions for collection, analysis and data exchange. There are now 49 FIUs around the world that have

deployed goAML.[13] In addition, some FIUs are also exploring advanced analytics on their own to support their work. FIC, for example, is looking at the possibility of supplementing the goAML application with an AI algorithm. FINTRAC has a Data Exploitation Laboratory, which explores and implements data analytics solutions. AUSTRAC, on the other hand, has analytics, innovation, development, information and infrastructure teams that aim to ensure that advanced analytics are in place to enable AUSTRAC to innovate.

35. **For some AML/CFT authorities, taking advantage of ready solutions in the market may be more efficient.** This is the case at ROSFIN, which bought ready solutions involving three external vendors.[14] UIF, meanwhile, relies on an external vendor and open source tools to build dashboard and visualisation tools that can analyse a large amount of data in a few seconds and can combine data from different sources. CNBV also relies on an external vendor to build several solutions.

36. **Some AML/CFT authorities are actively collaborating with the academic community and promoting research in this field.** Two projects at AUSTRAC are being undertaken in collaboration with two universities: (i) AML alert indicators are being developed with the Australian National University; and (ii) reporting based on smart contracts is being developed with Swinburne University. ROSFIN, on the other hand, is collaborating with the Lebedev Physical Institute for a project on network analysis to find hidden relationships in ROSFIN's databases as well as application of AI techniques, including recognition of patterns and detection of subgraphs within a multigraph. Moreover, UIF organises an annual workshop in collaboration with Bocconi University, banks' compliance officers, law enforcement agencies and judicial authorities to share ideas on quantitative methods to counter economic crime.

Issues associated with the development and use of these tools

Computational capacity

37. **Computational capacity may be an issue since these tools deal with large volumes of data.** AML/CFT authorities, however, are looking for ways to deal with this issue. UIF, for example, relies on the work done by the IT department of Bank of Italy, whose infrastructure employs Apache Spark, an open source layer. Spark bundles together the computational power of a group of heterogeneous computers, providing the user a platform that seamlessly combines all the computing resources of every computer in the cluster. A few authorities are also looking at using public cloud, but data residency requirements as well as data security concerns prevent them from doing so.

Data privacy and confidentiality

38. **Data privacy and confidentiality requirements provide safeguards that authorities must consider, including with regard to which information they can use.** For instance, the General Data Protection Regulation in Europe could limit the use of some detailed personal information for AML/CFT authorities' profiling activities. Authorities have to check with their legal units which information they can or cannot use in developing their data analytics tools before proceeding with the project. This is particularly the case for information from non-traditional sources.

39. **Data privacy and confidentiality concerns also constrains the use of external resources in developing data analytics tools.** Specialised skills are required to develop these tools, and sometimes these are not available internally. AML/CFT authorities may need to resort to external parties, but that

[13] https://unite.un.org/goaml/.

[14] One solution is an offline, fully functional version that is used without any support from the first vendor. Another solution was purchased as a baseline solution from the second vendor, and was then integrated into ROSFIN's system by the third vendor, which has a long-standing partnership with ROSFIN and has been cleared to work with restricted data.

could mean giving them access to confidential information. In order to address data privacy and confidentiality concerns, an external party hired by one of the authorities covered in this paper has to go through lengthy clearance and control processes. This added to the rigid procurement process and caused a significant delay in the project. Technologies that require permission from data owners when external developers need to access confidential data may contribute to dealing with this issue. The creation of a testing environment using only testing data is also another common process used to address this issue.

Efficiency vs effectiveness

40. **Assessing the efficiency of the tools is more straightforward and can be more easily quantified than effectiveness.** Efficiency refers to the amount of resources needed to fulfil a certain task, while effectiveness refers to the degree to which the desired outcome is achieved. Efficiency, as represented by the gains in terms of time savings provided by the new tools, is easily measured. This has been mentioned, for example, in the cases of ROSFIN and UIF. ROSFIN's new database management system, which allows strategic analyses of detailed data and the detection of large-scale schemes, results in some processes running 43 times faster. Meanwhile, the tool that UIF uses to analyse wire transfers resulted in an 80% reduction in the time required to perform analyses.

41. **Assessing the effectiveness of some of the tools used by AML/CFT supervisors may be feasible.** This is especially the case for tools used for risk scoring purposes. Assessment of these tools can be done by comparing the outcome of the tool with that of on-site inspection or of other technology. This occurs, for example, at FINTRAC, where the output of its risk model is compared against the outcome of inspections to assess the model's effectiveness. FINTRAC is also developing a challenger model using machine learning and heuristic methods and comparing its performance against that of the model in production.

42. **Assessing the effectiveness of the tools used by FIUs, however, is more challenging.** The process of proving that a money laundering activity has occurred takes time. Once an FIU detects that a suspicious activity warrants more investigation, it passes this information on to the law enforcement agency. The law enforcement agency, in turn, refers the case to the courts if it deems that it warrants prosecution. Court proceedings can be protracted, so that it could take years to determine whether the tool has correctly identified ML.

43. **In general, the effectiveness of these tools is also generally measured in terms of the reduction in the number of false positives, which in theory could lead to more false negatives.** False negatives, however, are more damaging in the fight against ML/TF. This could be mitigated by focusing on the bigger picture through network analysis, rather than analysis of individual transactions/entities.

44. **Moreover, tools based on supervised machine learning could lose their effectiveness over time if not regularly updated with new data set.** These tools might be effective initially, but criminal organisations could change their behaviour in order to avoid detection. This would render the tools ineffective unless they are constantly updated using new training data. Unsupervised machine learning solutions could also be considered to address this problem.

Section 4 – Conclusions

45. **It is important for AML/CFT authorities to understand the potential of innovative technology in supporting their data analytics work.** More data are now readily available and easily accessible to AML/CFT authorities. But more data alone will not really be useful to authorities if they do not unlock the relevant information that the data contain. Advanced data analytics tools that can sift through large volumes of data from different sources are helpful in this regard. They can assist AML/CFT authorities in pursuing their objectives, particularly in identifying, assessing and understanding the ML/TF risks in their jurisdictions.

46. **The exploration or development of advanced data analytics tools by AML/CFT authorities could be organised in different ways.** A number of authorities have established dedicated units to develop their own tools. Other authorities find it more efficient to either buy ready solutions or have external consultants build bespoke solutions for them. A number of authorities are also partnering with academic institutions either to develop data analytics tools or simply just to encourage more research in this field. There is no single approach that works best for all authorities. While most of the authorities covered in this paper use a combination of approaches, there are a number of factors that need to be considered in choosing the right approach. These include the profile of the authority (eg the size of the authority and the resources available to it), the characteristics of the financial system that the authority oversees (eg the number of supervised entities and the ML/TF risk in the system) and the legal framework in which the authority operates (eg data privacy laws, which could influence which data could or could not be used as well as the extent of reliance on external parties for developing data analytics tools).

47. **Similar technologies or methodologies underlie the data analytics tools AML/CFT authorities are developing or using, so there is scope for learning from counterparts in other authorities.** In general, AML/CFT authorities analyse either individual or networks of transactions, with a view to identifying, respectively, individual transactions that should be subject to further scrutiny, and interconnections among them (as well as entities involved in them) that may represent a coordinated ML/TF risk. Although the data analytics tools used by AML/CFT authorities are tweaked to reflect their mandates, the underlying methodologies of these tools are quite similar. There are therefore opportunities for peer learning through regular exchange of information and sharing of experiences on the development and use of these tools.

48. **Efficiency gains seem to be the number one benefit of advanced data analytics tools, which could help capacity-constrained AML/CFT authorities.** AML/CFT authorities have highlighted the gains in terms of time savings they achieved in using these advanced data analytics tools. This could translate into a reallocation of resources or capacity from more manual work to more judgment-based work. The paradox, however, is that capacity-constrained authorities typically do not have the specialised resources needed to develop, understand and use these tools. That is, they need to build capacity first before they can reap the tools' resource-saving benefits. In this regard, exploring available solutions in the market could be a starting point while these authorities develop specialised capacity.

49. **The benefits that these advanced data analytics tools bring are particularly important for jurisdictions that have been struggling to improve their AML/CFT capabilities and have, therefore, been heavily impacted by the unintended consequences of international standards in this area.** As mentioned, the jurisdictions most frequently exited by global correspondent banks seem to be those with weak AML/CFT supervisory and regulatory frameworks. That being so, domestic capacity building forms part of the FSB-coordinated action plan to assess and address the decline in correspondent banking. In this regard, international organisations have been providing extensive technical assistance to these jurisdictions in order to enhance their supervisory and regulatory frameworks. In most cases, the technical assistance may relate to fundamental issues such as putting in place basic supervision and sanctions regimes. However, where appropriate, laying the ground work for authorities to be able to explore

advanced data analytics tools (eg development of necessary skills in this area) could be part of this assistance.

50. **AML/CFT authorities that are just starting to develop their data infrastructure have a "late mover" advantage and may find it easier to integrate advanced data analytics tools.** These authorities have the advantage of developing their data infrastructure from scratch without the burden of legacy systems. They can design it in a way that makes the data collection, validation and management processes seamless, while more easily enabling the integration of newly developed analytical tools. For authorities with mature data infrastructure in place, new processes and tools need to be compatible with the existing infrastructure. This in turn limits the possible new processes and tools that these authorities can explore.

51. **ML/TF risks have international reach, so development of data analytics tools that are international in scope should be considered.** The tools discussed in this paper are all national in scope. ML/TF, however, is an international issue, and criminal organisations tend to exploit loopholes anywhere in the world. Therefore, a strong argument could be made for international cooperation and collaboration in terms of developing data analytics tools with an international coverage. We note that there are already efforts under way that may have this objective in mind (eg the Egmont Group initiative mentioned above). We encourage similar international initiatives towards that end.

52. **Despite the potential benefits of innovative technologies, human judgment continues to be the foundation of the success of AML/CFT authorities.** Having sufficient staff with the right skill sets, expertise and qualifications is critical for such authorities to achieve their objectives. Given the challenges associated with assessing the effectiveness of AML/CFT suptech tools, expert human judgment continues to be necessary to interpret and validate the outcomes of such tools. Therefore, suptech tools should be seen as a complement to the work of experienced staff at AML/CFT authorities and not a substitute.

References

AUSTRAC (2018): "2018–2022 AUSTRAC corporate plan".

Basel Institute of Governance (2018): "Basel AML Index 2018", October.

Broeders, D and J Prenio (2018): "Innovative technology in financial supervision (suptech) – the experience of early users", *FSI Insights on policy implementation*, no 9, July.

di Castri, S, M Grasser and A Kulenkampff (2018): "An AML suptech solution for the Mexican National Banking and Securities Commission (CNBV)", CNBV and Regtech for Regulators Accelerator, August.

Egmont Group (2017): "Operational guidance for FIU activities and the exchange of information", February.

__________ (2018): "IEWG plan on a page".

Financial Action Task Force (2012): "The FATF recommendations: international standards on combating money laundering and the financing of terrorism and proliferation", February.

Financial Intelligence Centre (2018): Annual Report 2017/18, South Africa.

Financial Stability Board (2015): "Report to the G20 on actions taken to assess and address the decline in correspondent banking", November.

____________ (2017): "Artificial intelligence and machine learning in financial services: market developments and financial stability implications", 1 November.

Kraft, O (2018): "Sharpening the money laundering risk picture: how data analytics can support financial intelligence, supervision and enforcement", *Royal United Services Institute Occasional Paper*, November.

LexisNexis Risk Solutions (2018): "2018 True Cost of Compliance Study", October.

MONEYVAL (2018): 2017 Annual Report.

Saaradeey, S, D Ghosh, R Ray, S Ganesan and R Rajagopalan (2019): "Disrupting status quo in AML compliance", *ORACLE White Paper*, 25 March.

ZDNet (2018a): "AUSTRAC looks to extend Human Services data-matching project to law enforcement", 11 May.

_______ (2018b): "Australia's Fintel Alliance combining data to thwart criminal activity", 23 October.

_______ (2019): "AUSTRAC trialling blockchain to automate funds transfer instructions", 24 February.

Glossary[15]

artificial intelligence: the theory and development of computer systems able to perform tasks that traditionally have required human intelligence.

big data: a generic term that designates the massive volume of data that is generated by the increasing use of digital tools and information systems.

cloud computing: an innovation in computing that allows for the use of an online network ("cloud") of hosting processors so as to increase the scale and flexibility of computing capacity. Cloud computing has made possible the analysis of very large datasets (big data), and a number of specific fintech applications.

machine learning: a method of designing a sequence of actions to solve a problem that optimise automatically through experience and with limited or no human intervention.

natural language processing: an interdisciplinary field of computer science, artificial intelligence and computation linguistics that focuses on programming computers and algorithms to parse, process and understand human language.

network analysis: the process of investigating structures through the use of networks and graph theory.

self-organising maps: a type of artificial neural network that is trained using unsupervised learning to produce a low-dimensional, discretised representation of the input space of the training samples, called a map, and is therefore a method of performing dimensionality reduction.

supervised learning: the machine learning task of learning a function that maps an input to an output based on example input-output pairs.

smart contracts: programmable distributed applications that can trigger financial flows or changes of ownership if specific events occur.

text mining: the process of exploring and analysing large amounts of unstructured text data aided by software that can identify concepts, patterns, topics, keywords and other attributes in the data.

unsupervised learning: a type of machine learning method that helps find previously unknown patterns in a data set without pre-existing labels.

[15] FSB (2017).