

TECHNICAL NOTE

The Next Wave of Suptech Innovation

Suptech Solutions for Market Conduct Supervision

MARCH 2021



Finance, Competitiveness & Innovation Global Practice

© 2021 International Bank for Reconstruction and Development / The World Bank

1818 H Street NW
Washington DC 20433
Telephone: 202-473-1000
Internet: www.worldbank.org

This work is a product of the staff of the World Bank with external contributions. The findings, interpretations, and conclusions expressed in this work do not necessarily reflect the views of the World Bank, its Board of Executive Directors, or the governments they represent.

The World Bank does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of the World Bank concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

RIGHTS AND PERMISSIONS

The material in this work is subject to copyright. Because the World Bank encourages dissemination of its knowledge, this work may be reproduced, in whole or in part, for noncommercial purposes as long as full attribution to this work is given.

Any queries on rights and licenses, including subsidiary rights, should be addressed to the Office of the Publisher, The World Bank, 1818 H Street NW, Washington, DC 20433, USA; fax: 202-522-2422; e-mail: pubrights@worldbank.org.



CONTENTS

Acknowledgments	iii
Acronyms and Abbreviations	iv
EXECUTIVE SUMMARY	1
1. INTRODUCTION	4
2. CATEGORIES OF SUPTECH SOLUTIONS FOR MARKET CONDUCT SUPERVISION	7
3. SUPTECH SOLUTIONS FOR MARKET CONDUCT SUPERVISION	10
3.1. Regulatory Reporting	10
3.1.1. Supervision Information Systems	10
3.1.2. Automated Data Submission via API	13
3.1.3. Web Portal Data Upload with Central Database	14
3.2. Collection and Processing of Complaints Data	15
3.2.1. Complaints Management System	15
3.2.2. Analysis of Unstructured Complaints Data	16
3.3. Non-traditional Market Monitoring	17
3.3.1. Web Scraping	17
3.3.2. Social Media Monitoring	18
3.3.3. Consumer Sentiment Analysis	19
3.3.4. Reputational Analysis	19
3.3.5. Dark Web Monitoring	20
3.4. Document and Business Analysis	20
3.4.1. Document Analysis for Regulatory Compliance	20
3.4.2. Document Analysis for Examination of FSPs	21
3.4.3. Document Analysis for Peer Group Comparison	22
3.4.4. Validation of Terms and Conditions	22
3.4.5. Automated Review of New Provider Registrations	23
3.4.6. Predictive Modeling of Financial Statements	23
3.4.7. Business Intelligence and Geospatial Analysis	24
3.4.8. Managed Data Platform	24

4. PEOPLE, PROCESS, AND IT INFRASTRUCTURE: THREE KEY ENABLERS FOR SUPTECH IMPLEMENTATION	25
4.1. People: Culture and Skillsets	25
4.2. Process: Internal Champions and Strong Governance	26
4.3. Underlying IT Infrastructure	27
5. IMPLEMENTATION CONSIDERATIONS	28
5.1. Key Decisions in Suptech Implementation	28
5.2. Initiatives to Accelerate Suptech Implementation	31
5.3. Additional Challenges Encountered by Regulators	33
6. LOOKING FORWARD	34
REFERENCES	35
FIGURES	
1. Suptech Solutions for Market Conduct Supervision and Key Enablers for Implementation	3
2. Suptech Solutions for Market Conduct Supervision and Key Enablers for Implementation	6
3. A Function-Based Suptech Taxonomy with Suptech Use Cases	7
4. Results Framework for Market Conduct Suptech Solutions	9
5. Overview of Suptech Solutions for Market Conduct Supervision	11
6. Dataflow Diagram of SIS Solutions	13
7. CMS Case Workflow and Data Architecture	15
8. Dataflow Diagram for Social Media Monitoring	18
9. Example of Dataflow Diagram in Document Analysis Solutions	21
10. Considerations for In-House Development Versus Using a Third-Party Vendor	29
CASE STUDIES	
1. How BNR Designed Its SIS Solution	12
2. How NBR Develops Suptech Solutions	14
3. How Researchers at Princeton University and FSD Kenya Worked with the Central Bank of Kenya to Analyze Complaints Data	17
4. The FCA's Development of Sleuth, Its NLP Platform	21
5. How AFM Prioritized People within Its Transformation to Data-Driven Supervisors	25
6. How ASIC's Innovation Office Collaborates with Industry Stakeholders	32
BOX	
1. FinCoNet: SupTech Tools for Market Conduct Supervisors	33



ACKNOWLEDGMENTS

This technical note is a product of the Financial Inclusion and Consumer Protection Team in the World Bank Group's Finance, Competitiveness and Innovation Global Practice.

This note was prepared by Ligia Lopes (former Senior Financial Sector Specialist, World Bank), Jennifer Chien (Senior Financial Sector Specialist, World Bank), Mackenzie Wallace (Market Conduct Supervision Consultant), and Edoardo Totolo (Operations Officer, International Finance Corporation). Mahesh Uttamchandani (Practice Manager, World Bank) provided overall guidance. The team is grateful for the substantive feedback received from peer reviewers Douglas Randall (Financial Sector Specialist, World Bank) and Matei Dohotaru (Senior Financial Sector Specialist, World Bank), and from the International Financial Consumer Protection Organisation (FinCoNet). Editorial inputs were provided by Charles Hagner and design and layout assistance was provided by Debra Naylor of Naylor Design, Inc.

The team also gratefully acknowledges the generous contributions of time and expertise by financial authorities at the Australian Securities and Investments Commission, the Authority for the Financial Markets (Netherlands), Autorité des Marchés Financiers (Québec, Canada), Banco de Portugal, Bank of England, Bangko Sentral ng Pilipinas (Philippines), the Central Bank of Ireland, the European Insurance and Occupational Pensions Authority, the Financial Conduct Authority (United Kingdom), the National Bank of Rwanda, and Nepal Rastra Bank.

Finally, the team gratefully acknowledges the generous financial support of the Ministry of Foreign Affairs of the Kingdom of the Netherlands and the Bill & Melinda Gates Foundation under the Financial Inclusion Support Framework (FISF) program, without which preparation of this paper would not have been possible.



ACRONYMS AND ABBREVIATIONS

ADF	automated dataflow
AFM	Authority for the Financial Markets (Netherlands)
AMF	Autorité des Marchés Financiers (Québec, Canada)
API	application programming interface
ASIC	Australian Securities and Investments Commission
BdP	Banco de Portugal
BI	business intelligence
BNR	National Bank of Rwanda
BOE	Bank of England
BOL	Bank of Lithuania
BSP	Bangko Sentral ng Pilipinas (Philippines)
CBI	Central Bank of Ireland
CFPB	Consumer Financial Protection Bureau (United States)
CMS	complaints management system
CRM	customer relationship management
EDW	Electronic Data Warehouse
EIOPA	European Insurance and Occupational Pensions Authority
EU	European Union
FCA	Financial Conduct Authority (United Kingdom)
FSP	financial service provider
MVP	minimum viable product
NLP	natural language processing
NRB	Nepal Rastra Bank
SIS	supervisory information system
USD	United States dollar



EXECUTIVE SUMMARY

Around the world, financial sector supervisors are experiencing a profound shift to data-driven supervision enabled by the next wave of technology and data solutions.¹ While technology and data are not new to financial oversight, their specific application to financial consumer protection and market conduct supervision is a newer and welcome trend.

Supervisory technology, or supotech, refers to the use of technology to facilitate and enhance supervisory processes from the perspective of supervisory authorities. As highlighted in the World Bank's 2018 discussion note on supotech for market conduct supervision (World Bank 2018), examples of supotech for market conduct supervision were initially limited. In recent years, the application of supotech for market conduct supervisory purposes has become more widespread and sophisticated. Recent advancements, particularly in the realm of unstructured and text analysis, present opportunities for market conduct supervision where a greater reliance on qualitative assessments is required.

This technical note draws from a wide set of regulatory experiences to showcase new supotech solutions specific to market conduct supervision. The main objective of this note is to assist market conduct authorities, particularly those in low- and middle-income countries, to build and enhance supervisory capacity and efficiency by providing concrete examples where supervisory technology can be leveraged.

Four key insights for market conduct authorities can be drawn from this note:

INSIGHT 1: Increasing operational efficiency and enhancing supervisory effectiveness are two of the primary motivations for adopting supotech solutions for market conduct.

In implementing supotech, financial authorities are often driven by two different motivations: (1) increasing operational efficiency and (2) improving hypothesis-driven supervision. The former often involves automating business processes by replacing elements of the supervision decision framework with data and algorithms, bringing significant efficiencies to the process, while the latter involves helping supervisors to test and prove hypotheses using new sources of analyses or data.

Given limited capacity at many financial authorities, implementation of supotech for market conduct often focuses on solutions to increase operational efficiency.

The rationale is to make existing staff more productive and to enable them to focus on higher-value activities. Repetitive or time-consuming tasks such as data cleaning or document, data, or complaints intake and processing are prime candidates for supotech automation. From an initial focus of operational efficiency, some market conduct supervisors have since expanded their overall approach to include enhancing the effectiveness of their supervisory program.

1. *Solution* is used in this note to refer to an implementation of people, processes, information, and technology that supports a set of business or technical capabilities that solve one or more business problems.

INSIGHT 2: Suptech solutions for market conduct supervision can be grouped into four categories.

This technical note explores 18 suptech solutions for market conduct, grouped into the following four categories. These categories generally align with their respective supervisory activity, rather than groupings based on technological functionality (which is another approach for categorizing suptech solutions).

1. **Solutions for regulatory reporting by supervised institutions:** A primary method for regulators to identify market conduct risks and issues is to collect information directly from supervised institutions, but doing this can be time consuming and labor intensive. Web portals, application programming interfaces (APIs), automated dataflows (ADFs), and comprehensive supervision information systems (SISs) allow for automated and standardized regulatory reporting that collects, validates, transforms, and stores data in real time.
2. **Solutions for collection and processing of complaints data:** Complaints data is one of the most valued data sources for market conduct supervisors. A complaints management system (CMS) is key to the efficient processing of these complaints and capturing and managing data to maximize its accuracy and value for supervisory purposes. The application of advanced analytics to complaints data, particularly to unstructured text, represents the next step for market conduct supervisors to deduce new insights in a more efficient manner from complaints data.
3. **Solutions for non-traditional market monitoring:** The internet provides the opportunity to utilize a range of new, non-traditional methods for monitoring the market, another core activity for market conduct supervision. Monitoring social media, online news, websites, and so on can provide early warning signals of emerging consumer risks. Foundational to these types of solutions is web scraping, which provides the mechanism for collecting and gathering online text for analysis. Such text can be used for social media monitoring, reputational analysis in the news, consumer sentiment scoring, and dark web monitoring. Non-traditional market monitoring provides supervisors a useful complement to traditional market monitoring.
4. **Solutions for document and business analysis:** Advances in analytics have been most profound in the realm of unstructured text data. For example, natural language processing (NLP) solutions can ingest and analyze large quantities of documents, extracting

insights in seconds where previously it would have taken supervisors weeks (if even possible at all). Given the more qualitative nature of market conduct supervision, advancements in the analysis of text present a potentially significant breakthrough.

In each of the above categories, the suptech solutions described span the data life cycle of a specific supervisory activity. The solutions within each category present a collection of tools that enable supervisors to collect new forms of data or introduce new, more efficient methods for collecting such data. Suptech solutions can also be used to conduct richer analyses on an exponentially increasing amount of information with limited analytical resources. These collections of suptech solutions therefore provide market conduct supervisors with both gains in efficiency and the ability to extract new insights to allow for data-driven decision making.

INSIGHT 3: Suptech implementation is about more than just the technology.

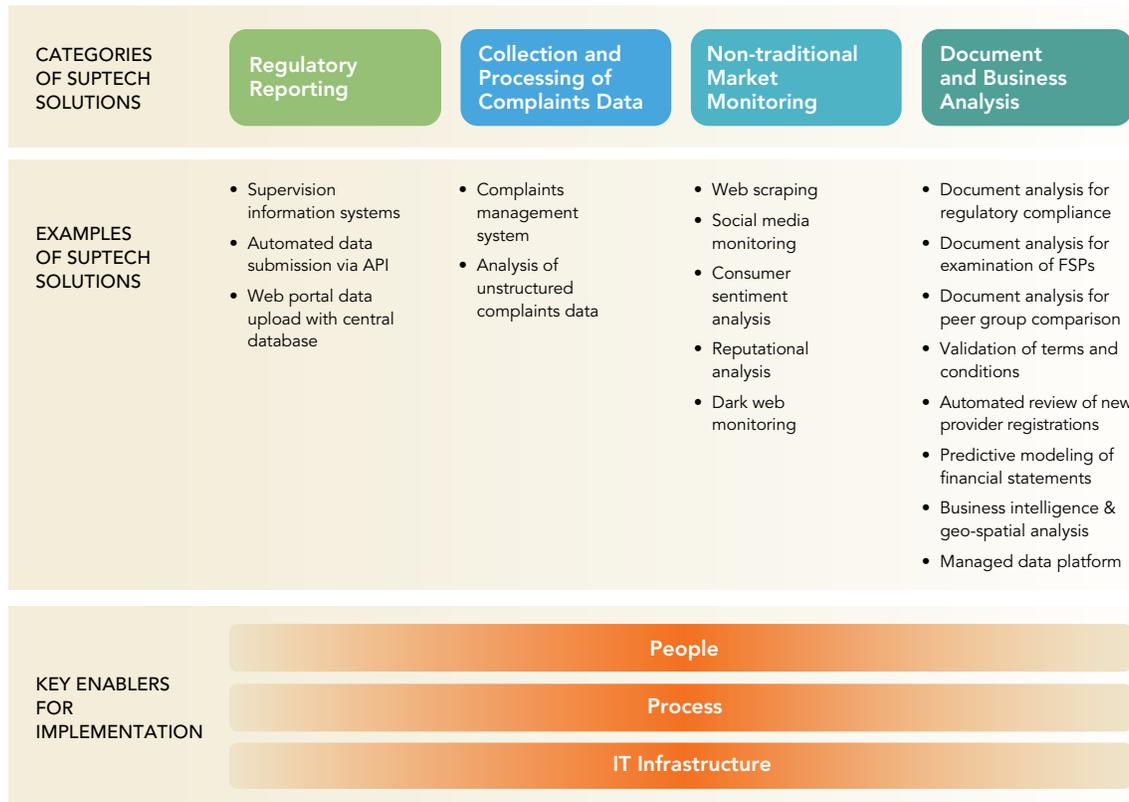
Embedding modern technology and data into the supervisory process is often an ongoing effort. Implementing suptech solutions requires more than just the solution. It requires making investments in three key enablers: people, process, and IT infrastructure. The culmination of broader efforts to implement suptech solutions and underlying enablers is organizational transformation into a data-driven supervisor.

- **People** refers to the talent, mindset, and skills of employees and the larger organizational culture toward data and technology.
- **Process** refers to how suptech ideas are supported from ideation to implementation, including how suptech is championed and governed.
- **IT Infrastructure** refers to the underlying IT infrastructure and capabilities needed to develop and operate suptech solutions internally.

INSIGHT 4: Various strategies can be used to help accelerate the development and implementation of suptech solutions.

Some financial authorities have benefited from the creation of formal, multiyear suptech or data strategies. Innovation offices can also be leveraged to provide a central place to encourage internal suptech ideation and learning, as well as improving dialogue with such external parties as fintechs or potential suptech solution providers.

FIGURE 1. Suptech Solutions for Market Conduct Supervision and Key Enablers for Implementation



In some instances, it is more appropriate to begin with an incremental, targeted approach, rather than a broader institutional strategy. Supervisors in low- and middle-income countries will inevitably face challenges during implementation. Common challenges include underdeveloped supervisory risk frameworks, staffing and resource constraints, and technology constraints among financial service providers (FSPs). However, successful implementation of suptech solutions in these contexts can provide more meaningful gains to efficiency and effectiveness in low-capacity countries. These constraints favor a targeted approach to suptech implementation that focuses scarce time, attention, and resources.

Utilizing experimentation and iteration in the technology-development process can be beneficial. In place of traditional approaches such as “waterfall,”² authorities now increasingly use design or tech sprints, proofs of concept, prototypes, pilots, “minimum viable products,” and agile approaches. Such approaches engage and validate capabilities with end users, both ensuring the utility of the solution when delivered and condensing the implementation timeline.

The expansion of digital activity prompted by the COVID-19 pandemic reemphasizes the necessity and value of suptech for financial authorities. This is true for all categories of suptech solutions for market conduct. The direct and automated collection of granular regulatory data from supervised institutions is critical to replacing on-site examinations, as is the ability of supervisors to engage directly with consumers and manage their complaints with providers digitally. Meanwhile, both non-traditional market monitoring and advanced text analysis allow supervisors to monitor fast-moving sentiment remotely and emerging risks to consumers on a more rapid basis.

Such tools that enable supervisors to oversee the financial sector with increased effectiveness and efficiency will only become more critical as digital transformation continues. The initial successes experienced by the authorities referenced in this technical note offer a glimpse of this future—one in which data and technology become core to the operations, identity, and culture of all supervisors. Such tools hold the promise to help empower financial authorities to meet the market conduct supervisory challenges of the next decade.

2. Waterfall software development methodology refers to a linear, sequential approach whereby customer and business requirements are gathered at the beginning of the project and the technology solution is developed following a sequential project plan to accommodate those requirements.



1. INTRODUCTION

Financial sector authorities around the world are experiencing a profound shift to data-driven supervision enabled by robust technology and data solutions. While technology is obviously not new to financial authorities, this new wave of digital solutions holds the promise to increase the efficiency and effectiveness of supervision in order to meet key regulatory objectives, including financial stability, financial integrity, and, increasingly, financial consumer protection. This technical note showcases new supervisory technology, or *suptech*, solutions specific to market conduct supervision that can assist financial sector authorities—including in low- and middle-income countries—to enhance and strengthen financial consumer protection and market conduct supervision.

While regulators have always leveraged data and technology for supervisory purposes, a marked increase in new and ambitious initiatives has occurred in recent years. Examples include the introduction of “TechSprints” at the Financial Conduct Authority (FCA) in the United Kingdom, development of the Electronic Data Warehouse (EDW) at the National Bank of Rwanda (BNR), and the launch of “Step 1” technology transformation at the Authority for the Financial Markets (AFM) in the Netherlands,³ among many other technological developments at financial authorities worldwide.

This latest wave of new *suptech* solutions builds on earlier generations of technology solutions. Most supervisory technology to date has focused primarily on data-management workflows and descriptive analytics. However, many of these solutions involve a certain degree of manual processing or had other limitations (BIS 2019).

A new generation of more advanced *suptech* solutions is currently emerging, driven by the latest technological innovations in big data architecture, machine learning (especially NLP), and automated data collection and management. In this note, the term *suptech* refers to the use of technology to facilitate and enhance supervisory processes from the perspective of supervisory authorities.

While technology solutions are not new to financial oversight, their specific application to market conduct is a newer and welcome trend. Historically, technology solutions for quantitative analysis have been more advanced than qualitative ones, with greater application for prudential supervision. Recent advancements in data and technology, such as NLP and other machine-learning applications, present new opportunities for market conduct supervisors by enabling greater qualitative analyses. As highlighted in the World Bank’s previous discussion note on *suptech* for market conduct supervision (World Bank 2018), examples of *suptech* for market conduct supervision were initially limited to complaints data collection and analyses. In the past few years, the application of *suptech* for purposes of market conduct supervision has become more widespread and sophisticated, as explored in this note.

***Suptech* solutions are increasingly critical given the digital transformation of the financial services industry in recent years.** Supervisors have often lagged behind in their capacity to monitor these growing and increasingly complex markets. However, supervisors can leverage the technological advances behind digital transformation to overcome resource constraints and make processes and

3. AFM (Netherlands) developed a multiyear, three-phase *suptech* transformation program: Build, Pilot, and Transform. The “Build” phase began with an assessment of AFM’s own data and analytics capacity.

procedures more effective and efficient. In the face of limited capacity and resources, a particular concern in lower-income economies, suptech can be used to leverage data and technology to supervise financial services more efficiently and effectively for market conduct.

While adoption has been most pronounced in high-income countries, suptech solutions are relevant and translatable to lower-capacity countries. The uneven global uptake of suptech solutions can be partly attributed to the additional logistical barriers that supervisors in low- and middle-income countries often face. However, the broadening landscape of suptech solutions presents such authorities with the opportunity to learn from technology examples in other countries. Many of these examples of technology and data solutions can still be translated and adapted to countries with lower capacity, where their potential for positive impacts on supervisory efficiency and effectiveness may be even more powerful.

Research Objectives and Key Audience

The main objective of this note is to assist market conduct authorities in their efforts to build and enhance supervisory capacity and efficiency by providing concrete examples of situations in which supervisory technology can be leveraged. An efficient market conduct supervision framework requires the collection of a wide range of data from disparate sources; doing this is challenging in many jurisdictions. Market conduct supervisors must also undertake complex qualitative analyses to determine compliance with legislation or regulation that is often principles-based or composed of judgement-based rules. These challenges are compounded when supervisors have under their jurisdiction a large diverse range of FSPs with unique or unfamiliar risk profiles. Consequently, market conduct supervision continues to be manual and labor-intensive in many countries. Suptech presents the opportunity to enhance both supervisory capacity and efficiency to tackle these inherent operational challenges, particularly important in light of growing and rapidly digitizing financial markets.

While collective knowledge on suptech has grown in recent years,⁴ the literature specific to market conduct supervision is limited. This note seeks to address this gap by providing financial authorities with (1) an array of concrete examples of suptech solutions that can be used

for market conduct supervision, and (2) practical considerations for successful implementation of a data-driven supervision program, such as investments and organizational changes required to support implementation.

The main audience for this note is market conduct authorities and other stakeholders in low- and middle-income countries. Considering that the potential for gains in supervisory efficiency and effectiveness is high in lower-capacity countries, this note highlights solutions that can be adapted to these contexts and practical considerations in doing so. In addition, the note should benefit development practitioners assisting financial authorities by informing the development and design of technology support programs.

Information Sources

This note draws from a wide set of regulatory experiences and is the result of primary and secondary research with 14 financial authorities. These financial authorities represent a diverse cross-section with varied levels of financial market development and internal capacity. Each authority was selected on account of its successful track record in developing suptech solutions for market conduct supervision. Research methods included interviews, demonstrations, questionnaires, and reviews of internal materials, external publications, and public-facing websites.

The following financial authorities contributed critical inputs to this note:

- Australian Securities and Investments Commission
- Authority for the Financial Markets (Netherlands)
- Autorité des Marchés Financiers (Québec, Canada)
- Banco de Portugal
- Bangko Sentral ng Pilipinas (Philippines)
- Bank of England
- Bank of Lithuania
- Central Bank of Brazil
- Central Bank of Ireland
- Consumer Financial Protection Bureau (United States)
- European Insurance and Occupational Pensions Authority
- Financial Conduct Authority (United Kingdom)
- National Bank of Rwanda
- Nepal Rastra Bank⁵

4. Since the World Bank published its note on suptech for market conduct supervision (World Bank 2018), organizations such as the Bank of International Settlements, International Financial Consumer Protection Organization, Toronto Center, Milken Institute, R2A, Consultative Group to Assist the Poor, Columbia University, and others have published on suptech.

5. The Central Bank of Brazil, Bank of Lithuania, and Consumer Financial Protection Bureau (United States) contributed to the 2018 World Bank discussion note on suptech (World Bank 2018).

Structure of Technical Note

The technical note is structured into the following chapters:

CHAPTER 2: Categories of Suptech Solutions for Market Conduct Supervision. Four main categories of suptech solutions are introduced: (1) solutions for regulatory reporting by supervised institutions, (2) solutions for collection and processing of complaints data, (3) solutions for non-traditional market monitoring, and (4) solutions for document and business analysis, especially of unstructured data.

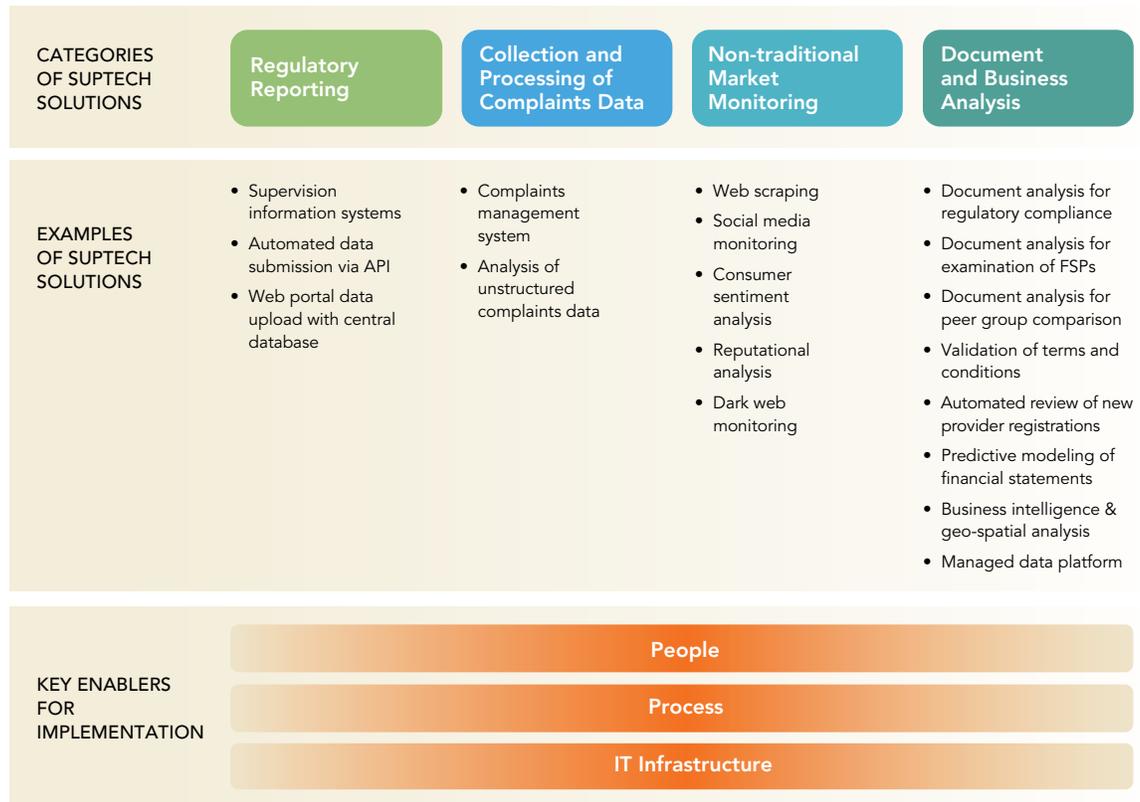
CHAPTER 3: Suptech Solutions for Market Conduct Supervision. Individual suptech solutions for market conduct are identified for each of the four main categories noted above, and a total of 18 solutions are presented. For each solution, there is a description of how the solution works, its benefits, and considerations for implementation, drawing from country experience and including detailed case studies.

CHAPTER 4: People, Process, and IT Infrastructure: Three Key Enablers for Suptech Implementation. Successful implementation of a suptech solution goes beyond the technology itself. Three suptech enablers are critical for implementation: (1) people, (2) process, and (3) underlying IT infrastructure.

CHAPTER 5: Implementation Considerations. Common considerations when implementing suptech solutions emerged across country examples. Authorities often face key decisions related to prioritization, determining whether to build a solution in-house or to buy from a vendor, and deciding how to organize data and technology staff. It is also useful to consider whether to undertake efforts to accelerate suptech adoption through formal suptech or data strategies, adaptive technology development, and internal innovation offices to liaise with external stakeholders.

CHAPTER 6: Looking Forward. This section includes brief final thoughts on the value of suptech solutions for market conduct supervisors operating in an increasingly complex environment.

FIGURE 2. Suptech Solutions for Market Conduct Supervision and Key Enablers for Implementation



2. CATEGORIES OF SUPTECH SOLUTIONS FOR MARKET CONDUCT SUPERVISION

No taxonomy of suptech solutions is widely accepted globally. To date, most existing taxonomies have taken a function-based approach toward describing suptech ecosystems (see Figure 3). Existing suptech taxonomies tend to categorize suptech solutions based on the flow of data from collection to validation, consolidation, and analysis.

This technical note takes a slightly different approach, categorizing suptech solutions for market conduct supervision by supervisory activity. Unlike the other taxonomies of suptech solutions, the categories employed in this note extend beyond dataflow to include engaging with supervised institutions and consumers as well as new types of non-traditional data collection and analysis. This categorization is not meant to be exhaustive for all possible suptech solutions for market conduct but reflective of the solutions described in this note.

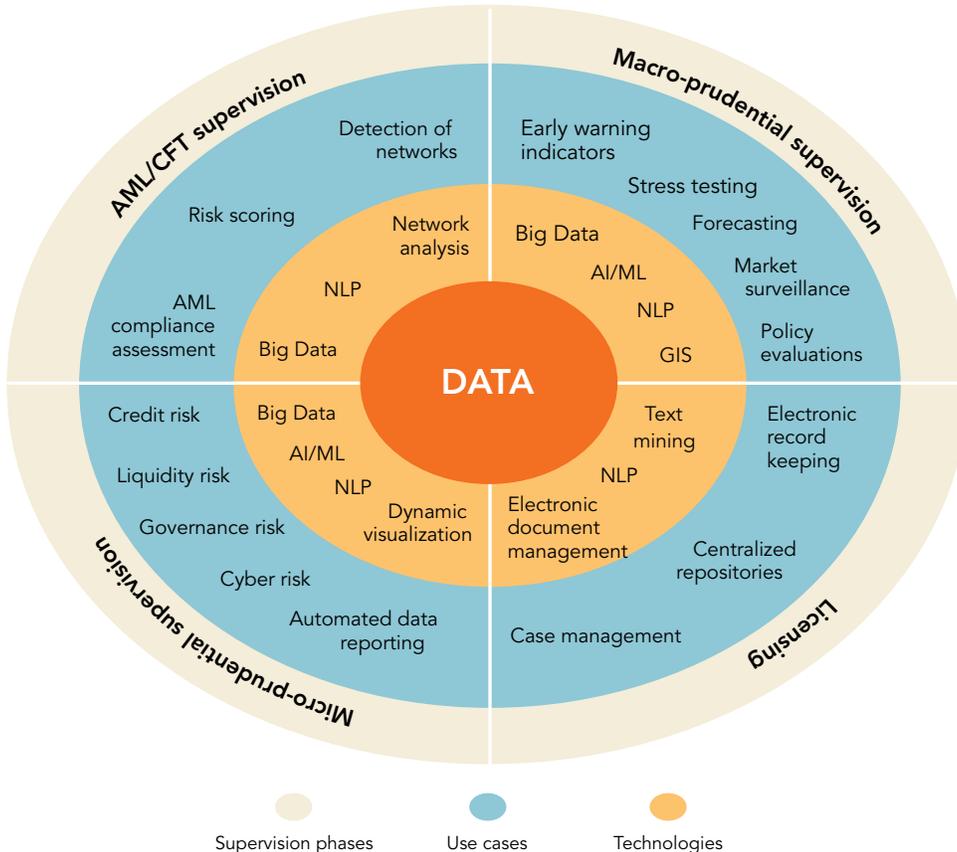
The four main categories of suptech solutions for market conduct supervision are as follows:

1. Solutions for regulatory reporting by supervised institutions
2. Solutions for the collection and processing of complaints data
3. Solutions for non-traditional market monitoring
4. Solutions for document and business analysis

1. Solutions for regulatory reporting by supervised institutions

A primary method for regulators to identify market conduct risks and issues is to collect information directly from supervised institutions. Historically, such submissions have been collected manually through reporting templates submitted by mail, email, or fax, resulting in a slower, inefficient, and more error-prone process.

FIGURE 3. A Function-Based Suptech Taxonomy with Suptech Use Cases



Source: World Bank (2020).

Today, web portals, APIs, ADFs, and comprehensive SIS allow for automated and standardized regulatory reporting that collects, validates, transforms, and stores data in real time. The most sophisticated solutions rely on machine-readable taxonomies, customer relationship management (CRM) systems, and data warehousing with permission-based datamarts. However, a solution need not be overly complex to deliver immense regulatory benefits for market conduct supervisors, including enhanced efficiency and increased analytical capability. Further, applying automated data analytics allows market conduct supervisors to support their supervisory framework and prioritize scarce supervisory resources toward areas of greatest risk.

Suptech solutions for regulatory reporting include the following:

- SIS⁶
- Automated data submission via API
- Web portal data upload with central database

2. Solutions for the collection and processing of complaints data

Complaints data is one of the most valued data sources for market conduct supervisors. Suptech solutions in complaints handling alleviate the operational burden through greater automation. Such solutions can also introduce new front-end digital channels to engage with consumers regarding their complaints and inquiries, such as via websites, mobile apps, text messaging, and chatbots. After initial setup, digital channels tend to be lower in cost to operate, expanding regulators' reach beyond urban areas. In addition, such solutions enhance the quality of the information collected about consumer complaints. Advancements in database management and analysis allow for supervisors to extract more understanding and insight from consumer-submitted complaints via CMSs, providing a critical resource for market conduct supervision.

Solutions for the collection and processing of consumer data include the following:

- CMS⁷
- Analysis of unstructured complaints data

3. Solutions for non-traditional market monitoring

The internet provides the opportunity to utilize a range of new, non-traditional methods for conducting market monitoring, another core activity for market conduct supervision. Monitoring social media, online news, websites, and so on can provide early warning signals of emerging consumer and reputational risks. By keeping a pulse on consumer sentiment in social media and web forums, these solutions provide the potential for more uninhibited, real-time access to the "voice of the consumer" and consumers' experiences with FSPs. Overall, these web monitoring solutions provide a useful, low-cost complement to traditional market monitoring to gather regulatory insights.

Solutions for non-traditional market monitoring include the following:

- Web scraping
- Social media monitoring
- Consumer sentiment analysis
- Reputational analysis
- Dark web monitoring

4. Solutions for document and business analysis

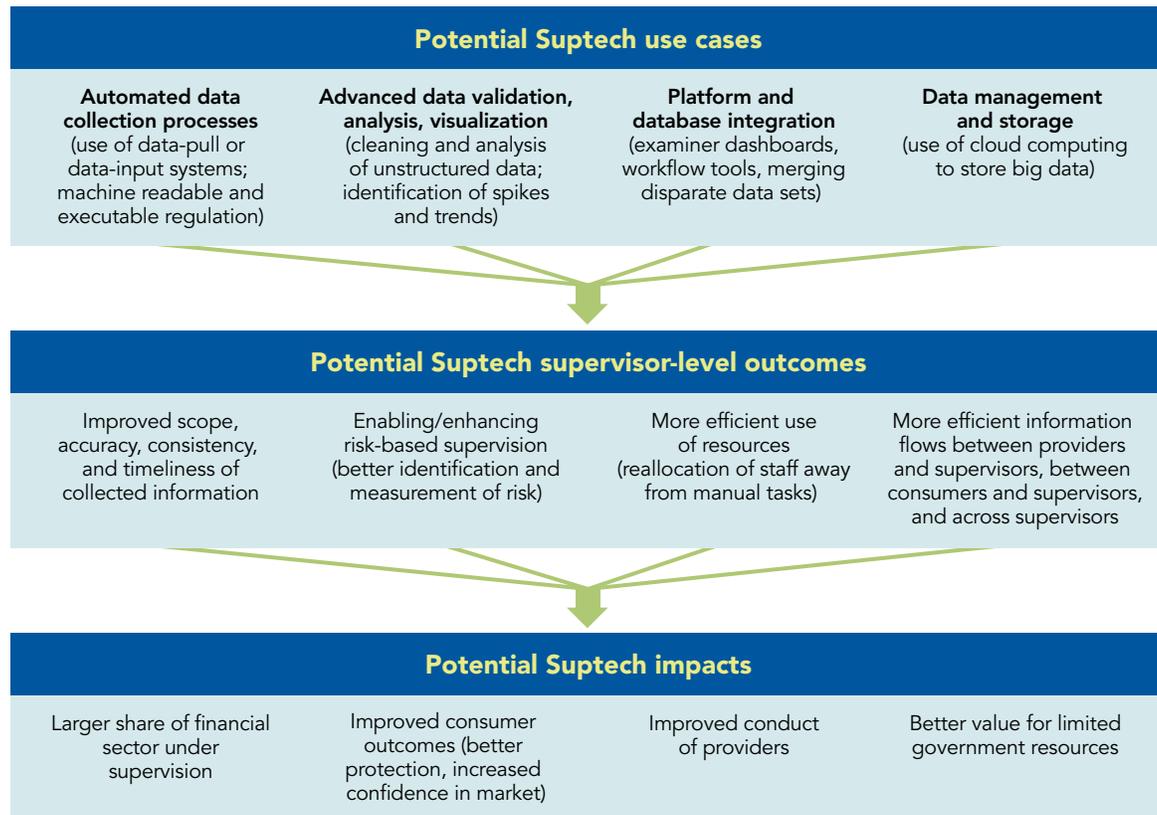
Advances in analytics have been most profound in the realm of unstructured text data. For example, NLP solutions can ingest and analyze large quantities of documents, extracting insights in seconds where previously it would have taken supervisors weeks (if even possible at all). Given the more qualitative nature of market conduct supervision, advancements in the analysis of text present a potentially significant breakthrough.

Suptech solutions for analysis also leverage automation and combine data sets together to produce a more holistic view. Some suptech tools also bring in new types of external data sets that were traditionally difficult to combine for analysis, such as geospatial data. Solutions for advanced analytics provide market conduct supervisors with both significant gains in efficiency and the ability to extract new insights from data to allow for data-driven decision making.

6. Solutions for regulatory reporting also include machine-readable taxonomies, data validation systems, and ad hoc transmission systems.

7. These solutions often include both case management interfaces for supervisory staff and digital user interfaces for consumers.

FIGURE 4: Results Framework for Market Conduct Suptech Solutions



Solutions for document and business analysis include the following:

- Document analysis for regulatory compliance
- Document analysis for examination of FSPs
- Document analysis for peer group comparison
- Validation of terms and conditions
- Automated review of new provider registrations
- Predictive modeling of financial statements
- Business intelligence and geospatial analysis
- Managed data platform

The four main categories of suptech solutions for market conduct supervision represent an update from the Suptech Conceptual Framework first introduced in the 2018 World Bank discussion note (World Bank 2018). As noted above, these suptech solutions drive both efficiency and effectiveness at the supervisor level and ultimately lead to potential beneficial impacts in the broader market, such as via improved consumer outcomes.

3. SUPTECH SOLUTIONS FOR MARKET CONDUCT SUPERVISION

Within the four main categories of supotech solutions for market conduct supervision, 18 individual solutions are described in this chapter. These supotech solutions are currently operational, in pilot, or were expected to be operational in 2020. For each solution, information is provided on what the solution is, the benefits it provides, and considerations for implementing the solution. Solutions are accompanied by country examples and select case studies.

It is worth noting that supotech solutions need not always be particularly “high-tech” or the most complex to have real, significant supervisory benefits. The complexity of supotech solutions varies within each category. What this means practically for financial authorities, especially in low- or middle-income countries, is that authorities have options. Financial authorities can focus on the solution(s) that best matches their needs, available resources, and existing capabilities. Figure 5 summarizes the level of implementation complexity across solutions. Supervisors in lower-capacity countries evaluating potential solutions from this list should first consider adding capabilities in a category(s) for which the authority does not currently have a solution. Once the authority has baseline capabilities within a category, authorities can opportunistically enhance their capabilities by implementing more sophisticated solutions, depending on supervisory need and available resources. As with any investment, authorities should evaluate the solution’s business case in context of supervisory goals and available resources.

Solutions within each category are interrelated and complementary. When viewed together, supotech solutions within each category span the data life cycle related to the specified supervisory activity. Individual solutions may allow authorities to collect new forms of data, introduce new methods for its collection, or conduct new or richer analyses of this information. This is particularly true as it relates to new types of analytics, whose functionality is common across all four categories of supotech solutions but can be employed to serve specific supervisory use cases requiring domain expertise.

In particular, the latest wave of advanced analytical solutions in multiple categories is enabled by NLP. NLP refers broadly to the ability of computers and algorithms

to analyze text and speech data. This includes the ability to infer topics in text, classify and categorize documents, and measure other text characteristics, such as sentiment. Common types of NLP algorithms found within supotech solutions include topic modeling, sentiment analysis, and text summarization. NLP has the advantage of being replicable, systematic, and more transparent, but challenges remain. NLP requires continuous fine-tuning and interpretation for its outputs to be accurate and regularly usable.

3.1. Regulatory Reporting

Data and reports submitted by supervised institutions are among the sources of information used most widely by market conduct supervisors to inform supervisory activities. In addition to market conduct, financial authorities regularly use technology solutions for regulatory reporting to support prudential, financial inclusion, or other goals. Solutions for regulatory reporting vary in their level of complexity and are presented here beginning from the most complex (3.1.1 “Supervision Information Systems”) to less complex (3.1.2 “Automated Data Submission via API”) to least complex (3.1.3 “Web Portal Data Upload with Central Database”).

3.1.1 Supervision Information Systems

SIS represent a comprehensive IT upgrade to the collection, validation, and analytics of reported information from supervised institutions. While the exact technical deployment can vary among authorities, SIS solutions share the following technical elements: ADFs to retrieve data from supervised institutions; a central data warehouse with a CRM system to store, manage, and secure documents and data; “datamarts” to facilitate permission-based access to different teams and departments within the authority; and business intelligence (BI) tools that equip supervisory staff to analyze and monitor data for trends and risks.

The solution’s high complexity requires a significant investment of organizational time and resources. This often includes external consultants and software vendors, in addition to in-house technology staff. Involvement of supervisory staff is also crucial to ensure the solution is designed appropriately to support an authority’s specific supervisory framework. This includes considering definitions of standards and reporting guidelines to supervised entities and the solution’s data validations.

FIGURE 5: Overview of Suptech Solutions for Market Conduct Supervision

CATEGORY	SOLUTION	DESCRIPTION	SUPERVISOR EXAMPLES	IMPLEMENTATION COMPLEXITY & COST ⁸
3.1 Regulatory Reporting	3.1.1 Supervision information systems (SIS)	Comprehensive IT upgrade to the collection, submission, and analytics of FSP reported data	BNR, AMF	Most sophisticated
	3.1.2 Automated data submission via API	FSPs prepare database extracts and share data via consolidated API transmission	BSP	Moderate sophistication
	3.1.3 Web portal data upload with central database	Low-complexity data sharing solution to replace manual data sharing over email, fax, or not at all.	NRB	Foundational capability, inexpensive
3.2 Collection & Processing of Complaints Data	3.2.1 Complaints management systems (CMS)	Automates complaints handling processes, improves data quality, and introduces digital interfaces for consumers and case workers	BOL, CFPB, BSP	Moderate sophistication
	3.2.2 Analysis of unstructured complaints data	Identifies topic, sentiment, and thematic patterns in consumer complaint text	FSD Kenya	Inexpensive, but requires analytics staff
3.3 Non-traditional Market Monitoring	3.3.1 Web scraping	Gathers text data from online sources (e.g., FSP website, social media, web forms, blogs, news)	FCA, AMF, CBI	Foundational capability, inexpensive
	3.3.2 Social media monitoring	Topical analysis of consumer posts on social media related to FSPs or financial products	FCA, AMF, CBI, EIOPA	Inexpensive, but requires analytics staff
	3.3.3 Consumer sentiment analysis	Analysis of consumers' tone and emotions in their interactions with FSPs online	BOE, AMF, CBI	Inexpensive, but requires analytics staff
	3.3.4 Reputational analysis	Analysis of news media's view of specified FSPs	AMF	Inexpensive, but requires analytics staff
	3.3.5 Dark web monitoring	Identify fraud, scam, etc. risks on the dark web	BOE	Moderate sophistication
3.4 Document and Business Analysis	3.4.1 Document analysis for regulatory compliance	Inspects FSP-provided documents to determine compliance with specified regulations	FCA	Inexpensive, but requires analytics staff
	3.4.2 Document analysis for examination of FSPs	Topical analysis of FSP-provided documents to scope and support supervisory examinations	AMF	Inexpensive, but requires analytics staff
	3.4.3 Document analysis for peer group comparison	Analysis of FSP-provided documents to spot risks and trends across a peer group	BOE	Inexpensive, but requires analytics staff
	3.4.4 Validation of terms and conditions	Automation of the review of product terms and conditions to identify compliance risks	BdP	Inexpensive, but requires analytics staff
	3.4.5 Automated review of new provider registrations	Evaluates and identifies new provider or product registrations that are higher-risk	AFM	Inexpensive, but requires analytics staff
	3.4.6 Predictive modeling of financial statements	Evaluates financial statements for misstatement or other risks	AFM	Inexpensive, but requires analytics staff
	3.4.7 Business intelligence (BI) & geo-spatial analysis	Supports analysis and interpretation of data, often a complement to other suptech solutions	AMF, BOE, FCA, NRB, AFM	Ranges from low to high complexity
	3.4.8 Managed data platform	Standardizes, centralizes, and makes accessible internal data from a multitude of sources	AFM	Most sophisticated

8. Implementation costs are from the authors' interpretation of anecdotal information.

The solution's benefits can be substantial. BNR designed its solution, called the Electronic Data Warehouse (EDW), to centralize data from across the authority into a single internal data store for comprehensive analysis, including data from the national payments system, credit reference bureaus, and the statistics department. Autorité des Marchés Financiers (AMF) in Québec, Canada designed its solution to serve as an Offsite Supervision System which streamlines many of the operational, cybersecurity, and data integrity challenges associated with collecting granular data from supervised institutions. Such granular data is typically contained in requests for supervisory information. Like BNR, AMF's solution also centralizes and compiles data sets from across the authority to create a richer, more holistic view to generate insights for data-driven decision making. The supervisory infrastructure to conduct off-site examinations has become increasingly important in 2020, as the logistics of on-site examinations are made more complex (or infeasible) due to the COVID-19 pandemic.

Pulling data directly from supervised institutions

An innovation of SIS solutions, ADFs allow supervisors to "pull" data directly from supervised institutions, rather than having supervised institutions "push" data to the authority. This data pull can be facilitated either through

a direct connection to the IT systems of supervised institutions or, more commonly, through "middleware" that serves as an intermediary between the SIS and the IT systems of supervised institutions. An advantage of middleware is its interoperability with the various types of databases used by supervised institutions (for example, Oracle, SQL, MySQL, and so on). This interoperability allows supervised institutions to continue to use their same provider and connect via the middleware using simple data-transfer protocols. The middleware also adapts data from different types of databases into a common readable format for the SIS. Finally, the middleware also provides supervised institutions with a buffer, as the SIS accesses only data that the supervised institution intentionally makes accessible. In this way, SIS solutions using middleware do not require access to the full database or core banking systems of supervised institutions.

Data analytics and reporting through datamarts

While central warehouses and CRM systems store, manage, and protect the data retrieved from supervised institutions, datamarts are used by supervisory staff to access and analyze the data. Datamarts are typically user-permissioned and facilitate access to the subsets of data within the central repository deemed appropriate based on job role, function, department, or other distinction of the user

CASE STUDY 1

How BNR Designed Its SIS Solution

BNR's EDW is an end-to-end regulatory reporting data platform with both prudential and market conduct applications. It was the culmination of a three-year IT effort from proof of concept to deployment and cost approximately USD 1M to implement. It overhauled previous data-management systems, requiring investments not only in hardware and software at BNR but also (and more importantly) in upgrading staff skills and coordination among the more than 600 institutions that it supervises.

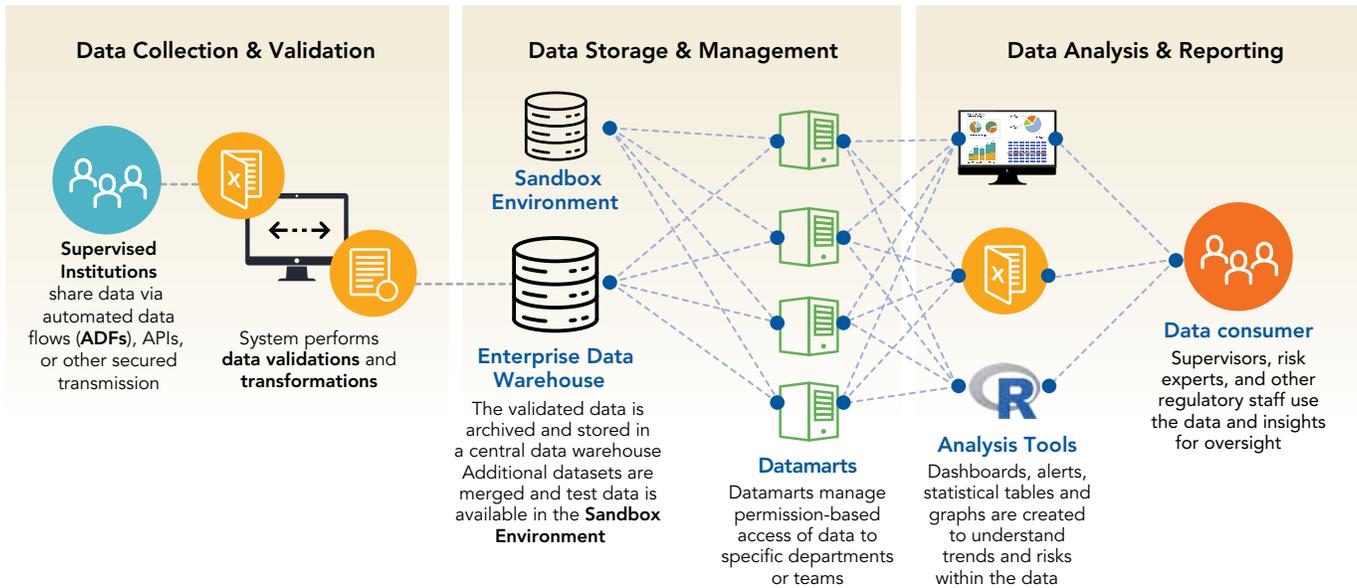
The EDW solution introduced three new dimensions to BNR's regulatory reporting infrastructure: (i) data-pull technology that allows supervisors to connect directly to the databases of FSPs and collect data from the source, rather than sharing data via Excel spreadsheets; (ii) the collection of account-level data that provides more granular data, provided daily, rather than aggregated by institution on a monthly or quarterly basis; and (iii) data analytics and reporting that are now automated and linked to interactive dashboards. Within BNR, the EDW was also designed to break down internal data siloes. As a central data warehouse, it integrates with other internal data sources, such as data from the national payments

system, credit reference bureaus, and the statistics department, among others.

The EDW imposed relatively little additional burden on FSPs. This is a result of its technical design for software interoperability. FSPs can continue using the same database provider (for example, Oracle, SQL, MySQL, and so on) and connect to the EDW using simple data-transfer protocols. Further, management at BNR reports that frequent engagement with FSPs, particularly relating to providers' concerns about the level, nature, and frequency of supervisor's access to their data, was key to its ultimate widespread adoption.

Throughout the three-year initiative, management at BNR indicated the importance of managing change within the financial authority. Supervisory staff accustomed to BNR's data-management processes initially met the changes introduced by the EDW with skepticism. Staff who performed manual data-cleaning or data-consolidation processes had to learn new skills to interact with the more sophisticated system. Many also were retrained to perform business analysis, focusing on the analysis and interpretation of the data (with greater value-add) rather than on such mechanical processes as consolidation and cleaning.

FIGURE 6. Data Flow Diagram of SIS Solutions



Source: Figure adapted from materials provided by the Autorité des Marchés Financiers (Québec, Canada)

requesting access. Datamart interfaces can also help users to link data sets together and produce automated reports.

In designing a SIS solution, supervisors should consider the nature of their supervisory framework. How the data is collected and maintained over time will partly depend on whether the authority takes a risk-based or institution-type focus to oversight.

3.1.2 Automated Data Submission via API

An API acts as a software intermediary that enables two or more systems to talk to each other. For regulatory reporting, supervised institutions can prepare database extracts and share their data with supervisors via API transmission. These data and report transmissions are most valuable in a machine-readable format to minimize the operational burden on supervisory authorities associated with manual processing, data cleaning, and validation and making the data readily available for market conduct analysis.

Direct machine-to-machine transmission via API has several benefits. The raw data extracted from supervised institutions' core banking systems is converted into a single encrypted XML file that is pushed directly to the supervisor. This single unified reporting scheme can replace dozens of previous reports submitted separately

by institutions. Further, this data can be validated in real time, as upward of thousands of validation rules are run in parallel. Together, the process typically averages 10 seconds per submission in the Philippines—a substantial improvement from the 30 minutes or more a submission via web portal upload might take for a supervisor to process and validate (di Castri, Grasser, and Kulenkampf 2020b). For supervisors, the solution reduces staff time spent on processing and managing data. This is especially true of the time spent on cross-validations, which grows as the number of items requiring reconciliation grows with every new report.

This solution provides benefits for supervised institutions as well, reducing reporting burden and compliance costs. In the case of Bangko Sentral ng Pilipinas (BSP) in collaboration with the RegTech for Regulators Accelerator (R²A),⁹ the number of reporting fields required of FSPs was cut in half, from 107,000 to 50,000, as duplicated or calculation fields were eliminated. Further, this consolidation allowed for the retirement of older reporting templates in the move to automated database extracts (in XSD format).

The desirability and feasibility of this solution is likely to vary among market conduct supervisors in low- and middle-income economies. Countries with larger digital finance

9. The RegTech for Regulators Accelerator, launched with support from the Bill and Melinda Gates Foundation, the Omidyar Network, and USAID, partners with financial sector authorities and technology firms to accelerate innovation in financial sector supervision, regulation, and policy analysis. See <https://www.r2accelerator.org/about>.

ecosystems are likely to have supervised institutions with relatively advanced levels of technical capability to support this solution. Countries with a varied ecosystem of supervised institutions, including those with more advanced and more limited technical capability, will likely need to maintain simpler means of regulatory reporting, such as reporting via a web portal, alongside more advanced solutions, such as API submission where feasible.

3.1.3 Web Portal Data Upload with Central Database

Regulatory reporting via web portal represents a low-complexity solution that can replace manual data sharing by email, fax, or mail. This solution can be developed relatively quickly and inexpensively because of the many off-the-shelf software solutions that have become commercially available.

A secure web portal allows for manual data entry, data upload via widely accepted formats like XBRL, CVS, or XML, or server-to-server transmission in more advanced instances between supervised institutions and the financial authority. The range of data-sharing methods is designed to accommodate supervised institutions that are likely to vary in their technical capabilities. The data

fields, formats, and frequency of reporting are prescribed by the supervisor depending on their needs.

Regulatory reporting via web portal provides many benefits for market conduct, though it is also helpful to understand its limits. The solution increases the efficiency of data collection, particularly beneficial where market conduct supervisors have limited capacity and must oversee a vastly greater number of institutions than prudential supervisors. In addition, web portals typically collect standardized reports at regular intervals. This data provides supervisors with a regularly updated, landscape-wide view that is crucial for market intelligence and informing a risk-based supervisory framework. However, these reports tend to be less helpful when supervisors need to investigate specific questions or risks requiring more detailed data in a follow-up information request.

This solution is used by financial authorities such as Nepal Rastra Bank (NRB), which standardized its reporting templates and introduced regulatory reporting via web portal and data upload in 2016. NRB now has plans to upgrade to a more comprehensive SIS that incorporates API-based submissions via XRB in 2021.

CASE STUDY 2

How NRB Develops Suptech Solutions

Financial authorities without a large IT department or dedicated innovation office may benefit from starting small, proving the concept, and making continuous technology improvements. At NRB, reporting involves uploading data via a web portal. This solution has been part of a natural technology evolution that has gradually leveraged increasing levels of technology and data to regulatory oversight. What started out with hard-copy submissions by FSPs prior to 2009 has evolved with the availability of new technologies. Over the past 10 years, fax, email, and a first version of its web portal were each deployed and retired. In 2016, NRB launched its current web portal.

NRB's web portal was designed to support the varied technical abilities of the country's FSPs. A web portal does not require supervised institutions to install or manage any new reporting or software in their systems (or to have any specialized systems). This avoids complex system integrations, the need to have in-house technical talent, or even a Microsoft Excel license, as users can upload or enter data directly into the web portal. Web portals, however, do require digital connectivity, which is a challenge for FSPs operating in rural regions.

The web portal acts as a single point of data submission. To achieve this, the solution's design required up-front work by NRB staff to consolidate its many different report types from throughout the authority into standard reporting templates with a consistent format. Consequently, a single set of validations (some of which are automated in the portal itself) are performed at the point of submission before the data is stored centrally and made available on a permissioned basis throughout the authority.

Digitizing its regulatory reporting infrastructure has streamlined how and what data NRB makes publicly available, including granular data on credit, deposits, branches, financial access, and financial inclusion. Sharing raw data and reports publicly has a multitude of benefits, including market intelligence for FSPs and overall intelligence on the state of financial stability, financial inclusion, and market conduct in Nepal.

For NRB, its technological evolution continues today. In 2021, NRB intends to launch the next generation of regulatory reporting—a SIS that will introduce enhancements such as API-based data collection and a stronger capability for business intelligence (BI) analysis.

3.2. Collection and Processing of Complaints Data

Consumer complaints data is one of the most valued data sources for market conduct supervisors. A CMS is key to the efficient processing of these complaints and capturing and managing complaints data to maximize its accuracy and value for supervisory purposes (see section 3.2.1). The application of advanced analytics to complaints data (see section 3.2.2), particularly to unstructured text, represents the next step for market conduct supervisors to deduce useful insights from complaints data in a more efficient manner.

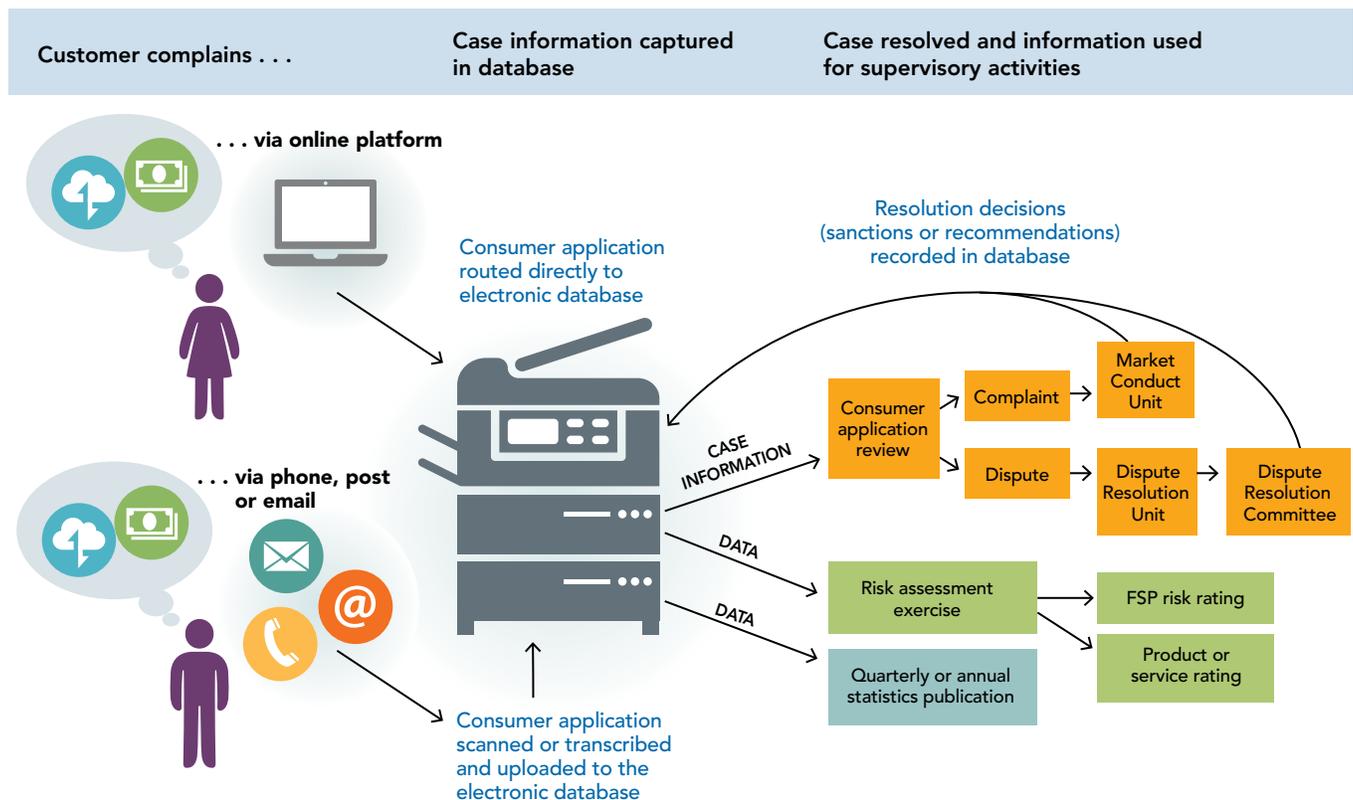
3.2.1 Complaints Management System

Traditional complaints handling often relies on manual processing and offers limited ability to interact with consumers, a result of outdated communication channels and accessibility issues. Complaints data may also suffer from the entry of incomplete, inaccurate, or inconsistent information that limits its value for supervisory purposes. This heavy reliance on manual complaints processing is both operationally intensive and prone to data-entry error. As a result, complaints departments at financial authorities are often overburdened and are not equipped to provide valuable data that can inform supervisory activities.

A CMS expands visibility and access to the public for submitting a complaint, introduces automation into complaints-handling processes, and optimizes data management for supervision. Complaints submission can be made further accessible to consumers through the introduction of new digital user interfaces such as SMS. More critically for market conduct supervisory purposes, centralizing manual submissions through a single “case manager” digital interface streamlines complaints submitted through other channels and helps to structure and categorize complaints data for supervisory analysis.

In 2020, BSP is expected to deploy its CMS solution in coordination with R²A and Sinitic, a software vendor. BSP’s CMS will include new digital interfaces for staff (called a case manager) and consumers (digital SMS submissions via API and chatbot) and a more robust complaints database. The CMSs at the US Consumer Financial Protection Bureau (CFPB) and Bank of Lithuania (BOL) were previously highlighted in the 2018 World Bank report *From Spreadsheets to Suptech*. The CFPB goes further than typical in sharing its complaints data publicly through an external interface that redacts personal information and is accessible to all. The CFPB also makes its complaint database available to other government agencies through a private portal.

FIGURE 7: CMS Case Workflow and Data Architecture



Source: World Bank (2018)

Standardized complaints capture

Modern CMS solutions introduce new digital channels for submitting complaints. For example, BSP's CMS is enabled by API and an NLP-powered text engine through which consumers can submit complaints using either smart or feature phones on a variety of messaging platforms, including Facebook Messenger, SMS, or a chatbot embedded on an FSP's website. Through each of these digital channels, an NLP-powered engine interprets and responds to consumer messages, leading consumers through a pre-defined conversation and complaints submission experience. In the background, a supervised machine-learning model identifies opportunities to improve the model's understanding and ability to interpret consumer messages accurately. CMS solutions employed at BSP, CFPB, and BOL also consolidate manual channels that require higher-touch human interaction, such as phone calls, mail, or emails, into a single standardized complaint-management interface. This provides operational benefits in the form of more streamlined human interaction, as well as the data integrity benefits of a single interface for intake.

Database storage

Incoming complaints data from new digital channels flows into a centralized complaints database. Past complaints data is also imported into the centralized database to deepen the historical record for analytics and to train machine-learning models for purposes of topical modeling and sentiment analysis.

Complaints processing via case manager interface

Supervisory staff can view analytics, configure complaints intake (such as chatbot logic in the case of BSP), and enter and process individual complaints, including complaint notes or attachments, from a central user interface. Further, the CMS automates the routing of complaints to the appropriate supervised entity for response with a portal for financial providers to manage their complaints.

The introduction of new, easily accessible digital channels also encourages more consumers to contribute their voice and enables the regulator to provide timely information on financial services to consumers, contributing to the enhancement of consumers' ability to make more informed choices. This ultimately reinforces the solution's effectiveness and relevance.

Analysis of complaints data

The process automation introduced by CMS solutions enables financial authorities to shift staff resources from time-consuming complaints processing to complaints analysis for supervisory purposes. Furthermore, CMS

solutions improve the quality of complaints data, aiding supervisory analyses and increasing the effectiveness of market monitoring. Analyses of complaints data allow supervisors to understand consumers' experiences with financial providers and identify emerging trends and risks.

In the case of BSP, the supervisor expects that new types of complaints analyses, such as topic modeling, will reveal previously hidden patterns of consumer and firm behavior to add to their market monitoring. This is particularly true as the CMS begins adding more complaints submissions from outside of the capital region of Manila, where the bulk of previous consumer complaints came from. New digital submission methods are expected to help the regulator reach harder-to-reach cities and rural regions of the country.

3.2.2 Analysis of Unstructured Complaints Data

As noted above, consumer complaints represent a rich data source for market conduct supervisors. However, such data is often unstructured, making it difficult for individual supervisors to identify patterns or emerging risks. This is particularly true where a consumer's description of a complaint comes in the form of a free-form narrative and in countries where complaints reporting is not standardized. Due to the large volume of data that results, supervisors often decide to sample random complaints or purposely select those that seem to pose higher market conduct risk. However, this can be a time-consuming process and result in missing key risks. Other common methods, such as keyword searches, can become biased if a systematic method is not followed.

NLP algorithms for topic, sentiment, and risk identification can help financial authorities automate analysis of unstructured complaints data and conduct such analyses on high volumes of data with greater accuracy. "Topic modeling" infers topics from raw text by analyzing word co-occurrence in the text of each complaint to identify the topics and proportion of topics within each. Co-occurrence is a measure of the frequency with which two topics or phrases are both present in a complaint. Topic models apply an inductive approach of inferring, rather than assuming, patterns in the text. It can also include metadata in the analysis and thereby link external variables such as industry, timing, or demographics to fine-tune the analysis.

The overall implementation complexity and investment level of these types of suptech solutions are generally low. However, these solutions require supervisors to have in-house expertise to conduct and maintain these analyses; this requirement may pose a limiting factor for financial authorities in lower-capacity countries.

CASE STUDY 3**How Researchers at Princeton University and FSD Kenya Worked with the Central Bank of Kenya to Analyze Complaints Data**

Researchers at Princeton University in partnership with FSD Kenya and the Central Bank of Kenya analyzed 37,000 consumer complaints, demonstrating how supervisors can apply these analytical tools to enhance their understanding of their own consumer complaints data.

The analysis consisted of the following five steps:

- i. **Data cleaning:** Ensures high-quality data for analysis. It can be time consuming, especially if the raw data is untidy and has not been used for quantitative analysis before.
- ii. **Algorithm selection:** Once the data and metadata are structured in tabular form, they can be fed into the algorithm. For this analysis, researchers used Structural Topic Modeling, a specific kind of topic-modeling algorithm that is open source and available in R—one of the most well-known free software environments for statistical computing and graphics—in a freely available package called “stm.” In order to run the algorithm, data scientists must make a number of initial arbitrary choices, the most important being the selection of k —the number of topics that the algorithm will identify in its first iteration.
- iii. **Model optimization:** Optimizing the algorithm to maximize accuracy requires an iterative process to ensure that the topics are relevant and not skewed by irrelevant factors.

iv. **Interpretation:** The most crucial step comes toward the end of the analytics process. Interpreting the findings requires a qualitative reading of the topics identified by the algorithm, often aided by visual tools and dashboards available in the software. Here, analysts identify the findings most pertinent to supervision and dig deeper where needed.

v. **Analysis and documentation:** As the NLP model and its outputs are finalized, the last step is to document patterns and relationships in the complaints data.

Once the final model is specified and analysts are confident with its results, the model can be used repeatedly for daily or weekly monitoring of incoming complaints from consumers. In Kenya, researchers identified three key learnings from the interpretation of the model results.

First, not all topics identified by the algorithm are useful. In fact, of the 60 most common topics, only 14 provided clear added value to the regulators, while an additional 12 topics provided potential added value.

Second, topic modeling is particularly well suited to identify emerging risks and unknown problems in the market. Confirming the capacity to discover topics without prior specifications or assumptions is one of the key strengths of topic modeling.

Lastly, topic modeling can help identify patterns in risks that are already known. By associating topics to the metadata, analysts discovered how complaints varied by bank or by time period, and whether there was a difference between complaints categorized as “open” or “closed.”

3.3. Non-traditional Market Monitoring

The internet provides the opportunity for supervisors to expand market monitoring activities to new, important pools of data in a more efficient and effective manner. The solutions described in this category span the data life cycle—enabling supervisors to collect the data contained within new, non-traditional text data, such as social media, blogs, forums, and news media, and apply new, advanced analytics methods. Foundational to all solutions in this category is web scraping (section 3.3.1), which provides the mechanism to collect online text data for analysis. Collecting data then allows for various forms of market monitoring analyses, such as social media monitoring (section 3.3.2) and consumer sentiment analysis (section 3.3.3), reputational analysis (section 3.3.4), and dark web monitoring (section 3.3.5). Overall, these solutions strengthen market conduct supervision and provide

a timely, useful, and low-cost complement to traditional market monitoring, such as complaints handling, particularly for resource-constrained financial authorities.

3.3.1 Web Scraping

Information on websites presents market conduct supervisors with a valuable, non-traditional data source. The primary method to gather information on websites is via web scraping, which pulls text, metadata, and other information from the websites of FSPs, or consumer posts shared on social media websites, online forums, and review websites. Web scraping allows supervisors to view and aggregate this information to identify emerging risks to consumers, such as how financial products are marketed or sold. Note that the actual analysis of web-scraped data is described in sections 3.3.2–3.3.5.

Prior to the advent of web scraping, market conduct supervisors would have to undertake this type of analysis manually via web-related search tools, such as Google, or manual review of individual websites. Besides being time intensive and inefficient, manual procedures proved to be quite ineffective for supervisory purposes given the practical difficulty of evaluating the whole market and spotting outliers.

Variations of this solution are currently in use or expected to be deployed by such authorities as the FCA, AMF, and the Central Bank of Ireland (CBI). While some authorities have built this tool in-house using APIs provided by Google and other third-party providers including social media companies, the predominant approach among authorities has been to engage an external software or data provider. Financial authorities that collect and store consumer data, such as social media data, cite anonymizing or redacting personal and sensitive information when web scraping as a best practice.

3.3.2 Social Media Monitoring

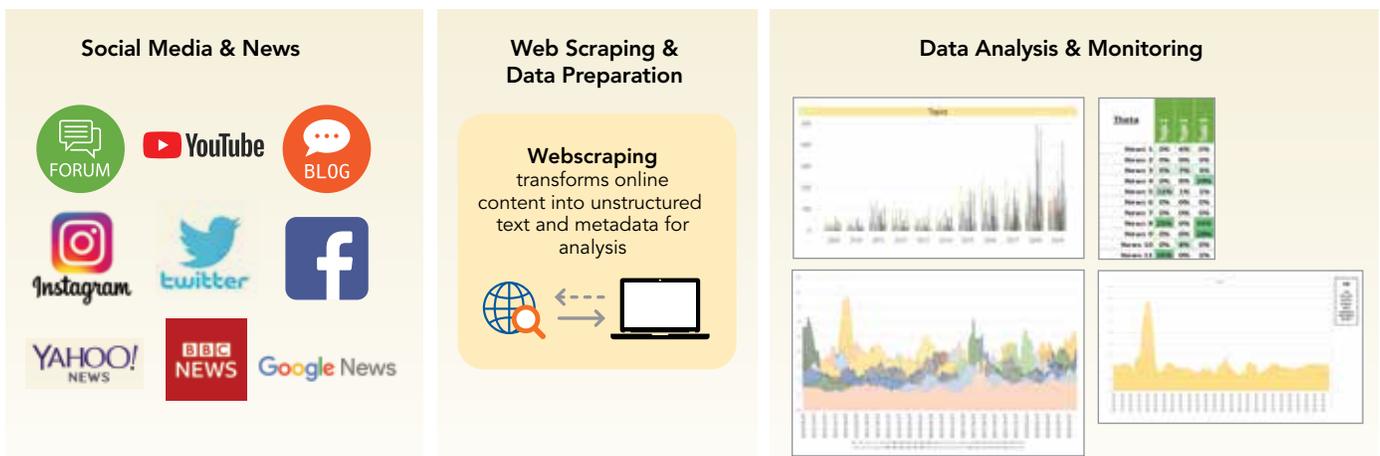
Social media monitoring is the most common solution employed for non-traditional market monitoring. Supervisors use this solution to attempt to listen directly to the “voice of the consumer” and understand consumers’ interactions with financial providers on a nearly real-time basis. This enables supervisors to identify and act on emerging consumer risks more quickly. As more and more commercial solutions have become available at competitive prices, more financial authorities have piloted or adopted such solutions.

Social media monitoring solutions analyze consumer posts collected via web scraping (see section 3.3.1) from web forums, social media websites such as Facebook, Twitter, and Instagram, and consumer-review websites. Consumer posts are categorized via keyword mining or topic modeling, tagging posts by such criteria as the associated FSP’s name, financial product (for example, insurance, mortgage, and so on), or stage in the product life cycle (opening an account, closing an account, and so on). Posts can also be categorized with a consumer sentiment score (see section 3.3.3).

Developing an appropriate topical categorization requires a taxonomy, which authorities indicated requires supervisory expertise and often takes three to six months to develop, refine, and measure. Even with additional refinement, some level of categorization errors will remain. Human review is recommended to correct for errors and make refinements to the solution to further enhance its accuracy.

Social media monitoring can provide supervisors with an early warning system. For example, the CBI discovered through social media that a new retail credit firm appeared to be operating without authorization in the Irish market. Investigation by the CBI determined that the firm was operating using disposable email addresses and fake names and addresses. The CBI moved to protect consumers by issuing a warning and publishing the name of the firm. Following these actions, no further activity involving this firm was seen in the market.

FIGURE 8: Dataflow Diagram for Social Media Monitoring



Source: Graphic developed by the World Bank based on research with the Autorité des Marchés Financiers, the Financial Conduct Authority, and the Central Bank of Ireland

In another example, the CBI became aware through social media monitoring of a sudden and substantial increase in the number of customer comments to one retail bank. Further investigation revealed that the complaints related to the availability of customer support services, with customers expressing their concern with having to wait up to 40 minutes to get through to a customer service representative. Detailed market intelligence and underlying data about call waiting times were used by CBI supervisors to confirm indications from social media monitoring, and CBI's concern was then relayed to the firm's senior management. Based on this information, the firm recruited additional customer service staff to address the lengthy call waiting times and also agreed to revisit staff scheduling to increase coverage at peak times.

Social media data can also be triangulated with the market conduct supervisor's own complaints data. For example, the CFPB began hearing of consumer issues with transaction failures at a prepaid card provider, RushCard, on social media in 2014, which was subsequently backed up by complaints reported directly to the regulator. The CFPB used social media data in tandem with complaints data as an early warning, allowing the regulator to position itself for action even before the company itself reported the outage issue to its customers or regulators.

However, though social media monitoring is a powerful new addition to the supervisory toolbox, it is important to be aware of its inherent limitations as a data source. First, social media data is self-reported by consumers. People on various social media platforms may differ from members of the general population in their education, wealth, age, financial literacy, and so on. Social media posts should not be taken as representative of the full population, particularly the most vulnerable segments.

Second, social media data is not statistically representative. For example, there may be significantly more social media posts regarding consumers' experiences with their banks than with their insurance providers because consumers interact with their banks regularly but may engage with their insurance providers only infrequently. A qualitative lens is needed to interpret and understand the voice and nature of consumer comments to generate useful insights for supervisory purposes.

Third, consumers provide varying details on social media, as there is no standard reporting template. Though a supervisor may want to understand consumer posts specific to a type of product, consumers tend to be quite general. For example, a consumer may complain generally about "mortgages," rather than about a specific mortgage product, such as interest-only mortgages, that may be of particular supervisory concern.

Lastly, supervisors should be aware of effects of language differences. Algorithms tend to be trained to analyze web content in a single language at a time. This represents a challenge in jurisdictions with multiple languages. For example, the European Insurance and Occupational Pensions Authority (EIOPA) oversees insurance regulation across the 27-member European Union (EU). In its own social media pilot, the authority found it difficult to draw conclusions due to variations in cultural norms and languages across the EU.

3.3.3 Consumer Sentiment Analysis

As more people have taken to expressing their opinions through social media, blogs, and online forums, market conduct supervisors now have timely access to consumer sentiment toward FSPs. This solution is often used to complement the trend and topical analyses described in section 3.3.2. Whereas complaints generally focus on negative experiences, this solution includes a broader window into both positive and negative experiences with FSPs. Further, this solution allows supervisors to hear from a wider universe of people than complaints data alone. This enables market conduct supervisors to enhance their understanding of consumers' experiences with financial providers.

Consumer sentiment analysis solutions monitor language sentiment toward FSPs across social media mentions, blogs, and online forums in a rapid, automated manner by using algorithms. The solution benefits from the use of open-source NLP libraries that detect and assign either a positive or negative tonal score to online posts. More advanced NLP libraries can be applied to identify more nuanced emotional cues. A consolidated report then benchmarks institutions over time and can compare peer institutions against each other.

Financial authorities such as the Bank of England (BOE) use such solutions to complement current supervisory analyses, rather than as a substitute for other data. It is also important to note that consumer sentiment analysis solutions share similar challenges as other NLP-derived solutions—principally, the misidentification of emotional tone, such as sarcasm. To date, sentiment algorithms are more often trained on an individual word basis, whereas jokes, humor, and sarcasm are embedded in sentence structure. As a result, human oversight and some level of manual review is recommended to correct for any miscategorizations.

3.3.4 Reputational Analysis

Given limited staff, financial authorities do not have the capacity to monitor the entire online ecosystem for changes in the financial marketplace. At the same time,

early detection of reputational or consumer risks in the marketplace would indicate where market conduct supervisors should focus their efforts. Reputational analysis tools are beneficial in providing real-time market intelligence on public perceptions of FSPs in the news media. Supervisors may look at insights from reputational analysis in combination with other market intelligence data sets, such as complaints data, to understand overall perceptions of a specific FSP.

Reputational analysis solutions analyze text from press releases and mentions in the financial news. Topic modeling or keyword mining is employed to categorize topical trends and assign sentiment scores to the text data (see section 3.3.3). Aggregated reports enable supervisors to understand changes in the topics discussed by financial institutions over time and compared to peer institutions, as well as to track shifts in market sentiment toward FSPs overall.

The reputational analysis solution is powered through topic modeling, which requires both up-front and ongoing refinement to maximize its accuracy. As noted previously, financial authorities report that setup can be a labor-intensive process lasting between three and six months to enable the tool to classify and analyze media mentions properly. The use of primarily open-source libraries in R and Python, such as the “lda” package for topic modeling, makes this solution more accessible to authorities and positions authorities to benefit from future developments and improvements in analysis software. This solution is typically used together with sentiment analysis. (See section 3.3.3.)

3.3.5 Dark Web Monitoring

Monitoring tools focused on the dark web can detect suspicious activity, identify risks, and enable proactive defense of the financial authority, FSPs, or consumers against threats posed by bad actors. For market conduct supervisors specifically, dark web monitoring can be beneficial in detecting identity theft, fraud, scams, or other activities that can lead to consumer harm.

The dark web refers broadly to an area of the internet where websites, networks, and content exist on private encrypted networks called “darknets” that are accessible only with specific browser software, such as Tor or I2P. Users can communicate and conduct business on the dark web anonymously. Its anonymous nature makes it attractive for use by bad actors. This particular suptech solution allows supervisors to curate intelligence reports and search both “open” and “closed” sources. “Open source” refers to information or derived information that is available to the general public and includes public

records and contact information. “Closed source” refers to information with restricted or private access. The solution presents topical information and sets up real-time alerts when new, relevant information appears.

While monitoring the dark web can help supervisors to identify emerging consumer risks proactively, it is important to note that the dark web is not the only place online where criminal activity occurs. Criminal activity can also happen through private transactions in closed networks that do not appear on the dark web or can occur outright in the public eye. This solution simply helps supervisors to monitor another potential source of harm to consumers.

While services for dark web monitoring are commercially available, many of these service providers are focused on servicing consumers and FSPs, rather than financial authorities. Partly as a result, BOE is piloting its own suptech solution for dark web monitoring.

3.4. Document and Business Analysis

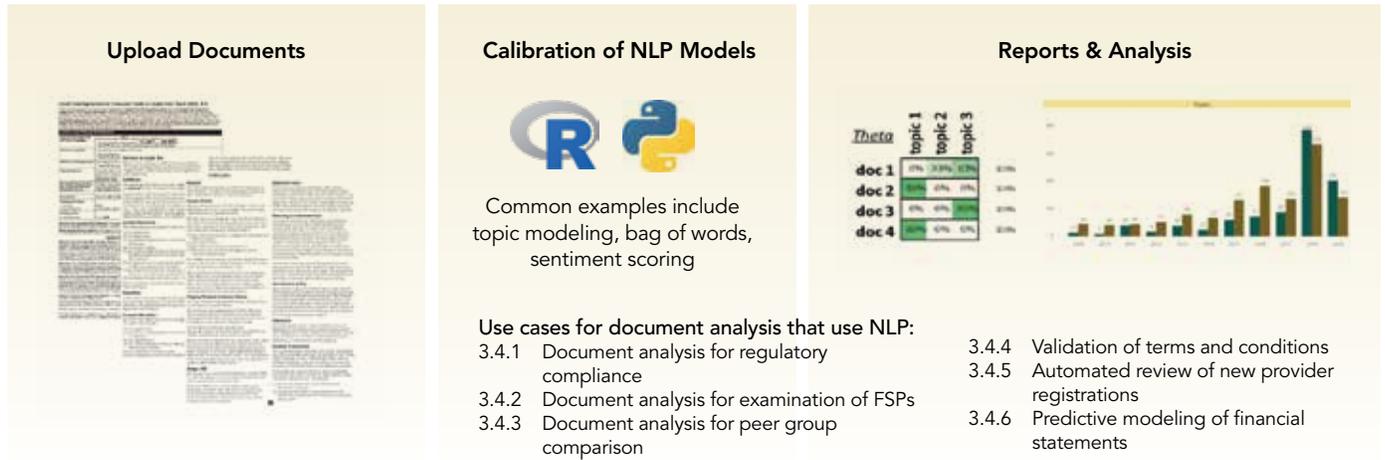
Market conduct supervisors have introduced a range of cutting-edge suptech solutions that focus specifically on leveraging NLP for document analysis. Many of these solutions share common NLP capabilities. For financial authorities, this means that the development of one of these solutions can assist in the development of others. NLP applied to document analysis can be used for a variety of market conduct use cases, including for regulatory compliance (section 3.4.1), institution monitoring (section 3.4.2), peer institution analysis (section 3.4.3), credit agreement reviews (section 3.4.4), new provider registrations (section 3.4.5), and public financial statements and investor report filings (section 3.4.6). In addition, business intelligence (section 3.4.7) and managed data platforms (section 3.4.8) represent suptech solutions that help supervisors manage and combine the data sets created by other solutions and to use the combined data for advanced supervisory analyses.

3.4.1 Document Analysis for Regulatory Compliance

Many financial authorities receive thousands of documents collected from FSPs, such as minutes of board meetings, internal audit reports, and other management reports. These documents contain key information for market conduct supervisors related to audit and compliance risks.

With this solution, supervisors can upload a set of documents and run a series of core analyses on document text. This solution does not aim to automate analyses and conclusions. Rather, it allows triaging of a large number of documents in a much faster manner in order to find relevant information, which can then be analyzed in greater

FIGURE 9: Example of Dataflow Diagram in Document Analysis Solutions



Source: Graphic developed by the World Bank based on research across multiple financial authorities

CASE STUDY 4

The FCA’s Development of Sleuth, Its NLP Platform

With the increasing power and accessibility of NLP to analyze text and speech data, NLP-enabled solutions are growing in popularity among market conduct supervisors. NLP platforms represent a new piece of IT infrastructure that is useful in developing such suptech solutions.

The FCA’s RegTech and Advanced Analytics team has received an increasing number of internal requests to assist supervisors in their analysis of the large volumes of documents received from supervised institutions and from publicly available data sources.

To meet the increasing supervisory need to analyze documents, the FCA developed its NLP platform, called Sleuth. This platform represents a common set of advanced IT and analytics infrastructure in cloud computing, database man-

agement, and NLP. Supervisors can upload a set of documents and run a series of core analyses. These analyses help with both understanding and navigation. Supervisors can understand common themes across and within documents, as well as where to navigate within the documents to learn more.

Sleuth was designed to scale and streamline subsequent investments in new, NLP-enabled suptech solutions by developing its platform to be able to support and be interoperable with future solutions. As a result, the resources required to implement subsequent NLP-enabled suptech solutions tend to be lower, given economies of scale, and faster to develop. The NLP platform also equips internal technology staff with the tools to develop their own tests and pilots of future suptech solutions.

detail for regulatory compliance purposes. The solution provides supervisors with the ability to determine the frequency and location of key search terms, as well as to make topical comparisons across documents.

While this solution facilitates the inspection of documents to determine compliance with specified regulations, it requires both legal and supervisory skills to determine the appropriate search terms to use in order to utilize the solution effectively. For example, in its use of a suptech solution for document analysis, the FCA drew search

terms directly from the EU General Data Protection Regulation (GDPR)¹⁰ and EU regulations.

3.4.2 Document Analysis for Examination of FSPs

Financial authorities receive a broad array of documents from supervised institutions, covering topics ranging from business strategy and financial performance to risk management. However, capacity is limited for staff to be able to review this high volume of documents in their entirety and to identify the most pertinent institution-specific patterns.

10. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation).

This solution can be conceptualized as a powerful “Ctrl-F” tool to find certain words and phrases, powered by two types of NLP algorithms: (1) “bag of words” and (2) “topic models.” “Bags of words” allow for “smart searches” to identify the location and frequency of words, phrases, or groupings of words within the document text. “Topic models” go beyond this and use pre-defined topic filters to detect patterns and prevalent themes within or across documents. This solution is designed to be both inductive and deductive—that is, supervisors can both inspect documents for a pre-determined theme, while the solution is also able to review and indicate back to the supervisor the key themes contained within a document.

While similar in functionality to the solution described in “Document Analysis for Regulatory Compliance” (section 3.4.1), this particular solution is broader in its scope and capabilities. Where the previous solution is focused on identifying regulatory compliance with a specific regulation, this solution inspects a broader range of documents to help identify institution-specific trends beyond those related to regulatory compliance. This is helpful for supervisors as they receive documents asked for as part of information requests sent to supervised institutions and to develop the scope of these examinations. When applied across institutions in a methodical way, this solution provides helpful benchmarking as described in the following section, “Document Analysis for Peer Group Comparison” (section 3.4.3).

For example, AMF implemented its document analysis solution to support the monitoring and scoping of examinations. Manual review of documents received from FSPs would consume considerable staff time. Further, due to the multitude of people involved in the document review process, comparisons across documents reviewed by different people were challenging. As a result of implementing their document analysis solution, supervisors can now efficiently extract insights and data that had previously remained contained within these documents.

This solution is inexpensive to develop using open-source R or Python software and libraries. However, it does require underlying IT infrastructure in database management for the documents and text data. It is also recommended that NLP algorithms be updated periodically, at least once per year.

3.4.3 Document Analysis for Peer Group Comparison

Much of the documentary information that supervisors receive is unstructured, contained in board presentations, financial reports, internal audit reports, meeting minutes, and other reports and documents. Due to its unstructured nature and the high volume of documents received, timely

comparisons across FSPs are challenging. Often, making comparisons requires an intensive process in which supervisory staff look through hundreds of pages of documents belonging to multiple institutions to spot risks and trends across a peer group.

The peer group comparison solution specifically helps market conduct supervisors analyze and compare documentary information across FSPs more efficiently. The solution’s core functionality is similar to solutions described in “Document Analysis for Regulatory Compliance” (section 3.4.1) and “Document Analysis for Examination of FSPs” (section 3.4.2), as it also uses NLP to analyze documentary text.

However, this solution’s capabilities go further in three areas. First, through topic modeling, the solution allows for thematic comparisons across peer groups and the ability to confirm the presence or absence of key concepts. This includes the ability to classify documents according to type, language, and other criteria. Second, the solution can detect unstructured text that is numerical or financial, such as tables, graphs, charts, or images, and export this data in a single machine-readable CSV or XLS file for enhanced quantitative analysis. Third, the solution can apply sentiment analysis to the documents and assign a sentiment score to the document text. These insights are presented through an interactive dashboard that permits supervisors to understand emerging themes, identify trends at the firm level, and conduct a cross-sectoral comparison across a peer group.

For example, the Data Innovation Team within BOE developed its solution to help supervisors analyze and compare unstructured information more efficiently. The new suptech solution has allowed supervisors at BOE to shift to more strategic analysis of the data and is currently being used by approximately 400 supervisors. The solution was the culmination of a four- to six-month proof of concept followed by a year of supplier selection and contracting. A key challenge faced by BOE was navigating legal constraints on sharing data with third parties when testing and selecting the right solution and supplier.

3.4.4 Validation of Terms and Conditions

The timely review of credit agreements, including terms and conditions in particular, can be challenging given limited supervisory staff and the volume of lengthy and ever-changing agreements to review. This solution introduces automation to the validation of credit agreements and their terms and conditions to identify compliance risks by leveraging NLP algorithms to analyze the text of the credit agreements and compare the text to legal and regulatory requirements. This solution could also be applied

to verify whether other consumer disclosure documents comply with requirements, as well as for other types of financial products and services.

For example, Banco de Portugal (BdP) receives over 300 credit agreement templates annually related to new products introduced in the market, and over 1,000 changes to credit agreement templates for existing products. Such a large volume of credit agreements poses challenges in the allocation of supervisory resources sufficient to manually review the agreements for compliance with market conduct rules. In 2019, BdP piloted the use of supertech solutions focused on the validation of 12 compliance rules of relatively low complexity for 20 types of credit agreement templates. The subset of rules covered by the solution represent around 20 percent of the total number of compliance rules for these products. BdP initially partnered with an external vendor that made its platform available, where the rules to be validated were designated and then applied to the sample of credit agreements either via the platform or by exposing APIs to be invoked by other applications.

This solution to validate terms and conditions comes with certain challenges. A lack of standardization in the language or format of agreements used by credit providers makes the solution more difficult to apply. Similarly, the addition of more complex legal and compliance checks on credit agreements will require higher levels of investment in both time and budget to define and refine the NLP algorithms appropriately.

3.4.5 Automated Review of New Provider Registrations

Financial authorities, and market conduct supervisors specifically, commonly license or register new financial institutions, such as investment funds. The registration or licensing process is often characterized by several manual steps.

This solution reviews and signals any risks associated with fund registrations. Automation is enabled through detection and classification functions, both types of NLP algorithms. A detection function reviews documents and information received from the fund registration. Classification functions then undertake an automated check and provide either approval or specific advice for further testing. These checks are based on predetermined decision and legal rules established by the user, in this case, the market conduct supervisor.

In its deployment of this solution, AFM estimates supervisory time savings of 25 to 50 percent in the process of reviewing fund registrations. It has also led to more robust assessment of new fund registrations given that supervisors can now focus on the review of higher-risk registrations more than on the manual review of each registration. These efficiencies were made possible through the implementation of over 16 decision rules that replaced checks previously done manually. AFM designed its solution as a central data platform built as part of its CRM, which provides a central place to view and configure decision rules and to make insights available through interactive dashboards to end users (that is, supervisors).

3.4.6 Predictive Modeling of Financial Statements

Ensuring the accuracy of public financial statements, investor report filings, or other financial statements is important to protect investors from errors, omissions, or mistakes. Solutions that allow for predictive modeling of financial statements are intended to facilitate supervision of listed public firms for purposes of consumer and investor protection.

The solution is best described as a “bot report” that checks for the percentage of mistakes in financial statements of listed firms. First, the solution imports the financial statements via XBRL¹¹ format. (No manual input is required to prevent data quality issues.) Quality checks are then carried out to verify, for example, if liabilities match assets, negative values that should have been positive, check for extreme differences, and so on. The next step is scanning all imported financial statements with a validating algorithm to estimate the risk of misstatements. From there, qualitative risk scores are assigned to each financial statement and a logistic regression model is used to predict the chance of misstatement per financial statement. Finally, a decision tree algorithm is applied to identify the greatest possible chance of misstatement. Financial statements with a high risk of error, based on the limits set by supervisors, are given higher priority and are fully reviewed by supervisors.

By employing this solution, AFM enabled its supervisory staff to increase the percentage of financial statements it reviewed from 10 percent to nearly 100 percent of the approximately 350 financial statements received each year. These efficiency gains have allowed their market conduct supervisory to scale up, enabling risk identification in a faster, more comprehensive, and higher-quality manner.

11. eXtensible Business Reporting Language (XBRL) is a global framework and standards-based way to communicate and exchange business information. The reporting language is XML-based. XBRL is commonly used to define and exchange financial information, such as financial statements.

3.4.7 Business Intelligence and Geospatial Analysis

A supervisor may be data rich but information poor if the financial authority lacks the ability to evaluate data for insights. Business intelligence (BI) refers to a technology-driven process for analyzing data and presenting actionable insights to help executives, managers, and supervisory teams make informed evidence-based decisions.

Whereas NLP and its associated solutions help supervisors make sense of unstructured text, BI solutions help supervisors make sense of large quantities of structured data. BI tools can be used on a single data source or can combine data sources together for a more holistic view. An enterprise-wide BI platform typically supports live data connections to a variety of underlying databases (either directly or through datamarts that act as an intermediary layer). A BI tool can also operate without a live data connection and use CSV or comparable data exports from a database instead.

Financial authorities typically use BI tools in combination with other suptech solutions to interpret information and data. For example, interactive, self-service dashboards can be created using BI visualization tools such as Tableau or Microsoft PowerBI, which are often used with solutions for regulatory reporting to interpret findings for non-technical supervisory staff. R or Python workbenches are used by technical users to synthesize and analyze large data sets, especially as they relate to the aforementioned document analysis suptech solutions with NLP. Meanwhile, relational database management and query tools, such as MySQL, and larger BI platforms from large tech providers, like SAP, Oracle, and IBM, support the retrieval and management of data contained within large databases.

BI can also bring in new data sources. For example, market conduct analysis can be enhanced through the addition of geospatial data to understand geographic-level trends. NRB uses geospatial data to understand access points to FSPs in rural Nepal and other trends in rural financial inclusion. Its solution uses an “E-Map Portal” that relies upon a GIS system to map financial access points across Nepal. NRB analytics use this solution to combine submitted data from supervised institutions to understand

financial data based on geography, region, municipality, and branch (Gurung and Perlman 2018).

BI tools benefit users through faster and more accurate reporting, analysis, or planning. At the same time, these solutions also come with challenges for their effective use. BI tools require training and successful socialization to be used within and across an organization. Many require initial setup and connections to the regulator’s underlying databases. In addition, any BI or analysis tool is only as useful and accurate as the integrity of the underlying data.

3.4.8 Managed Data Platform

For market conduct supervisors that receive a large quantity of data regularly from varying sources, a comprehensive, integrated data and analytics platform can support ongoing supervisory functions such as identifying, monitoring, assessing, and analyzing risks. A managed data platform standardizes, centralizes, and makes accessible data from across the financial authority for supervisory staff.

While a managed data platform can be designed in a multitude of ways, AFM designed an integrated platform, structured around three areas: (1) Source data is collected and stored in a persistent staging area. (2) Data is then processed and stored in an integration area and transformed (for example, validated, classified, organized, structured). (3) After applying specific rules or algorithms, the resulting data is finally made available through datamarts for use by staff via dashboards and reports. The platform collects and stores different types and formats of semi-structured data, including CSV, JSON, XML, databases, and so on.

AFM uses its managed data platform solution for all of its regulatory and supervisory activities of retail and wholesale financial markets, providing supervisors with data for around 90 percent of risk monitoring and 50 percent of risk assessments undertaken. This coverage across multiple markets was possible given AFM’s decision to design the platform using a modular approach with cloud-based infrastructure. Risk-mitigation activities continue to be conducted manually by supervisors, as they require very technical judgments and a strong legal background.

4. PEOPLE, PROCESS, AND IT INFRASTRUCTURE: THREE KEY ENABLERS FOR SUPTECH IMPLEMENTATION

The authorities who contributed to this note indicated that they view suptech adoption as an evolutionary process, rather than an end goal. Embedding modern technology and data into the supervisory process is an ongoing effort, and the desired end state is organizational transformation into a data-driven supervisor. Achieving this transformation requires making investments in three key enablers: (1) people, (2) process, and (3) IT infrastructure. These implementation enablers are broadly applicable across all types of suptech solutions, for both market conduct as well as other regulatory objectives.

- **People:** The talent, mindset, and skills of employees and the larger organizational culture toward data and technology
- **Process:** How suptech ideas are supported from ideation to implementation, including how suptech is championed and governed
- **IT infrastructure:** The underlying IT infrastructure and capabilities needed to develop and operate suptech solutions internally

4.1 People: Culture and Skillsets

Financial authorities often cite investment in their staff as the greatest contributor to success in suptech adoption. Skilled employees are important not just to process data and build solutions, but to ensure that the right questions are asked in the first place. It is important to invest

in people's capacity and skills so they can identify data needs and use data insights in their supervisory work to identify, assess, and mitigate supervisory risks.

Supervisory authorities are perennially resource constrained, resulting in the need to invest in internal capacity. Investments in internal capacity can involve two elements: (1) investing in people directly to add new skillsets either through hiring, retraining existing employees, or leveraging external parties; and (2) investing in technology to make employees more productive.

Supervisory authorities need new technology and data skillsets within their workforce, including data scientists, analysts, and programmers. These new skillsets may be added through training of existing employees or external hiring. As an example of the former, the FCA promotes internal analytics communities (to which a quarter of the staff belong) and leads a Community of Practice for analytics skills development and augmentation, which includes frequent training through internal workshops and seminars, internal publications and demonstrations, and seminars and interviews with external subject matter experts.

As a financial authority implements suptech solutions, it can be difficult to know in advance which types of talent and skills may be needed and in what quantity. In fact, all authorities who contributed to this note indicated that they employed the assistance of external parties, such as consultants, technology vendors, or technical assistance providers before building up in-house expertise. For example, the FCA, BOE, AFM, and CFPB all heavily relied on consulting firms during initial stages of suptech implementation. By hiring consultants with data,

CASE STUDY 5

How AFM Prioritized People within Its Transformation to Data-Driven Supervisors

In the case of AFM, its suptech journey can be grouped into three distinct steps: (1) build, (2) pilot, and (3) transform. Step 1, "Build" (2016–19), began with an assessment of its own in-house capacity to deal with and analyze data to support supervisory activities. Since most supervisors have a stronger legal background than a technology background, AFM realized that new investments would need to be made in its staff and culture through trainings and new hires, alongside data and technology solutions. Step 1 efforts concluded with a successful shift in the internal cultural mindset and developing the inter-

nal know-how to leverage data to enhance supervisory activities. Now underway, AFM's step 2, "Pilot" (2019–21), seeks to amplify these cultural changes further with new processes that support the piloting of new technologies. Lastly, AFM's step 3 "Transform" (from 2022 onward), seeks to transform the organization both culturally and operationally to become a data-driven regulator through fully distributed and decentralized responsibility for data teams throughout the financial authority. In each phase, AFM's focus is on its people and on cultivating the skills, culture, and structures to support its suptech strategy.

technology, and business analysis expertise, these regulators were able to move quickly to establish innovation, data-science, or other technology offices. Similarly, BSP worked with R²A to implement its suptech solution, while NRB worked with the United Nations Capital Development Fund.

These initial successes created the rationale for further investment. Regulators were also able to understand through firsthand experience which skillsets were needed internally to support desired supervisory outcomes. Only after learning this did regulators begin hiring or retraining staff to reflect the skillsets now recognized as needed to be suptech- and data-driven regulators. In both the Philippines and Nepal, regulators started with a small core team (supported with external technical assistance) and began expanding the data and technology skillsets of their own staff only after working with the external party to develop its suptech solution.

However, while external parties can help develop suptech solutions, regulators need to plan to have internal capacity to manage, use, and maintain the solution immediately at its launch. BNR experienced a gap of nearly one year between the completion of its infrastructure upgrade and the development of its internal analytics team. During this time, the authority was not able to make full use of the data available. This gap shows the need to have a clear plan for appropriate capacity to leverage suptech solutions to process and analyze large quantities of data over both the near and long term in order to reap the full benefits from such technologies. The same is true of the solution's maintenance. As suptech is a rapidly evolving area, continued maintenance and support are critical to keep the solution relevant. If an external vendor is used to build the solution, the authority may want to consider negotiating ongoing maintenance or for the vendor to train internal staff to manage the solution. Such training can present a capacity-building moment for staff.

4.2 Process: Internal Champions and Strong Governance

Across all country examples, an important factor for success was having an internal champion for suptech. The most frequent internal champions for suptech were innovation offices or a specific business office. In the case of both AMF and the FCA, dedicated innovation offices were created to support business units to develop proofs of concept for suptech solutions. Via a hub-and-spoke model, innovation offices centralized technical and data expertise across each authority, making the recruiting and retaining of technical talent easier. In addition, by making individual business units responsible for the deployment

of their suptech solutions, business units were motivated to identify the supervisory problem they sought to solve and to own ultimate deployment of the suptech solution.

Other offices can also play the role of internal champion. In several authorities, IT departments led suptech initiatives. A specific business office may champion the adoption of specific suptech solutions, such as the supervisory department, complaints department, or the office responsible for reporting by FSPs. For example, the complaints-handling department at BSP led investments in increased automation for complaints handling as described above, supported by external software vendors and technical assistance.

In addition to an identifiable champion for suptech, buy-in from top management is also important. Top management support is helpful both in setting the strategic, high-level goals for data and technology and in the tactical implementation of specific solutions. Top management also provides crucial advisors to the identification and implementation of solutions. Senior management can identify data and skills gaps and serve as advocates for suptech generally and for broader investments or changes to people, process, or technology that represent the key implementation enablers for suptech. Sometimes, this support takes the form of a public statement. For example, BOE committed in 2019 to developing a “world-class regtech and data strategy” with the aim of making data collection more efficient for firms while also improving BOE's technology and data capabilities (BOE 2019). Support of senior management is also key to secure adequate funding and to cultivate relationships with industry players, including both providers of potential suptech solutions and supervised institutions.

Financial authorities also indicated that strong governance processes are necessary to complement the introduction of suptech solutions, in particular regarding access to the new data collected through these solutions. Governance processes create mechanisms for the authority to manage the suptech solution and align it with evolving regulatory objectives. The insights derived from these solutions are important inputs for data-driven policy making. For example, AMF's supervision information system is managed by a collection of internal regulatory and IT teams to support its off-site supervision framework. Solution governance includes a supervisory team focused on standards and reporting guidelines to supervised institutions and a second group of risk experts focused on determining and maintaining data validations. Together, these two teams govern the solution's data collection, analysis, and reporting to support AMF's supervision framework.

It is also common for authorities to develop formal procedures to govern data access to minimize data mismanagement, cybersecurity, and information security risks. Each department has specific permission-based access to view data that is proportional to the need of the role or department. For example, the research department at BNR has access to data determined to support its policy-making objectives, whereas supervisors reviewing a specific FSP have a more granular view. The CFPB and Australian Securities and Investments Commission (ASIC) go further and are among the authorities that have introduced a central department with a chief data officer to govern data access. ASIC's Chief Data Office has developed a Data and Information Governance Framework that has introduced dedicated forums, such as a digital governance board, data governance council, and data-analyst network, to govern data access and management.

Good governance also helps ensure that the insights derived from these solutions are fed as inputs into data-driven policy making. For example, clear processes are put in place to ensure that the outputs of the CBI's social media monitoring solution are distributed and used across the authority. This is facilitated by the Market Monitoring Working Group, which convenes people from across the authority to understand developments in the marketplace and discuss implications across internal departments.

4.3 Underlying IT Infrastructure

All suptech solutions rely to varying degrees on the financial authority's underlying IT infrastructure. Suptech solutions are often built as applications on top of existing IT infrastructure. Where modern IT infrastructure exists, the time, cost, and effort to implement the suptech solution is likely to be lessened. Relevant underlying technology and competencies include database warehousing and management, big-data computing, CRM systems, information security protocols, and licenses to BI and analysis software. The absence of modern or existing

IT infrastructure may be a particular challenge for low-income countries. These authorities may need to devote dedicated attention and resources to IT infrastructure, especially to implement solutions of higher complexity. When upgraded IT infrastructure is available and accessible across a financial authority, any internal "entrepreneur" can run experiments, test concepts, and propose new suptech solutions.

Successful implementation of suptech solutions, particularly for lower-income countries, also depends on the technical capacity, readiness, and involvement of external users, such as supervised institutions. It is important to assess if FSPs will be able to implement and use certain suptech solutions, such as with respect to regulatory reporting. This consideration is particularly relevant for market conduct supervisors who may be required to oversee a large number of smaller and less sophisticated institutions. For example, BNR suffered operational delays due to the complex nature of its EDW. Through active collaboration and dialogue, BNR refined its solution to accommodate the varied technical capabilities of its supervised institutions. FSPs with strong technical capabilities report regulatory data through ADFs,¹² while FSPs with lower technical capabilities can report data via simpler methods, such as by means of APIs or web portal uploads, depending on their capabilities.

Lastly, when considering a large investment in underlying IT infrastructure, it is important to consider the near-term and long-term suptech application it supports. IT infrastructure investments vary in their complexity and required timeline for implementation. For example, AFM refers to different "horizons" when considering suptech and IT infrastructure investments—near term (one to two years), medium term (three to five years), and long term (up to 10 years out). Through this lens, authorities can take an approach that considers and sequences anticipated projects in a portfolio and balances the delivery of both near-term wins and planning for large, longer-term investments.

12. As of 2020, 16 banks, 16 insurance companies, and several forex bureaus and microfinance institutions participate in this manner.

5. IMPLEMENTATION CONSIDERATIONS

Common considerations when implementing suptech solutions emerged across the country examples included in this note. In the process of implementing suptech solutions, supervisors often face key decisions related to determining the business case for a suptech solution, whether to build a solution in-house or buy from a vendor, and how to organize their data and technology staff. Financial authorities that have implemented suptech solutions for market conduct have also benefited from initiatives that accelerate adoption, including formal suptech or data strategies, innovation offices, and liaising with stakeholders.

5.1 Key Decisions in Suptech Implementation

Building the Business Case for Suptech Adoption

Investment in any suptech solution should have a clear rationale and business case. Suptech solutions should be supervisor-centric—that is, they should represent clear enhancements to support established supervisory objectives. Value is delivered through a combination of greater operational efficiency and enhanced supervisory effectiveness. While seemingly obvious, authorities who participated in this technical note emphasized the importance of beginning with defining a specific supervisory problem, rather than elevating the technology solution before the problem.

Suptech solutions can serve multiple types of supervision. In pursuing a suptech solution, market conduct supervisors can partner with other departments to build a stronger business case for a solution that supports multiple departments and supervisory objectives. Many suptech solutions used for prudential or financial inclusion supervision can be adapted to serve market conduct supervisors or vice versa. For example, the data and insights gathered directly from supervised institutions and consumers, or from non-traditional market monitoring sources, can be used to support market conduct as well as prudential or financial inclusion supervision. Solutions for document and business analysis can also be adapted to support other types of supervision beyond market conduct.

The motivation to explore suptech solutions is often driven by two major forces: (1) supervisory demand and (2) technological innovation. On the one hand, demands on supervisory staff with limited resources drives efforts to boost operational efficiency using technology. On the other hand, the availability of new technologies present authorities with the opportunity to enhance the supervisory process.

Suptech can be driven by “pull” factors like supervisory demand. This is especially the case where supervisors face limited resources and seek increases in employee productivity. For example, regulators in Nepal, Rwanda, and the Philippines developed interactive dashboards for supervisors to view and query supervisor data in real time. Interactive dashboards are not a new technological innovation; they have existed for several years. The new dashboards were developed to enable employees to analyze the increasing volumes of data faster.

Industry can also lead to demand for suptech adoption.

Pressure exists to make market conduct supervision more efficient for both financial authorities and supervised institutions. For example, solutions for regulatory reporting, such as the APIs in the Philippines and ADFs in Rwanda, were driven by promises made to industry by regulators to reduce the regulatory burden on supervised institutions.

Other suptech solutions are driven by “push” factors such as the availability of new technologies in data analysis, IT infrastructure, and computer science.

These innovations provide supervisors new applications to perform new functions or enhance existing ones. In this context, supervisors are rarely driving the primary research and development of new algorithms or technologies. More often, they leverage cutting-edge innovations, new technologies, and advancements in IT infrastructure and configure them to their own needs and use cases. This is particularly true regarding advancements in machine learning and NLP. Advanced algorithms were until recently restricted to small teams of highly specialized computer scientists, often in university departments or high-tech firms. Now, they are in reach for financial sector regulators, due to the popularization of data-science software and platforms. These are enabled by widely used programming languages, such as R and Python, and made accessible with open-source machine-learning libraries and APIs. Though programmers still need to know how these algorithms work and their underlying statistical processes and parameters, this knowledge has become far more accessible and easier to learn.

Addressing Legacy IT Systems

As supervisors seek to develop new suptech solutions, authorities must decide what to do with existing, older IT systems. Supervisors must choose whether to focus their energies and budget on replacing older systems with new infrastructure or on building new solutions using existing IT systems, and what the appropriate balance between these two competing demands is. Across authorities who contributed to this note, the typical approach was to choose aspects of both, depending on the regulatory need and the solution’s context.

For new supotech areas with comparatively limited legacy IT systems, such as NLP analysis, regulators can focus on building new systems and solutions. For example, the FCA developed a new in-house NLP platform that supports several supotech applications.

For supotech areas with legacy IT systems, such as regulatory reporting solutions or complaints management, authorities often focus on improving the existing IT infrastructure. For example, NRB is maintaining its legacy regulatory reporting database and data-submission infrastructure while introducing a next-generation solution in 2021. This solution will replace and eventually sunset its web portal with data upload with a new API-based reporting method. Likewise, AFM developed a supotech roadmap that prioritizes both near-term supotech development with longer-term infrastructure needs, as exemplified by its managed data platform that will be expanded across its supervisory department as the solution expands to include additional markets under its jurisdiction.

Deciding Between In-House Solutions Versus Third-Party Vendors

Supervision departments must make the strategic decision on whether to invest the staff time and resources to build in-house tech solutions for market conduct or to rely on commercially available solutions. This decision depends primarily on the complexity of the solution and the optimal allocation of financial resources and staff skills. At the same time, it is also a decision that reflects the business culture of an authority and the available solutions in the market.

Buying solutions from software vendors is often best for complex investments in IT infrastructure, reporting,

or analytics. For example, ASIC has partnered with the Sydney-based intelligence, analytics, and cybersecurity firm Nuix since 2005 in order to extend ASIC’s capacity to become a more data-driven, intelligence-led law enforcement agency (ASIC 2017b). In Rwanda, BNR partnered with Sunoida Solutions to develop an Electronic Data Warehouse to automate and streamline data submission and analytics.

Sourcing supotech from industry vendors can be faster and usually benefits from economies of scale and the availability of dedicated staff with specialized technical expertise. Widely available commercial solutions benefit from economies of scale, leading to demonstrated track records for quality and efficient timelines for implementation. External sourcing also reduces the challenge of hiring and retaining technical staff, at least initially. Both CBI and AMF contract with external vendors for specific solutions for non-traditional market monitoring. These arrangements are particularly efficient when the needs of regulators can be readily adapted from commercial solutions and they do not change frequently over time.

However, there are limits to relying on external parties; the market for specialized private sector supotech solutions is growing but not yet mature. Few firms specialize in serving financial authorities, though this is evolving with the emergence of innovation offices within authorities to collaborate with vendors on solutions. Most vendors have predominately developed tech solutions predominately for private sector firms, which is the larger market. For example, in the context of non-traditional market monitoring, banks or insurance firms contract vendors for brand protection on social media, whereas supervisors aim to detect market conduct risks such as product mis-selling

FIGURE 10: Considerations for In-House Development Versus Using a Third-Party Vendor

	Building a Custom Supotech Solution In-House	Buying a Supotech Solution from a Vendor
PROS	<ul style="list-style-type: none"> • Absolute control over the solution • Ability to customize or adapt to authority’s needs • Build internal tech capacity and technology skillsets 	<ul style="list-style-type: none"> • Brings external technical and data expertise and perspective • Stability and quality of widely available tools are typically high, known
CONS	<ul style="list-style-type: none"> • Requires ongoing resources to maintain or upgrade solution • Challenge of finding, hiring, and retaining skilled staff 	<ul style="list-style-type: none"> • Vendor lock-in (for example, a long-term contract is typically required) • Difficult to discern how readily available a solution is out of the box (for example, navigating a vendor’s marketing of the solution) • Solution may require adaptation from an industry use case to regulatory use case • Procurement process may be burdensome, or process may preclude hiring the preferred vendor with a cheaper, less preferred alternative • Requires a strong vendor oversight process • Constraints on data sharing may exist

and customer complaints. While the two are related, they require customization and adaption, nonetheless.

Depending on the legal and regulatory framework, limits on the collection, sharing, and management of data with external parties can also be a barrier. In some countries, privacy or data-sharing laws limit regulators' ability to collect data or share data with others. There may also be restrictions on how data can be housed. Legal requirements may limit viable solutions available commercially.

The "build vs. buy" decision should be made on a solution-by-solution basis after regulatory goals, benefits, and costs are considered properly. Examples of the factors to consider are included in figure 10.

Organizing Internal Staff Working on Data and Technology Management of internal talent on data science and analysis is essential for a data-driven supervisory strategy to take effect fully. Authorities such as the FCA, ASIC, and AFM view investments in building data-science awareness and capacity among internal staff as crucial to supporting their suptech strategy. In low- or middle-income countries such as Rwanda, the Philippines, and Nepal, recruiting to increase data-science capacity is often cited as a key challenge, making effective management of employees even more important.

Across the county experiences included in this note, authorities faced a common decision point regarding how to integrate supervisors and data scientists. Three main approaches to organizing data-science and analysis teams were observed, as described below. Most authorities adopt hybrid approaches, establishing centralized teams but encouraging the formation of interdisciplinary working groups to tackle specific business questions. As authorities expand data-science capacity overall, some envision transitioning to decentralized data-science teams or hub-and-spoke models.

- 1. Centralized data-science team:** All data scientists sit together in the same department and manage analytics across the authority. This is also often called the "center of excellence" model, as data-science teams are expected to maintain a channel of communication with business units but maintain a degree of independence. The main benefit of centralization is that it helps data-science teams focus on their own initiatives. However, this independence from business units limits their visibility into business needs.
- 2. Decentralized data-science team:** Data scientists are embedded across different business units. The main benefit of decentralization is that it promotes strong

collaboration between supervisors and data scientists and enables data scientists to be more aware of the day-to-day challenges that business units face. However, this often creates silos between departments, reducing the opportunity for spillovers and cross-fertilization of ideas. Moreover, a decentralized function requires highly talented managers capable of managing both supervisory and data-science staff of different backgrounds within the same team.

- 3. Hub-and-spoke model:** This increasingly popular model combines elements of the two previous approaches. In this approach, a centralized data office coordinates with data-analytics staff within the business units, maintaining a constant flow of communication and spillover within the teams. This approach is particularly effective at avoiding silos as well as isolation of data-science teams. However, it requires higher coordination and managerial capacity to implement properly. This approach normally makes sense in larger organizations.

Utilizing Adaptive Approaches for Suptech Development
The purpose of adaptive approaches is to engage and validate a suptech solution's design and functionality by involving end users, which maximizes the utility of the solution when delivered. People are the ultimate end users of these solutions, and solutions must be designed to serve their specific needs. A lack of involvement by end users in developing these solutions is likely to hamper the solutions' ability to meet organizational needs. For example, "Agile," a common adaptive approach to software development, prioritizes delivery of smaller, more frequent pieces of functionality for users. Earlier user interaction with the solution results in more engagement with end users and the opportunity for those users to provide more timely feedback. Adaptive approaches can replace more traditional software-development processes, such as "waterfall" methods. Under waterfall, extensive definition of requirements and development of the proposed solution happen before interaction with end users. This results in limited flexibility for changing requirements once final, or expensive changes that may come up only at the end stages.

Authorities can utilize experimentation and iteration in the technology-development process. These processes validate a solution's concept, establish a solution's feasibility, and build the investment case internally for a larger outlay of resources to build and deploy a full solution. Iteration is key to the development of appropriate suptech solutions and is present in many of the experimental approaches of financial authorities. Possible approaches include the following:

- *Design or tech sprints, including hackathons:* These occur early in the technology-development process, typically in the ideation phase. A small project team comes together to develop a product idea in a confined amount of time, most frequently over a few days to a week. During the “sprint,” the team develops a potential solution, typically presented as an initial solution concept or mockup. This process allows for ambiguities to be identified and addressed before proceeding with the idea further. Dummy data and mock visualizations are common at this stage.
- *Proof of concept:* A small application, experiment, or exercise to test the feasibility of a suptech solution. It is usually conducted to demonstrate that a solution can be implemented, but without exploring the full implementation details.
- *Prototype:* An approach to developing a tangible model to test the desirability of a suptech solution. The objective of a prototype is to share the solution with users for feedback to identify errors, inefficiencies, or other issues.
- *Pilot:* As the prototype of a suptech solution becomes higher fidelity (that is, more closely resembling the final product), it is common to pilot the solution with a limited population or subset of the final users. The goal is to understand how the solution works in a real context and what changes may be needed to scale to full production.
- *Minimum viable product (MVP):* The suptech solution is implemented with a core set of capabilities or features. It is important to note that when introduced, an MVP solution will not be fully featured and will have aspects of its end-state functionality missing. The objective is to minimize time to implementation and delivery of value to the authority and for end users. From this core set of features, additional capabilities are then added to improve the solution over time.

5.2 Initiatives to Accelerate Suptech Implementation

Formal Suptech or Data Strategies

The creation of a formal suptech or data strategy can accelerate data capacities and suptech adoption within a financial authority. Formal strategies allow for a structured approach to increasing data capacities or suptech adoption that has institution-wide buy-in, and coordinate action and the support of an authority’s senior management. For example, ASIC developed a three-year data strategy (2017–20) outlining a vision for suptech and its approach to capture, share, and use data (ASIC 2017a). ASIC’s data strategy has been accompanied by a series

of initiatives to accelerate data-driven supervision. This includes creating a Chief Data Office to develop and manage internal data policies, establishing a Data Governance Council and Digital Governance Board as dedicated forums for data governance, and setting up a data-analyst network. In the United Kingdom, the FCA developed its first data strategy in 2013 (FCA 2013) and updated it in 2020 to include new technologies and capabilities. Similarly, BOE published the 2020 report *Transforming Data Collection from the UK Financial Sector*, by which it initiated a dialogue with supervised institutions and solution vendors to shape data reporting over the next 5–10 years (BOE 2020). In each instance, the development of a formal institutional strategy on data served as a catalyst to align internal departments around a multiyear suptech program to foster technology development and inform data-driven supervision.

While a formal suptech or data strategy can be helpful, such a formal plan is not mandatory in order to begin implementing suptech solutions. Successful regulators were just as likely to employ an incremental approach that focused on introducing targeted solutions to address specific supervisory problems, as opposed to developing a comprehensive institutional strategy.

Innovation Offices and Liaising with Stakeholders

Innovation offices can be helpful for financial authorities implementing formal data or suptech strategies. Also referred to as “labs” or “hubs,” innovation offices provide a dedicated forum and often dedicated IT, data-science, and regulatory staff to test and develop a variety of data and suptech solutions. Innovation offices allow for an experimental mindset and a data-driven culture of suptech innovation within the constraints of a regulatory authority. For example, BdP launched its innovation lab (called *inov[®]*) in 2019 to streamline current supervisory processes by exploring new capabilities, with a focus on NLP use cases.

Innovation offices also create a central place for learning and development opportunities for employees to gain stronger data and technology skills. For example, the FCA’s innovation office has organized analytics communities. It has also designed a comprehensive Data Training Programme tailored to different roles, ranging from a member of the FCA Executive Committee to supervisory analysts. An informal innovation community has developed around its innovation office through an organized calendar of events and seminars including “Data Week,” a week-long program of data-focused events with over 50 sessions; a “reverse mentoring” scheme that matches senior leaders with data scientists; and competitions to expand appetite for analytics.

Liaising with stakeholders is also key to accelerating successful supotech implementation. Authorities will need to engage with industry, both as a source of supotech solutions and as supervised institutions. Innovative ideas can often come from unexpected areas. Financial authorities can benefit from bringing together regulators, academics, vendors, and developers to promote an environment for innovation, share experiences and information, and enable new developments that can benefit the varied participants.

Innovation offices can help improve dialogue among diverse types of experts, industry, and fintech startups. For example, CBI, ASIC, BdP, and AFM each established an innovation hub¹³ whose primary purpose is to engage

fintech or regtech¹⁴ firms more closely. Through collaboration, the authorities intend to build awareness about the regulatory landscape and encourage industry innovation in regtech and supotech solutions. By opening a dedicated office to collaborate with industry, authorities can signal greater priority and demand by the regulator for such solutions and encourage solution providers to bring new supotech and regtech solutions to market.

Lastly, international networks also promote financial innovation and provide regulators with platforms for interaction. For example, the Global Financial Innovation Network, formally launched in January 2019, comprises an international group of financial sector regulators and related organizations committed to supporting financial

CASE STUDY 6

How ASIC's Innovation Office Collaborates with Industry Stakeholders

ASIC launched its innovation hub¹⁵ in 2015 to promote the development of fintech and regtech solutions that improve outcomes for consumers and market integrity.

As part of ASIC's research efforts, the innovation hub aims to keep up to date on the latest developments in regtech and supotech. ASIC often convenes roundtables, liaison forums, national and international network events, and regulator meet-ups, and it promotes training events for its proofs of concept. Through these activities, ASIC has learned about trends related to consumer and industry demand, identified appropriate technology use cases in the market, and clarified roles of industry and regulators regarding the use of regtech and supotech.

ASIC is keen to monitor and understand market developments related to technology innovation for purposes of its own internal use. Since 2018, ASIC has led a series of regtech and supotech research-and-development initiatives, including research to understand commercially available solutions and promoting internal trials of new, emerging technologies. In order to participate, ASIC generally requires supotech and regtech providers to demonstrate how their technology solution(s) can potentially promote better outcomes for investors, financial consumers, or markets, or how their technology solution(s) promotes ASIC's regulatory objectives, improved risk management, or compliance.

For example, ASIC trialed NLP solutions in 2018 to understand the relevancy of the technology to key supervisory use

cases. ASIC began by releasing a set of problem statements to solution providers and invited applications on how NLP could solve for each of them.¹⁶ The trials explored potential efficiencies in supervision, including through automation and prediction. The results of these trials gave ASIC insights into data-availability issues, data-annotation work, and areas on which to focus internal capacity in the future. Subsequent pilots focused on monitoring financial promotions, financial advice, voice analytics and voice to text, and technology-assisted guidance tools (ASIC 2019). With respect to financial promotions, trial demonstrations accurately detected potential breaches of mandatory disclosure requirements in over 90 percent of specific cases. Similarly, demonstrations identified compliance issues in financial advice files at accuracy rates of around 90 percent in the sample data set.

The innovation hub also provides informal assistance to start up and scale up businesses to navigate the regulatory framework and share how regulation may affect them. Fintech and regtech providers receive access to (1) practical regulatory support and informal assistance from senior ASIC staff, (2) options relating to ASIC's relief powers, such as the regulatory sandbox, and (3) events focused on fostering cross-sector engagement, including regtech quarterly liaisons, demonstrations, and a role as an observer in trials.

13. As an example, see www.centralbank.ie/regulation/innovation-hub.

14. Regtech refers to the use of new technologies by FSPs to meet their regulatory requirements. See FSB (2020).

15. <https://asic.gov.au/for-business/innovation-hub/>.

16. The initial pilots focused on (1) identifying promotions of concern for financial and credit services, (2) phone sales practices of insurance providers, (3) review of managed-fund product-disclosure statements, (4) review of financial advice files, (5) review of financial reporting in company announcements, and (6) review of prospectuses.

innovation in the interests of consumers. It seeks to provide a more efficient way for innovative firms to interact with regulators, enabling pilots for firms wishing to test innovative products, services, or business models across jurisdictions. It also aims to create a framework for cooperation between financial sector regulators on innovation-related topics and to facilitate the sharing of different experiences and approaches.¹⁷ The International Financial Consumer Protection Organisation (FinCoNet), an international organization of market conduct supervisory authorities, similarly provides a platform for market conduct supervisors to share knowledge and experiences and learn from one another. Financial sector authorities have also developed bilateral relationships to learn from one another, share data, and work more closely together on suptech solutions and market conduct supervision. For example, the FCA and AFM announced a bilateral partnership in 2019.

5.3 Additional Challenges Encountered by Regulators

Where regulators still need to develop their basic supervisory processes or framework for market conduct, suptech implementation should likely take secondary priority. As noted previously, supervisors commonly encounter the challenge of how to design and implement suptech solutions that support the authority's supervisory processes and are tailored to meet their supervisory needs and capacity. However, in some jurisdictions, market conduct supervision may be a brand-new or just-emerging function. Implementing suptech solutions is particularly challenging in jurisdictions that have not developed their core market conduct supervisory framework yet. In such cases, there are several caveats

BOX 1

FinCoNet: SupTech Tools for Market Conduct Supervisors

FinCoNet is an international organization of financial consumer protection supervisory authorities. The goal of FinCoNet is to promote sound market conduct and to enhance financial consumer protection through efficient and effective financial market conduct supervision, with a focus on banking and credit.

In recent years, the effects of digital transformation on financial consumer protection have become a priority item on the agenda of supervisory authorities. In this vein, FinCoNet has stated that the shift from traditional financial sector delivery channels to online and mobile technology has important implications. These include supervisory authorities' ability to identify emerging consumer risks arising from digitization and to have appropriate tools to mitigate such risks. Consequently, the FinCoNet Governing Council decided to include a standing committee to develop further work in these areas in FinCoNet's Programme of Work for 2017/2018.

This work led to the report *Practices and Tools Required to Support Risk-Based Supervision in the Digital Age* (FinCoNet 2018), published in November 2018. This report reflects the experiences of various supervisory authorities as they tackled the challenges stemming from the need to ensure proper consumer protection in the framework of digitization and the ways they are adapting supervisory

tools to the challenges of digital financial product and services, and it highlights a series of useful takeaways to be considered by the supervisory community.

Following the release of this report, the standing committee then initiated a project to review the most innovative tools carried out by the supervisors' community, summarized in the report *SupTech Tools for Market Conduct Supervisors* (FinCoNet 2020), published in November 2020. A workshop on this topic took place at FinCoNet's annual general meeting in November 2019, with the participation of the World Bank, the Organisation for Economic Co-operation and Development, and the Bank for International Settlement, and other organizations.

This new report, based on a survey of a wide range of authorities, aims at capturing the general framework for the development of suptech, detailing strategies, supervisory needs, levels of use, operational readiness and areas of application, thus complementing the research work hereby presented by the World Bank. Moreover, FinCoNet's report contains a description of the most relevant SupTech tools, how such tools are applied in market conduct supervision, how they enhance supervisory processes, and the main challenges encountered.

17. <https://www.thegfin.com/>.

to consider in focusing on technology and data, without first putting in place the fundamentals of a framework for proper risk indicators and strategic metrics for market conduct supervision. For such jurisdictions, supervisors likely will first need to dedicate more effort to developing supervisory frameworks for market conduct before beginning to explore suptech solutions to automate processes and procedures and collect the types of data that support their supervisory needs.

Regulators may also open themselves to new operational and reputational risks through suptech. The collection of more, and new, types of data requires regulators to assume a new role as responsible data managers and to act appropriately to protect this data from accidental disclosure or from bad actors. Data privacy, information security, and cybersecurity represent new competencies for regulators, who will need to mitigate these risks. In a rush to collect more and more data without a commensurate investment in internal skills and capacity, regulators also open themselves to the risk of having more information than can be processed or analyzed in a proper or timely manner. This presents a potential reputational risk, whereby the regulator has had access to the proper information to identify the market conduct issues but ultimately lacked the proper capacity to do so in a timely manner.

6. LOOKING FORWARD

Over the last decade, consumer finance has been transformed due to digital technology. Traditional players have undergone digital transformations, while a range of new, non-traditional players have entered the market, including fintechs, mobile network operators, and technology firms. While this transformation has led to beneficial innovations for financial consumers, it also presents market conduct supervisors with new challenges. They must safeguard markets burgeoning with new users, new institutions, new technologies, and new business models, balancing protection of consumers with innovation.

It may be tempting for regulators to step on the brakes to hold back these forces. The solution, of course, is not to limit their development but to ensure that regulators have the tools they need to oversee these rapidly evolving markets efficiently and effectively. Greater regulatory confidence and capacity reinforce healthier, inclusive financial markets. This is precisely where suptech can play such a critical role.

Already in 2020, financial markets and authorities faced a huge new test due to the COVID-19 pandemic.

The crisis has prompted a significant increase and reliance on digital financial services, further accelerating the digital transformation. It is likely that many consumers will become comfortable with digital financial services during this crisis and will stick to these new behaviors even as the crisis lessens. What started over a decade ago in small pockets may soon become the default way of banking around the world.

Within this landscape of digital transformation, suptech becomes an invaluable tool for financial authorities.

This is true of every category of suptech for market conduct supervision. The direct and automated collection of granular regulatory data from supervised institutions obviates the need for on-site examinations, while digital interfaces for complaints handling enable regulators to engage directly with consumers online while automating complaints data collection and analyses. Non-traditional market monitoring, meanwhile, provides a real-time pulse on fast-moving sentiment and consumer risks with FSPs, while solutions for text analysis can extract insights from documents in seconds where previously it would have taken weeks or longer. The real-world solutions presented in this note provide authorities with the tools to oversee increasingly complex markets with increased effectiveness and efficiency.

To be sure, suptech for market conduct is not a silver bullet.

It certainly does not replace having the fundamentals of market conduct supervision in place, and there are many pre-conditions and challenges to successful implementation of suptech solutions. However, when combined with the right mix of smart, competent staff and a comprehensive market conduct supervision framework, suptech solutions can better position market conduct supervisors for the challenges and opportunities ahead. The initial successes of regulators highlighted in this note offer a glimpse of the future—one in which data and technology become core to the operations, identity, and culture of financial authorities to enable them to achieve their regulatory mandates.



REFERENCES

- ASIC (Australian Securities and Investments Commission). 2017a. *ASIC's Data Strategy 2017–2020*. <https://download.asic.gov.au/media/4511295/asic-data-strategy-2017-20-published-11-october-2017.pdf>.
- . 2017b. "ASIC Extends Partnership for Digital Investigation and Analytics Software." <https://asic.gov.au/about-asic/news-centre/find-a-media-release/2017-releases/17-262mr-asic-extends-partnership-for-digital-investigation-and-analytics-software/>.
- . 2019. *ASIC's Regtech Initiatives 2018–19* (Report 653). <https://download.asic.gov.au/media/5424092/rep653-published-20-december-2019.pdf>.
- BCBS (Basel Committee on Banking Supervision). 2012. *Core Principles for Effective Banking Supervision*. Washington, DC: Bank for International Settlements. <https://www.bis.org/publ/bcbs230.pdf>.
- . 2017. *Sound Practices: Implications of Fintech Developments for Banks and Bank Supervisors*. Washington, DC: Bank for International Settlements. <https://www.bis.org/bcbs/publ/d415.pdf>.
- BIS (Bank of International Settlements). 2018. *Innovative Technology in Financial Supervision (Suptech)—the Experience of Early Users*. <https://www.bis.org/fsi/publ/insights9.htm>.
- . 2019. *The Suptech Generations*. <https://www.bis.org/fsi/publ/insights19.pdf>.
- BOE (Bank of England). 2019. *Future of Finance*. <https://www.bankofengland.co.uk/report/2019/future-of-finance>.
- . 2020. *Transforming Data Collection from the UK Financial Sector*. <https://www.bankofengland.co.uk/-/media/boe/files/paper/2020/transforming-data-collection-from-the-uk-financial-sector.pdf?la=en&hash=6E6132B4F7AF681CCB425B0171B4CF43D82E7779>.
- CGAP (Consultative Group to Assist the Poor). 2017. *Data Collection by Supervisors of Digital Financial Services*. <https://www.cgap.org/research/publication/data-collection-supervisors-digital-financial-services>.
- di Castri, Simone, Matt Grasser, and Arend Kulenkampff. 2018. *Financial Authorities in the Era of Data Abundance: Regtech for Regulators and SupTech Solutions*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3249283.
- . 2020a. *A Chatbot Application and Complaints Management System for the Bangko Sentral ng Pilipinas (BSP): R²A Project Retrospective and Lessons Learned*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3596268.
- . 2020b. *An API-Based Prudential Reporting System for Bangko Sentral ng Pilipinas (BSP): R²A Project Retrospective and Lessons Learned*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3596276.
- FCA (Financial Conduct Authority). 2020. "Data Strategy" (web page). <https://www.fca.org.uk/publications/corporate-documents/data-strategy>.
- . 2013. *The FCA Data Strategy—How We Will Manage and Use the Data We Collect*. <https://www.fca.org.uk/publication/corporate/fca-data-strategy.pdf>.
- FSB (Financial Stability Board). 2020. *The Use of Supervisory and Regulatory Technology by Authorities and Regulated Institutions: Market Developments and Financial Stability Implications*. <https://www.fsb.org/2020/10/the-use-of-supervisory-and-regulatory-technology-by-authorities-and-regulated-institutions-market-developments-and-financial-stability-implications/>.

FinCoNet (International Financial Consumer Protection Organisation). 2018. *Practices and Tools Required to Support Risk-Based Supervision in the Digital Age*. http://www.finconet.org/Finconet_Report_Practices-tools-for-risk-based-supervision-digital-age_November_2018.pdf.

———. 2020. *SupTech Tools for Market Conduct Supervisors*.

http://www.finconet.org/FinCoNet-Report-SupTech-Tools_Final.pdf

Gurung, Nora, and Leon Perlman. 2018. "Use of Regtech by Central Banks and Its Impact on Financial Inclusion." <https://dfsobservatory.com/publication/use-regtech-central-banks-and-its-impact-financial-inclusion>.

IMF (International Monetary Fund). 2015. *Can Financial Inclusion Meet Multiple Macroeconomic Goals?* <https://www.imf.org/external/pubs/ft/sdn/2015/sdn1517.pdf>.

Murphy, Dan, and Jackson Mueller. 2018. *RegTech: Opportunities for More Efficient and Effective Regulatory Supervision and Compliance*. <https://milkeninstitute.org/reports/regtech-opportunities-more-efficient-and-effective-regulatory-supervision-and-compliance>.

Toronto Center. 2018. *SupTech: Leveraging Technology for Better Supervision*. <https://res.torontocentre.org/guidedocs/SupTech%20-%20Leveraging%20Technology%20for%20Better%20Supervision%20FINAL.pdf>.

World Bank. 2018. *From Spreadsheets to Suptech: Technology Solutions for Market Conduct Supervision*. <https://openknowledge.worldbank.org/handle/10986/29952>.

———. 2020. *A Roadmap to Suptech Solutions for Low Income (IDA) Countries*. <https://documents.worldbank.org/en/>

[publication/documents-reports/documentdetail/108411602047902677/a-roadmap-to-suptech-solutions-for-low-income-ida-countries](https://documents.worldbank.org/en/publication/documents-reports/documentdetail/108411602047902677/a-roadmap-to-suptech-solutions-for-low-income-ida-countries).

