Financial Stability Institute

FSI Insights
on policy implementation
No 9

Innovative technology in financial supervision (suptech) – the experience of early users

By Dirk Broeders and Jermy Prenio

July 2018

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 978-92-9259-183-0 (print)

ISSN 2522-249X (online)
ISBN 978-92-9259-184-7 (online)

Contents

Innovative technology in financial supervision (suptech) – the experience of early users[1]

Executive summary

Supervisory technology (suptech) is the use of innovative technology by supervisory agencies to support supervision. It helps supervisory agencies to digitise reporting and regulatory processes, resulting in more efficient and proactive monitoring of risk and compliance at financial institutions. A number of supervisory agencies are already using innovative ways to effectively implement a risk-based approach to supervision. Now, technological progress as well as data availability offers the potential to radically improve existing supervisory tools or develop better ones through suptech applications.

Suptech is currently found in two areas of applications: data collection and data analytics. Within data collection, applications are used for supervisory reporting, data management and virtual assistance. Examples include the ability to pull data directly from banks' IT systems, automated data validation and consolidation, and chatbots to answer consumer complaints while collecting information that could signal potential areas of concern. Within data analytics, applications are used for market surveillance, misconduct analysis as well as microprudential and macroprudential supervision. Examples include detecting insider trading activities, money laundering identification, monitoring supervised entities' liquidity risks and forecasting housing market conditions. These applications are in different stages of development and implementation, ranging from academic research questions through proofs-of-concept and use-cases to fully operational.

Supervisory agencies initiate and organise their suptech activities in several ways. Applications used for data collection tend to be management-initiated projects, while those used for data analytics usually start out as research questions but in a few cases may also be suggested by supervision units. A number of supervisory agencies, particularly those active in exploring data analytics applications, have recently created dedicated units. A few others leverage their existing research units. Supervisory agencies also use both internal and external resources in developing suptech applications. In addition, some are partnering with academic institutions, particularly in the area of data analytics, to keep track of the latest developments and learn how to build state-of-the art algorithms.

Expected benefits motivate supervisory agencies to use or explore suptech applications. These benefits include enhanced effectiveness, reduced costs and increased capability. Suptech applications, particularly in the area of data analytics, are seen as capable of turning risk and compliance monitoring from a backward-looking into a predictive and proactive process.

Agencies face a number of challenges in developing or using suptech applications. Some of these issues relate to computational capacity constraints, increased operational risks, including cyber-risk, data quality, finding the right talent, management support and buy-in from supervision units, and rigid rules in project management. Lack of transparency in some of the data analytics applications is also a critical issue. Hence, human intervention in the form of supervisory expertise is still viewed as indispensable in the supervisory process, particularly in further investigating the results of analyses and deciding on a course of action.

[1] Dirk Broeders, Netherlands Bank and Jermy Prenio, Bank for International Settlements.

The authors are grateful to the representatives from the organisations interviewed in Annex 1; to the participants of the FSI meeting on the use of innovative technology for financial supervision held in Basel on 29–31 May 2018 for the insightful discussions; and to Thomas Beretti, Jon Frost, Dirk Grolleman, David Jutrsa, Fabiana Melo, Greg Sutton, Joy Wann and Christopher Wilson for helpful comments. We are also grateful to Esther Künzi for valuable support with this paper.

The benefits of suptech applications can extend to supervised institutions. They can lead to reduced compliance costs and contribute to enhanced risk management effectiveness. This is particularly the case for automated reporting. At the same time, however, a few supervisory agencies recognise the risk that their use of suptech might lead to market participants adjusting their behaviour in order to "game" the technology.

Supervisory agencies would be best placed to explore the potential benefits of suptech applications if they have a well defined suptech strategy. A suptech strategy should comprise the following three key elements, at a minimum: first, ambitious, but achievable, targets (eg which technology will be used, in which area of supervision, and how will it be funded); second, an assessment of today's data availability, data quality and availability of analytical resources; and third, a step-by-step action plan on how the supervisory agency will get from the current situation to full implementation of its suptech strategy.

The experience of early suptech users, as discussed here, yields some useful insights on how to develop such a strategy. Some specific considerations for supervisory agencies are:

(a) **The overall approach to supervision should adapt to the digitisation of the activities of supervised entities.** As finance becomes increasingly digitised, financial supervision needs to keep up.

(b) **Management support is critical in exploring the opportunities and benefits of suptech.** For this to happen, management needs to appreciate the potentials of suptech, while keeping in mind its limitations and risks.

(c) **Supervisory agencies engaged in suptech need specialised human resources.** Supervisory agencies should carefully consider their strategy in attracting and retaining suptech staff, as well as in ensuring that institutional knowledge is maintained should there be a high rate of staff turnover.

(d) **The buy-in of supervision or enforcement units helps to fully embed suptech in supervision work.** Input from supervision or enforcement units should be considered in developing suptech applications.

(e) **Supervisory agencies can benefit from partnerships with the academic community.** To keep up with fast-moving technical developments, supervisory agencies need to explore and stay attuned to new ideas emerging in academia.

(f) **The use of suptech reinforces the case for further improving risk management at supervisory agencies.** The increasing use of suptech exposes supervisory agencies to more risks, such as legal risk, operational risk, including cyber-risk, and reputational risk, which must be mitigated if the benefits of suptech are to be maximised.

(g) **As supervisory agencies can learn from each other, it is important to seek opportunities for collaboration.** The key to growing or enhancing suptech capabilities is for supervisory agencies to continuously exchange knowledge and experience at a global level.

Section 1 – Introduction

1. **Supervisory technology (suptech) is the use of innovative technology by supervisory agencies to support supervision[2].** It helps supervisory agencies digitise reporting and regulatory processes. Suptech could be a game-changer in efficient reporting and proactively monitoring the risk and compliance of financial institutions. It could turn risk and compliance monitoring from a backward-looking into a predictive process. Furthermore, suptech could have significant organisational impact and may raise uncharted legal and ethical issues.

2. **Suptech is the result of the emergence of advanced technologies in the economy.** These technologies have led to new technology-oriented business models in the financial industry, including new products and services (fintech). The efficiencies that these technologies offer can also be harnessed in support of compliance with financial regulation (regtech) and conduct of financial supervision (suptech). Regtech is already a familiar word in the financial sector. It refers to applications of innovative technologies that support compliance with regulatory and reporting requirements by regulated financial institutions.[3] Suptech, on the other hand, refers to technologies used by supervisory agencies themselves.

3. **A number of financial sector supervisors are already using innovative ways to effectively implement a risk-based approach to supervision.** Risk indicator dashboards, centralised data warehouses for supervisory reports and early warning systems are just a few examples of tools that are now entrenched in a number of supervisory agencies around the world. Suptech offers the potential to either radically improve these existing tools or develop considerably better ones.

4. **There are many reasons why the emergence of suptech is accelerating.** Post-crisis regulatory reforms have led to an upsurge in reporting requirements. This increases the need for efficient and effective monitoring to benefit from the resulting boost in data availability. Next to more data, better data are also a catalyst for suptech. Better data are created through the harmonisation of definitions in data dictionaries.[4] Also, growth in storage capacity and computing power open the way for suptech. Furthermore, advances in data science have led to the development of new technologies that can be applied in supervision.

5. **Suptech supports the core objectives of regulation and supervision.** Suptech helps supervisory agencies achieve their objective of promoting trust by economic agents in financial institutions and markets. Suptech is likely to lead to supervision that can adapt more quickly in response to a constantly evolving environment. Innovative technologies can be applied at the design stage of new regulations to assess the potential impact of policy proposals. Currently, suptech applications do not replace human judgment but serve as input for supervisors as they evaluate whether further investigation or enforcement is necessary.

6. **As an emerging field in supervision, suptech is likely to come with relevant challenges.** Data standardisation, data quality and data completeness are all necessary conditions for effective suptech applications. Understanding the capabilities and limitations of innovative technologies is key to assessing their value added in supervision work. These technologies may, for example, detect spurious rather than meaningful correlations. Use of suptech may also expose supervisors to more risks, including legal, operational and reputational risks.

7. **Overcoming these challenges requires a number of considerations.** Suptech requires a tech-oriented approach to supervision, as opposed to a purely finance- or legal-oriented one. Having talented

[2] BCBS (2018) defines suptech as the use of technologically enabled innovation by supervisory authorities; our definition specifies that the purpose of such use is to support supervision.

[3] See FSB (2017b).

[4] The BIRD initiative of the ESCB's Statistics Committee is a good example, see http://banks-integrated-reporting-dictionary.eu/.

data scientists and developers is critical in introducing this tech-oriented approach. Buy-in of senior management and supervision staff, as well as implementing robust risk management frameworks are also required to fully embed and maximise the benefits of suptech. Lastly, collaboration or partnership with other supervisory agencies and academic institutions will also help in meeting the challenges of suptech.

8. **This paper aims to provide an overview of the experiences of some early suptech users, particularly those with key applications under development or in use, highlighting their benefits and challenges, and the implications for supervisors.** The paper is based on interviews with a number of organisations (see Annex 1) as well as a review of publicly available literature.[5] Early work on suptech can be found in FSB (2017a,b), Toronto Centre (2017, 2018), Accenture (2018) and BCBS (2018). The key contribution of this paper is a more detailed take on the hands-on experience of supervisory agencies. This provides useful insights on how to develop a suptech strategy.

9. **The remainder of the paper is as follows.** Section 2 sketches the areas of financial supervision in which the technologies are being used. Section 3 outlines the experience of early suptech users. Section 4 concludes with some practical considerations.

Figure 1: "Wordcloud" of popular words related to suptech

Section 2 – Uses of suptech

In which areas of financial supervision is suptech commonly used?

10. **There are already various areas where innovative technology supports supervision work.** This section highlights some of the areas in which suptech applications can be found. Table 1 shows which supervisory agency uses which technology. The table is based on publicly disclosed activity, and does not attempt a complete overview of all the innovative technologies currently in use. Figure 2 summarises the

[5] The "wordcloud" in Figure 1 represents the relative popularity of words found in the papers listed under "References".

areas of financial supervision in which suptech is used. Within data collection, suptech applications can be found in reporting, data management and virtual assistance. Within data analytics, four key areas are distinguished: market surveillance, misconduct analysis, microprudential supervision and macroprudential supervision.[6]

Early adopters and their innovative suptech technologies

This table shows some examples of the technologies currently used by supervisory agencies or under development. The table is indicative only, based on publicly disclosed activity and makes no attempt to provide a complete overview of all suptech applications.

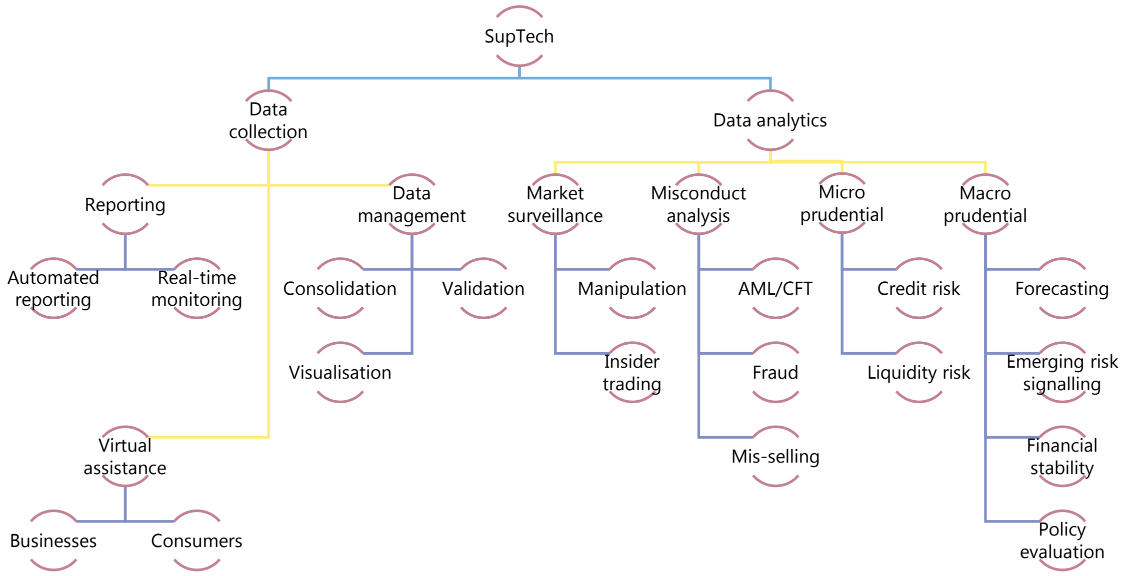
Table 1

Technology	Supervisory agency									
Data collection										
API	ASIC			BSP						
Data input approach	ASIC								OeNB	SEC
Data pull approach	ASIC		BNR	BSP			FCA			
Machine-readable regulation							FCA	MAS		
Cloud computing	ASIC				CNBV	DNB	FCA			SEC
Chatbots				BSP			FCA			
Data analytics										
Big data	ASIC	BoI			CNBV	DNB	FCA	MAS		SEC
Artificial intelligence					CNBV	DNB	FCA	MAS		SEC
NLP	ASIC	BoI			CNBV		FCA	MAS		SEC
Machine learning	ASIC	BoI			CNBV	DNB	FCA	MAS	OeNB	SEC
Supervised learning	ASIC	BoI				DNB	FCA			SEC
Unsupervised learning	ASIC					DNB	FCA		OeNB	SEC
Topic modelling							FCA			SEC
Random forest	ASIC	BoI					FCA			SEC
Image recognition							FCA			
Neural networks						DNB			OeNB	SEC

Note: based on interviews and public sources. Definitions of each technology can be found in the Annex.

[6] For the purpose of this paper, a distinction is made between market surveillance and misconduct analysis as the suptech applications may differ in each area. Market surveillance applications use data from trading or market transactions, while misconduct analysis applications use other sources of data, such as regulatory filings, advertisements etc.

Figure 2: Areas of financial supervision in which suptech applications are used



Data collection

11. **Suptech applications in data collection focus on reporting, data management and virtual assistance.** Within reporting, suptech applications include various forms of automated reporting and real-time monitoring. The key applications in data management are data validation, consolidation, visualisation and cloud computing. Suptech applications, such as chatbots and machine-readable regulations, can be found in the area of virtual assistance.

Automated reporting

12. **A key suptech solution for automated reporting is a data push approach.** In a joint effort with supervised entities, the Central Bank of the Republic of Austria (OeNB) has developed a reporting platform that bridges the gap between the IT systems of supervised entities and the supervisory agency. The system allows the banking sector to send critical information to the OeNB without increasing the administrative burden for data providers. The platform is housed in a buffer company, Austrian Reporting Services GmbH (AuRep) that was set up in 2014 and is co-owned by seven of the largest Austrian banking groups.[7] Banks feed data into a so-called basic datacube at AuRep. This represents a single, complete description of the reporting data that is redundancy-free and contains harmonised definitions. Consecutively, a set of standardised transformation rules convert the data in the basic datacube in such a way that it can be "pushed" to the OeNB. AuRep covers about 90% of the Austrian banking market in terms of balance sheet totals as well as the number of credit institutions. All reports specified in the OeNB's reporting data model run via AuRep. At the moment, almost all statistical and financial stability reports as well as some regulatory reports run on the basis of this data model. This approach not only offers economies of scale but also allows risk-sharing in the financial industry.

13. **Closely related is a data pull approach for automated reporting.** The National Bank of Rwanda (BNR) is one of the first supervisory agencies to use a data pull approach. This technology extracts data directly from the IT systems of supervised institutions. Data-pulling is done automatically every 24 hours or, in some cases, every 15 minutes. For other data, the frequency is monthly. Combined with data from

[7] See BearingPoint (2015).

BNR's internal systems, the suptech application streamlines the reporting process and produces meaningful information for supervisors and policymakers. The BNR developed this electronic data warehouse (EDW) system together with Sunoida Solutions (see Box 1 for more detail).[8] The system has been operational since 2017. The EDW also includes information that can be used for financial inclusion measurement. Bangko Sentral ng Pilipinas (BSP), with the support of RegTech for Regulators Accelerator (R^2A),[9] is also developing a prototype for an API-based data input approach to extract regulatory reports directly from banks. The aim is that Philippine banks will no longer have to submit templated reports to BSP, a process that is prone to human error and thus subject to multiple validation iterations. Instead, the data input solution will enable better and more timely analytics and a more flexible reporting format that, in the long run, will be more efficient to maintain and to manage.

Box 1

Data pull approach at the National Bank of Rwanda

An electronic data warehouse that pulls data directly from the IT systems of financial institutions

The National Bank of Rwanda (BNR) uses an electronic data warehouse (EDW) to "pull" data directly from the IT systems of more than 600 supervised financial institutions, including commercial banks, insurance companies, microfinance institutions, pension funds, forex bureaus, telecom operators and money transfer operators.① Data are automatically pulled from these institutions every 24 hours or even every 15 minutes in the case of mobile money and money transfer operators. For this purpose, a data dictionary was developed and each financial institution was required to write data scripts that would map the data dictionary to the information in its own systems. This, however, was a one-off investment. The mapped information is then put in a "staging area" where BNR can pull the information it needs. The encrypted data run over a VPN channel and through data integrity check mechanisms. In this way, the data pull approach delivers timely, consistent and reliable data to the BNR. It has also led to improvements in financial institutions' data, which they now also use for internal risk management purposes. The EDW reduces errors and inconsistencies. To this end, the BNR has built quality and integrity rules into the system. If data do not meet certain standards, they are rejected and an automatic email alert is sent to bank examiners within the BNR and the supervised institution. Furthermore, historical data cleaning from the data supply side is currently in progress. Next to improving data quality, the EDW also offers flexibility and the ability to quickly analyse large amounts of data. The BNR can adapt its supervisory processes and methodologies to fully leverage the collected data and allocate supervisory resources more efficiently. The BNR has not ended manual reporting completely due to data gaps on the part of external stakeholders. The BNR is also streamlining internal business processes to ensure that information is completely captured.②

① BNR (2017). ② Kamali and Randall (2017).

14. **Suptech applications can enable real-time monitoring.** The Market Analysis and Intelligence (MAI) system of the Australian Securities and Investments Commission (ASIC), for example, allows for real-time monitoring of the Australian primary and secondary capital markets (ASX and Chi-X). The MAI system ingests real-time data feeds from all equity and equity derivatives products and transactions. The MAI system provides real-time alerts, identifying anomalies within markets that may be investigated or detected upon execution. These real-time alerts are integrated with daily operations and staff workflows. Alerts trigger a workflow process that leads to further investigation and analysis to identify likely root causes. The results feed into a triage process that determines prioritisation and triggers in-depth investigations where appropriate (see also Box 2).

8 See Kamal and Randall (2017).

9 https://www.r2accelerator.org/.

Box 2

Market surveillance at the Australian Securities and Investments Commission

Real-time alerts identifying anomalies within markets

The Market Analysis and Intelligence (MAI) system is the suptech platform for market surveillance of the Australian Securities and Investments Commission (ASIC).① It has real-time data feeds from all Australian primary and secondary capital markets (ASX and Chi-X) for equity and equity derivatives products and transactions. These real-time feeds are supplemented by daily overnight data feeds for listed futures and derivatives products. The technology underpinning the MAI system is a KDB/Q column-oriented in-memory database. To complement the view of Australia's capital market activity, information on OTC derivatives transactions is received daily from designated trade repositories and fed into a post-trade analytics environment. The technology underpinning the post-trade analytics environment comprises KDB/Q, R, Python and MS-Excel. This is being extended into a big data platform (using Elasticsearch/SPARK/Kibana) with a set of complementary technologies that enables the visualisation (reports, dashboards, networks etc) of time-series data and provides inherent support for machine learning, alert generation and extensive searching across big data sets. The MAI system has two sets of outputs. First, real-time alerts identifying anomalies within markets that may be investigated or detected upon execution. These real-time alerts are integrated into daily operations and staff workflows and may lead to in-depth investigation and analysis. Second, big-data historical analytical capabilities that can provide market-complete reporting and assessment of large and complex thematic risks. The historical analytics can be executed within an integrated environment linked to other cases or to matters prioritised by surveillance staff. The post-trade analytics environment provides the ASIC with an evolving picture of the Australian financial market, and continues to be enriched with more data for meaningful insights. The ASIC can access a number of algorithms, analytics and reports based on this data set that highlight entities or trades of interest. In the near future, the ASIC will switch on machine-learning algorithms across the data set to identify anomalies in trading behaviour.

① Based on an interview with an ASIC representative.

Data management

15. **Data validation is another key area for suptech applications.** Automated data validation checks may include: checks for receipt of data, checks for data completeness, checks for data correctness and plausibility and consistency checks.[10] The Monetary Authority of Singapore (MAS), for example, uses technologies for data validation, including data-cleaning and data quality checks. These increase efficiency and save time, allowing supervisors to focus more on investigations. The importance of data quality cannot be overemphasised. A good model cannot overcome bad data. And good data are better than more data.[11] Machine learning (ML) can help improve data quality by automatically flagging anomalies as potential errors to the statistician or the data-providing source.[12] The OeNB has also developed a prototype for data validation purposes based on ML and unsupervised learning.

16. **Data consolidation is an important part of many suptech applications.** Within reporting, suptech allows for the smooth creation of macro data by aggregating micro data, such as risk exposures and interconnections between financial institutions. Suptech applications are capable of combining multiple data sources to support analytical work. Often this involves connecting structured data and unstructured data. A good example is found at the Bank of Italy (BoI), which combines suspicious transactions reports (structured data) with press reviews (unstructured data) for anti-money laundering (AML) detection. The BNR combines regulatory data with data from internal systems to produce meaningful information for supervisors and policymakers.

[10] See Dias and Staschen (2017).

[11] See Bauguess (2017).

[12] See FSB (2017b).

17. **Many supervisory agencies apply technology for data visualisation.** Data after all do not equal information. Powerful visualisation tools are required, given the quantity, density and complexity of data, to present information to supervisors in a readily comprehensible way. IBM i2 and associated iBase information schemes are applications for data and network visualisation analytics used by ASIC to represent temporal, associative and causal relationships from structured data sources. The Netherlands Bank (DNB) puts significant effort into transforming data output into logical indicators, eg traffic lights and dashboards developed in-house. MAS uses interactive dashboards and network graphs for imaging data.

18. **One relevant suptech application in the area of data management is cloud computing.** Cloud computing allows for greater and more flexible storage, mobility capacity and computing power. As an example, the UK Financial Conduct Authority (FCA) has cloud solutions for collecting, storing and processing market data. This involves "billions of data elements each day". Auto-scaling cloud facilities can deal flexibly with peaks in market data. Mexico's National Banking and Securities Commission (CNBV), as well as the DNB, MAS and the US Securities and Exchange Commission (SEC), also report the use of cloud computing to process large data volumes. Cloud computing lowers costs and increases storage capacity at supervisory agencies. But storing regulatory data in the cloud may require stronger oversight.[13]

Virtual assistance

19. **Several supervisory agencies use chatbots to answer consumer complaints automatically.** The BSP, with support from R^2A, is developing a prototype chatbot to answer consumer complaints. The system will be able to triage the concerns received, answer simple questions and appropriately direct those that need to be first addressed to supervised institutions. The data received in connection with the consumer complaints will also allow BSP to analyse potential areas of concern. Complaints can signal unlawful behaviour by a supervised entity. Data on consumer and investor complaints are also used by the BoI, FCA and SEC to signal concerns at supervised entities, although in some cases this is still at an experimental stage.

20. **Supervisory agencies can also use chatbots to provide virtual assistance to supervised entities.** The FCA, for example, is engaging in a proof-of-concept for the use of chatbots to interact with supervised entities to efficiently answer straightforward day-to-day questions. Chatbots can help supervised entities better understand the requirements of a particular rulebook or item of legislation.[14]

21. **Machine-readable regulations could also help to facilitate compliance.** The FCA is exploring the potential to implement machine-readable regulations. Converting regulatory text to a machine-readable format using natural language processing (NLP) leads to greater consistency and improved compliance. It can help to narrow the gap between regulatory intent and interpretation.[15] Machine-readable regulations could also help supervisory agencies to efficiently assess the impact of regulatory changes, consult on regulatory reforms, and reduce regulatory complexity.[16] This is an area where suptech and regtech interact.

Data analytics

22. **In the area of data analytics, the four key areas are market surveillance, misconduct analysis, microprudential supervision and macroprudential supervision.** Market surveillance focuses on suspicious trading, such as market manipulation and insider trading. Suptech applications in

[13] See Dias and Staschen (2017).

[14] See Accenture (2018).

[15] See Accenture (2018) and FCA (2018).

[16] See FCA (2018) and Toronto Centre (2017).

misconduct analysis focus on AML/combating the financing of terrorism (CFT) detection, fraud detection and mis-selling. Within microprudential supervision, applications can be found for credit risk evaluation and liquidity risk detection. Important areas of applications in macroprudential supervision are identifying macro-financial risks, including identifying signals of risks emerging in the financial system, and policy evaluations.

Market surveillance

23. **Suptech applications allow huge amounts of data to be analysed for market surveillance and the detection of suspicious trading.** Financial markets produce astronomical quantities of data every trading day. Securities supervisors are therefore typically experienced in the processing of huge transaction data volumes. ASIC, FCA and SEC all apply innovative technologies in transforming enormous data sets into usable patterns for market surveillance and suspicious trading detection. Box 2 explains how suptech applications at ASIC allow market anomalies to be identified in real time. The SEC uses analytics that harness the power of big data to drive its surveillance programmes and foster innovation in market risk assessment initiatives.[17]

24. **Examples of suspicious trading that can be detected by suptech include insider trading and market manipulation.** To detect insider trading, the FCA receives details of over 20 million equity market transactions per day. Supervised learning ML tools analyse these data to produce signals of market manipulation. The FCA Market Oversight team can monitor the normal behaviour of traders and detect any deviations that might flag insider trading.[18] The historical reviews done by ASIC's suptech application MAI (see Box 2) can provide quantifiable metrics that flag the scale of any insider trading activity. This is done through profit and loss analytics or another assessment measure for market-wide harm such as market manipulation effects.

Misconduct analysis

25. **Many suptech applications concentrate on the detection of possible AML and CFT infringements.** Smart technologies can detect unusual transactions, relations and networks that are not apparent to human supervisors. Innovative technologies in the area of AML/CFT are in the planning stages or applied at many supervisory agencies, such as the BoI, BNR, BSP, MAS and CNBV. MAS uses NLP and ML to analyse suspicious transaction reports to find potential money laundering networks (see Box 3). The BoI's Financial Intelligence Unit applies big data in AML detection, using structured data with a five-year history of all transactions in excess of EUR 15,000 in combination with unstructured data such as press reviews. Alongside a remarkable reduction in the analysis time required, an additional advantage mentioned by the BoI is the scope for performing the analysis in real time. In an experimental project, the BoI is also testing the use of machine learning and deep learning techniques to classify incoming suspicious transaction reports.

[17] See Bauguess (2017).

[18] See Hunt (2017).

Box 3

Natural language processing and machine learning at the Monetary Authority of Singapore

Detecting suspicious networks for potential money laundering issues

One suptech innovation in the area of data analytics for detecting anti-money laundering (AML) violations can be found at the Monetary Authority of Singapore (MAS).① MAS is creating a data analytics system to search through the 3,000 monthly Suspicious Transaction Reports (STR) on money laundering and terrorist financing risks that financial institutions file with MAS. The STR Network Analysis solution will use natural language processing and machine-learning technologies to analyse the reports. It will produce a suspicious money laundering sub-STR-network, which is a part of the overall STR network generated by all STRs. Supervisors will use the detected suspicious money laundering sub-STR-network for further investigations. The sub-STR-network includes the information generated from the original STRs such as the entities in the STRs and the relationship between these entities. Based on this information, supervisors will be able to look for more data, such as transactions from the entities under suspicion. The technology will dramatically increase efficiency and effectiveness. Manually creating a network for identifying potential AML violations takes about two years. Using artificial intelligence (AI)/machine learning (ML) to do the same thing will require only a few minutes. Furthermore, AI/ML can pick up patterns in data that humans cannot. Currently under development at MAS, the suptech application is expected to be operational by the end of 2018.

① Menon (2017).

26. **AML/CFT detection often builds upon the analysis of networks of market participants or events.** The CNBV developed a prototype NLP application to discover what a suspicious AML/CFT network is "talking about". It has a proof-of-concept on how to use big data to reveal not only relationships between people but also between events, with the aim of detecting potential money laundering issues. In monitoring market activity, the FCA is experimenting with the deployment of graph learning to identify networks of market participants and potential collusive behaviour based on order and execution data.[19] Graph learning is a technique that automatically generates graphs from data based on the inferred or known relationships between data points.

27. **ML algorithms can help to identify possible fraud.** The SEC uses a sequential approach in some of its misconduct detection methods. First, it employs unsupervised learning to detect patterns and anomalies in the data, such as SEC filings. Topic modelling can be of help here by providing insights and recognising the key theme embedded in the data. This technology is used by the SEC to produce groups of "like" documents that identify both common and outlier behaviours among market participants. Second, it injects human direction and judgment into the process to help interpret the ML output. For example, human findings from registrant examinations can be used as a form of supervised learning, by training an algorithm to understand what kind of pattern, trend or language in the underlying examination data may indicate possible fraud or misconduct. The successive algorithms can then be applied to new SEC filings to predict the likelihood of fraud.[20]

28. **Predicting mis-selling.** The FCA is experimenting with the use of supervised learning and "random forest" techniques to predict the probability of an adviser mis-selling financial products. To prevent financial product mis-selling, the FCA is particularly interested in knowing where such activity is likely to occur.[21] Visual analytics can, for example, be used by supervisory agencies to identify potentially misleading advertisements.[22]

[19] See Accenture (2018).

[20] See Bauguess (2017).

[21] See Hunt (2017).

[22] See Accenture (2018).

Microprudential supervision

29. **ML can be used for credit risk evaluation.** The BoI is starting to explore how loan default forecasting can benefit from the use of ML algorithms, merging different data sources for this purpose (eg the Central Credit Register, the balance sheet data of non-financial firms and other firm-level data). This blend feeds into an ML tool that produces forecasts of loan defaults for comparison with the standard models.

30. **Neural networks can be used to detect liquidity risks.** The DNB is researching an auto-encoder to detect anomalies, ie unusual liquidity flows, in payment data derived from a real-time gross settlement system. An auto-encoder is a neural network, ie an unsupervised learning method that learns the main features from the data. Experimental results on real-world payment data show that the auto-encoder can detect liquidity problems at a bank in anticipation of a bank run.[23]

Macroprudential supervision[24]

31. **Innovative techniques can be found in the identification of macro-financial risks.** BoI researchers use a variety of techniques in forecasting housing prices and inflation. Using a popular online portal for real estate services, the researchers use ML techniques to first eliminate multiple ads for the same property. After that, evidence suggests that online interest in a particular area is a leading indicator for prices. In another research paper, the BoI shows that information extracted from tweets gives a meaningful signal for inflation expectations.[25] Researchers at the DNB use daily figures to define network indicators, operational indicators and liquidity flows between TARGET2[26] and other financial market infrastructures (FMIs). They identify cyclical patterns as a basis for forecasting risk indicators. The forecasts are then compared with the actual observed values. A strong deviation may be a sign of increased risk.[27]

32. **Suptech applications may also be useful in identifying signals of emerging risks in the financial system.** Technology can be used to identify such signals using massive amounts of data from FMIs, such as payment systems. The data that these systems produce are ideal for big data applications. For this purpose, researchers at the DNB convert the massive amount of transactions processed in TARGET2 into risk indicators. They do this by applying traditional econometric methods to their data, developing algorithms to pick up relevant transaction types (eg unsecured interbank money market loans), developing indicators based on globally defined principles for financial market infrastructures and machine learning. The different stages of both the Great Financial Crisis and the sovereign debt crisis are clearly visible in the interbank money market.[28]

33. **One important use of NLP is to measure sentiment.** For example, the BoI studies the sentiment expressed in tweets in order to nowcast retail deposits. A negative sentiment corresponds to lower retail deposit growth rates. In the same study, the BoI also measures the interconnectedness between banks based on the occurrence of two banks in the same tweet.[29] The SEC performs sentiment analysis using NLP to assess the tonality of registrant filings. Together with topic modelling, the tonality signals are

[23] See Triepels et al (2017).

[24] For more examples of applications for financial stability purposes, see Financial Stability Board (2017b).

[25] See Loberto et al (2017) and Angelico et al (2018).

[26] Trans-European Automated Real-time Gross Settlement Express Transfer System.

[27] See Timmermans et al (2018).

[28] See Berndsen and Heijmans (2017).

[29] See Accornero and Moscatelli (2018).

converted into risk measures using ML algorithms. Tonality analysis measures the negativity of a text by counting terms with a negative connotation.[30]

34. **Other possible applications are in financial stability and policy evaluations.** The Federal Reserve, the European Central Bank and the Bank of England, for example, use "heat maps" to highlight potential financial stability issues. The heat maps are derived from automated analyses of daily and other data (such as stress tests) being produced by supervised entities.[31] Data analytics support policy development and objectives. The CNBV, with support from R^2A, uses AML compliance data to produce customised reports for policy development purposes.

Section 3 – User experiences

35. **This section covers the experience of the supervisory agencies with suptech applications, as mentioned in the previous section.** Some background is provided on how and why supervisory agencies develop suptech applications. Also outlined are the challenges for supervisory agencies in developing suptech applications and the implications for supervised entities.

How do supervisory agencies develop suptech applications?

Governance

36. **Suptech applications for data collection purposes tend to be management-initiated.** Table 2 shows how suptech applications are identified in various supervisory agencies. As an example, in 2011 the senior managements of the OeNB and commercial banks started a dialogue on how to improve data quality and reporting efficiency. This led to a cooperation agreement in 2013 to work together on an integrated reporting data model and to create AuRep.

How suptech applications are identified

Practices in different supervisory agencies — Table 2

Approach	Supervisory agency									
Management-initiated	ASIC	BoI	BNR	BSP	CNBV		FCA		OeNB	
Identified by dedicated unit	ASIC	BoI				DNB	FCA	MAS		SEC
Identified by supervision units	ASIC						FCA	MAS		SEC

Note: based on interviews and public sources.

37. **Data analytics solutions, by contrast, often originate as research questions.** For the DNB and BoI, ideas emerge and start out as academic or policy questions that are then further explored by their researchers. It is only at a later stage, when found feasible, that they are adopted at the management level and used by relevant supervisory units. The need for a particular data analytics solution therefore is not necessarily identified with input from supervision units. At the FCA, through a combination of suggestions from its supervision and enforcement units, and centrally initiated research projects, the RegTech and Advanced Analytics department is helping to drive experimentation with new suptech solutions and their

[30] See Bauguess (2016).

[31] See Arner et al (2016).

adoption. The team also conducts internal tech sessions and "hackathons" (referred to as TechSprints) that are helpful in showcasing the possibilities of suptech/regtech to the wider organisation.

38. **In other supervisory agencies, there may be input from supervision units as to what suptech applications might be useful.** At the SEC, the need for a specific suptech solution is sometimes identified and suggested by the SEC's enforcement unit. This is more likely if the enforcement unit itself is tech-oriented. The same is true in MAS. Once a specific request is received from the supervision unit, the SupTech Office assesses what data are available and whether a solution can be developed.

39. **Supervisory agencies that are exploring suptech applications for data analytics have recently created dedicated units to develop these solutions.** A good example is MAS with its SupTech Office mandated to "conduct data analyses on supervisory and financial sector data in partnership with MAS departments". It forms part of MAS's Data Analytics Group established in 2017.[32] Within the DNB, a Data Science Hub is experimenting with the use of innovative technologies throughout the organisation (see Box 4). ASIC, on the other hand, established a Chief Data Office responsible for data strategy, and a Data Governance Council that is tasked to ensure appropriate governance of data used in the institution. The FCA, meanwhile, has a RegTech and Advanced Analytics team for the adoption of suptech and related experimentation.[33]

Box 4

Data Science Hub at the Netherlands Bank

Mastering data

The Netherlands Bank (DNB) recently initiated a Data Science Hub initiative① as part of its bank-wide data governance programme. The programme's purpose is to work more efficiently with data, by answering key questions such as the governance of data ("who owns it?"), data storage and pricing, and data protection and sharing. The Data Science Hub is a workstream within this broader initiative. The challenges that the hub is designed to meet are (i) enhancing the IT infrastructure by exploring cloud solutions for data storage and computing; (ii) researching several proofs-of-concept (PoC) in the area of data analytics, including credit risk analysis, central clearing risk indicators, contagion detection and text analytics; and (iii) creating a community of data scientists via knowledge networks, informal lunches and seminars. The teams working on the PoCs operate agile and full-time in sprints of, at most, two or three weeks. Emphasis is put on working towards responsible data use through a "coding hygiene" document and a "Git", a code repository for documenting and coordinating the work of several staff on the same file. In this way, the Data Science Hub lowers the entry barriers for DNB employees who want to work on suptech applications.

① Based on an interview with a DNB representative.

40. **Other supervisory agencies active in suptech applications for data analytics leverage existing units for exploring and developing these solutions.** For example, since the Great Financial Crisis, the SEC's Division of Economic and Risk Analysis has been looking at innovative ways of analysing information from financial institutions to identify potential compliance issues.[34] BoI relies on three existing units to explore the use of innovative technologies – Research and Statistics, Supervision and Financial Intelligence (see Box 5).

[32] "MAS Sets up Data Analytics Group", MAS press release, 13 February 2017.

[33] The Bank of England (BoE) has also established a FinTech Accelerator. The objective is to be able to work with innovative firms for the BoE to understand new technologies and for these firms to understand BoE's emerging policy issues and needs. While broader in scope, the BoE FinTech Accelerator has also produced some proofs-of-concept that explore suptech applications.

[34] See Bauguess (2017).

Box 5

Big Data & Machine-Learning Team at the Bank of Italy

A multidisciplinary team on Big Data

The Bank of Italy (BoI) has recently created a multidisciplinary team on big data. Its staff includes economists, statisticians and computer scientists from various departments.① The team has built a hardware and software infrastructure to deal with various types of big data for both macroeconomic and microeconomic issues. The infrastructure provides statistical tools such as Matlab, Python, R and Stata as well as the open-source in-memory software layer, Spark. This software bundles together the computing power of a series of computers, creating a platform that combines the memory capacity of every computer in the cluster. BoI analysis takes advantage of both structured data, from a classical data warehouse, and unstructured textual data from social media. Some studies focus on Twitter, with the aim of measuring inflation expectations or gauging retail depositors' trust toward their banks. Social media are also used to assess customer sentiment towards specific companies and the effect on stock returns, volatility and trading volumes. Twitter and news channels are used to measure economic policy uncertainty and to investigate payment card scams, with the aim of relating them to trends in consumer payments. The Financial Intelligence Unit combines structured data with unstructured textual data to detect AML activities. Another strand of research uses data from online real estate ads extracted from the web with a view to elucidating the structure of the Italian real estate market.

① Based on an interview with a BoI representative.

Relationship with third parties

41. **Suptech applications for data collection tend to be developed by external service providers.** For example, the BNR contracted the development of its Data Warehouse to Sunoida Solutions, a business intelligence and analytics company in emerging market economies (EMEs). For its part, the BSP is supported by the R^{2}A, which provided funding for the services of a solution provider to build and pilot a prototype solution. The AuRep software was developed externally by BearingPoint.[35]

42. **In the area of data analytics, supervisory agencies tend to use both internal and external resources in developing suptech applications.** One supervisory agency chose to develop solutions entirely in-house, highlighting the need for the developer to understand the needs of internal users and collaborate with them. For the other supervisory agencies, collaboration between in-house experts and external developers is quite common. In other cases, the market is screened for existing ("off-the-shelf") solutions and, if one is not available, a global tender will be issued. Some agencies, such as ASIC and MAS, engage with other supervisory agencies to exchange views on applications.

43. **Some supervisory agencies have partnered with academic institutions to explore potential data analytics applications for supervision purposes.** There are several ways to do this. The academic institution may collaborate with the research departments of supervisory agencies, or send visiting scholars to supervisory agencies to do research on-site, or serve as a sounding board for new ideas. Academics can also be hired as consultants. A number of supervisory agencies also organise research seminars with the participation of academic institutions as a way of showcasing and discovering potential use cases for new technologies. For academics, cooperation is an attractive way of gaining access to regulatory data. For supervisory agencies, this cooperation helps them to keep up to date with new technology and ideas emerging in academia.

[35] See http://sunoida.com/; https://www.r2accelerator.org/; and www.bearingpoint.com/, respectively.

Why are supervisory agencies developing suptech applications?

44. **Enhanced effectiveness, reduced costs and improved capabilities are the most often cited motivations for developing suptech applications.** Supervisory agencies mention a combination of these benefits as the main motivation for using or developing suptech applications. For solutions related to data collection, all these benefits certainly apply. This is particularly the case if these solutions can be applied in EMEs with a simpler and still growing financial system that may not yet be burdened with legacy IT systems. However, in the light of increasing reporting requirements, a complete overhaul is also meaningful in advanced economies, as the Austrian case shows. For solutions related to data analytics, the emergence of innovative technologies allows supervisory agencies to significantly improve their analytical capabilities while enhancing effectiveness. Some supervisory agencies also anecdotally mention cost reductions coming from data analytics solutions. Supervisory agencies are at different stages of developing suptech applications (see Table 3).

How far advanced are supervisory agencies with suptech applications?

This table shows the different stages of suptech application development at various supervisory agencies. The table is indicative only and based on publicly disclosed activity. Table 3

	Supervisory agency								
Supervisory area	ASIC	BoI	BNR	BSP	CNBV	DNB	MAS	OeNB	SEC
Automated reporting			Operational	In development		In development		Operational	
Real-time monitoring	Operational	Experimental stage				Experimental stage			
Validation		Experimental stage		In development	In development		Operational	Experimental stage	Operational
Consolidation	Operational						Operational		Operational
Visualisation	Operational				In development	In development	Operational	Operational	Operational
Virtual assistance		Experimental stage		In development			In development		
Machine-readable regulations									
Manipulation	Operational				In development		In development		Operational
Insider trading	Operational						In development		Operational
AML/CFT		Operational	Experimental stage		In development		In development		
Fraud	In development						In development		In development
Mis-selling									Experimental stage
Credit risk evaluation		Experimental stage							
Liquidity risk evaluation						Operational			
Macro-financial risks	In development	Experimental stage				In development			In development
Emerging risks signalling						In development			
Policy evaluation			Experimental stage		Experimental stage	Experimental stage			
Financial stability		In development				Experimental stage			

Note: Experimental stage; In development; Operational

Enhancing effectiveness

45. **Suptech applications enhance effectiveness by improving on traditional or manual processes, thereby allowing for faster supervisory action.** In the area of data collection, suptech supports a faster and more flexible data capture from supervised firms compared with the traditional template-based approach. This, in turn, lets supervisors further improve off-site monitoring and allows for

better and earlier detection of potential risks. In the area of data analytics, suptech drastically cuts the time needed for analysis. For example, identifying potential money laundering issues in payments transactions data could now be done in minutes instead of months.

Reducing cost

46. **Suptech applications reduce costs by automating processes that typically used to involve several people.** Cost efficiency certainly motivates solutions related to data collection. The embedded data dictionary or data scripts and data integrity mechanisms ensure that data from supervised entities are consistent and reliable.[36] This reduces the need for the traditional data quality checks and the ensuing communication between supervisor and bank to correct reporting mistakes. Data analyses, particularly those that involve a significant amount of data, also benefit from cost reduction since these now require less staff time. At one supervisory agency, an AML solution is estimated to have reduced costs by 80%.

Increasing capability

47. **Suptech applications enhance supervisory capability by making possible the humanly impossible.** Securities markets supervisors, for example, receive thousands of regulatory filings from supervised entities. It is impossible for supervisors to review each one closely. Suptech applications can sift through all these regulatory filings and identify potential supervisory issues. At the SEC, for example, back-testing analyses show that algorithms are five times better than random at identifying "red flag" language in investment adviser regulatory filings.[37] Suptech also allows structured and unstructured data to be integrated, thus making analyses richer. In particular, the use of AI/ML could pick up patterns in data that may not be apparent to humans.

What challenges do supervisory agencies meet in developing suptech applications?

48. **Alongside these benefits, supervisory agencies have encountered various issues and challenges in the course of developing and using suptech applications.** These include technical, data quality, legal, operational, reputational, resource, internal support and practical issues.

Technical issues

49. **Technical issues relate to computational capacity constraints and lack of transparency on how some technologies work.** One supervisory agency notes that computational capacity constraints limited the amount of training data that could be used in developing an analytics solution. Moreover, the "black-box" nature of the results of some solutions is a challenge, as supervisors need to understand them. The expert opinions of supervisors are thus indispensable when deciding on any further action on the output of the innovative technologies.

50. **Suptech applications may deliver better but obviously not perfect predictors of vulnerabilities or wrong-doing.** Solutions may produce false positives or false negatives. False positives are results that indicate culpable behaviour exists when it in fact does not. False negatives occur if the technology is unable to detect actual cases. All supervisory agencies agree that suptech output will need to be evaluated by a human supervisor before any follow-up action is taken.

Data quality issues

51. **Supervisory agencies working with suptech encountered data quality issues.** Data quality and completeness can be an issue for non-traditional sources of information (eg social media). In addition,

[36] In theory though, supervised entities could report consistently false data.

[37] See Bauguess (2017).

a number of supervisors also mention data size as a key issue. This is the case, for example, where data sources are too big to handle (eg equity and derivatives markets transactions).

Legal risk

52. **Supervisory agencies face legal issues related to suptech.** Supervisors need to be aware of the legal implications of the use of suptech, particularly in the area of data collection. This keen awareness of potential legal issues is reflected in OeNB's reporting solution. The solution is designed in such a way that OeNB has no direct access to the commercially sensitive raw data of the supervised banks since there is no legal basis for supervisors to have access to this data. Legal issues may also arise from either intended or unintended breach of data privacy laws. This is particularly the case in using alternative sources of data such as social media.[38] Therefore, supervisors need to ensure that they have the necessary legal permissions in place to use data for supervisory purposes.

Operational risk

53. **Heightened operational risks, including cyber-risk, were mentioned underscoring the need for improved risk management in supervisory agencies when using suptech applications.** While supervisory agencies note that they have controls in place to ensure data security, some note that this is a particular risk in relation to open source and cloud applications. Increased third-party risks related to cloud computing and algorithm providers can result when data is transmitted online or is handled by third parties. Data security issues may also arise in the context of supervisory reporting applications where the supervisors' and the banks' systems are interconnected. Indeed, data security including cyber-attacks is a key threat in an automated environment, where the threats include data losses and interruption of supervisory activities.[39] A robust risk management and control framework should therefore accompany the use of suptech.

Reputational risk

54. **Reputational risk is also a key concern in the use of suptech.** Suptech applications have the potential to mitigate reputational risk of supervisors by picking up early-stage signals of, for example, fraud. However, these applications may also expose users to reputational risk. The risk of false positives or false negatives from poor-quality algorithms or data may affect the reputation of supervisors. In addition, the lack of transparency in so-called black-box algorithms could compromise the accountability of a supervisory agency.[40]

Resource issues

55. **Finding the right talent is a key challenge and key person risk is high when it comes to suptech.** The ideal candidates for suptech support work should be knowledgeable in data science, computer science and supervision. One supervisory agency admits that, as people with knowledge in all three areas are "outpriced by the market", it focuses on hiring people with a background in two of the three areas. Another supervisory agency highlights the difficulty of attracting experts with senior experience. The agency attracts young talents and develops them internally. Because of the scarcity of staff with the right background, each suptech solution may be dependent on just one or two key persons. This raises continuity issues if the key person leaves. Supervisory agencies can mitigate this risk by ensuring that solutions are well documented; assigning at least two staff to each project; running internal seminars so that the whole team is familiar with what each member is doing; and encouraging open discussions on

[38] See Toronto Centre (2018).

[39] See Toronto Centre (2018).

[40] See Toronto Centre (2018).

career development. Even so, retaining qualified staff for the long term is likely to become increasingly difficult.

Internal support issues

56. **There is management support for suptech projects, but supervisory agencies concede that room for improvement exists.** One supervisory agency notes that people working on suptech need to be persuasive to get senior management approval for their projects. The key to persuading management is presenting technology output in a meaningful and clear way. This is also necessary to get access to government funding in some cases. Another agency reports that it may take between five and 10 years for suptech to be fully embraced and integrated into the organisation. This is especially true if the supervisory agency has less affinity for technology. In such organisations, education is needed to develop the technology mindset at the top. The "black-box" nature of some technologies also is a challenge in getting management support. However, efforts are being made to make deep-learning results more transparent.[41]

57. **It is also crucial to get buy-in from the supervision units that are the ultimate users of these new technologies.** In doing so, it is important not to oversell the benefits of suptech. Data analytics tools are just one part of the supervisory toolkit. Qualitative assessments of business models and on-site inspections, for example, are also important. More importantly, it must be recognised that while suptech may be able to significantly improve the identification of potential issues and weaknesses in supervised entities, it does not tell supervisors what the appropriate actions are.

Practical issues

58. **Supervisory agencies face specific constraints that may also affect the implementation of suptech applications.** Supervisory agencies need to adhere to the standard rules and policies applicable in their jurisdictions when undertaking projects. Hence, procurement, building, testing and implementing technology solutions may take longer than it would in a private enterprise. While this may also apply to existing supervisory tools, it may be exacerbated in the case of suptech by the novelty of the applications, the unfamiliarity of supervisory agencies' procurement offices with these new technologies, and the inexperience of some technology vendors with these procurement processes.

What are the implications of suptech for supervised entities?

59. **Supervised institutions also stand to benefit when supervisors use suptech applications.** This is especially true of automated reporting, which enhances efficiency at supervised entities since they no longer have to submit error-prone template-based reports that take up a substantial amount of staff time. In addition, ad hoc data requests from supervisors are practically eliminated because supervisors now have access to flexible data formats that lend themselves well to different analytical needs. These solutions also enhance effectiveness of risk management in supervised firms. The BNR, for example, said that the overhaul of its reporting framework enabled supervised entities to benefit from improved data for risk management purposes. By doing away with unwieldy template-based reports to supervisors, these solutions also reduce reporting costs for supervised firms. Machine-readable regulations, on the other hand, should also promote improved compliance at a lower cost.

60. **At the same time, supervisory agencies recognise the risk that their use of suptech might induce market participants to adjust their behaviour accordingly.** One supervisory agency, for example, is weighing the pros and cons of disclosing information on its suptech applications. On the one hand, disclosure would be good practice and would help compliance units in supervised firms. On the

[41] See Knight (2017).

other hand, it might enable market participants to gain enough knowledge of the suptech solution to game the technology.

Section 4 – Concluding remarks

61. **Innovative technologies, together with increased data availability, create scope to strengthen financial supervision.** Supervisory agencies around the world recognise this and are now either using or exploring a wide variety of innovative technologies to support their work. These technologies are applied in a broad range of supervisory activities, which can be grouped into two broad areas: data collection and data analytics.

62. **Suptech for data collection purposes is likely to benefit both supervisors and supervised entities.** This will be driven by the need for cost-efficient and flexible reporting formats. Suptech also brings real-time monitoring of markets within reach. The best approach to successful suptech applications in data collection is to focus on long-term value as opposed to short-term cost efficiencies.

63. **Suptech applications for data analytics could potentially transform risk and compliance monitoring from a backward-looking into a more predictive and proactive process.** As technology becomes more developed, suptech applications will be increasingly used to anticipate the behaviour of supervised entities or their risk exposures.

64. **However, supervisory agencies need to appreciate the risks and limitations of suptech.** Use of suptech without taking the necessary measures to address technical, data quality, legal, operational, reputational, resource, internal support and practical issues may expose supervisors to undue risks. Moreover, although suptech can help identify potential issues and problems, human intervention is necessary to pursue further investigations and decide on a suitable course of action.

65. **Supervisory agencies also need to be cautious of a growing data-knowledge gap.** On the one hand, data availability, data quality and data storage facilities are improving rapidly, as are techniques for combining different data sources. On the other hand, data analytics may not be advancing at the same pace. It takes time to learn, develop and implement new technologies in supervision work. Agencies could make an assessment of data availability and to what extent data is being fully used in supervision work.

66. **Supervisory agencies would be best placed to explore the potential benefits of suptech applications if they have a well defined suptech strategy.** A suptech strategy should comprise, at a minimum, the following three key elements: first, ambitious, but achievable, targets (eg which technology will be used in which area of supervision three to five years from now, how this technology will be embedded in the organisation and how it will be funded); second, an assessment of today's data availability, data quality and analytical resources; and third, a step-by-step action plan on how the supervisory agency will get from the current situation to full implementation of its strategy. The experience of early users of suptech discussed in this paper provides some useful insights as to how to develop such a strategy. The following are some specific considerations that could help supervisory agencies take advantage of suptech developments.

(a) **The overall approach to supervision should adapt to the digitisation of the activities of supervised entities.** There are several key success factors to a further digitalisation of supervision. These are based on a technology-centred approach with clear data standards and reporting formats that are useful for users and make it easier for developers to build data analytics applications; high-quality data; and educating supervisors about the benefits and use of technology.[42]

[42] Or, as one interviewee put it: "every supervisor should become a little bit of a data scientist".

(b) **Management support is critical in exploring the opportunities and benefits of suptech.** For this to happen, management needs to appreciate the potential of suptech, while remaining aware of its limitations and risks. Suptech needs to be made more approachable. Hence, technology experts within supervisory agencies should educate and engage management in this area.

(c) **Supervisory agencies engaged in suptech need specialised human resources.** Based on their existing staff complement, supervisory agencies should consider whether to establish new units or use existing units for suptech work. In any case, diversity through a multidisciplinary approach is needed to create the best suptech teams. Assembling the right mix of talent will be a challenge. Thus, supervisory agencies need to carefully weigh how to attract and retain suptech staff, as well as ensure that institutional knowledge is maintained in the face of a potentially high rate of staff turnover.

(d) **The buy-in of supervision or enforcement units helps to fully embed suptech in supervision work.** Input from supervision or enforcement units should be considered in developing suptech applications. Supervision or enforcement units need to be on board from the development stage for two reasons: (1) as these units will be using the applications, their buy-in is crucial if development efforts are to proceed beyond the study phase; and (2) their expertise is indispensable to fully understand and further examine the output of suptech applications.

(e) **Supervisory agencies can benefit from partnerships with the academic community.** To keep up with fast-moving technical developments, particularly in cutting-edge technologies that may have suptech applications, supervisory agencies need to stay attuned to new ideas emerging in academia.

(f) **The use of suptech reinforces the case for further improving risk management at supervisory agencies.** Improving risk management has historically been less of a concern for supervisory agencies than it is for supervised entities. But the increasing use of suptech exposes supervisory agencies to additional risks, such as legal and operational risk, including cyber-risk, and reputational risk, which must be mitigated to maximise the benefits of suptech.

(g) **As supervisory agencies can learn from each other, it is important to seek opportunities for collaboration.[43]** The key to growing or enhancing suptech capabilities is for supervisory agencies to continuously exchange knowledge and experience on a global level. This collaboration will push forward the technological progress needed to support better global financial supervision.

[43] See BCBS (2018).

References

Accenture (2018): *RegTech for regulators: Re-architect the system for better regulation*, February.

Accornero, M and M Moscatelli (2018): "Listening to the buzz: Social media sentiment and retail depositors' trust", Bank of Italy, working paper, no 1165.

Angelico, C, J Marcucci, M Miccoli and F Quarta (2018): "Can we measure inflation expectations using Twitter?", Bank of Italy, March.

Arner, D, J Barberis and R Buckley (2016): "FinTech, RegTech and the reconceptualisation of financial regulation", *Northwestern Journal of International Law and Business*, October.

Basel Committee on Banking Supervision (2018): "Sound Practices: implications of fintech developments for banks and bank supervisors", February.

Bauguess, S (2016): "Has Big Data made us lazy?", Midwest Region Meeting – American Accounting Association, October.

——— (2017): "The role of Big Data, Machine Learning, and AI in assessing risks: A regulatory perspective", keynote address, OpRisk North America, New York, June.

Berndsen, R and R Heijmans (2017): "Near real-time monitoring in real-time gross settlement systems: A traffic light approach", Netherlands Bank.

Copeland, M (2017): "What's the difference between Artificial Intelligence, Machine Learning, and Deep Learning?", July.

Dias, D and S Staschen (2017): "Data collection by supervisors of Digital Financial Services", *CGAP Working Papers*.

Financial Conduct Authority (2018): "Call for Input: Using technology to achieve smarter regulatory reporting".

Financial Stability Board (2017a): "Financial stability implications from FinTech: Supervisory and regulatory issues that merit authorities' attention", 27 June.

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

Hunt, S (2017): "From maps to apps: The power of Machine Learning and Artificial Intelligence for regulators", Beesley Lecture Series on regulatory economics, 19 October.

Kamali, W and D Randall (2017): "Leveraging 'suptech' for financial inclusion in Rwanda", Private Sector Development (blog), World Bank.

Knight, W (2017): "The dark secret at the heart of AI", *MIT Technology Review*, 11 April.

Loberto, M, A Luciani and M Pangallo (2018): "The potential of big housing data: An application to the Italian real-estate market", Bank of Italy, 27 March.

Lux, J and M Piechocki (2015): "Reforming regulatory reporting: Are we headed toward real-time?", BearingPoint, *Institute Report*, no 006.

Marr, B (2016): "What is the difference between artificial intelligence and machine learning?", 6 December.

Menon, R (2017): "Remarks by Mr Ravi Menon", *Singapore FinTech Journey 2.0*, Singapore FinTech Festival, 14 November.

National Bank of Rwanda (2017): *Annual Report*, 1 July 2016–30 June 2017.

Timmermans, M, R Heijmans and H Daniels (2017): "Cyclical patterns in risk indicators based on financial market infrastructure transaction data", *DNB Working Papers*, no 558, Netherlands Bank, June.

Toronto Centre (2017): "FinTech, RegTech and SupTech: What they mean for financial supervision", *TC Notes,* August.

________ (2018): "SupTech: Leveraging technology for better supervision", *TC Notes*, July.

Triepels, R, H Daniels and R Heijmans (2017): "Detection and explanation of anomalous payment behavior in real-time gross settlement systems", Netherlands Bank.

Wall, L (2017): "Some financial regulatory implications of Artificial Intelligence", working paper, Federal Reserve Bank of Atlanta, September.

Annex 1 – List of organisations interviewed

Australian Securities and Investments Commission (ASIC)

Bank of Italy (BoI)

Bangko Sentral ng Pilipinas (BSP)

BearingPoint

Boston Consulting Group (BCG)/Expand Fintech Control Tower

Central Bank of the Republic of Austria (OeNB)

National Banking and Securities Commission (CNBV)

Netherlands Bank (DNB)

European Securities and Markets Authority (ESMA)

Federal Reserve Bank of San Francisco

Financial Conduct Authority (FCA)

Monetary Authority of Singapore (MAS)

National Bank of Rwanda (BNR)

RegTech for Regulators Accelerator (R^2A)

Securities and Exchange Commission (SEC)

Suade Labs

World Bank (WB)

World Economic Forum (WEF)

Annex 2 – List of definitions[44]

Application programming interfaces (APIs): allow different software to interact. They comprise a set of rules and specifications that software programmes use to communicate with each other, and an interface between different software programmes that facilitates this interaction. APIs offer an efficient and flexible way of collecting and disseminating data.

Data push approach: a reporting framework in which supervised entities "push" data to a supervisory agency. The supervised entities prepare their data in a standard format in a series of basic datacubes as prescribed by the supervisory agency. The basic datacubes are uploaded to the supervisory agency and then transformed into a series of smart datacubes formatted to the agency's requirements. The data input approach combines transparency with confidentiality as the supervisory agency does not have direct access to commercially sensitive data.[45]

Data pull approach: a reporting framework in which supervisory agencies "pull" data directly from the IT systems of supervised entities. The supervisory agency controls and initiates this fully automated reporting process.[46]

Machine-readable regulations: regulations issued as programming codes that can be assimilated immediately by supervised entities' operational systems, without the need for a human to interpret them.[47]

Cloud computing: refers to the use of an online network ("cloud") of hosting processors to increase the scale and flexibility of computing capacity. Cloud computing is easily scalable and generates cost savings over internal IT systems.[48]

Chatbots: are virtual assistance programmes that interact with users in natural language. Financial services firms are introducing chatbots on a large scale as a cost-efficient way of helping customers make financial decisions.

Big data: refers to the voluminous amounts of structured or unstructured data that can be generated, analysed and increasingly used by digital tools and information systems. This capability is driven by the increased availability of structured data, the ability to process unstructured data, increased data storage capabilities and advances in computing power. "Big data" sometimes refers to situations in which the amount of data challenges the limits of computer processing power.[49]

Artificial intelligence (AI): is intelligence demonstrated by machines. It is defined as IT systems that can perform functions that would otherwise require human capabilities. AI can ask questions, discover and test hypotheses, and make decisions automatically based on advanced analytics operating on extensive data sets.

Natural language processing (NLP): the capacity of computer programs to process human language. As a component of artificial intelligence, NLP is an interdisciplinary field comprising computer science, AI, and computational linguistics. It focuses on programming computers and algorithms to parse, process, and

[44] Definitions are taken from FSB (2017b) or BCBS (2018) unless otherwise stated.

[45] See BearingPoint (2015).

[46] See Toronto Centre (2017).

[47] See Toronto Centre (2017).

[48] See Toronto Centre (2017).

[49] See Bauguess (2016).

"understand" human language. NLP applications are designed to understand natural human communication, either written or spoken, and to respond using natural language.

Machine learning (ML): entails computers learning from data without human intervention. It is a method of designing problem-solving rules that improve automatically through experience. Machine-learning algorithms give computers the ability to learn without specifying all the knowledge a computer would need to perform the desired task, as well as study and build algorithms that can learn from and make predictions based on data and experience. Machine-learning algorithms are used to identify patterns that are correlated with other events or patterns.[50]

Supervised learning: involves computers learning from examples of correct input-output pairs. It is a subset of ML in which an algorithm is fed a set of training data with labelled observations. Supervised learning can be used to categorise items (eg whether something is a cat) and to predict numerical values (eg stock returns).[51] Supervised learning algorithms are invested with human knowledge. Algorithms can, for example, identify relationships with new variables or discover previously undetected interactions among variables.

Unsupervised learning: involves computers discovering the hidden structure in unlabelled data. It is a subset of ML in which the data provided to the algorithm do not contain labels. A common method of unsupervised learning is that of clustering, ie to find patterns in the data by identifying clusters of observations that depend on similar underlying characteristics.[52]

Topic modelling: method of unsupervised learning that lets the data define key themes in a text. Topic modelling can efficiently identify hidden trends in large amounts of unstructured financial information.

Random forest: combines multiple ML algorithms, allowing for overall better performance. It is a supervised learning algorithm that can be used for both classification and regression tasks. Random forest techniques can be used with historical data to develop new predictive models.[53]

Deep learning: an algorithm that can, independently, learn new skills. This subset of ML refers to a method that uses algorithms inspired by the structure and function of the brain, known as artificial neural networks. The more computation time it gets, the better the algorithm becomes. The process by which deep learning techniques come to predictions or decisions is unclear.

Neural networks: are the base concept for deep learning algorithms and can be used for supervised and unsupervised learning. Like a brain, a neural network contains a large number of nodes and typically learns by training on real data in which the correct answer is already known.[54]

Image recognition: a form of deep learning that can be applied to many image-processing and computer vision problems such as categorising handwritten numerals within an image.

[50] See GRCI (2017).

[51] See Wall (2017).

[52] See Wall (2017).

[53] See Wall (2017) and Accenture (2018).

[54] See Wall (2017).