Cambridge
Centre
for Alternative
Finance

UNIVERSITY OF CAMBRIDGE
Judge Business School

CAMBRIDGE SUPTECH LAB

STATE OF SUPTECH REPORT 2022

SUPPORTED BY

BILL & MELINDA GATES *foundation*

UKaid
from the British people

Foreign & Commonwealth Office

ABOUT THE CAMBRIDGE **SUPTECH LAB**

The Cambridge SupTech Lab at the Cambridge Centre for Alternative Finance, the University of Cambridge Judge Business School, accelerates the digital transformation of financial supervision.

While financial services are becoming increasingly global, digital and complex, analogue processing and antiquated technologies in data gathering, validation, storage and analysis erode the analytical capabilities of supervisory agencies, who are often too late in protecting consumers from fraud and seeing signs of stress in the financial system or miss the underlying causes. This is all happening while financial crime remains a trillion-dollar issue, and public agencies face new challenges, such as the regulation and supervision of crypto assets, and monitoring environmental, social and governance (ESG) aspects of the financial industry's business.

The Lab aims to meet financial sector supervisors' needs by working with them to develop new methodologies and processes that further market oversight and empower consumers, and deploy suptech applications that generate relevant, reliable, timely insights to inform their decisions.

From research to executive education to technical assistance to crafting production-grade suptech solutions, we are committed to supporting the emergence of the suptech ecosystem and to empowering a new generation of innovation leaders seeking to digitally transform financial supervision.

We invite you to find out more at

www.cambridgesuptechlab.org

@Cambridesuptechlab

@CamSupTechLab

Cambridge Centre for Alternative Finance

UNIVERSITY OF CAMBRIDGE Judge Business School

Suggested citation:

Cambridge SupTech Lab (2022), State of SupTech Report 2022, Cambridge: Cambridge Centre for Alternative Finance (CCAF), University of Cambridge. Available at www.cambridgesuptechlab.org/SOS

The mention of specific companies, manufacturers or software does not imply that they are endorsed or recommended by the Cambridge SupTech Lab in preference to others of a similar nature that are not mentioned.

All graphics and charts can be downloaded at www.cambridgesuptechlab.org/SOS

AUTHORS

Simone di Castri
Matt Grasser
Juliet Ongwae
Jose Miguel Mestanza

DESIGN

Marta Lopera
Emily Duong
Jessica Ali

COPY EDITING

Alpa Somaiya

Adebola Daramola, Alexander Apostolides, Kyriakos Christofi, Philip Rowan, Yue Wu and Bryan Zhang of the Cambridge Centre for Alternative Finance (CCAF) contributed to the analysis.

The Cambridge SupTech Lab is supported by

UKaid from the British people

Foreign & Commonwealth Office

BILL & MELINDA GATES foundation

CAMBRIDGE SUPTECH LAB

TABLE OF CONTENTS

EXECUTIVE SUMMARY

THE CAMBRIDGE SUPTECH LAB STATE OF SUPTECH REPORT 2022 PRESENTS INSIGHTS ON THE CURRENT STATE OF THE DIGITAL TRANSFORMATION OF FINANCIAL SUPERVISION WORLDWIDE.

The Report provides a global snapshot across several facets of suptech, including underpinning digital infrastructure and technologies, supported supervisory use cases, approaches employed for developing and deploying suptech applications, and the related challenges and risks.

The State of SupTech Report 2022 focuses on how financial authorities are developing and implementing supervisory technologies (suptech), and establishes a baseline from which to track the progress and impact of suptech adoption allowing financial authorities across the world to benchmark the progress of their suptech initiatives.

To facilitate granular analysis of these macro trends, the Report introduces a novel version of the SupTech Taxonomy adopted by the Bank for International Settlements (BIS) (BIS 2018, BIS 2019), classifying supervisory use cases, technologies, and data science tools in a standardised and structured manner. To complement the analysis and ground the findings in a practical context, the Report also provides a timeline of disruptions and innovations in supervision, and a set of six case studies of suptech applications.

The Report is based on the insights that **146 financial authorities** shared through (both or either):

- A survey of 134 financial authorities from 108 jurisdictions
- A questionnaire on data models with 74 individual supervisors representing 46 agencies and 35 jurisdictions.

The analysis also advances the understanding of the suptech marketplace from the supply side, providing critical insights from the nascent but rapidly growing industry of suptech vendors through in-depth qualitative research of key vendors sampled from the Cambridge SupTech Lab's SupTech Marketplace and highlighting their perspectives on the business case for suptech, the primary use cases they focus on and the challenges they face in commercialising suptech solutions.

HIGHLIGHTS FROM THE STATE OF SUPTECH REPORT 2022

- **Suptech 'is happening'. Most financial authorities have already engaged in suptech initiatives.**

  While suptech development is still at a nascent stage with room for growth, the survey results indicate that 71% of financial authorities are rising to the challenge as we see the adoption of suptech solutions, strategies and roadmaps increasing.

- **Suptech efforts remain in the experimentation stage, primarily focused on improving data collection and basic analysis.**

  Based on the classification provided by the Bank for International Settlements (BIS 2019) and revised by the Lab in this Report, the technologies deployed by financial supervisors mostly fall into the first or second generation of data architecture, and mainly support data collection as well as descriptive and diagnostic analytics.

- **Most suptech use cases centre around consumer protection and prudential supervision.**

  59% of financial authorities report their suptech applications focussing on consumer protection supervision, while 58% report their suptech applications support prudential supervision use cases.

- **Significant challenges to suptech adoption remain to be addressed.**

  Limitations in budget, data quality and technical skills remain the most significant barriers to implementing suptech. There is a remarkable mismatch between the experience of financial authorities and vendors when it comes to procurement, with tech providers urging public agencies to address legacy procurement processes.

  Financial authorities also express an unmet need for data teams, data sharing and data synthesis as a foundational part of their modernisation.

- **There are significant distinctions in the state of suptech in emerging markets and developing economies (EMDEs) as compared to advanced economies (AEs).**

  Financial authorities in AEs are early adopters of suptech, more often have sufficient digital infrastructure, more often assign dedicated suptech roles and departments, have seen more substantial internal outcomes than those in EMDEs, and seek funding primarily to grow their teams. EMDEs agecies tend to run suptech initiatives within the supervision department itself, are more interested in trainings, technical assistance, digital tools, and seek funding primarily for solutions design and development.

- **Financial authorities in EMDEs and in AEs face very similar challenges in the digital transformation of their supervisory process and capabilities.**

  Agencies in EMDEs and AEs report lack of budget to be the main constraint to the development and deployment of suptech.

- **Centralised data office models to accelerate suptech development and implementation are emerging.**

  35% of the surveyed financial authorities have a dedicated centralised office reporting to a Chief Data Officer who is either solely responsible for the suptech initiatives or works with other functions to develop and deploy suptech.

- **Funding to accelerate the suptech market is a key area of focus.**

  Although suptech vendors report some secondary support from grants, funding for financial authorities' suptech initiatives comes primarily from the financial authorities themselves. Most suptech solutions are provided by external sources like contracted vendors and purchased off-the-shelf software, yet these vendors also report challenges in funding and an ability

to deeply understand financial authorities' prioritized needs.

- **The top suptech challenges differ between agency types.**
  For central banks, the challenges are primarily related to internal culture and strategic buy-in. For capital markets, securities, and investment instruments supervisors, challenges tend to be related to upgrading their existing systems and processes. For other supervisors, the uniquely prominent challenges are with IT systems.

- **Most authorities still do not have a gender data strategy**.
  Only 21% have a currently operating strategy, 9% have one in development, 70% report no strategy at all.

- **Suptech is enabling new supervisory use cases that would not otherwise be possible.**
  While suptech solutions use chatbots and APIs to optimise existing processes and augment legacy tools, others are opening completely new opportunities for supervisors. The ability to ingest massive online datasets like social media streams to conduct sentiment analysis, parse online reviews to assess risks or identify fraudulent fintech apps, and conduct real-time, on-chain analysis for digital assets supervision are just a few of many examples.

Taken whole, these insights frame a suptech space that is relatively nascent, but rapidly and necessarily accelerating to address the needs of supervisors in the face of novel and newly magnified risks introduced by a financial sector that is digitalising and generating supervisory data at an exponential rate. Addressing the needs of the ecosystem in an effective and equitable manner will require close collaboration between financial authorities, vendors, funders, educators, researchers, technologists, data scientists, and the rest of the suptech ecosystem.

This inaugural annual State of SupTech Report aims to feed that conversation and support collaboration, building a baseline against which to conduct agency and regional benchmarking, methodically tracking year-on-year trends and the growth of the suptech marketplace.

1.

SAMPLE, METHODOLOGY, AND TAXONOMY

1.1. Research methods

Three primary data sources were used to compile this report:

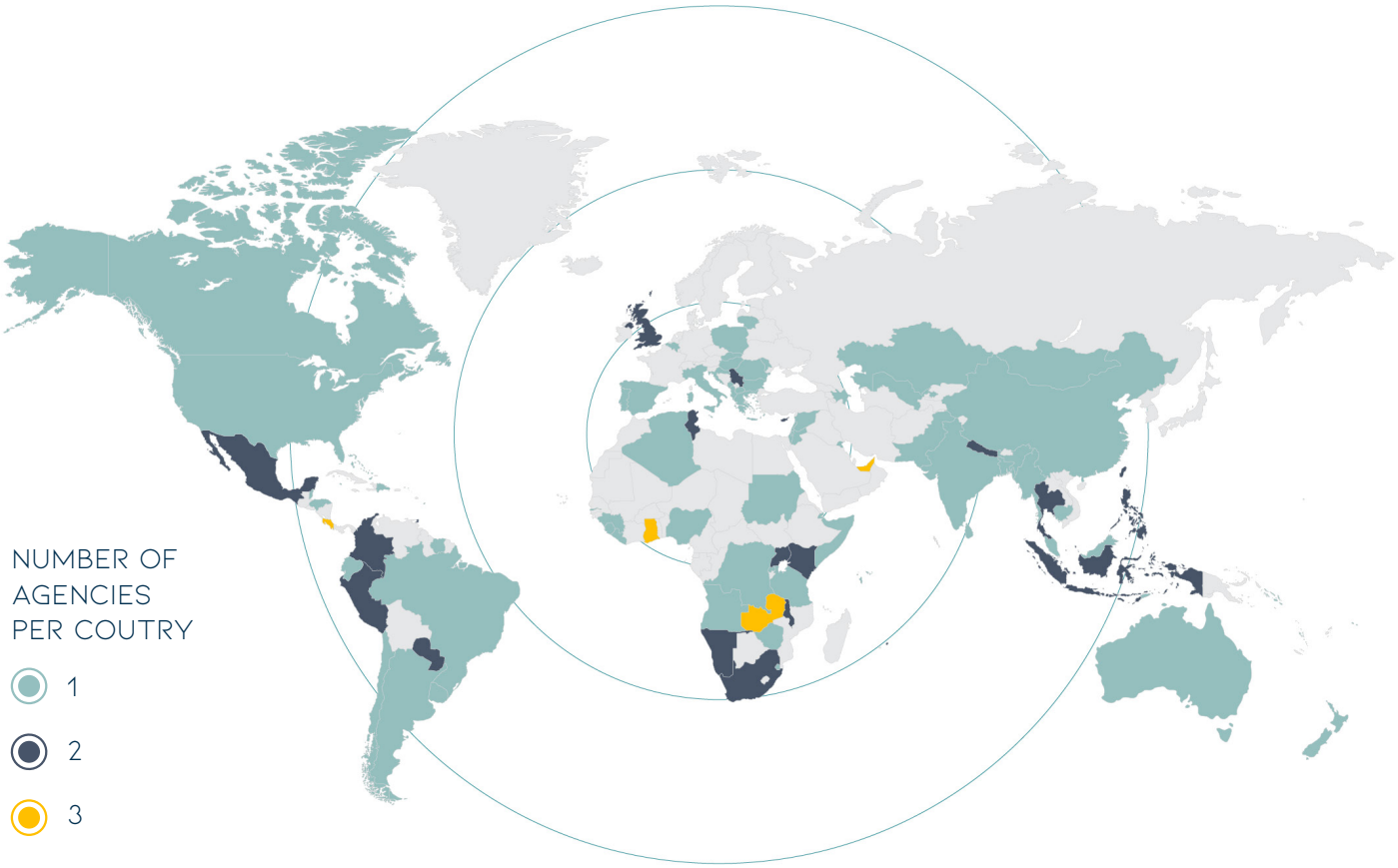
- A survey of 134 financial authorities from 108 jurisdictions
- A questionnaire for 74 individual supervisors (representing 46 agencies and 35 jurisdictions) on the specifics of supervisory data
- A questionnaire for six selected suptech vendors.

In addition, the Lab complemented these resources with qualitative interviews and case studies to further develop and test hypotheses arising from the quantitative data and more deeply understand the challenges and opportunities in adopting suptech applications.

1.1.1. Sample of financial authorities by geography and income classification

Most of the data presented in this Report were collected between May and October 2022 through a global survey conducted by the Cambridge SupTech Lab. The respondents include financial authorities such as central banks, securities and capital market authorities, financial conduct authorities, and insurance regulators. Of the 134 responses, 81 are from central banks, representing 60% of the total sample. 92 responses were received from agencies in emerging markets and developing economies (EMDEs), representing 67% of the responses. The remainder were from advanced economies (AEs).

FIGURE 1.
GEOGRAPHICAL DISTRIBUTION OF SURVEY RESPONDENTS



The final respondent sample is geographically diverse and representative of the World Bank country income groups.

Table 1 maps 108 geographic jurisdictions of the 134 financial authorities who responded to the survey. The complete list is available in Appendix 1.

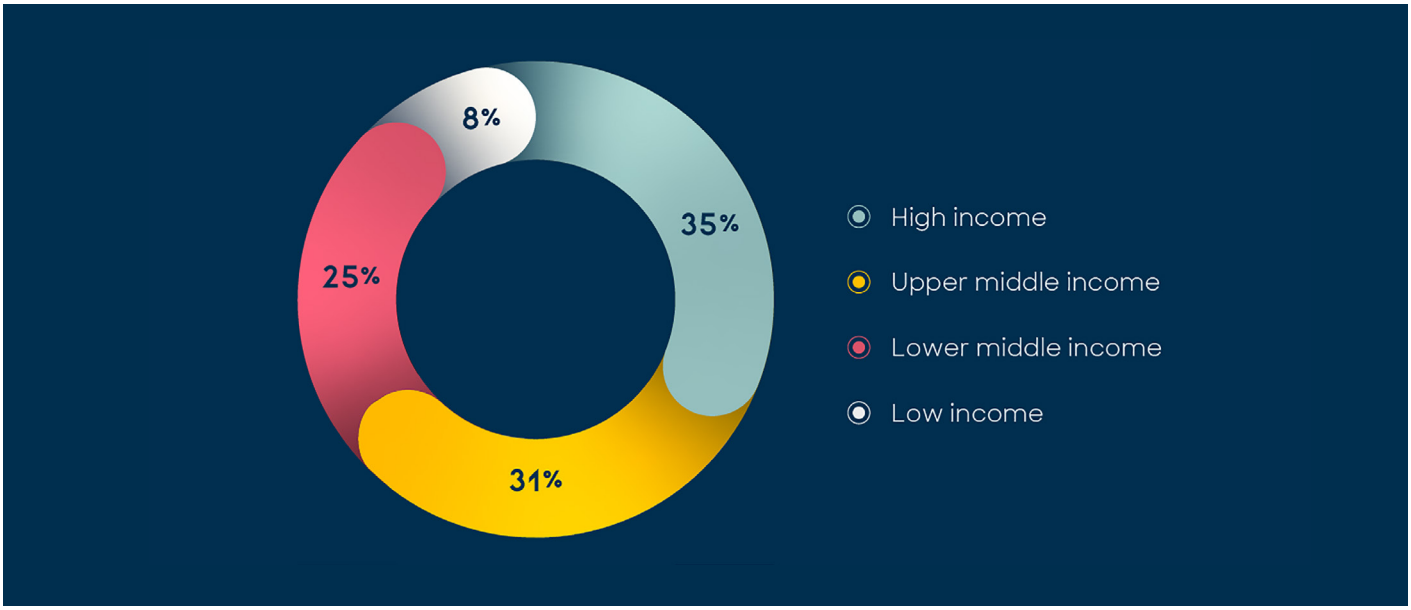
Figure 2 illustrates the response distribution according to the World Bank's classification by income level. The sample contains responses from jurisdictions across all four income classifications, with 55 responses from either low or lower-middle-income jurisdictions. In some areas of the analysis, we group these categories into EMDEs (low, lower-middle and upper-middle income) and AEs (high income).

TABLE 1.
GEOGRAPHICAL DISTRIBUTION OF RESPONDENTS BY REGION

REGION	NUMBER OF RESPONDENTS	PERCENTAGE OF SAMPLE BY REGION	PERCENTAGE OF JURISDICTIONS COVERED WITHIN REGION
East Asia and the Pacific	22	16%	46%
Europe and Central Asia	29	22%	41%
Latin America and the Caribbean	27	20%	44%
The Middle East and North Africa	14	10%	46%
North America	3	2%	100%
South Asia	6	5%	63%
Sub-Saharan Africa	33	25%	48%
Total	**134**		

* Income and region are based on the World Bank country classification. If a jurisdiction was not listed geographically, its classification was based on neighbouring jurisdictions.

FIGURE 2.
BREAKDOWN OF RESPONDENTS BY INCOME GROUP (N=134)



1.1.2. Questionnaire for financial authorities on specifics of supervisory data

In November 2022, we asked individual supervisors four questions on the specifics of supervisory data to further assess the state of data collection for financial supervision:

1. **Thematic areas:** the supervisory areas for which data is collected
2. **Channels:** the mechanisms and channels through which is collected
3. **Formats:** the digital format and structure of data that is collected
4. **Challenges:** the specific challenges faced at each layer of the supervisory data lifecycle stack

We received information from 74 supervisors representing 46 agencies and 35 jurisdictions. This sample included some agencies that did not participate in the primary survey.

1.1.3. Questionnaire for suptech vendors

To complement the insights shared by the demand side of the suptech market and develop a deeper understanding of the broader suptech ecosystem, we also engaged directly with six suptech vendors to discuss ten questions that characterise the opportunities, challenges and other qualitative characteristics of the market. The vendors were selected from the Cambridge SupTech Lab's SupTech Marketplace Vendor Database based on the following criteria:

- **Centricity of suptech in strategic focus**
  While some vendors provide suptech solutions as a small part of a broader portfolio of products and services, others focus primarily on suptech solutions. For this set of interviews, we prioritised the latter.
- **Maturity of offering**
  The sample prioritised vendors with a mature product or service to ensure actual experiences of operating in the market, rather than hypothetical or early-stage ideas based only on pilots or experiments.
- **Diversity of market position**
  The sample aimed to incorporate a range of market perspectives, including relatively new entrants (those who have only recently adapted their mature offering to address supervisory use cases) and those who have been working with supervisors since before the inception of the word 'suptech'.
- **Diversity of geographies where solutions are deployed**
  The sample aimed to capture experiences across a range of jurisdictions to avoid sample bias towards any one set of cultural norms or localised market restrictions.

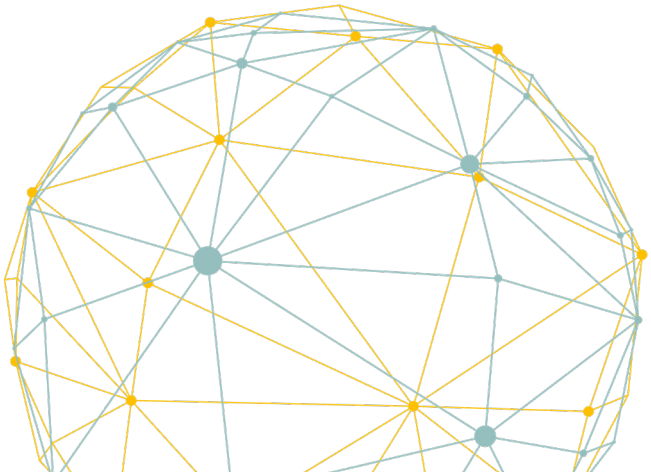
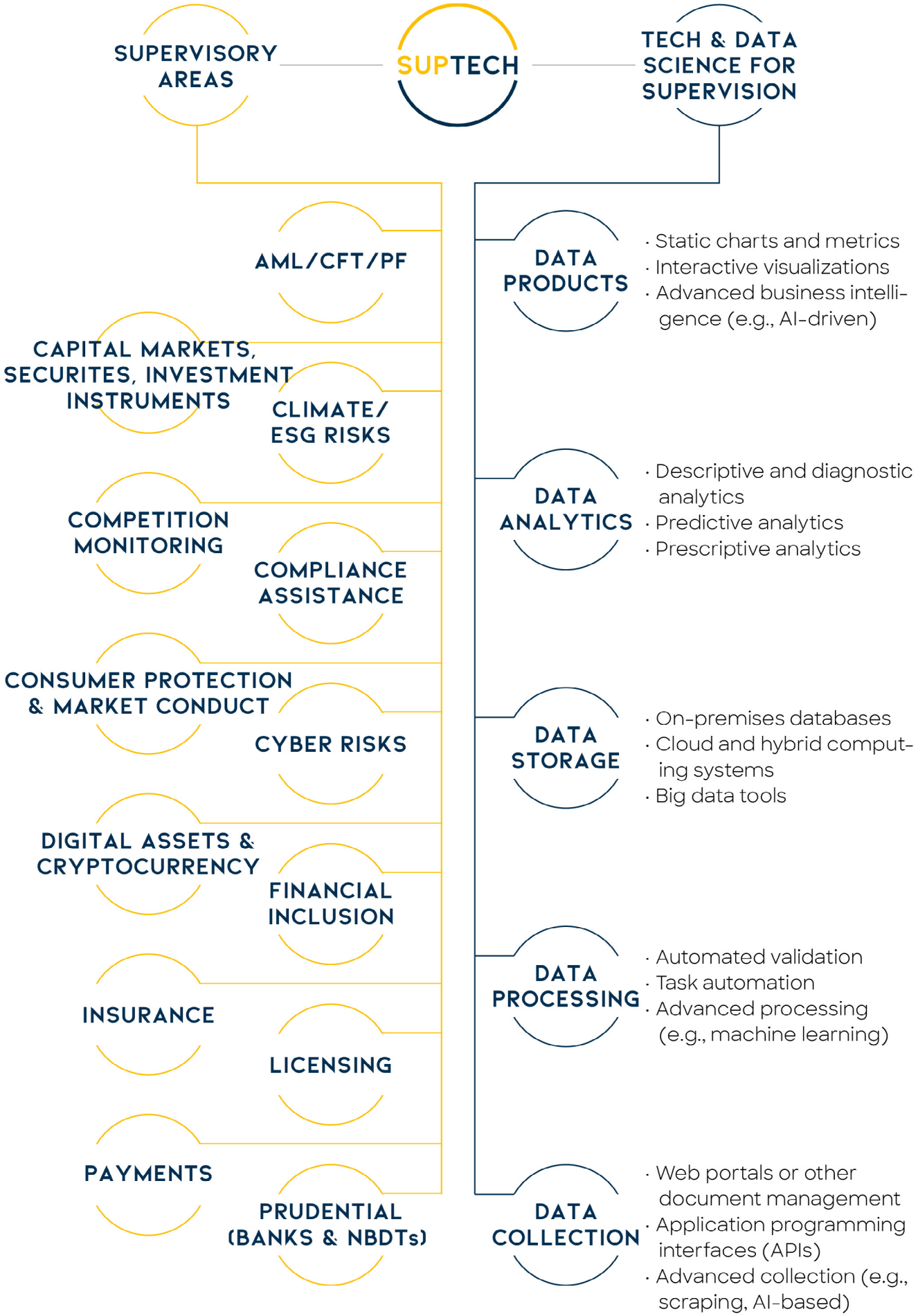


FIGURE 3:
SUPTECH TAXONOMY



1.2. Suptech taxonomy

The Cambridge SupTech Lab has developed a comprehensive classification system to consistently organise various entities – namely, suptech vendors, suptech solutions and suptech diagnostics – by supervisory use case (the 'sup' in suptech) and by the technologies and data science tools used (the 'tech').

This taxonomy is based on past efforts to map the space (BIS 2018, BIS 2019) and explicitly differentiates between the 'sup' and the 'tech'. This disaggregation affords a novel opportunity to systematically map the needs of supervisors, classify the tools serving those needs and ultimately serve as an ontology for strategically and intentionally connecting the solutions to needs. It was refined and validated through desk research, review of deployed suptech applications (see the Lab's SupTech Marketplace), and input from over 130 financial supervisors and leading suptech experts. The taxonomy will be periodically revised, based on internal research and external feedback, to reflect the suptech space's dynamic nature.

1.2.1. Supervisory areas and use cases

This first iteration of the taxonomy covers 13 broad supervisory categories subdivided into 87 use cases. The structure of the classification system is hierarchical and built on a conceptual framework that groups use cases according to the activities conducted by supervisory functions within authorities.

While thematic focus areas refer to policy or supervisory areas/activities, use cases refer to more specific tasks supported by identified suptech tools.

The 13 thematic focus areas are:

Anti-Money Laundering/Countering the Financing of Terrorism/Financing the Proliferation of Weapons of Mass Destruction (AML/CFT/PF) supervision
Suptech allows financial authorities to identify potentially suspicious customers or activities (for example, through customer due diligence and suspicious transactions detection) and enhances data analytics to monitor institutions' compliance and AML/CFT/PF risk management (for example, assisted/automated examination, metadata analytics, and text analytics).

Capital markets, securities and investments supervision
Suptech equips financial authorities to detect potential misconduct (for example, insider trading, market manipulation and poor disclosure) and enhances data analytics to monitor the capital markets (for example, automated examination, peer-group/risk classification and text analytics). Securities and investments use cases focus on empowering securities commissions and other financial authorities with a securities mandate to augment their capabilities by generating improved data-driven insights and detecting insider trading and market manipulation.

Climate/ESG risk supervision
Suptech enables financial authorities to enhance data collection and analytics to assess institutions' climate and environment, social and governance (ESG) risk management (for example, green market monitoring, peer-group/risk classification and stress testing).

Competition monitoring
Suptech focuses on monitoring market competition dynamics and rates and fees.

Compliance assistance
Suptech makes available automated compliance auditing and guidance for compliance queries.

Consumer protection and market conduct supervision (now referred to as **consumer protection**)
Suptech empowers financial authorities to enhance data collection (for example, advanced/real-time monitoring and data consolidation) and improve data analytics to monitor consumer risks and supervise market conduct (for example, assisted/automated examination, misconduct detection, peer-group/risk classification and text analytics). In addition, these use cases also support authorities in providing consumers with virtual assistance (for example, complaints handling and credit bureau rectification).

Cyber risk supervision
Suptech improves data analytics to monitor institutions' compliance and cyber risk management (for example, automated examination, assessment of vulnerabilities and compliance monitoring).

Digital assets supervision
Suptech is deployed to supervise cryptoassets or DLT-based protocols, platforms or systems (for example, cross-jurisdictional intelligence checks and information-sharing capacity, embedded supervision and on-chain analysis).

Financial inclusion
Suptech is used by financial authorities to monitor the access and use of financial services (for example, gender-based and geospatial analysis). These use cases can also collect consumer data (for example, consumer satisfaction analysis) and provide virtual assistance (for example, financial education tools).

Insurance supervision
Suptech serves some prudential supervision use cases that enable insurance supervisors to enhance data collection and data analytics. In addition, it covers use cases that allow insurance supervisors to provide virtual assistance to firms for procedures often required in the insurance sector (for example, registration of intermediaries and product registration).

Licensing
Suptech supports financial authorities providing virtual assistance to firms requesting a license or authorisation to operate within the regulatory perimeter (for example, automated guidance and automated processing of requests).

Payments oversight
Suptech assists financial authorities in monitoring and testing the performance of payments infrastructures, networks and systems (for example, advanced/real-time monitoring and stress testing).

Prudential supervision of banks and non-bank deposit-taking institutions (now referred to as **prudential supervision**)
Allows financial authorities to enhance data collection (for example, automated reporting, automated validation and data consolidation) and data analytics for both macroprudential and microprudential supervision (for example, assisted/automated examination, peer-group/risk classification and stress testing).

The complete list of suptech use cases grouped by thematic focus area is available in Appendix 2.

1.2.2. Technologies and data science tools in the supervisory stack

On the other side of the taxonomy in Figure 3 are the technologies and data science tools deployed to address authorities' challenges and realise the aspirations within the aforementioned supervisory areas and use cases. These technologies are classified by their applications within the context of the five layers of a supervisory 'stack' (R²A 2020):

Data collection
This is the layer where data is gathered. It is collected from entities, including supervised financial service providers, consumers of financial technologies and the general public, into the supervisor's domain. Examples of data collection

mechanisms used for supervision include web portals and other document management, application programming interfaces (APIs), advanced collection techniques including AI-based tools like chatbots, embedded supervision of distributed ledger technologies (DLT), and automated data gathering like web scraping and data streaming.

Data processing

As the data is being gathered, it should be validated, cleaned and consolidated using data processing tools to maximise its utility. Examples of data processing technologies in the supervisory context include integrated validation techniques like rules on APIs that send errors back to the submitting party in real time, task automation techniques like those programmed in scripting languages or recorded and replayed via robotic process automation (RPA), and advanced processing tools such as machine learning-based computer vision and natural language processing models to extract structured supervisory data from less structured sources.

Data storage

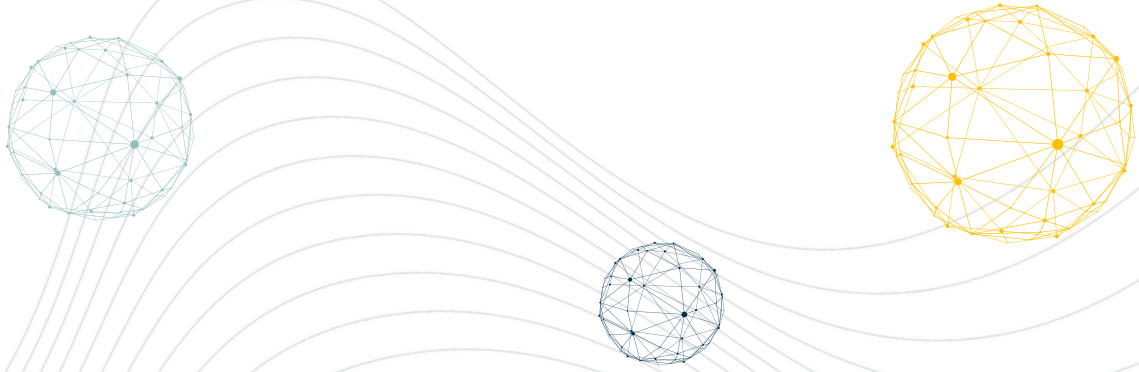
Once the data has been collected and processed, it needs to be stored in a manner that ensures security and ease of access across supervisory areas. Examples of storage methods for supervisory data include databases hosted and managed onsite by the financial authority itself, cloud and hybrid computing technologies that introduce the benefits of virtualisation, and big data tools such as data lakes and data warehouses.

Data analytics

With the data suitably stored, extracting insights can begin a process enabled by data analytics technologies. Examples of data analytics tools used by supervisors include descriptive and diagnostic analytics that summarise the current moment in time, predictive analytics that create statistical models from historical data to infer the most likely outcome in the future, and prescriptive analytics tools that use those predictions to recommend the most effective action the financial authority can take to optimise achieving their mandates and goals.

Data products

At the top of the stack are the products and interfaces that directly connect supervisors to the insights derived from the analytics. Examples of data products for financial authorities include charts and key metrics from static reporting tools, interactive visualisations and dashboards that allow deeper exploration and combinations of data, and advanced business intelligence tools that leverage artificial intelligence (AI) to deliver alerts proactively.



2.

EVOLUTION OF THE SUPTECH LANDSCAPE

2.1. Timeline of the digital transformation of financial supervision

The use of technology and data science for financial supervision and market monitoring has rapidly evolved over the past two decades.

In part, this evolution has been a conjunctural phenomenon, a response to events - sometimes endogenous, other times exogenous to the financial system - that have reshaped financial supervision. Such events include international terrorism in the 2000s, major financial scandals at the beginning of the same decade, the global financial crisis in 2008, and, more recently, the Covid-19 pandemic. Moreover, this evolution also reflects a structural shift connected to the digitisation of the financial market and the exploitation of big data by financial firms, driven by progress in technology and computing power, and their increased availability and affordability. Along with this progress comes the introduction and magnification of risks, such as cybersecurity and data privacy, which become ever more prominent with the advent of this digital era and proliferation of abundant digital financial data.

In this context, financial authorities have increasingly experienced a digital flood of supervisory data without being able to distill more intelligence to govern the financial sector. Therefore, supervisory agencies have started to re-engineer their institutional arrangements, rescope their mandates, review their risk management frameworks, readjust their methodologies, step up their data management and governance approaches, and enhance their competencies and capabilities to further their digital transformation. Notably, the suptech era appears to be only the most recent chapter in the broader anthology of tech-enabled innovation in financial supervision. This section frames suptech in that broader context, highlighting some of the key milestones along the ongoing journey towards a suptech-augmented, responsible and resilient approach to financial supervision.

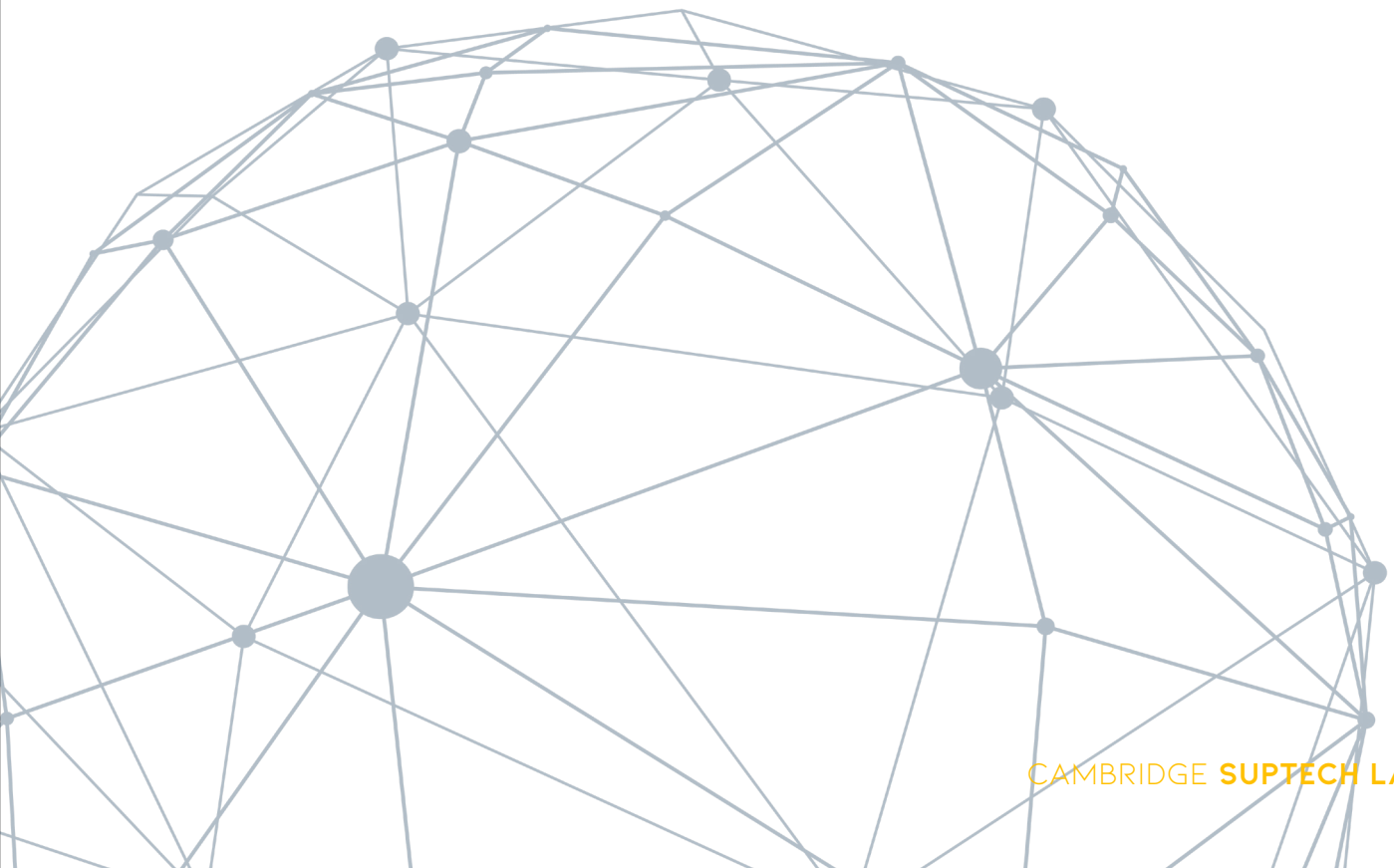
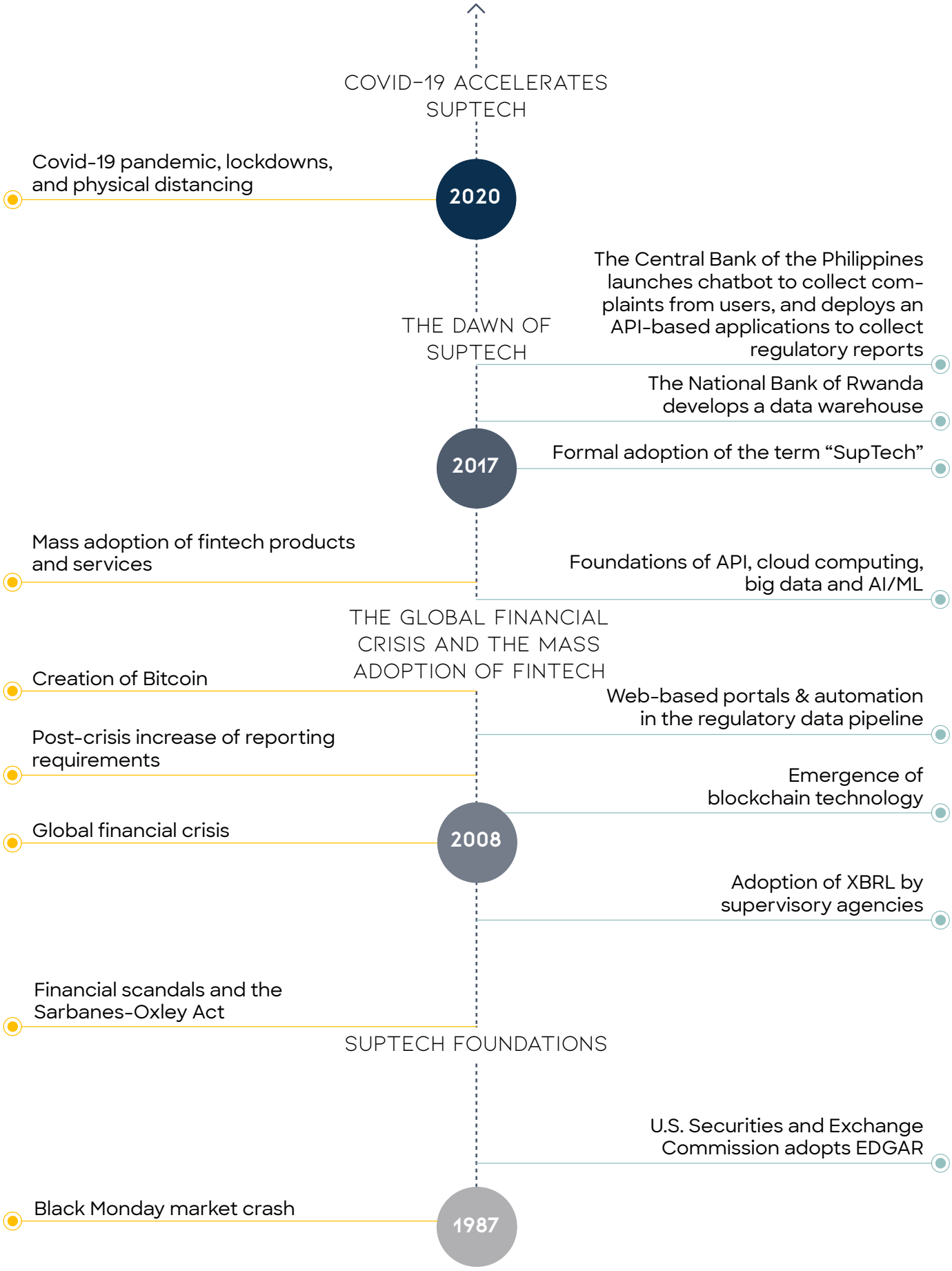


FIGURE 4.
TIMELINE: THE EVOLUTION OF SUPTECH



2.1.1. 1987–2007: Suptech foundations

After the Black Monday market crash of 1987, regulators and supervisors began digitising their operations to improve transparency and risk management in the financial markets. In 1993, the United States Securities and Exchange Commission (SEC) mandated electronic filing through its Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system to integrate digital disclosure requirements by supervised entities and enable monitoring by supervisors (ADB 2022). In 1994, a risk management tool called Value at Risk (VAR) was developed to compute capital requirements and assess market risks per the Basel Framework.

Through the 1990s and 2000s, financial authorities' data management workflows were still heavily manual. Data collection relied on paper forms, then emails, portals and Excel files. Data was stored in disjointed databases or folders. Validation occurred manually through spot checks, and the analysis primarily generated static management reports.

In the United States, the Sarbanes-Oxley Act, signed into law in 2002, marked an important moment because – following the crash of Enron, Tyco and WorldCom – after two decades of deregulation, the U.S. congress reaffirmed the need for financial accountability. The reform applied to every public company in the United States and had a huge impact on the acceleration of solutions for the storage and management of corporate electronic records.

In the second half of the 2000s, a wave of modernisation began when standard data reporting formats started being incorporated into supervisory processes and technologies. The incremental adoption of the XBRL format by supervisory authorities on a global scale - as reported by the United States SEC, the Capital Market Authority in Chile (XBRL 2009a), the Bank of Japan (XBRL 2009b), the Reserve Bank of India (XBRL 2009c) and others - marked a critical moment in the suptech timeline.

2.1.2. 2008–2016: The global financial crisis and the mass adoption of fintech

The 2008 global financial crisis triggered a seismic shift in the financial landscape. The reporting and supervisory mechanisms were insufficient and did not allow financial authorities to detect the irresponsible actions of financial institutions through their predatory lending and subprime mortgage practices. As a result, many people were left financially exposed and dejected while taxpayers had to bear the cost of bank bailouts.

Authorities responded with stringent regulatory reforms that increased reporting requirements (BIS 2018). This encouraged the industry to develop new technologies to streamline regulatory reporting. On the other end, supervisory agencies increasingly automated their procedures for data collection and management, adopting web-based portals or bulk uploads that allowed financial firms to file regulatory returns accompanied by inbuilt automated validation checks when uploading. The post 2008 crisis years earmarked another heavy increase in the number and complexity of templates to be reported for prudential purposes, with banks required to report layers upon layers of harmonised reporting templates in digital formats (after the 2014 reforms, 8,000 European banks were reporting up to 700,000 data points quarterly, while in 2018 HSBC announced it was linking up 10 petabytes of data from over 300 data sources in a data lake, which is bigger than the entire internet was at the turn of the century) and therefore making major investments in the development of regulatory technologies (regtech) solutions to handle compliance more efficiently.

Groundbreaking approaches to applying advanced technologies to large financial datasets also emerged during this period, such as business intelligence through integrated management of

micro-databases (BIS 2014), social media sentiment analysis to monitor consumer confidence (ECB 2014), and extensive set of publications around the big data strategies for central banks as an explicit reaction to the data revolution (BOE 2015, Riksbank 2015). The resulting larger volume of collected data was translated into dynamic data visualisation through business intelligence dashboards and diagnostic analytics (for example, via scorecards), enabling richer insights.

In addition, the mass adoption of fintech in the aftermath of the crisis pushed supervisory agencies to readjust their supervisory processes to keep pace with innovations in the financial sector. Mass adoption of mobile phones in EMDEs and online shopping in AEs propelled the success of innovative products such as mobile money products in Africa (e.g., mPesa) and mobile banking (e.g., PayPal) in the United States. Advances in technology allowed start-ups and firms outside the traditional financial sector to develop disruptive business models, such as crowdfunding and peer-to-peer lending. Furthermore, decentralised finance, a new paradigm based on blockchain technology, emerged with the creation of bitcoin. The substantial rise in the volume and availability of data produced concerning digital and traditional markets prompted financial authorities to seek technological solutions that could support them

2.1.3. 2017–2019: The dawn of suptech

While the use of innovative technologies for supervisory purposes accelerated, it was only in 2017 that the term 'suptech' was introduced more formally into the conversation by Ravi Menon, Managing Director at the Monetary Authority of Singapore (MAS), to refer to supervisory technologies. In his view, technological innovation was necessary for financial authorities to reduce inefficiencies and make supervision more effective (Menon 2017). Like the MAS, many other authorities began to adopt an institutional approach toward suptech, which became an object of interest to global standard-setting bodies.

In 2017, the Basel Committee on Banking Supervision (BCBS) recommended that supervisors should consider exploring the potential of new technologies, such as artificial intelligence, machine learning, distributed ledger technology, cloud computing and application programming interface (API), to improve their methods and processes (BCBS 2017). In 2018 and 2019, the BIS' FSI published two seminal papers that provided an overview of the developments in the suptech ecosystem. The first (BIS 2018) explored the experiences of ten early suptech users, highlighting the benefits and challenges and the implications for supervisors. They also proposed a taxonomy of areas of financial supervision in which suptech applications are used. This suptech taxonomy was slightly expanded in the second (BIS 2019), which examined a sample of 99 suptech initiatives and traced the evolution of the different generations of technology used by financial authorities, namely the "SupTech Generations".

A new generation of suptech applications went into production in this period, powered by API and natural language processing (NLP), and under the leadership of EMDE financial authorities. In 2018, the National Bank of Rwanda deployed an electronic data warehouse that pulls data directly from the IT systems of 600 supervised financial institutions, including commercial banks, insurance companies, microfinance institutions, pension funds, forex bureaus, telecom operators and money transfer operators. Meanwhile, the RegTech for Regulators Accelerator (R²A) partnered with the Bangko Sentral ng Pilipinas (BSP) and the Mexican Comisión Nacional Bancaria y de Valores (CNBV) and three technology firms to test and develop three suptech prototypes, namely an API-based prudential reporting system (R²A 2018a), a chatbot application and consumer complaint management system (R²A 2018b), and

an AML data infrastructure and advanced analytics solution (R²A 2018c).

2.1.4. 2020-present: Covid-19 accelerates suptech

Building on the experiments of the previous decade, and consistently with the digital transformation of most industries and sectors of the economy, financial authorities are investing in the development of suptech applications. Some recent innovations are quickly scaling and being adopted in multiple countries - e.g., chatbots for consumer protection were deployed in 2022 in Rwanda and Ghana while a number of other countries have put them into production, while APIs to collect and validate industry's data will soon become a standard. The budgetary section of the European Central Bank (ECB) Annual Report on supervisory activities (ECB 2021) is indicative of the increasing pace and magnitude of the suptech transformation (di Castri 2022): *"The 7.9% increase in expenditure compared with 2020 mainly reflects the onboarding of new IT systems dedicated to banking supervision. […] With respect to the developments in IT systems, the increases in expenditure seen in the policy, advisory and regulatory functions [16.1%] as well as macroprudential tasks [80.6%] relate to the SSM information management system (IMAS) and the Stress Test Account Reporting platform (STAR). The main increase in expenditure in the Supervisory Board, secretariat and supervisory law section [43%] resulted from significant investment in 2021 in supervisory technologies (suptech), which exploit the potential of artificial intelligence and other pioneering suptech in the context of banking supervision, for internal and external stakeholders."*

Covid-19 has further accelerated this modernisation push. During the pandemic, many jurisdictions implemented measures such as lockdowns and physical distancing to reduce physical contact, which necessitated the urgent acceleration of digital financial services for the payment of goods and services, welfare transfers, etc. Digital transactions increased dramatically and, unlike in 2008 when financial innovation was at the centre of the crisis, technology and finance were instrumental in responding to the crisis and supporting the recovery (ADB 2022).

One of the most significant challenges for financial authorities during the pandemic was adapting their processes to remote working. On-site supervisory activities and interactions with financial institutions and their staff had to move online. This made necessary the upgrade of agencies' infrastructure to allow for remote access to databases, and accelerated the development of new suptech tools for qualitative scrutiny and risk assessments (BIS 2021). The use of NLP helped automate the review of voluminous documents to identify corporate governance and credit risks.

Additionally, analysis of responses to our 2022 survey shows that financial authorities' perception of increased consumer risk due to Covid-19 is correlated with their adoption and development of suptech applications. 62% of those who had reported increased risks had an initiative or were developing an application, compared to just 44% of those who had not reported increased risks. A further 26% of respondents who identified increasing risks had taken no action to improve their supervisory capabilities using suptech versus 41% who did not identify increased consumer risk.

This is in line with previous research findings that suptech adoption was accelerated by the Covid-19 pandemic, which necessitated a shift towards off-site supervision (CCAF and World Bank 2020, World Bank 2020) and an increased focus on using suptech applications to better address increased consumer vulnerabilities.

3.

THE STATE OF SUPTECH

3.1. Demand: Financial authorities

3.1.1. Adoption

71% of financial authorities have suptech initiatives.

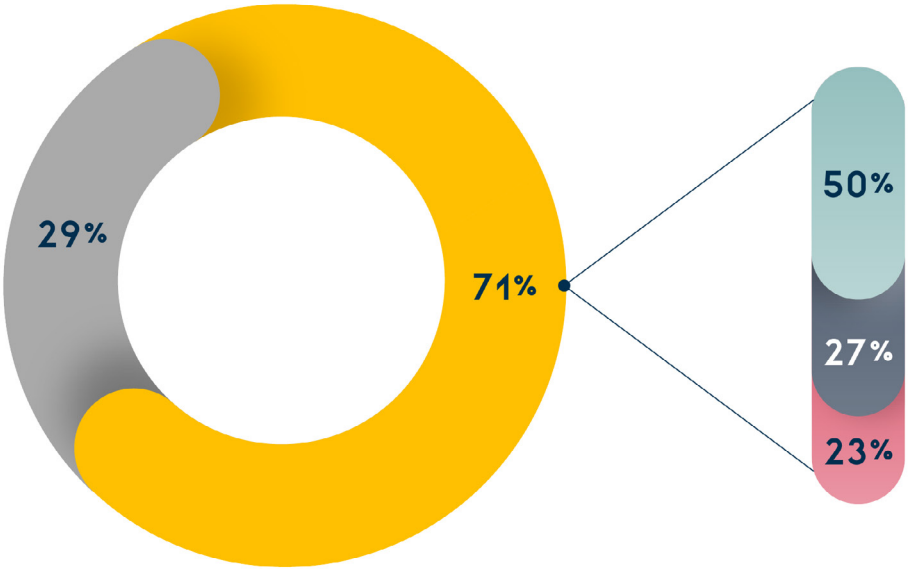
While suptech development is still at a nascent stage with room for growth, the survey results indicate that financial authorities are rising to the challenge as we observe increased adoption and use of suptech solutions.

71% of the financial authorities surveyed indicated that they have already engaged in different suptech efforts. Half of them have one or more suptech applications in operation (50%), and the other half have either an application in development (27%) or a suptech strategy or roadmap (23%).

Financial authorities in advanced economies are early adopters.

While advances in suptech adoption are evident worldwide, financial authorities from AEs are early adopters of suptech applications. 50% of financial authorities who stated they had already deployed suptech applications were from AEs, compared to 31% from EMDEs.

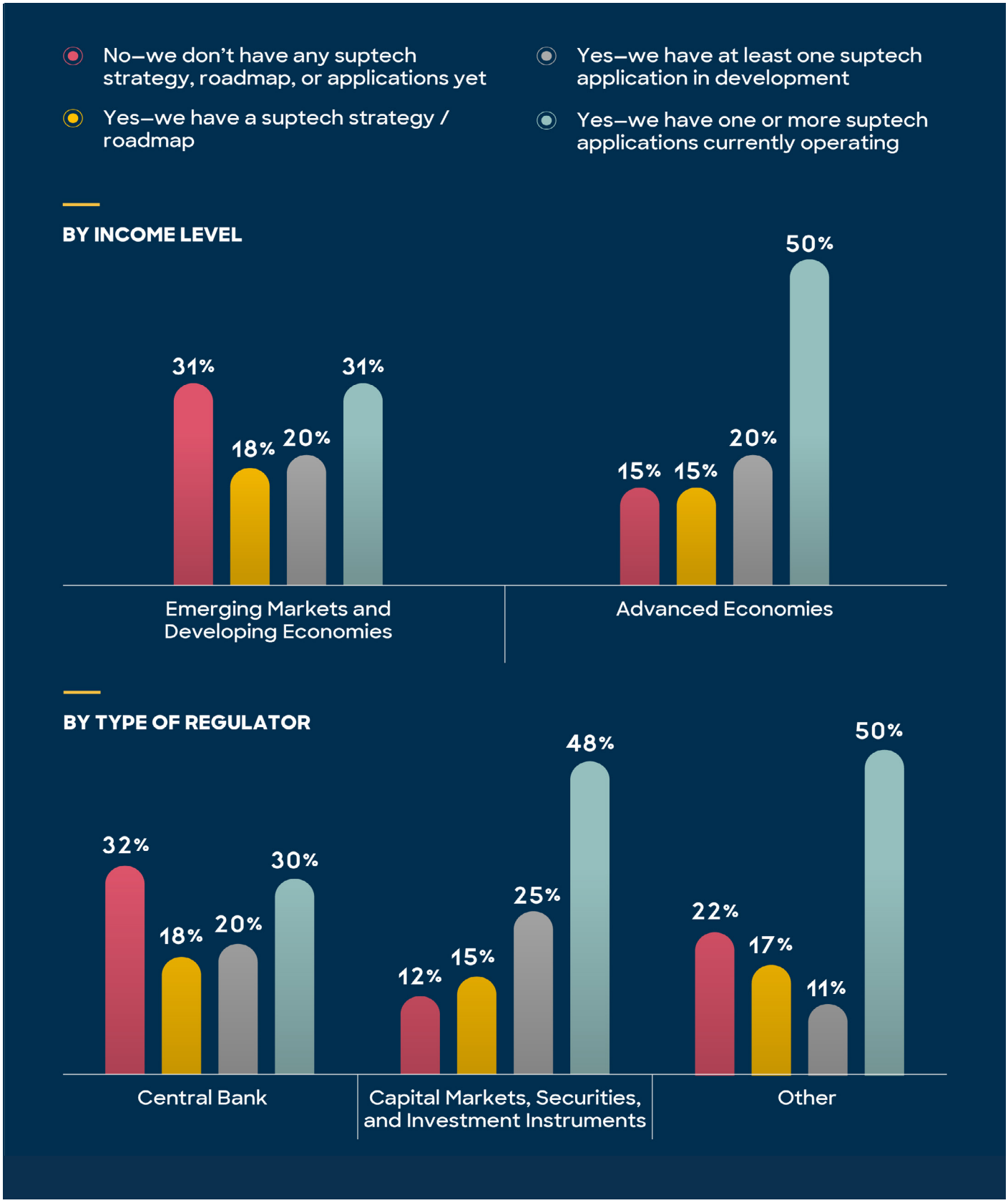
FIGURE 5.
RESPONSES TO THE QUESTION 'DO YOU HAVE ANY INITIATIVES IN YOUR ORGANISATION THAT YOU WOULD CONSIDER AS SUPTECH?' (N=134)



- Yes
  - Yes–we have one or more suptech applications currently operating
  - Yes–we have at least one suptech application in development
  - Yes–we have a suptech strategy/roadmap
- No–we don't have any suptech strategy, roadmap, or applications yet

FIGURE 6.

RESPONSES TO THE QUESTION 'DO YOU HAVE ANY INITIATIVES IN YOUR ORGANISATION THAT YOU WOULD CONSIDER AS SUPTECH?', SEGMENTED BY INCOME LEVEL (N=134)



Securities supervisors lead in suptech adoption. Central banks lag behind.

Securities industry supervisors take the lead in suptech adoption, while central banks lag behind. The reason might be that the formers rely more on off-site supervision than central banks, and up-to-date analytics software is needed to analyse massive volumes of transaction data. For example, the Australian Securities and Investments Commission (ASIC) has adopted suptech to transform datasets into usable patterns for market surveillance and suspicious trading detection.

Securities supervisors' core risks are of conduct nature rather than prudential. Often, these risks are evidenced in the processing of large unstructured information documents such as compliance manuals and internal policies, where new technologies allow supervisors to extrapolate content from different firms to obtain an industry-level view that was previously just sensed by them.

The three most adopted supervisory technologies are descriptive and diagnostic analytics, web portals and document management, and APIs.

Investment in suptech applications primarily occurs at the data-collection layer of the supervisory stack, followed by data analytics and data storage.

FIGURE 7.

UNDERPINNING TECHNOLOGIES *USED* BY AGENCIES TO ENABLE SUPERVISORY PROCESSES (N=134)

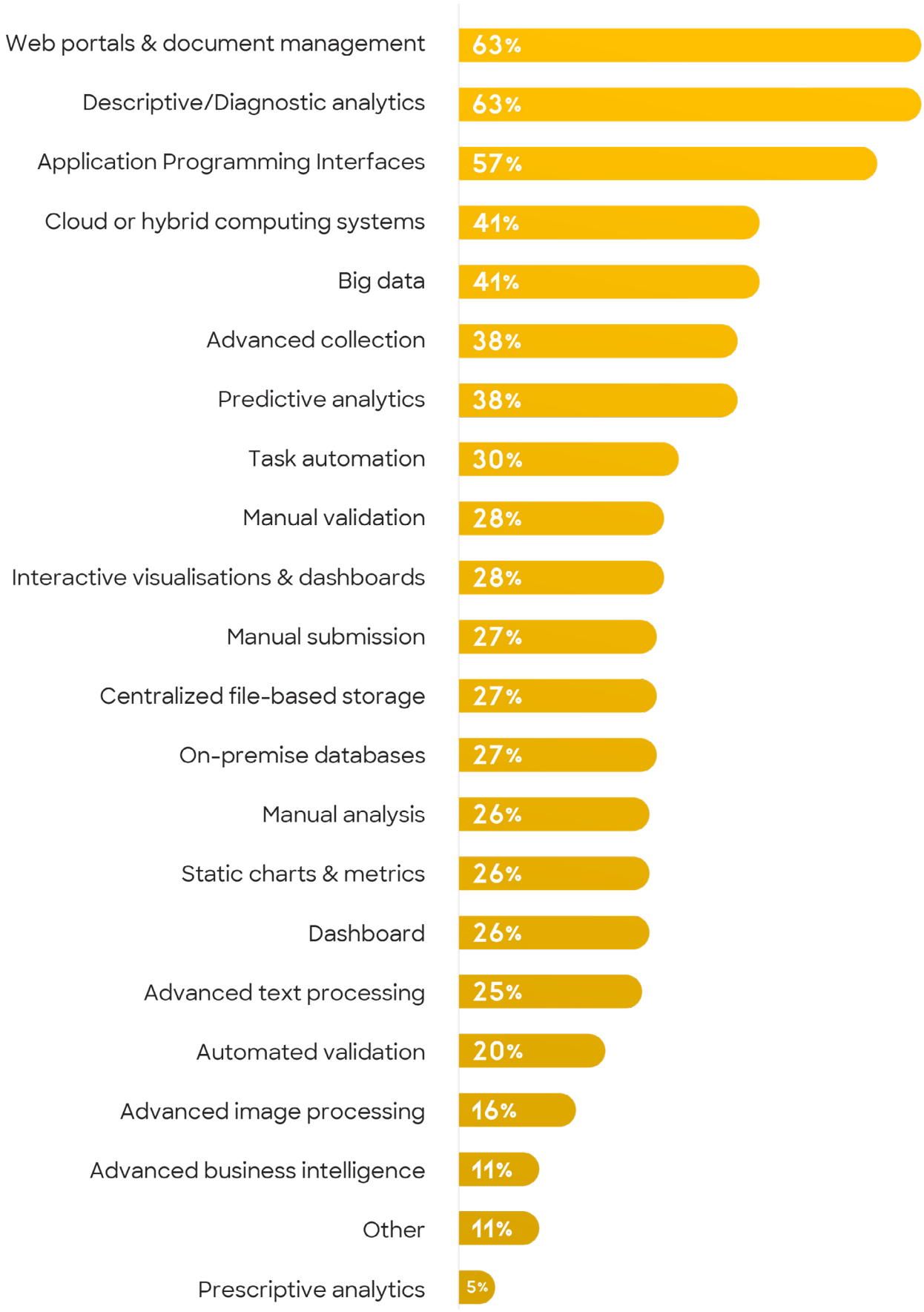
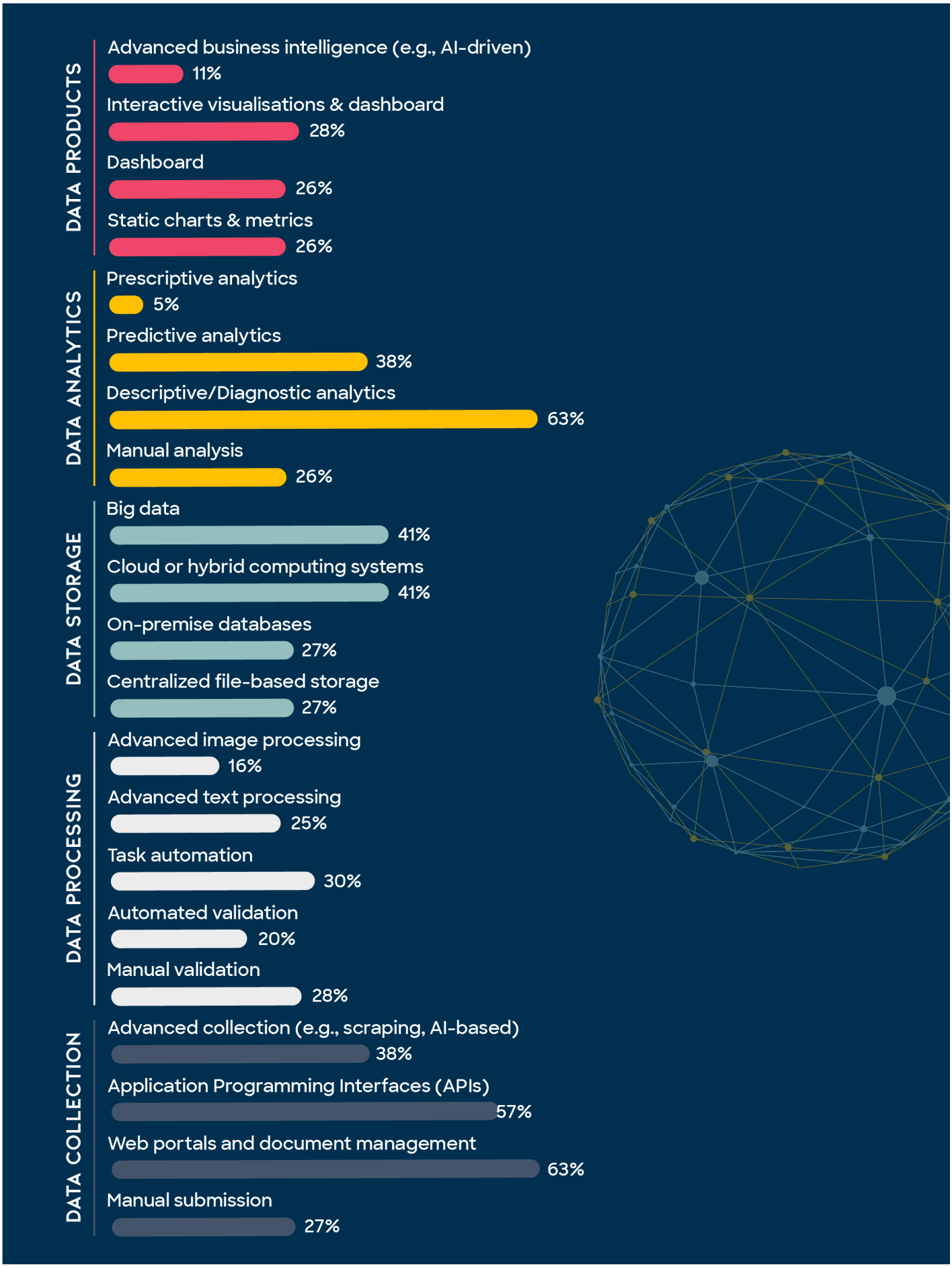


FIGURE 8.

UNDERPINNING TECHNOLOGIES *USED* BY AGENCIES TO ENABLE SUPERVISORY PROCESSES, ORGANISED BY LAYER OF THE SUPERVISORY STACK OF THE SUPTECH TAXONOMY (N=134)



3.1.2. Gaps

Even with the adoption trends noted above, there is a substantial demand for improvement that arises from factors such as:

COMPOSITION

Each suptech application is but one component of a larger system, and as such the whole supervisory stack can be composed of suptech building blocks. In this sense, one suptech solution often begets the next. For example:

- A suptech data collection mechanism that processes higher volumes of data, then demands a storage mechanism within which to store them.
- In turn, a more advanced suptech data storage solution offers more accessible and robust datasets, which unlocks a stronger demand for analysis.

EVOLUTION

An agency may choose to undergo digital transformation using less advanced technologies to soften the learning curve and reduce the perceived costs of change management. Once the technology is adopted and a cultural shift to digital-first occurs, demand for more advanced technologies may arise. This evolution is reflected in more detail in the "SupTech Generations" section below.

Financial authorities seek to push the envelope, with desire for prescriptive analytics, task automation, advanced image processing, predictive analytics and advanced collection techniques.

86%

want prescriptive analytics, which use data to guide them on what actions to take based on historical data.

81%

want task automation to record and replay tasks on a supervisor's behalf to save time.

79%

want advanced image processing, such as computer vision in general, or more specific components like optical character recognition.

78%

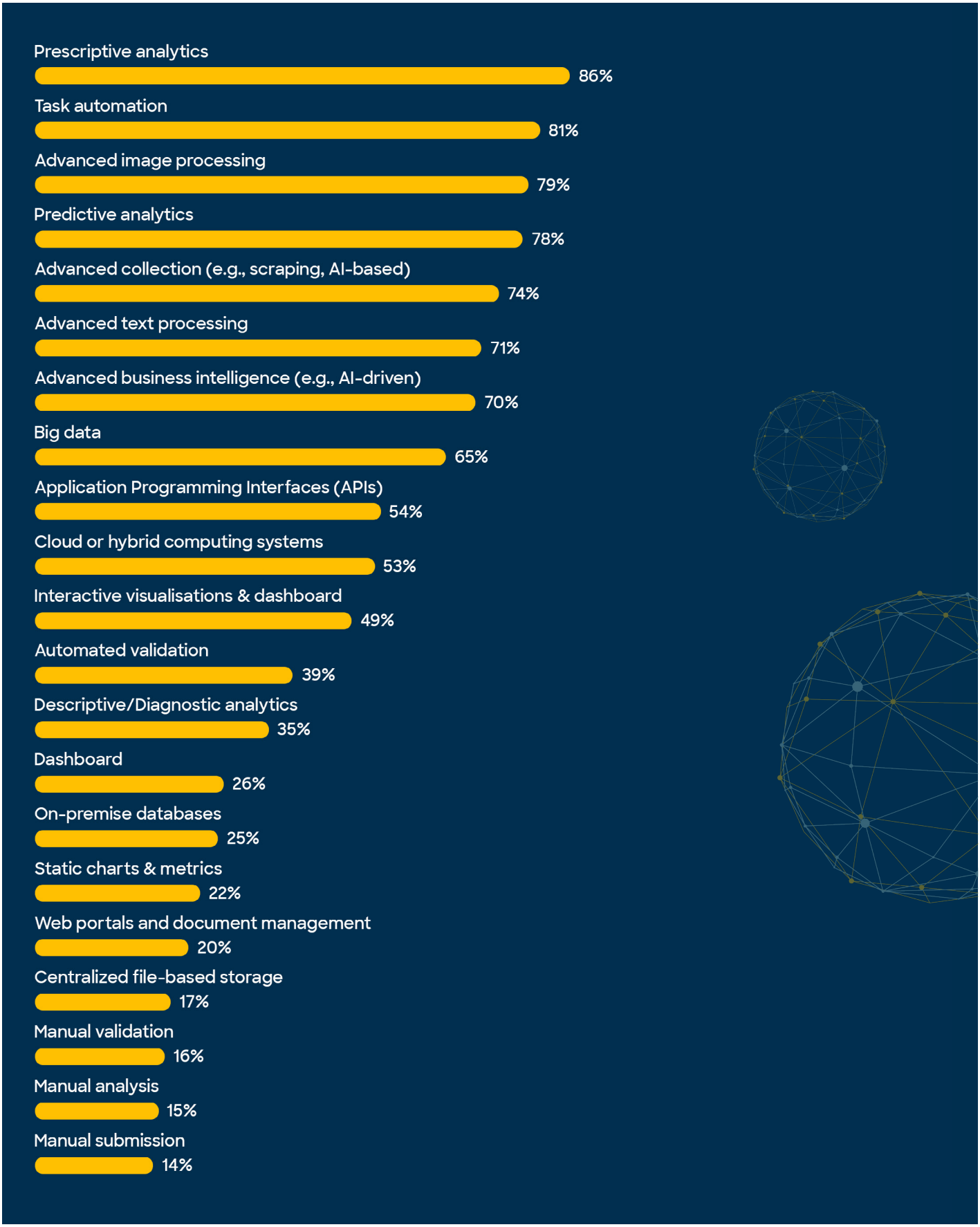
want predictive analytics, which can take past trends and forecast what's to come.

74%

want advanced data collection capabilities, such as web scraping or AI-based collection tools like chatbots.

FIGURE 9.

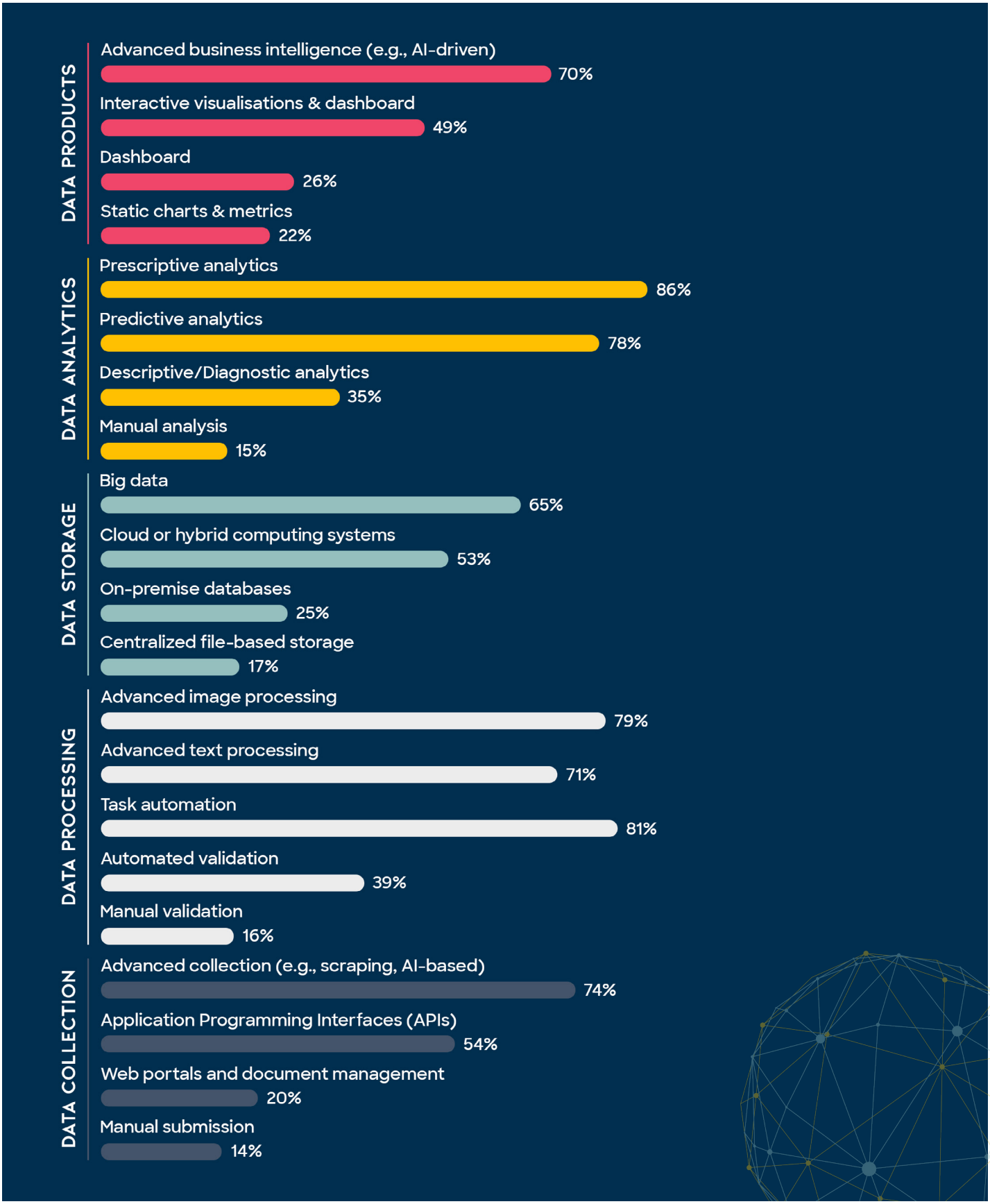
UNDERPINNING TECHNOLOGIES CURRENTLY *DESIRED* BY FINANCIAL AUTHORITIES TO ENABLE AND ENHANCE SUPERVISORY PROCESSES (N=134)



The supervisory stack layer where demand for new suptech is highest is the data processing layer, followed by advanced analytics/collection and advanced business intelligence.

FIGURE 10.

UNDERPINNING TECHNOLOGIES *DESIRED* BY AGENCIES TO ENABLE SUPERVISORY PROCESSES, ORGANISED BY LAYER OF THE SUPERVISORY STACK OF THE SUPTECH TAXONOMY (N=134)



Given the lack of existing adoption of data processing tools and data products noted in the previous section, it is unsurprising to see a high demand to fill these gaps with advanced image processing (79%), advanced text processing (71%), task automation (81%), and advanced business intelligence tools (70%). Notably, authorities also want to build on existing technologies, introducing more advanced versions of their analytics and collection tools.

3.1.3. SupTech Generations 2.0

A long-term plan for building suptech capabilities inevitably requires that the organisation sets out the detailed steps by which its current capabilities will be expanded; these must be incremental, realistic transitions that leave the organisation with greater supervisory capabilities than before, even in the relatively short term. BIS (2019) provided a 'four generations' framework to demonstrate how each technology-enabled element of the supervisory process might evolve through time.

This framework now serves as an ontology to frame the primary research here on enabling suptech, wherein we extended the suptech generations framework to include a 'generation zero' of manual processes and an additional layer of data products, as presented in Figure 11 and referenced throughout the survey results in this section.

FIGURE 11.
SUPTECH GENERATIONS 2.0 (CAMBRIDGE SUPTECH LAB 2022 VERSION)

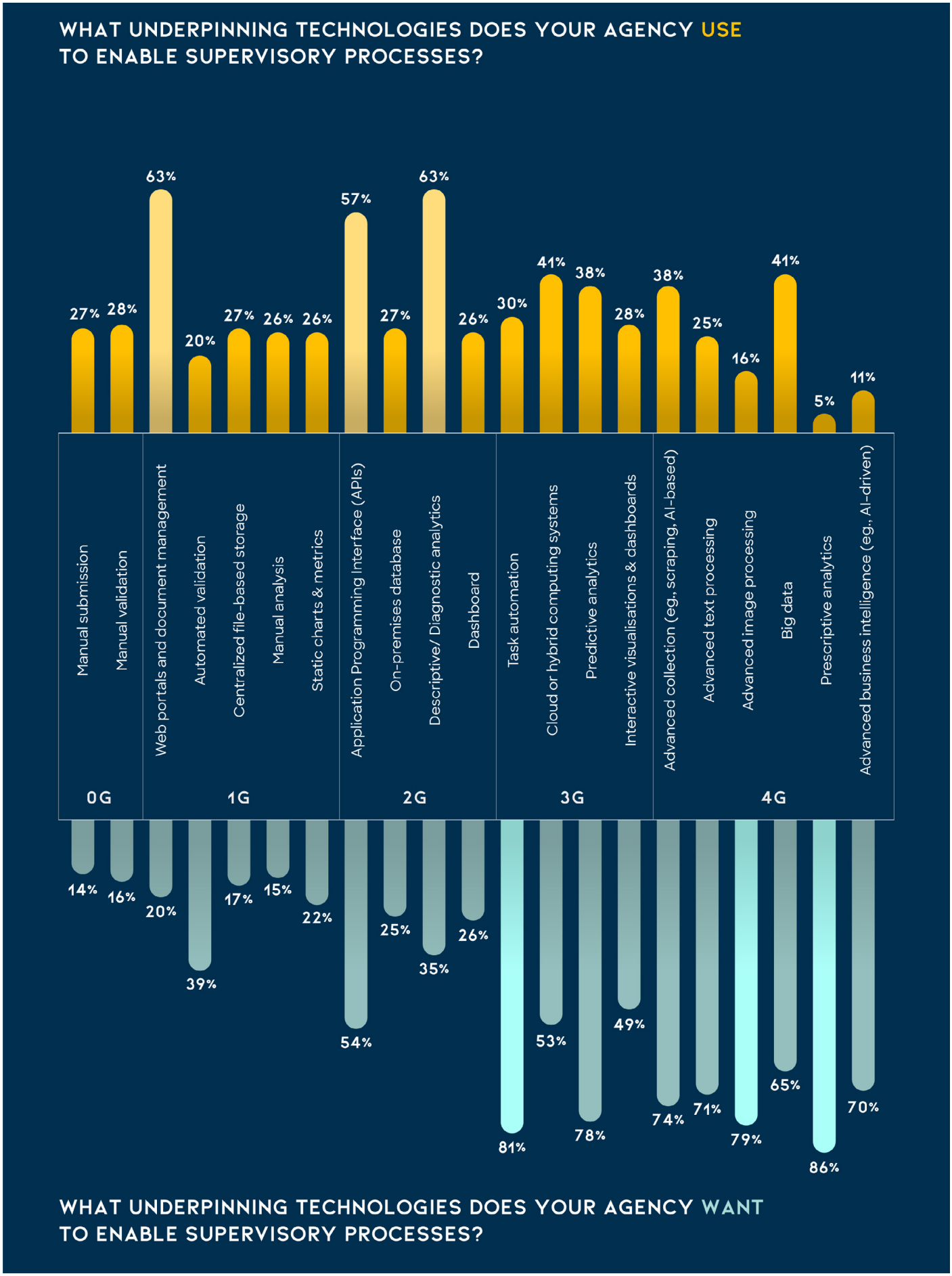
Suptech efforts remain in the experimentation stage. A vast majority of technologies used in the supervisory stack of surveyed authorities fall into the first (1G) or second (2G) generations of suptech, which mainly support descriptive and diagnostic analytics.

Demand for new suptech is highest for the most advanced 3G and 4G technologies, and decreases with each lower generation.

This clear trend validates the adapted SupTech Generations framework not only as a descriptor for the current state of suptech but equally as a roadmap for future suptech adoption. Supervisors with existing technology express a clear desire to upgrade, and those without advanced suptech may have opportunities to leapfrog by skipping earlier generations.

FIGURE 12:

UNDERPINNING TECHNOLOGIES USED VERSUS DESIRED BY AGENCIES TO ENABLE SUPERVISORY PROCESSES, ORGANISED BY SUPTECH GENERATION (N=134)



3.1.3.1. Data collection

Regulatory reporting can be challenging and resource-intensive for supervised institutions and financial authorities. As regulatory reporting has become increasingly complex, authorities face challenges in collecting delayed and poor-quality reporting data, which can, in turn, impact their supervisory ability (FCA 2020).

63% of authorities collect data through 1G web-based portals or bulk uploads.

Financial authorities collect this data periodically in the form of standard reporting templates. Hence, their focus has been on creating templates rather than using the data to construct analytical reports. A large proportion of the agencies indicated they use a combination of manual, bulk (web) uploads and automated reporting to collect data.

57% of the surveyed financial authorities have 2G application programming interfaces (APIs).

In a move to facilitate more efficient data flow between supervised institutions and supervisors and thus lower the costs, more than half indicated having automated their data collection process and have developed an API that allows institutions to submit data. While web-based portals and APIs can support the submission of large amounts of structured data, they are ineffective for unstructured data, such as social media and annual reports data.

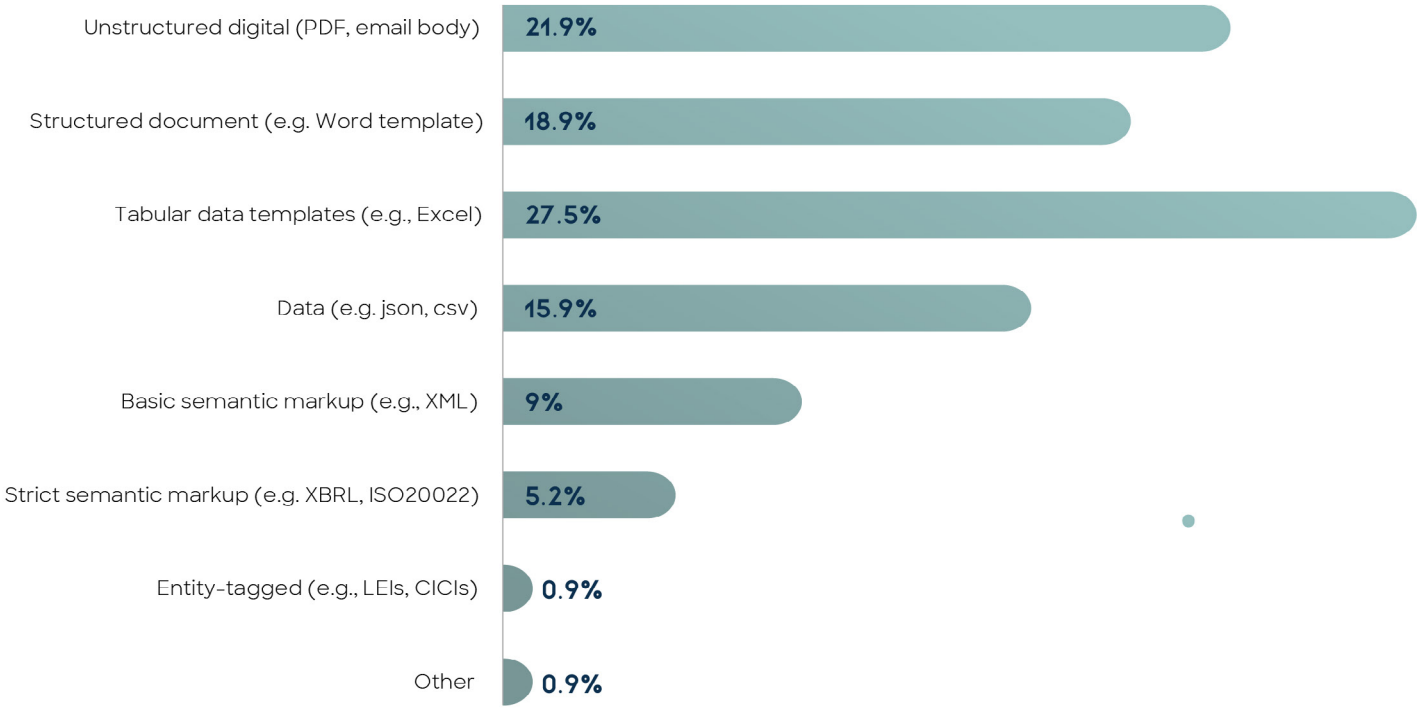
Only 38% of the respondents have 4G advanced collection techniques.

This suggests an opportunity for the development of AI-based chatbots, web scraping and data streaming technologies.

Most supervisors collect data in formats that necessitate manual processing.

Supervisors who collect data via unstructured digital formats (21.9%), structured flat files (18.9%) and tabular data templates (27.5%) are required to do at least some manual processing, whereas less than one-third of respondents (30.1%) use formats that are machine-readable by default.

FIGURE 13.
DATA FORMATS IN WHICH SUPERVISORY DATA IS COLLECTED BY FINANCIAL AUTHORITIES, ORDERED FROM LEAST (TOP) TO MOST (BOTTOM) MACHINE-READABLE (N=74)



3.1.3.2. Data processing

Efficient and reliable mechanisms for ensuring quality in data management are fundamental to the supervisory process. The data management cycle has two main tasks related to data processing (BIS 2019):

- **VALIDATION**
  The quality control checks of completeness, correctness and consistency of formatting and calculation as per reporting rules.

- **CONSOLIDATION**
  The integration of data from multiple sources and in varying formats.

Many financial authorities still validate data manually through time-consuming and error-prone 0G manual 'spot checks' or 1G spreadsheet-based formulas.

The effectiveness of financial supervision relies on efficient data management to provide timely, adequate and accurate data, covering various facets of financial institutions' business as well as integrating, where necessary, relevant macro data about the markets and the economy.

A large proportion of the survey respondents still validate data manually through time-consuming and error-prone manual 'spot checks' (28%) or spreadsheet-based formula (20%), which are unsuited for working with large datasets. The importance of upgrading these methodologies and tools to deliver higher-quality data cannot be overemphasised. A good model cannot overcome inaccurate data, and good-quality data is better than more data. It is, therefore, critical that financial authorities transition to automated data validation as this ultimately reduces human errors, builds up databases needed for performing analytical work, and enables scarce human resources to be reallocated to more judgement-based work.

3.1.3.3. Data storage

Increased regulatory requirements and the growth of the digital economy have led to a huge increase in available data. The integration of 4G suptech to aid in processing this big data has been noted as a challenge for supervisors – both to prepare for and integrate beyond the pilot phase – for some time (BIS 2015b). This trend continues today, and is likely driven by the challenges (with mitigation strategies) detailed in Section 4 of this Report.

An intermediate 3G solution to address the rising cost of big data storage is to use cloud technology, enabling greater and more flexible storage, mobility capacity and computing power. Financial authorities have subsequently increased the type of data collected for supervisory purposes and reap the benefits of lower costs and increased storage capacity.

However, challenges remain for supervisors in the adoption of cloud computing. First, financial authorities perceive fundamental limitations and risks when weighed against the current policy, organisational, technical, and legal structures (BIS 2018). Secondly, there are perceived operational risks, cyber risks, dependency and vendor lock-in, data sovereignty and concentration (TC 2020). Finally, even if those perceived risks and limitations are sufficiently mitigated, there is a need for stronger oversight of regulatory data stored in the cloud (FSB 2020).

Financial authorities must put these risks in context and compare not against perfection or an idealised state, but rather against the status quo (HBS 2020). Notably, when treated in this manner, the risks of not moving to the cloud could be higher than making the move (TC 2020).

Many agencies use 1G file-based storage mechanisms or 2G on-premise relational databases.

Despite the benefits of cloud computing, only a minority of respondents have adopted scalable storage solutions such as cloud computing (41%) or advanced document management systems such as data lakes and other warehouse techniques (41%). The majority of financial authorities still store their data centrally, or have fragmented, disjointed data management through spreadsheets, desktop databases or paper records..

3.1.3.4. Data analytics

Suptech applications support financial authorities in analysing data that have increased in volume and variety, streamlining processes to drive efficiency and generating intelligence to identify risks, trends and outliers that might have been missed previously.

26% of respondents conduct 1G manual analysis of supervisory data.

These financial authorities rely predominantly on relatively rigid and simplified spreadsheet models for data analysis.

Analytical activity is dominated by 2G descriptive and diagnostic analysis.

63% of financial authorities leverage these applications that are used to search and summarise historical data to identify patterns or meaning, including automated statistical summaries and the data feeding into dashboards and data visualisation tools.

3G predictive analytics is present for 38% of financial authorities, but 4G prescriptive analytics are only adopted by 5%.

These tools enable advanced analysis of historical data to create statistical models to predict future events, values, facts or characteristics and then prescribe an optimal response. This process may include recommendation engines (tools where the prediction is an optimal value or action) and employ machine learning (computerised, iterative optimisation of the aforementioned statistical models).

3.1.3.5. Data products

26% of respondents employ 1G static charts and metrics, while 28% have basic 2G dynamic dashboards.

This investment in basic reporting tools to report statistical snapshots and key performance indicators (KPIs) is consistent with the substantial investment in diagnostic and descriptive tools at the analytics layer.

Only 11% have adopted 4G advanced business intelligence tools.

To extract the most meaningful and actionable insights from data, authorities have started to invest in AI-enabled dashboards that leverage big data tools to allow numerous analytical operations.

3.1.4. Supervisory areas

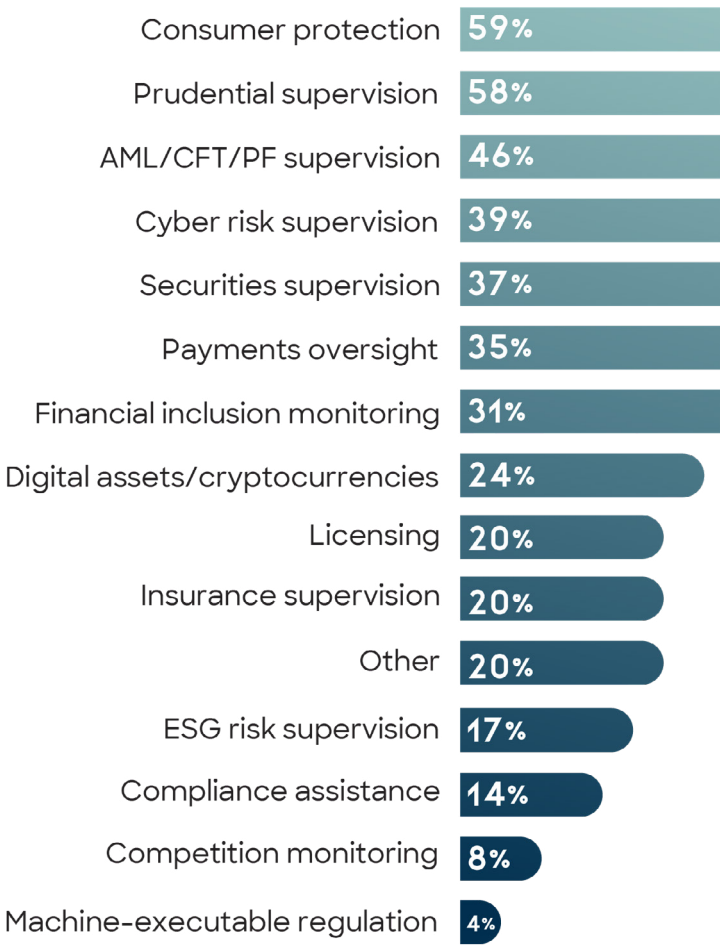
Suptech initiatives cluster mainly around the areas of consumer protection and prudential supervision.

Consumer protection supervision is a relatively new area of focus for financial authorities. While the growth of fintech and its offerings bring new opportunities, especially for the underserved and marginalised market, it may unintentionally place a greater financial burden and risks on these vulnerable customers. Financial authorities are mindful of these inherent risks. They are making deliberate efforts to build trust and confidence in these products, such as expanding their mandates to include responsibilities that some once considered conflicting with the stability mandate, for example, consumer protection, competition and financial inclusion. Recent advancements in data and technology, such as NLP and real-time monitoring, also support market conduct supervision as they present new opportunities for supervisors by enabling greater qualitative analysis.

To the contrary, prudential supervision – needed to ensure the safety and soundness of the banking system – is a traditional, core component of many financial authorities' mandates. As highlighted above, numerous efforts are being dedicated to upgrade the tools for data collection and validation that are foundational for prudential supervision.

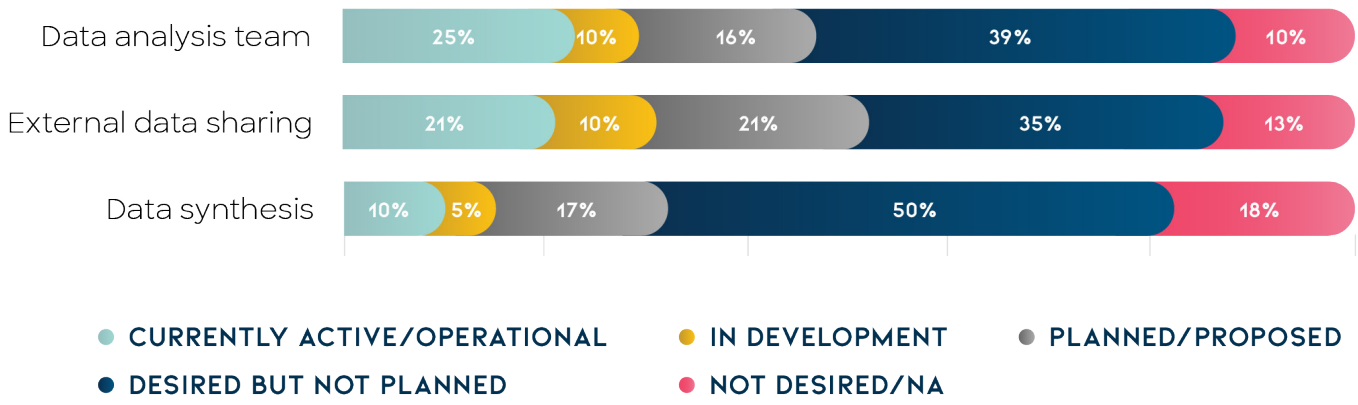
The analysis of the regional thematic focus revealed that suptech initiatives that support financial inclusion monitoring are particularly prevalent in sub-Saharan Africa (59.4%) compared with other regions.

FIGURE 14.
THEMATIC AREAS OF FINANCIAL AUTHORITIES' SUPTECH INITIATIVES (N=134)



3.1.5. Enabling factors

FIGURE 15.
ENABLERS OF DATA CAPABILITES FOR SUPERVISORY AGENCIES (N=119)



There is an unmet desire for data teams, data sharing and data synthesis as a foundational part of the data capabilities of supervisory agencies.

Only 25% of respondents who felt a data analysis team was important had a team that was currently active/operational. An additional 65% expressed a desire, plan, or ongoing development of a dedicated team.

Of the respondents who identified external data sharing as necessary for enhancing regulation and supervision capabilities, only 21% had this core component of their digital regulatory and supervisory infrastructure in place and operational.

Notably, 50% of respondents who considers data synthesis to be necessary said the technology was desired but not planned, while only 10% had this technology currently operational in their organisation. This high ratio of expressed desire to operation solutions was similar across EMDEs and

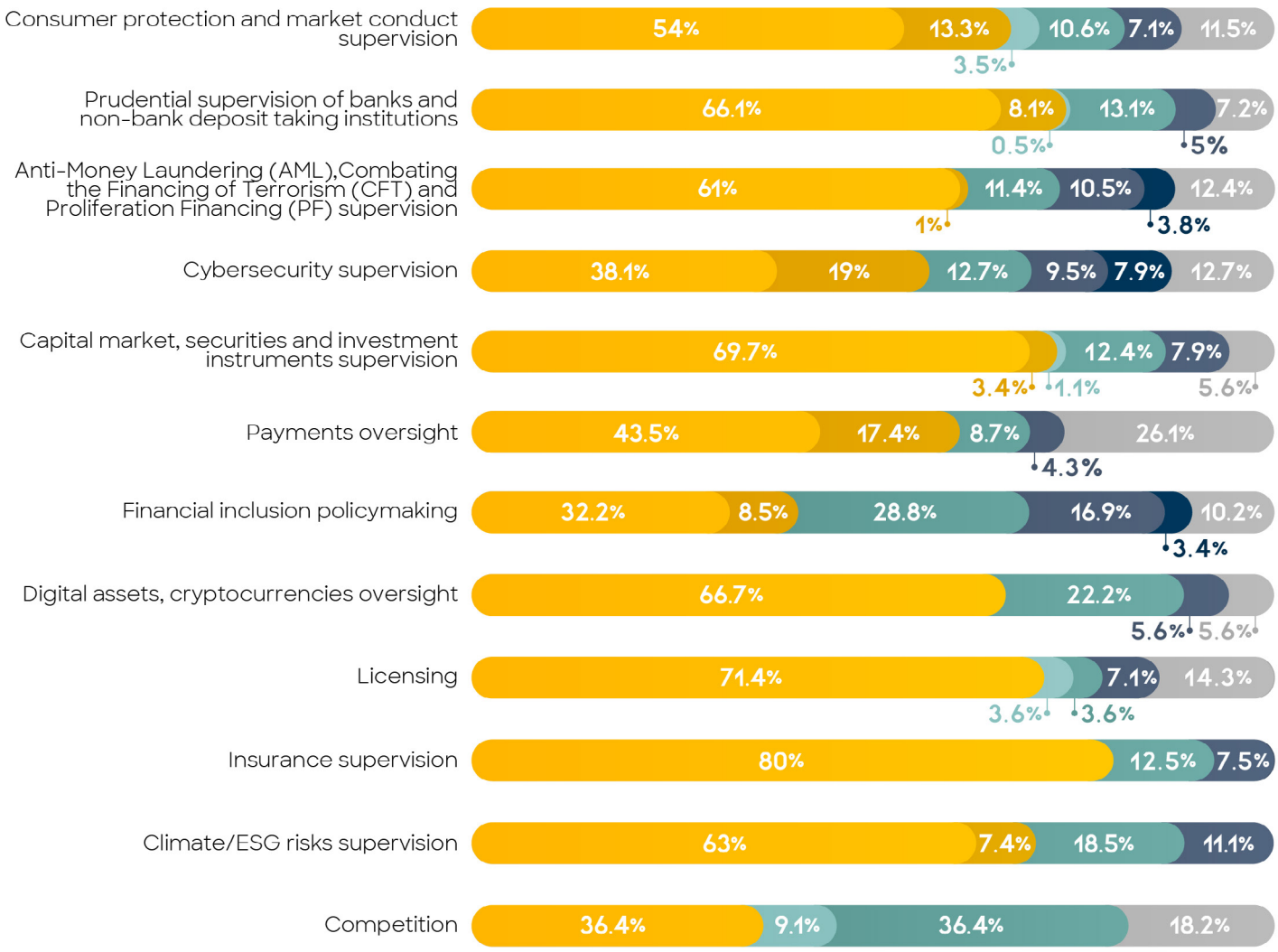
3.1.6. Funding

Funding for financial authorities' suptech initiatives comes mainly from the financial authorities themselves.

Suptech solutions for insurance supervision (80%) and prudential supervision (74.1%) are the least dependent on external funding sources. Only financial inclusion (40.7%) and competition supervision (36.4%) are funded primarily by external sources.

FIGURE 16.

SUPTECH FUNDING SOURCES, WITH SUPERVISORY AREAS PRESENTED IN ORDER OF MOST TO LEAST PREVALENT (N=48)



3.1.7. Governance

With the advent of suptech, financial authorities are placing significant focus on evolving from ad hoc initiatives driven from within the supervision department and implemented by the IT team to more strategic investment in roles such as Chief Data Officer (CDO), Chief Technology Officer (CTO), and centralised data science units to drive or support suptech implementation.

Predictably, the survey responses show that suptech initiatives are driven by supervision departments (58%) or IT departments (38%). In a few instances, multiple departments or functions shared the responsibilities for suptech initiatives, a common combination being the supervision department, IT department and data analytics office leading different aspects of the suptech initiative. For example, while a given suptech initiative may be owned and driven by the business unit (supervision department), the data analytics department may be responsible for data strategy, quality and governance, and the IT department for the technology infrastructure.

FIGURE 17.
WHO IS LEADING THE SUPTECH INITIATIVES (N=40)

Over one-third of the surveyed financial services authorities have designated Chief Data Officer leading suptech efforts.

Of note, more than one-third of the survey respondents (35%) have a dedicated centralised office reporting to a CDO who is either solely responsible for the suptech efforts or works with other functions to support the suptech initiatives. This signals the financial authorities' growing interest in adopting data-driven approaches to support the supervisory process.

Suptech leadership differs across income level and type of financial authority

Respondents from financial authorities in AEs reported that suptech efforts are primarily led by a CDO or a data analytics department. In EMDEs, this is still mainly led by the supervision department.

Similarly, the leadership of suptech within capital markets, securities and investment instruments primarily lies with a CTO, while for most central banks, this charge is led by the supervision department.

Data is a strategic priority for several financial authorities.

Creating a formal data strategy can accelerate data capacities, increase institution-wide buy-in and coordinate action and the support of an authority's senior management. For example, the United Kingdom Financial Conduct Authority (FCA) developed its first data strategy in 2013 and later updated it in 2020 and 2022 as its data journey transformed over the years, from focusing on how they collect and manage data to supporting its continued transformation journey towards becoming a digital and intelligence-led institution.

Data strategies are typically followed by creating a CDO who leads the development and management of internal

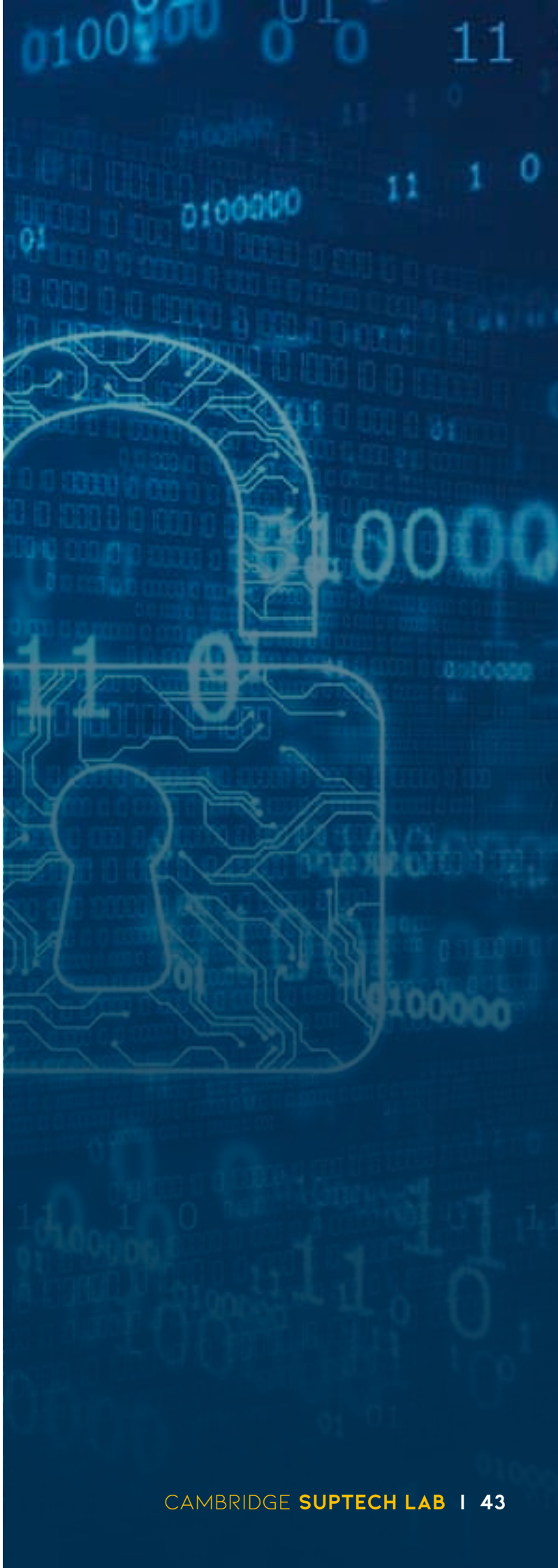
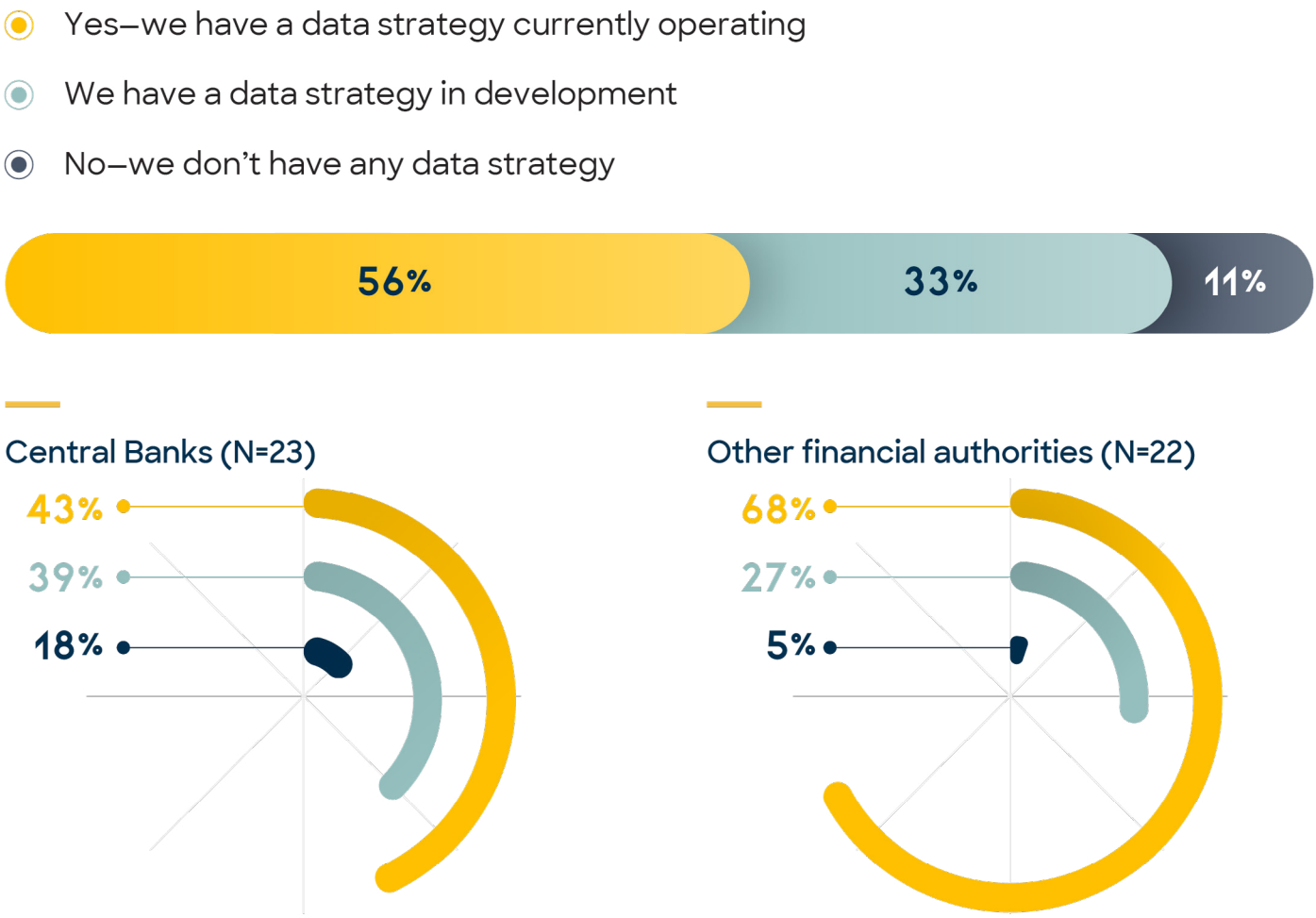


FIGURE 18.
DATA STRATEGY MATURITY, BY FINANCIAL AUTHORITY TYPE (N=45)



data policies and governs data access and management.

Based on the survey responses, data is a strategic priority for an increasing number of authorities, with most respondents either having a data strategy in place (56%) or in development (33%).

While data is a critical factor for central banks, only 43% have a data strategy in place, and a further 39% are currently developing their data strategy.

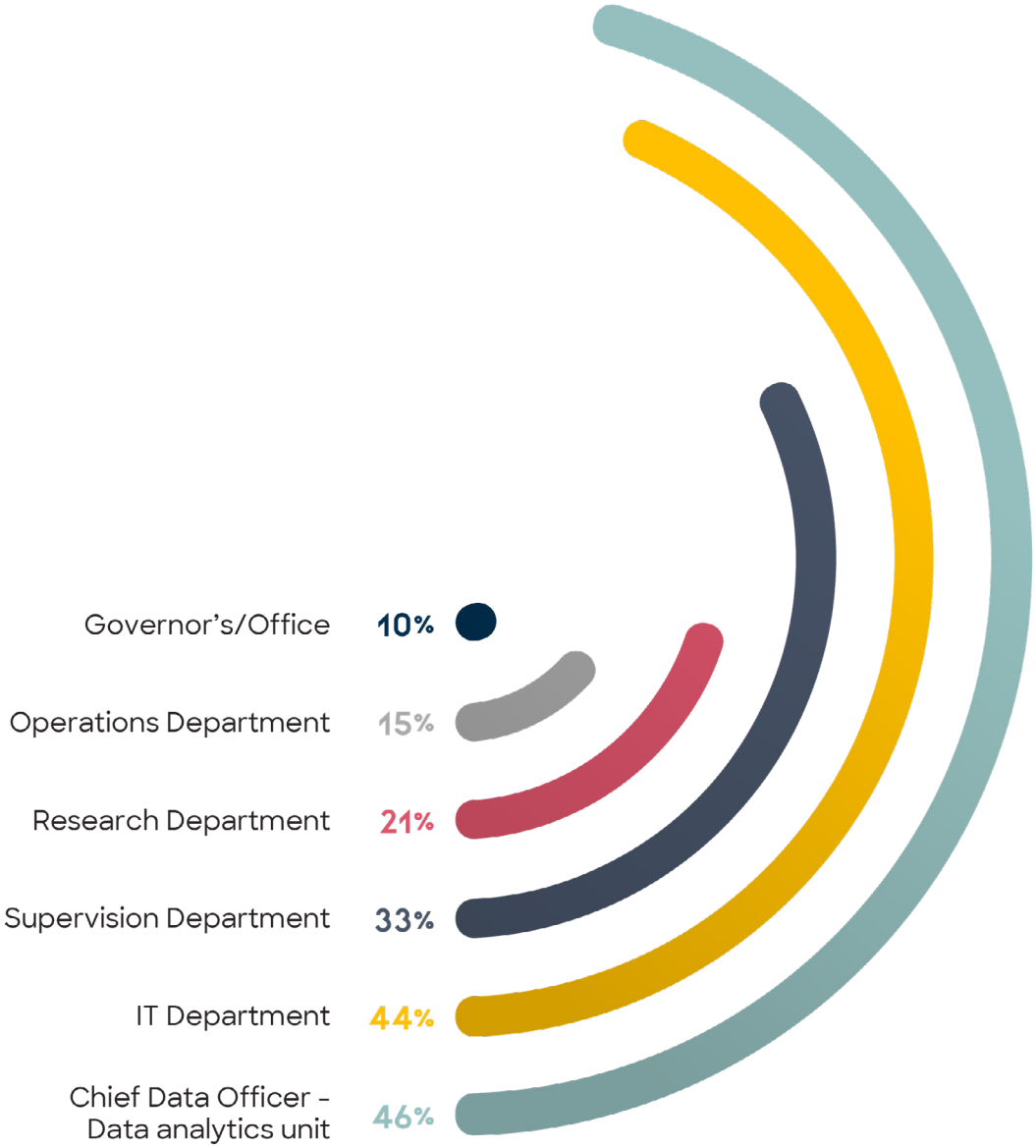
Financial authorities are actively interested in creating dedicated units to lead strategic data initiatives.

The development of formal procedures to govern data access is especially critical for financial authorities as it minimises data mismanagement, cybersecurity and information security risks and ensures that the insights derived from data analytics (suptech) solutions support data-driven policymaking.

As such, about 46% of the respondents have set up a central unit with a CDO who is responsible for their data strategy. This group primarily consisted of authorities from EMDEs (61%). The remaining respondents leverage existing units, such as the IT department (44%), to lead their data analytics initiatives.

FIGURE 19.

WHO IS LEADING DATA STRATEGY INITIATIVES (N=40)



3.1.8. Gender

There is a great opportunity for suptech solutions to support financial authorities in collecting and analysing granular/ disaggregated gender data.

Without the collection of national sex-disaggregated, supply-side operational data offering a clear picture of the situation and comparison between men and women on their access, usage and quality indicators of inclusive finance, it will be extremely challenging to close the gender gap in financial inclusion and achieve the related economic benefits (AFI 2020).

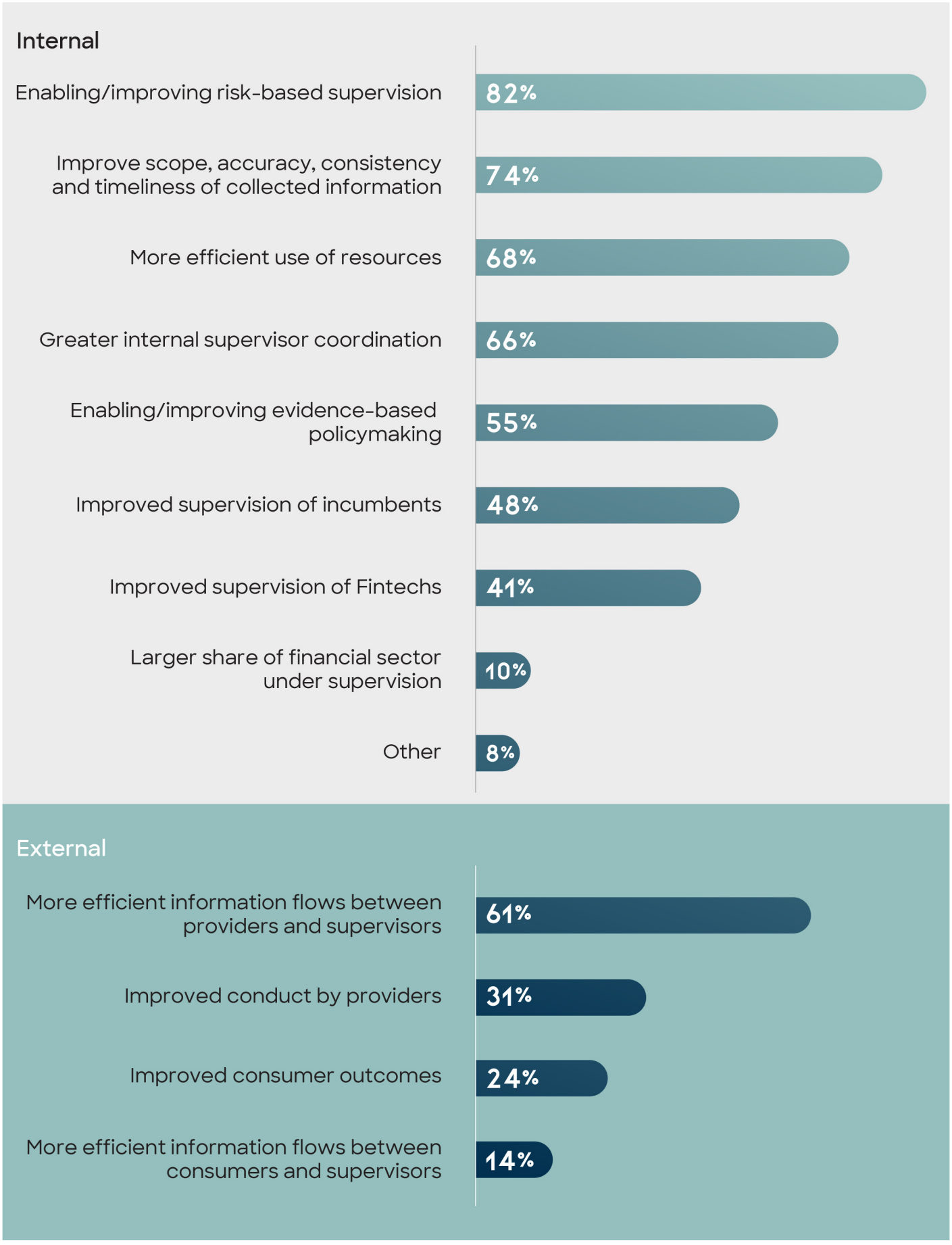
Globally, financial institutions are at different stages of collecting sex-disaggregated data. Very few institutions currently use this data to identify and highlight barriers to women's financial inclusion, whether it be policy-related or awareness and understanding of available financial service products.

Survey results are consistent with research finding as most respondents (70%) stated that they do not have a strategy around financial gender-disaggregated data. In breaking down the overall 21% that are currently operating gender data strategy by type, central banks report slightly greater progress (24%) than other types of authorities (17%).

FIGURE 20.
GENDER STRATEGY MATURITY, FOR EXAMPLE, AS A STANDALONE WORKSTREAM, AS PART OF OVERALL DATA STRATEGY, AS PART OF FINANCIAL INCLUSION STRATEGY (N=43)

FIGURE 21.

OUTCOMES THAT FINANCIAL AUTHORITIES' SUPTECH INITIATIVES HAVE SUPPORTED (N=88)



3.1.9. Outcomes

Suptech can catalyse a risk-based supervisory approach that can adapt quickly to a constantly evolving environment.

The post-financial crisis reforms, the impact of Covid-19 and the advent of new technologies require that supervisors establish proportionate, risk-based approaches underpinned by efficient data management.

Risk indicator dashboards, centralised data warehouses for supervisory reports and early warning systems are some of the tools that are now entrenched in several supervisory agencies worldwide (BIS 2018).

Aligned with the above, when asked what they considered the primary outcomes of embracing successful suptech initiatives, most respondents pointed to enabling/improved risk-based supervision leading to better identification and measurement of risk (82%), improved and consistent data collection (74%), and increased efficiencies in the use of resources by the reallocation of staff away from manual tasks (68%).

At the same time, data and reports submitted by supervised institutions are among the sources of information used most widely by supervisors to inform supervisory activities.

Based on the survey results, supervisors also acknowledged suptech's potential to enhance regulatory reporting, with 61% of respondents stating that a key 'external' outcome of suptech was more efficient information flow between providers and supervisors.

Capital market, securities and investment instruments supervisors have seen more substantial internal outcomes than central banks and other supervisory agency types.

Respondents in this category have seen more impact in using internal resources more efficiently (82%, versus 63% for central banks and 58% for others) and in greater internal supervisor coordination and information flow (74.1%, versus 61.2% for central banks and 66.7% for others).

Financial authorities of all types in advanced economies have seen more substantial internal outcomes than those in emerging markets and developing economies.

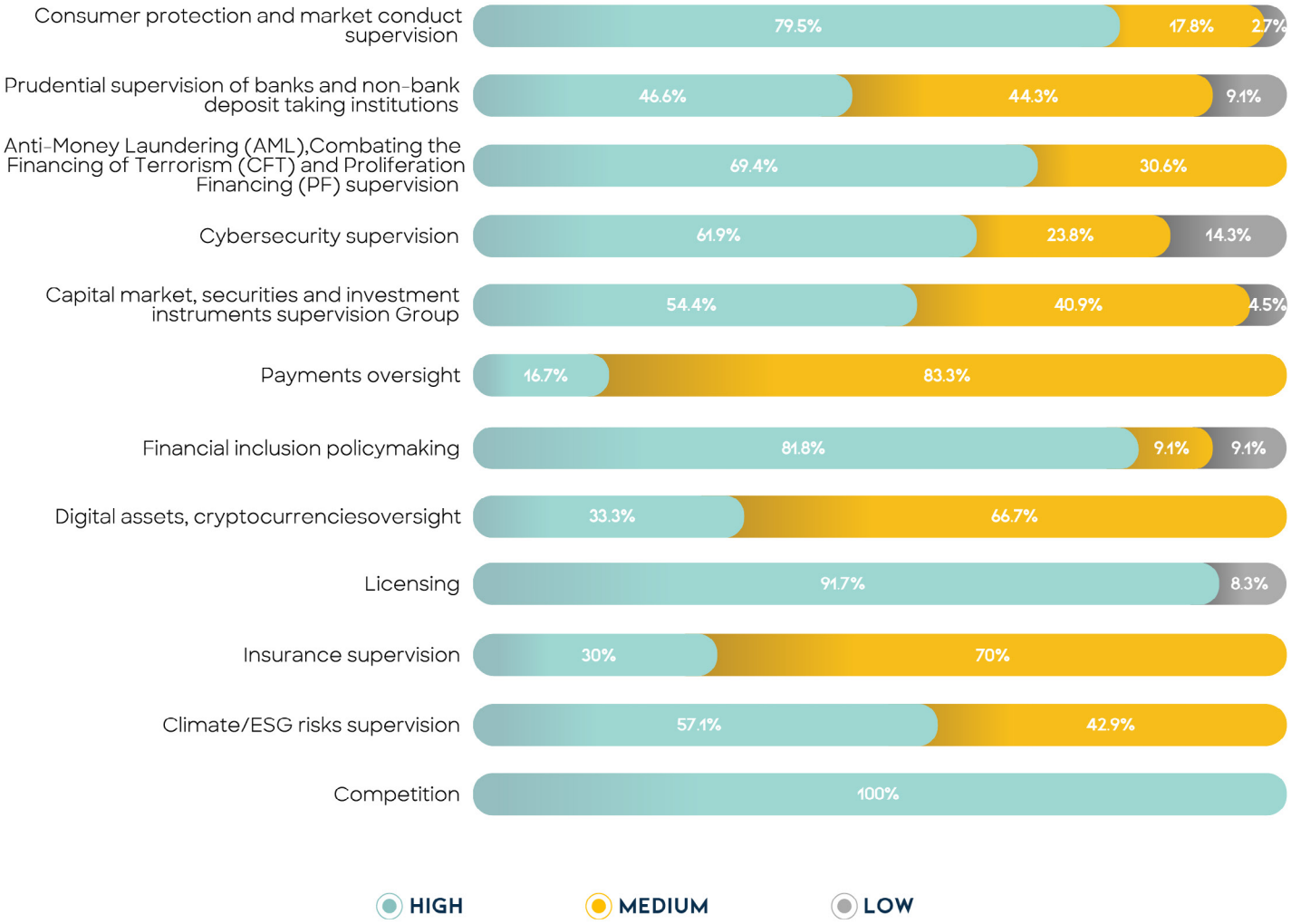
These authorities have noted greater improvement in scope, accuracy, consistency and timeliness of collected information (85% in AEs versus 67% in EMDEs) as well as greater internal supervisor coordination and information flow (79% in AEs versus 58% in EMDEs).

The impact of suptech solutions generally meets expectations, with few exceptions.

Cybersecurity supervision is the only area where respondents report significant levels of low impact (14.3%). The top categories for high impact are competition monitoring (100%), consumer protection and market conduct supervision (79.5%) and AML/CFT/PF (69%).

FIGURE 22.

ACTUAL VS EXPECTED IMPACT OF SUPTECH SOLUTIONS, WITH SUPERVISORY AREAS PRESENTED IN ORDER OF MOST (TOP) TO LEAST (BOTTOM) PREVALENT (N=88)



3.2. Supply: Sourcing solutions

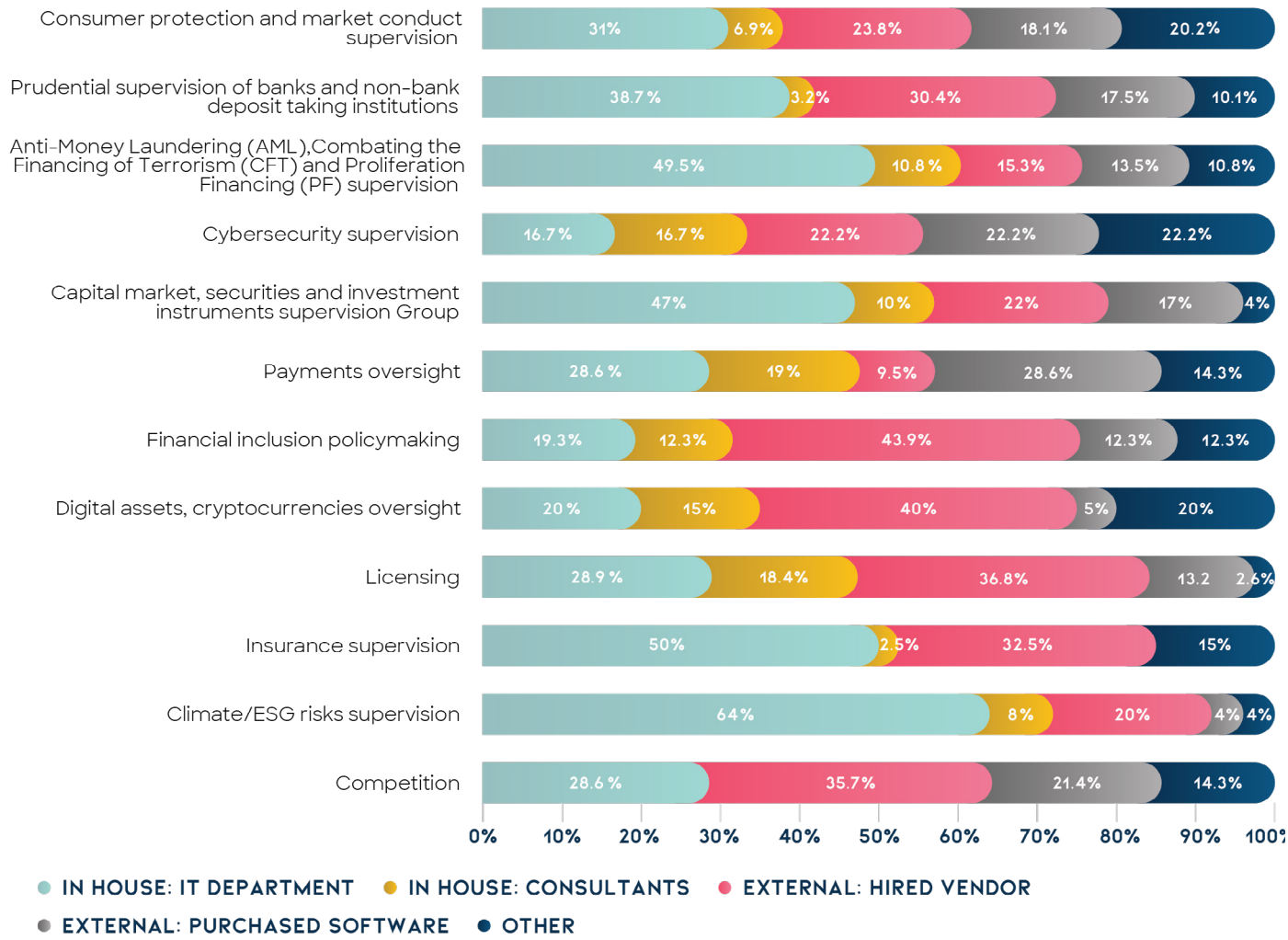
Suptech projects can be developed in-house by a team of the financial authority itself, collaboratively with external vendors, or a combination of the two. This strategic decision is usually based on whether the financial authority has the resources or technical capacity to invest in developing the solutions in-house. When these prerequisites to build in-house are not met, financial authorities must engage with vendors: doing market research to source, running competitions, engaging in pilots, and ultimately conducting formal procurement.

3.2.1. Sources of suptech apps

Most suptech solutions are sourced externally, either from vendors that develop custom solutions or as purchased off-the-shelf software.

Areas where off-the-shelf and other purchased software were most prevalent are payments (28.6%), cybersecurity (22.2%) and competition monitoring (21.4%). The only supervisory areas where suptech is primarily built internally

FIGURE 23.
HOW SUPTECH IS BUILT, WITH SUPERVISORY AREAS PRESENTED IN ORDER OF MOST (TOP) TO LEAST (BOTTOM) PREVALENT (N=91)



or with internal consultants are climate/ESG risks (72.0%), AML/CFT/PF (60.2%), and capital markets, securities, and investment instruments (57.0%).

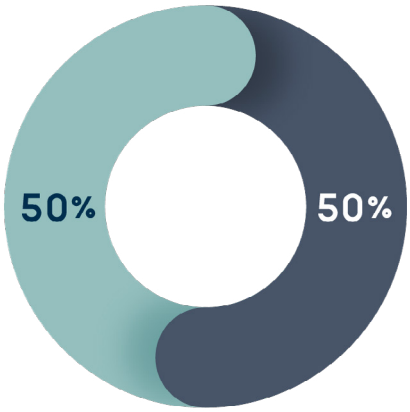
3.2.2. The vendor's business case

Based on what we have observed through desk research and direct engagement with many of the 73 vendors listed in the Cambridge SupTech Lab's online Vendor Database, for most of them suptech is not the main business, but rather an activity pursued as complementary to the provision of regulatory technology (regtech) solutions to the more scalable and profitable financial industry, or as part of their broader provision of technologies to various industries. So far, we have identified only a few who have made suptech central in their offering and business model.

Suptech is currently a strong business case for the suptech vendors we interviewed, who expect it to grow over the next two years.

When asked how strongly suptech factors in their business model, all the vendors responded that suptech was either their core business or a key business offering. This suggests that the vendors see suptech's value proposition. They further reported that they foresee a growing demand for suptech applications in the coming years.

FIGURE 24.
HOW STRONGLY SUPTECH FACTORS INTO SURVEYED VENDORS' BUSINESS MODELS (N=6)



- Very strongly - Suptech is at the very core of our offering
- Strongly - Suptech is an important component among several lines of business
- Somewhat - We pursue suptech opportunities as part of a broader strategy
- Minimally - Suptech act as a loss leader for new markets, when necessary, but is not an active pursuit

FIGURE 25.
HOW DO YOU EXPECT THIS SUPTECH COMPONENT OF YOUR BUSINESS MODEL WILL HAVE SHIFTED TWO YEARS FROM NOW? (N=6)

3.2.3. Offerings by focus area

Suptech solutions provided by surveyed vendors focus primarily on the data collection, data processing and data analytics layers of the supervisory tech stack.

3.2.4. Funding

From a vendor perspective, suptech applications are mainly funded by the financial authorities themselves but are also often supplemented by grant funding.

FIGURE 26.
LAYERS OF THE SUPERVISORY STACK WITHIN WHICH VENDORS ENGAGE WITH FINANCIAL AUTHORITIES (N=6)

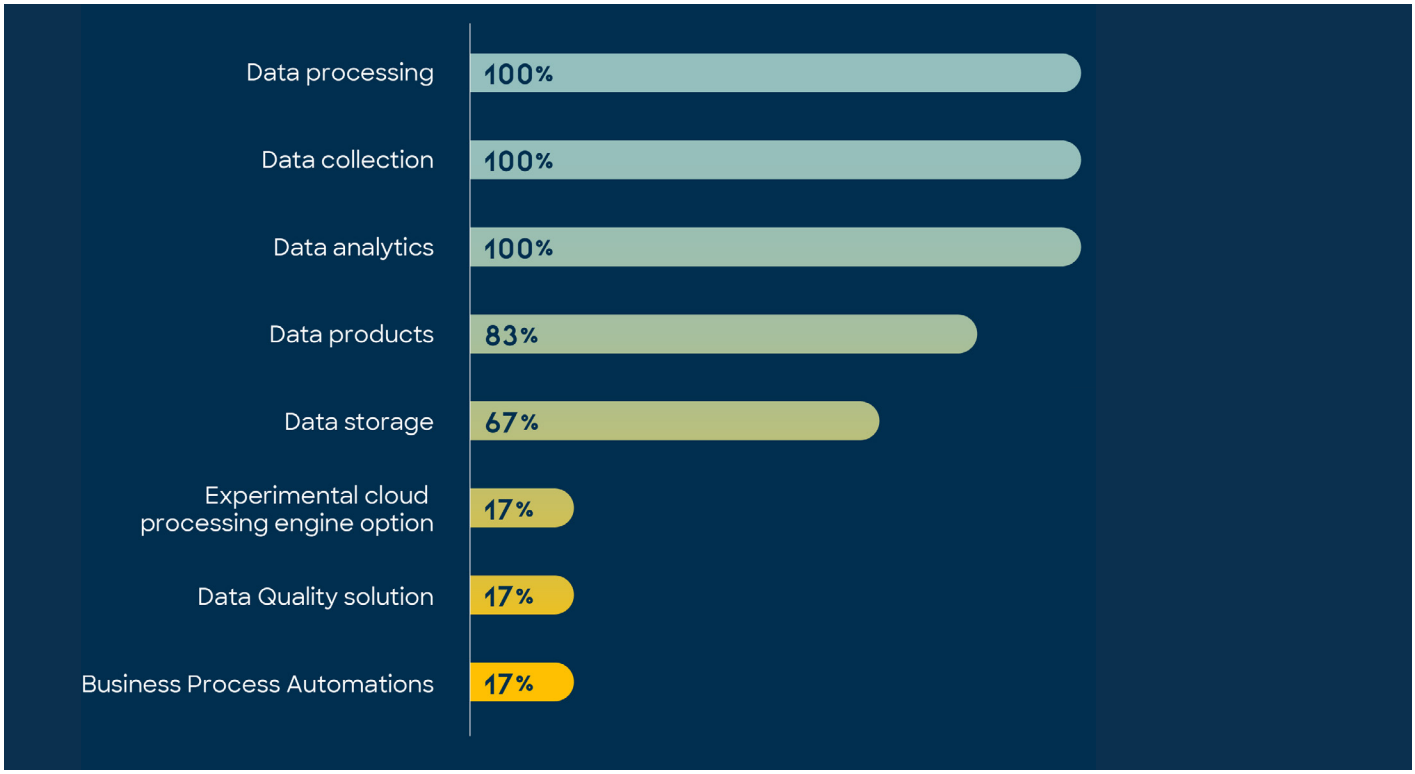
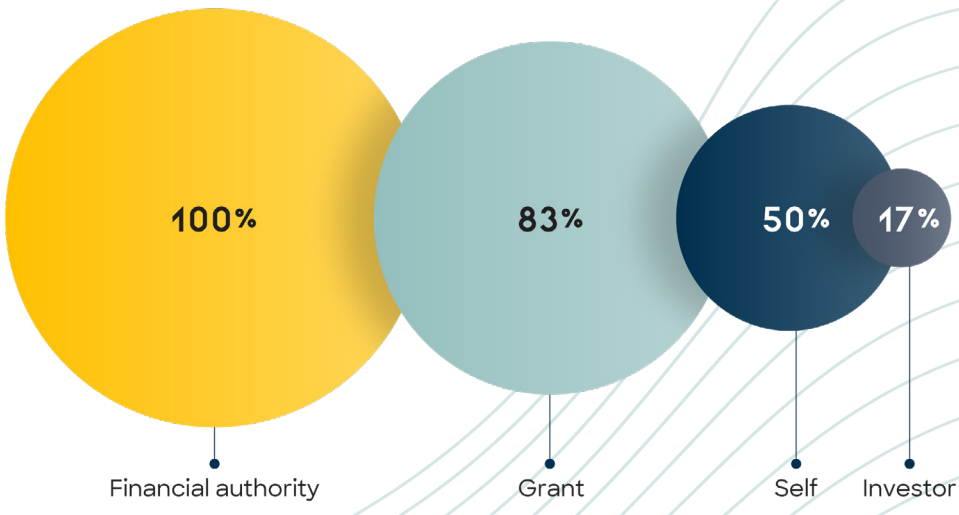
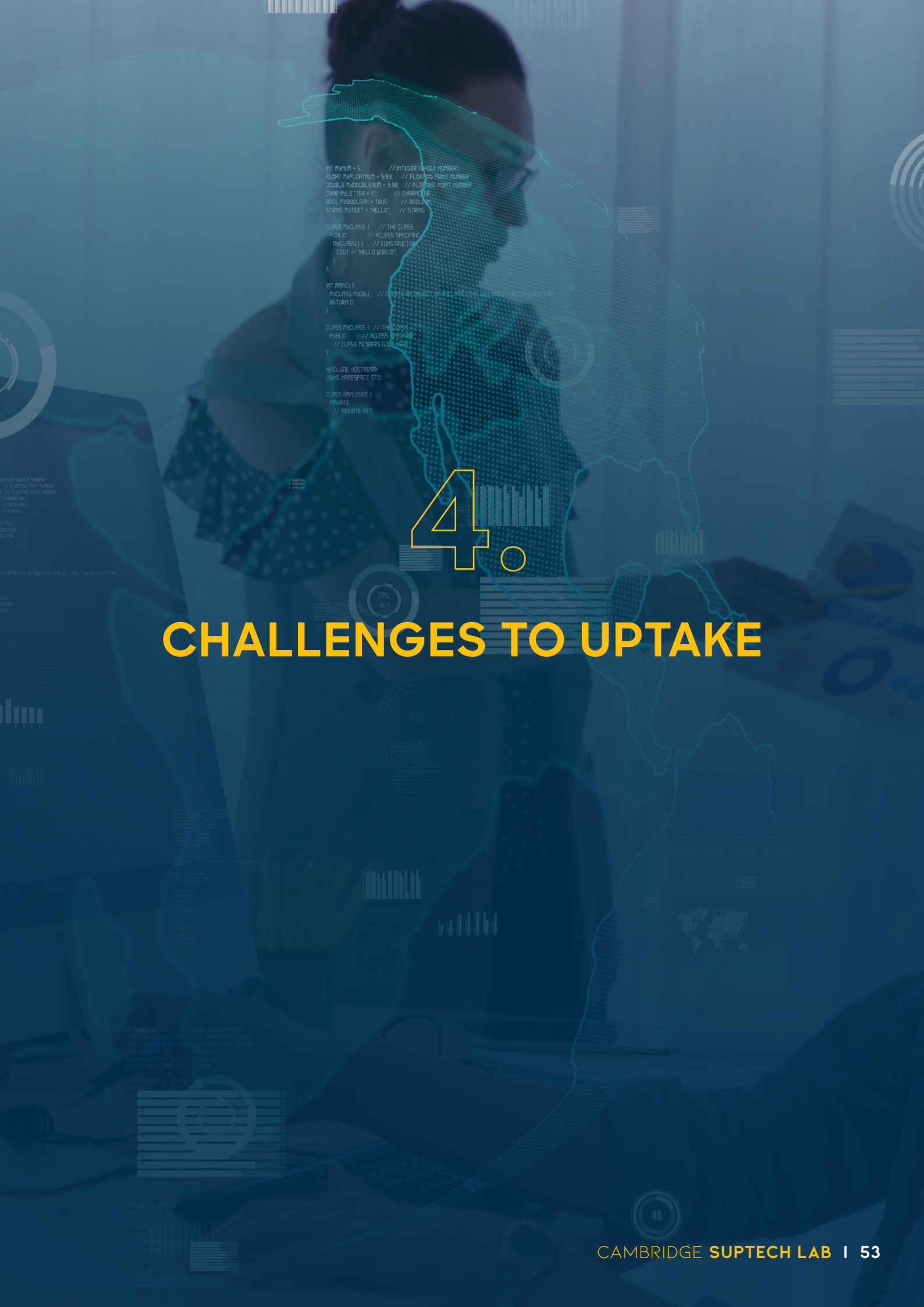


FIGURE 27.
SUPTECH FUNDING SOURCES FOR VENDORS (N=6)





4.

CHALLENGES TO UPTAKE

4.1. Challenges: financial authorities

4.1.1. Implementation

Limitations in budget, data quality and technical skills remain significant barriers to implementing suptech.

Despite the efforts of supervisory authorities to enhance supervisory processes through technology, various challenges have been encountered in developing and using suptech applications. As they embark on their modernisation journey, authorities are becoming cognizant of the challenges associated with the digitisation of their processes and methodologies.

Research has outlined some of these challenges, including limitations in data quality, lack of transparency in data, lack of management support and buy-in, increase in cybersecurity risks in an automated suptech environment, lack of adequate expertise, algorithmic biases, third-party dependencies and legacy systems (BIS 2018, BIS 2019).

The Report sheds light on the perception and experience of the agencies dealing with these issues.

Supervisory authorities reported that budgetary constraints (58%), data quality issues (57%), limited staff with data analytics capability (54%), legacy IT systems (49%) and limited staff IT skills top the list of internal challenges they encounter when developing, deploying, and maintaining suptech solutions. They clearly pointed at challenges related to data analytics and tech development, and the need for capacity building in those two areas.

In addition, when it comes to the external factors, one third of the respondents - led by agencies in EDMEs - highlighted challenges coordinating data sharing with

external stakeholders (29% average, 38% in EDMEs).

Only 7% of these authorities with suptech initiatives lack management buy-in, and only 14% lack a suptech strategy or roadmap, further highlighting these as prerequisites to engaging with suptech.

The analysis of responses across different income levels shows that financial authorities in AEs and EDMEs report facing similar challenges in the digital transformation of their supervisory processes and capabilities.

Different types of financial authorities face different kinds of challenges.

While the same challenges appear across all types of financial authorities, their prevalence differs.

Across agencies, budget is the main challenge, but capital markets, securities and investment instruments supervisory authorities are more impacted.

75% of capital markets, securities and investment instruments supervisory authorities report internal budgetary constraints (versus 49% of central banks and 58% of others).

FIGURE 28.

CHALLENGES FACED BY FINANCIAL AUTHORITIES IN DEVELOPING SUPTECH, GROUPED BY INTERNAL AND EXTERNAL FACTORS (N=95)

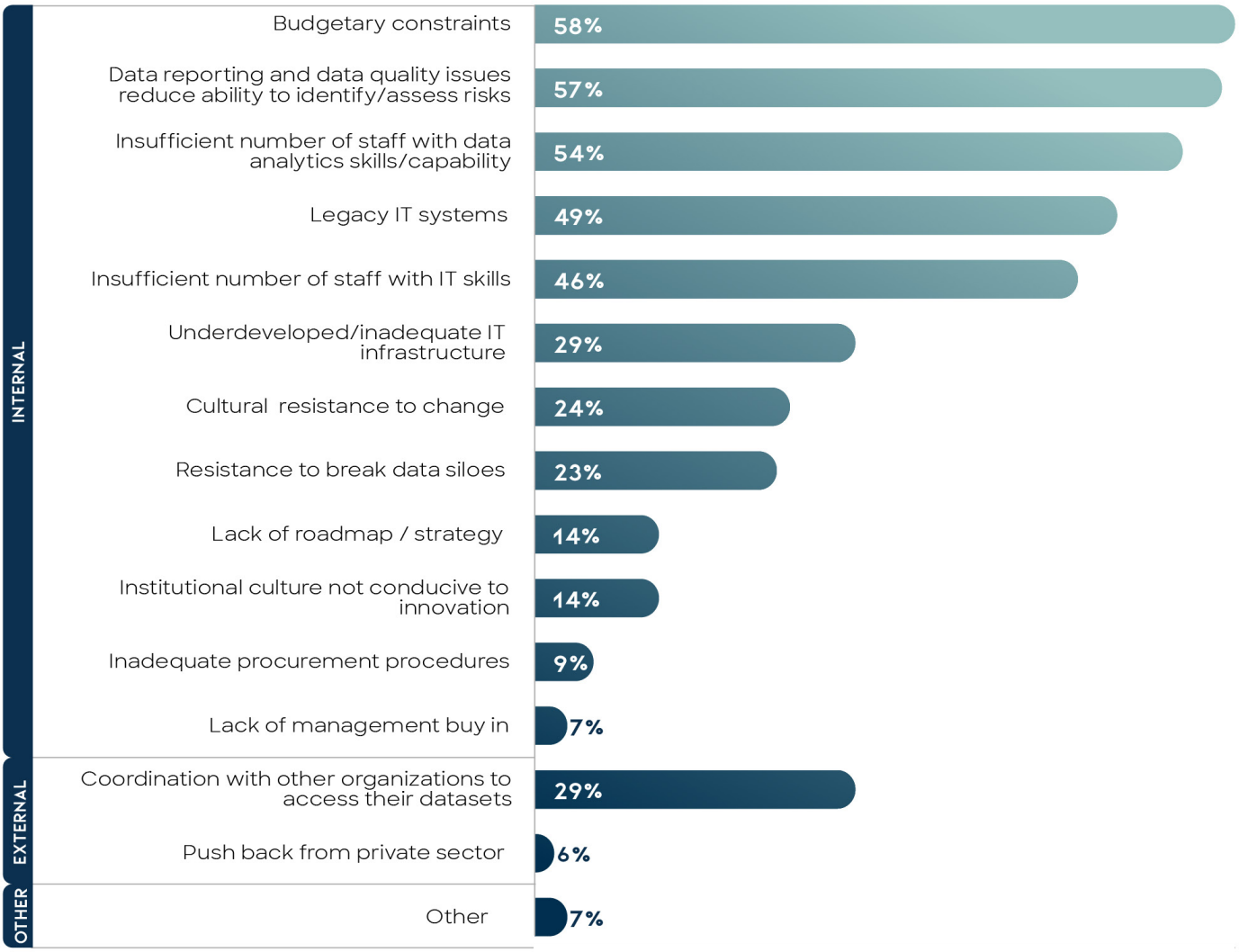


FIGURE 29.

CHALLENGES FACED BY FINANCIAL AUTHORITIES IN DEVELOPING SUPTECH, GROUPED BY INCOME LEVEL (N=95)

For central banks, the challenges are primarily related to internal resistance to change.

29% of central banks reported internal resistance to breaking data siloes (versus 21% for capital markets, securities and investment instruments supervision and 8% for others). Earlier generations of technology tend to capture and store data in siloes, making it difficult for authorities to gather insights from the data. While big data architecture, such as AI, has the potential to address these challenges, they require data expertise and a data-driven culture (BIS 2019). A lack of this expertise can lead to resistance and pushback from staff.

24% of central banks reported internal cultural resistance to change (versus 21% for capital markets, securities, and investment instruments supervisory agencies, and 8% for others). Often, one of the main reasons for suptech implementation is a lack of stakeholder engagement and poorly planned change management. This ultimately leads to resistance by staff. As the financial authorities adopt even more sophisticated technologies or suptech applications, they might find themselves lacking the capacity or skills required, for example, for data analytics or AI/ML applications or lack of understanding of the new processes. The absence of this understanding could lead to a lack of trust in the application results, leading to pushback by staff. This challenge calls for a culture change within the agencies to enable teams to work together and adopt an agile approach. This is especially critical as they begin to engage with the new entrants in the financial sectors who have a different culture and operate in different ways.

For capital markets, securities, and investment instruments supervisors, challenges tend to be related to upgrading their existing systems and processes.

39% have challenges with external coordination with other organisations to access their datasets (versus 25% of central banks and others). 36% report underdeveloped/inadequate internal IT infrastructure, like the inability to use the cloud (versus 27% of central banks and 25% of others).

For other supervisors, the uniquely prominent challenges are with IT systems.

75% report challenges related to internal legacy IT systems (versus 40% of central banks and 57% of capital markets, securities, and investment instruments supervisors). 75% report insufficient staff with IT skills (versus 38% of central banks and 50% of capital markets, securities, and investment instruments supervisors).

Overcoming the common challenges of all financial authorities and the particular challenges for financial authorities of each income level and agency type will help them realise the benefits of wider suptech adoption.

4.1.2. Data lifecycle

The top challenges for financial authorities are data collection, followed by a lack of analytical skills and automated data processing.

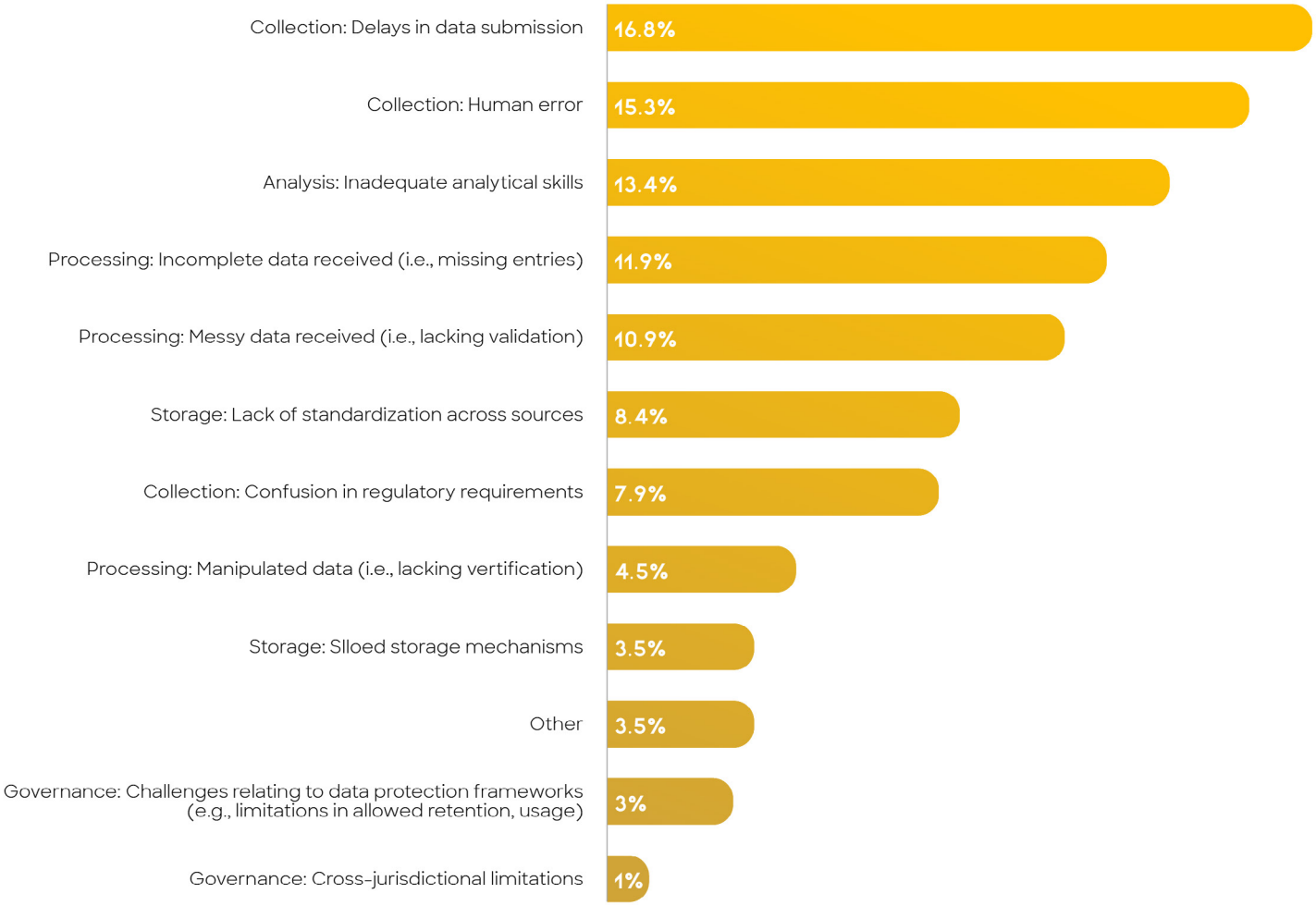
The top five challenges are delays in data submission (16.8%), human error in data collection (15.3%), inadequate analytical skills (13.4%) and incomplete (11.9%) or messy (10.9%) data received due to lack of automated processing,

4.1.3. Resources

Financial authorities seek support to build staff skills required to develop and implement suptech solutions.

FIGURE 30.

DATA LIFECYCLE CHALLENGES ALONG THE SUPERVISORY STACK (N=74)



In recognition of the challenges they face in accessing talent with adequate skillsets and ensuring that their staff have the right skills to use any suptech applications, financial authorities are investing in building their internal capacity.

36% of the surveyed respondents have already undertaken capacity-building activities to support the implementation and use of suptech. A further 45% are interested in supporting their staff to build their skills and knowledge on suptech.

Supervisors in emerging markets and developing economies express more interest in capacity-building programmes than those in advanced economies.

60% of authorities from EMDEs responded that they were interested in undertaking capacity-building programmes to support their suptech initiatives, versus only 14% of authorities in AEs.

Activities that authorities are undertaking to improve their staff capacity include training to enhance technical and digital skills and investments in building and fostering a digital culture within the agencies.

For example, the ECB has introduced a comprehensive digital training curriculum to promote a culture of innovation and build knowledge and understanding of suptech. Over 600 supervisors across Europe recently completed an introductory six-week training programme on AI and how it relates to supervisory work. (ECB 2022)

FIGURE 31.

ENGAGEMENT IN CAPACITY-BUILDING PROGRAMMES IN THE CONTEXT OF SUPTECH (N=44)

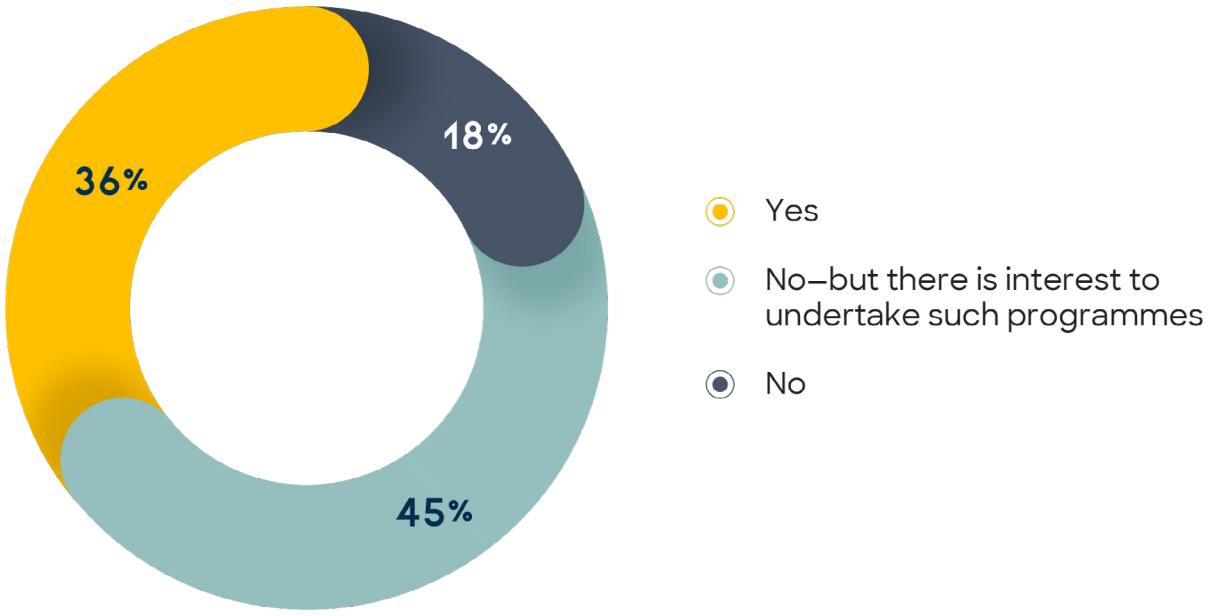
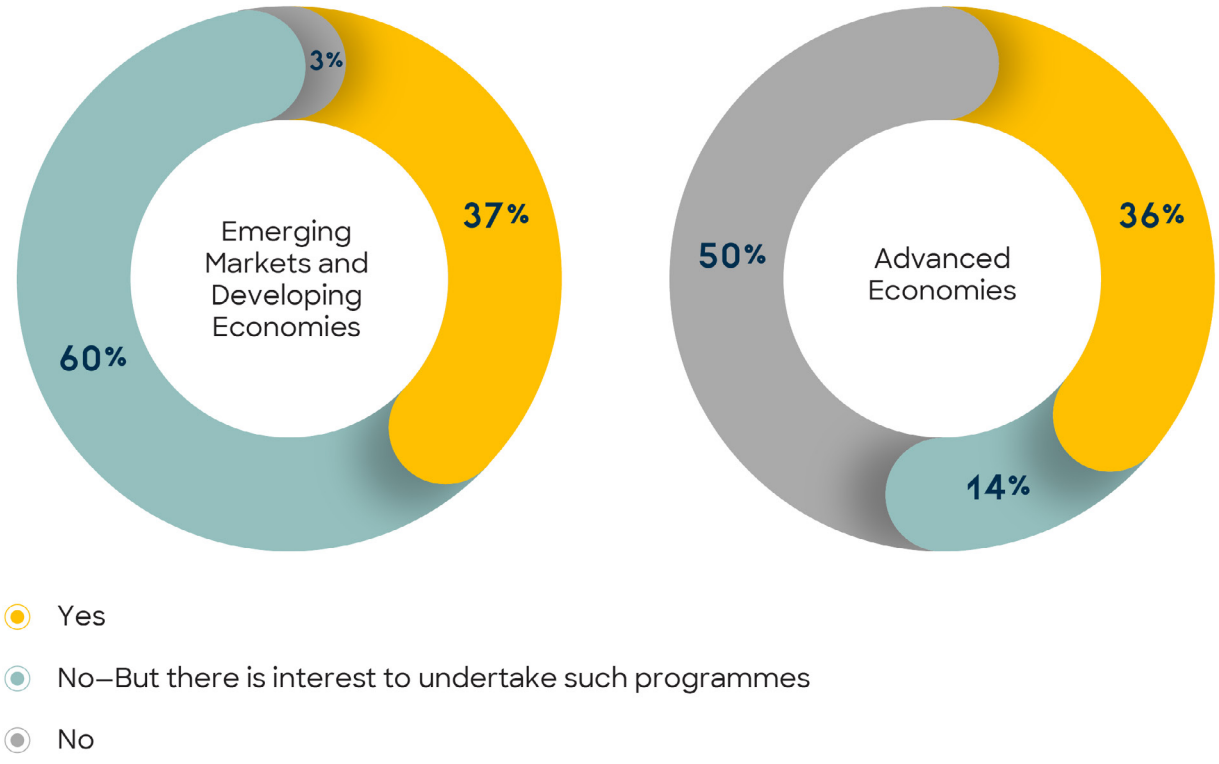


FIGURE 32.

ENGAGEMENT IN CAPACITY-BUILDING PROGRAMMES IN THE CONTEXT OF SUPTECH, SEGMENTED BY INCOME LEVEL (N=44)



This is part of a broader hub-and-spoke innovation model to foster agile collaboration and the joint development of suptech solutions where innovation teams from the ECB and national supervisors pool their knowledge and contribute to the overall goal of digital transformation. These teams are composed of experts from various functions (for example, IT, supervision and statistics) with diverse skill sets.

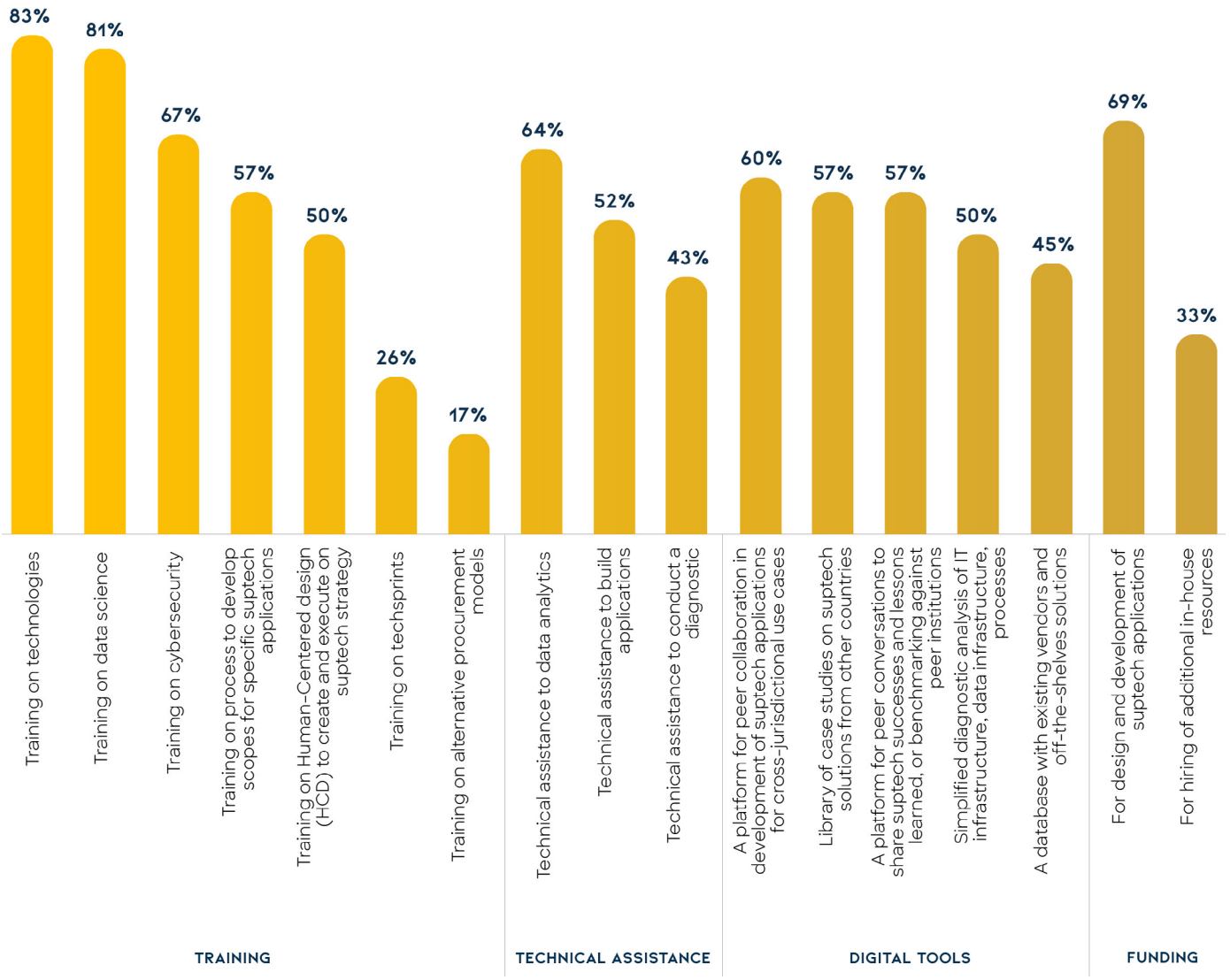
Assessing existing capabilities to undertake and successfully implement suptech solutions is an important step. This not only enables financial authorities to map existing resources and gaps but also helps them understand and estimate future capability requirements.

Top areas of support sought by financial authorities are training, funding and technical assistance.

When asked what areas of support they required, most respondents pointed to the same top five, including technology training (83%), data science training (81%), funding for design and development of suptech tools (69%), cybersecurity training (67%) and technical assistance for data analytics (64%). A key consideration for financial authorities as they adopt and adapt suptech is their capability level regarding skills, talents and resources.

Nearly all central banks primarily seek training on technologies, while

FIGURE 33.
AREAS OF SUPPORT SOUGHT BY FINANCIAL AUTHORITIES (N=42)



securities supervisors seek a mixture of training on cybersecurity, a digital platform for conversations, and technical assistance and funding for suptech development.

Responses were segmented by agency type to understand the top areas of support for each type. Central banks tend to seek training more actively, most prominently on technologies (92%). Capital market, securities and investment instruments supervisors seek training on cybersecurity (83%), technical assistance to build applications (67%), a digital platform for peer conversations (75%), and funding for the design and development of suptech (92%).

Financial authorities in emerging markets and developing economies express more need for support than those in advanced economies.

In particular, tech training (93%), data science (90%), suptech process (70%) and technical assistance on data analysis (80%) and conducting a diagnostic (60%) were in demand for EMDEs. The demand for digital tools was similarly strong across both categories. In terms of funding, financial authorities in EMDEs, in particular, focused on funding for the design and development of suptech solutions (77%), while those in AEs more often focused on funding for hiring (42%).

4.1.4. Infrastructure

Financial authorities face considerable challenges related to the digital infrastructure in their jurisdictions. The top two challenges are limited knowledge/expertise (cited by 63% of respondents) and funding/resource constraints (57%). Legacy IT systems (reported by 49%), a lack of capabilities (48%), poor quality or insufficient data (44%), and the availability of technology (42%) are also common challenges.

4.2. Challenges: vendors

While there is an increasing demand for suptech applications from financial authorities, vendors face obstacles due to the procurement process (83%), dealing with siloed teams and multiple stakeholders within the financial authorities (50%), lack of visibility into the needs of financial authorities (50%) and lack of adequate funding sources (50%).

Additional challenges in providing suptech solutions relate to engaging with financial authorities (in particular technical capacity of supervisors and long sales cycles), developing suptech applications (compliance with data protection mechanisms, digital infrastructure limitations, insufficient access to historical data, compliance with cybersecurity requirements, lack of global data standards and limitations to transfer technologies

As financial authorities develop strategies on how to upskill, train and build capacity internally for data collection, analysis and management, they can also consider the following strategies:

- Develop comprehensive (digital) curricula to help staff build both technical and soft skills required to thrive in an innovative environment
- Recruit and retain digitally skilled staff such as data scientists
- Collaborate with different stakeholders to facilitate knowledge transfer and peer-to-peer learning
- Tap into external technology solutions vendors with vast knowledge of IT and suptech systems, who can assist financial authorities in dealing with rapid changes in technology and overcoming limited in-house technical skills and resources.

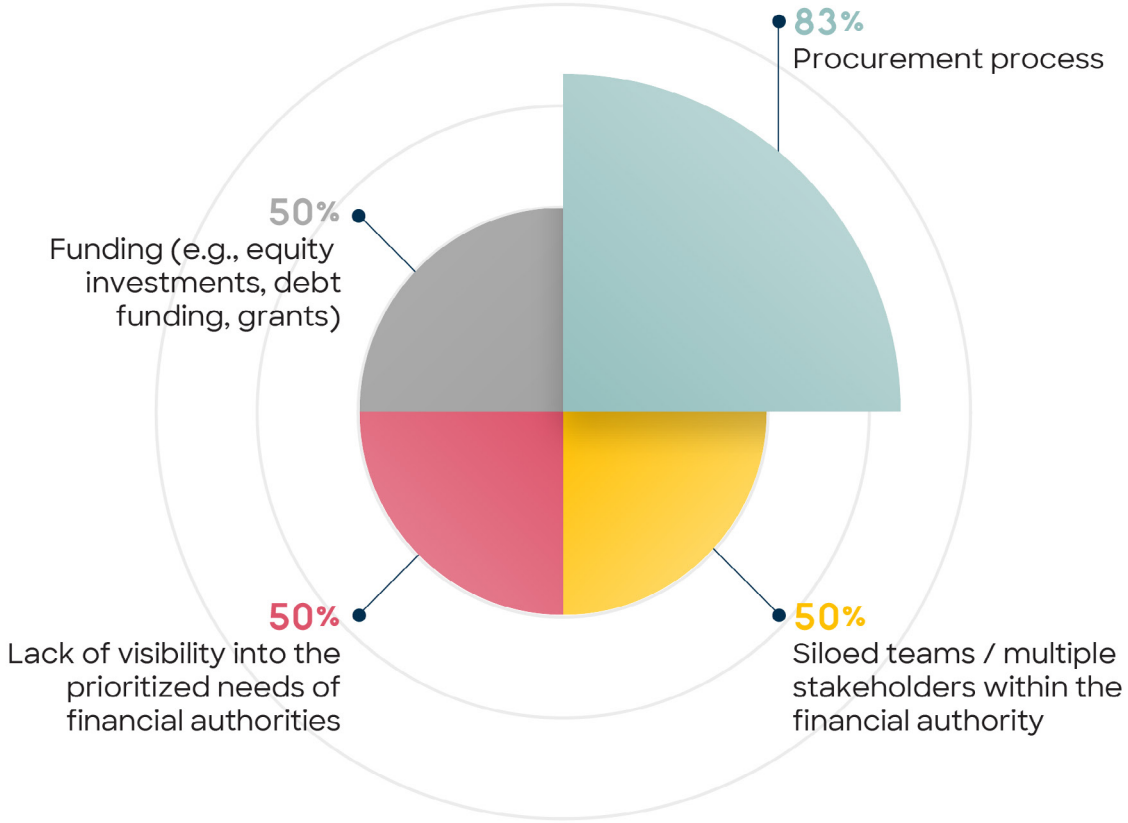
from early adopters to the rest of the authorities) and expanding the vendor's suptech portfolio (mapping their technological offerings to supervisory use cases, the lack of global scalability of their solutions, legal and regulatory restrictions, and the relatively small size of the suptech market).

Procurement process

For financial authorities to achieve their goals of transforming supervisory processes and tools, they must access the right capabilities, capacities, services and products. These are sometimes provided internally by civil servants, but in many cases need to be obtained from the private sector through public procurement processes, which are flagged by most suptech vendors (83%) as the main challenge they face to engage with the agencies.

Public procurement is a sensitive domain that must be carried out efficiently and to high standards to safeguard the public interest. Designed to promote accountability, integrity and effectiveness in the management of public budgets, procurement processes often "enact arduous procedures regulated by long and complex legal frameworks, which may limit the capacity for innovative ideas to be implemented, or even considered. While the strict procedures surrounding public sector procurement aim to protect public money, they often generate perverse incentives, delay processes and could ultimately compromise the quality of service delivery. Such complex public procurement systems and processes represent a major hurdle to SME participation in public procurement markets, as such companies are disproportionately affected by these factors, due to limited financial, technical and administrative capacities." (OECD 2019) Numerous obstacles deter smaller, innovative outlets to participate in public bids, including "a lengthy and overly complex contracting process, a lack of clarity on how to connect with agencies, and a sense that newcomers have little chance to win contracts over incumbents." (BCG and Eastern Foundry 2017)

FIGURE 34.
TOP FOUR CHALLENGES FACED BY VENDORS IN EXPANDING A SUPTECH PORTFOLIO (N=6)



It is important to note that most (88%) suptech vendors listed in the SupTech Marketplace are micro, small, and medium enterprises (MSMEs). Micro (1-10 employees) and small (11-50 employees) enterprises make up 45% of the marketplace, while medium enterprises (51-500 employees) constitute another 43%. Only 12% of vendors are large firms (501+ employees).

Moreover, during the past several years there has been notable consolidation occurring within the nascent suptech market. Of the 74 suptech vendors listed in 2018 by the R²A, have now been acquired by larger firms. Rather than maturing independently and providing competitive solutions, these vendors - who represented 27% of the recorded market in 2018 - were since absorbed into larger competitors. Given the expressed procurement challenges and the asymmetric effect these challenges have on MSMEs, one plausible inference is that the challenges faced by small vendors, primarely due to procurement processes, may be impacting competition and innovation in the suptech space, and conversely resolving procurement issues could increase the volume and variety of competitive offerings available to financial authorities seeking suptech solutions.

Financial authorities seem to underestimate the negative impact of public procurement procedures on vendors. (see Table 29)

In fact, the concern expressed by interviewed vendors resonates with our experience, and is supported by an extensive body of literature that identifies the rigidities in government procurement rules among the main causes of failure of public sector digital technology projects (Dunleavy and Carrera 2013, World Bank 2016).

Siloed teams/multiple stakeholders within financial authorities

Another relevant challenge for vendors is dealing with siloed teams and multiple stakeholders within financial authorities. When it is not clear "who the client is" and vendors are not aware of all the departments and units involved in a modernization project, they end up losing time and burning budget while dealing with redundancies and delays in engaging with the appropriate people.

Lack of visibility into the needs of financial authorities

More and better visibility into the needs of financial authorities would allow vendors to tweak their products to serve more specific needs, and develop a business pipeline that would make their businesses more viable and attract investments.

Insufficient funding

Vendors may face challenges in covering the costs of business development in a market that, at first, may offer little scale, and where payments can be deferred and delayed as part of the procurement cycle.

The Cambridge SupTech Lab partners with financial authorities and technology vendors to co-create and deploy cutting-edge, scalable suptech applications. The Lab furthers the capacity of financial authorities to drive their engagement with technology providers by helping them develop proofs of concept and technical specifications (see the Digital SupTech Diagnostic Tool), identify off-the=shelves solutions and developers outlets that are serving this market (see the SupTech Vendor Database) and also acting as a broker that curates and facilitates the collaboration between agencies and vendors.

The Lab's Application Foundry - which builds largely on the successful experience of the R²A - accelerates the development of ground-breaking suptech applications by detailing the technical specifications, de-risking procurement for all parties, providing project management support and hands-on technical assistance, introducing an agile approach to the collaboration between financial authorities and technologists.

5.

CASE STUDIES

The following suptech case studies serve to complement the insights derived from the survey and presented in the analyses above, and to ground the findings in a practical context. This select set of case studies has been systematically drawn from across the supervisory stack, representing each layer of the "tech" side of the suptech taxonomy presented in Section 1.2.2. While each case study involves multiple layers (e.g., a data collection solution often involves data processing), they are classified by the layer at which the *primary* innovation occurred. A final case study includes a solution that focuses on a holistic approach, innovating at all layers of the stack in parallel. A database of solutions is available in the Cambridge SupTech Lab's **Suptech Marketplace**.

5.1. Data collection: Bank of England transforming data collection from the UK financial sector

In 2019, the Bank of England (BOE) conducted a review of the future of the UK's financial system, and what it might mean for the BOE's agenda, toolkit and capabilities over the coming decade. The Future of Finance report (BOE 2019) identified some critical issues in relation to the efficiency and effectiveness of data collection, pointing to policies and processes that made the collection of data from regulated firms costly, time consuming, relatively inflexible, and often redundant for both the BOE and the industry. The report suggested that a number of **underlying factors** may contribute to these issues:

i. Heterogeneity in firms' data

For any given product or transaction, different firms might hold and describe equivalent data differently. This makes it hard for the BOE to write a set of reporting instructions that are unambiguous to all firms. In turn, this can lead to 'pain points' for firms in interpreting instructions and locating data, which has the potential to cause long timelines and quality issues for the BOE.

ii. Heterogeneity of the Bank's data needs

Reports are designed to address specific use cases. For instance, the BOE often requires data to be aggregated in ways that make reports hard to repurpose. This leads to more requests for new reports or breakdowns of existing reports than would otherwise be the case. It also leads to redundancy in the reporting process, as firms need to re-assemble the same underlying building blocks in different ways for different reports.

iii. Duplication of processes across firms

Many elements of the production of reports are common across firms. This raises the possibility that further centralising some processes could reduce duplication and improve the system's efficiency.

The review sought ways to decrease the burden on industry and to increase the timeliness and effectiveness of data in supporting supervisory judgements.

Because the BOE determines what information is required of regulated firms, it has a significant influence on their data governance and management. Therefore, the report pointed to the potential for the BOE to better support the firms' own use of data, making them more productive and competitive.

The BOE committed to conduct - in consultation with the industry - a review to explore the transformation of the collection, hosting and use of regulatory data over the next decade, identifying ways to decrease the burden on the firms and increase the timeliness and effectiveness of data supporting supervision.

Five **key challenges** relating to the efficiency and effectiveness of how data is collected, which can be found in different jurisdictions across the world, were identified (BOE 2021):

1. Complexity, legacy, and strategic planning

The two main sources of complexity are: i) on the firm side, the data for reports can come from various types of legal entities, business lines, and operational systems, and ii) on the government side, different authorities ask for similar data with slightly different definitions, across multiple reports, at different breakdowns. In addition, the legacy of decisions made in the past has created a complex reporting landscape that has not been adapted to current data needs. This is reflected in a legacy of manual, siloed processes, and outdated, fragmented operational systems. If the authorities are overly focused on meeting short-term objectives and lack strategic planning, they will not tackle these legacy issues.

2. Value and collection rationale

Data collection contributes to improving decision-making and making it evidence-informed. However, an agency can find it hard to estimate the value of reporting.

3. Interpretation

Industry participants expressed that understanding reporting instructions was one of the greatest sources of (avoidable) cost of the data collection process. The difficulties they mentioned include finding the latest version of the instructions, locating all the relevant documents, navigating the BOE website hosting the instructions, and understanding the instructions as they are written in over-complex legal language.

4. Finding and sourcing data

Due to their legacy systems and the complexity of data requested, at times industry participants found it challenging to locate or source the required data.

5. Reconciliation and data quality

Complexity and legacy issues made resolving data quality problems unduly difficult.

The BOE formulated its vision for data collection, which is that "***The Bank gets the data it needs to fulfil its mission, at the lowest possible cost to industry***", and developed its transformation plan for data collection addressing **three areas**:

1

COMMON DATA INPUTS

Developing common data inputs at a more granular level would provide a defined way for firms to record certain data (for example, data elements for individual mortgages) or capture the key elements in a common input layer. This could provide a more consistent cross-firm foundation from which to build reports, reducing costs and improving speed and quality.

2

MODERNISING REPORTING INSTRUCTIONS

Common data inputs could also form the basis of a move to modernise how the Bank writes reporting instructions. This could include moving from our current natural language approach towards more precise instructions for selecting and transforming the data of interest. Doing so could reduce the cost and time it takes for firms to respond to new requests.

3

CHANGES TO THE REPORTING ARCHITECTURE

Common data inputs could also support different architecture solutions, such as a 'pull' data collection model. A 'pull model' would allow the Bank to query certain data held within firms and generate reports on demand. This could improve the speed and flexibility of reporting while reducing the marginal cost to firms of responding to new questions.

The BOE-FCA Joint Transformation Programme

To help deliver this reform, in 2021 the BOE and FCA set up a 'Joint Transformation Programme', in collaboration with the industry. During each phase, with an iterative and pragmatic approach, the Programme aims to deliver a series of use cases, defined as a collection, set of related collections or an aspect of a data collection. Within each phase, each use case, in turn, passes through a 'discovery and design' stage, and then an 'implementation stage' (beta), where solutions are developed and tested for delivery.

Phase 1: Discovery and design

The use cases selected for this phase were:

- **Commercial real estate (CRE) data** (BOE use case), with a focus on improving the quality and coverage of commercial real estate data available to the Prudential Regulatory Authority (PRA) and the BOE's directorate for Financial Stability Strategy and Risk.

- **Quarterly derivatives statistical return (Form DQ)** (BOE use case), aiming to improve data on the derivative asset and liability positions of the UK financial sector. This data ultimately feeds into the UK's balance sheet compiled by the Office for National Statistics.

- **Financial resilience survey** (FCA use case), looking to formalise a post-pandemic ad-hoc collection of select data points from FCA firms, used for prudential risk monitoring.

Through a process of workshops and discussions between subject matter experts from the industry and regulators, the delivery teams identified similar issues across the different use cases with seemingly similar root causes. Some of these issues were:

- User experience challenges (such as users finding it difficult to find the right information they need to prepare reports, including the context and rational for the collection)

- Difficulties firms face in understanding and interpreting reporting requirements

- Issues firms and regulators face in providing and receiving feedback on data quality

- A lack of understanding by firms on the impact the data has on the regulators/supervisors (such as how the data is used to improve decision-making)

- Concerns that the same or similar data was being collected across multiple collections.

The identification of similar issues across different use cases supported the Programme hypothesis that data collection processes can be redesigned to achieve better effectiveness and efficiency.

Phase 2: Implementation (ongoing)

The use cases for this phase are:

- **CRE data**, building on the findings of phase 1, which confirmed that the current CRE data the agency receives is inadequate, fragmented, and burdensome to collect. Phase 2 will focus on exploring business practices and processes, mapping user journeys, creating a project roadmap and creating problem statements.

- **Strategic Review of Prudential Data Collection (SRPDC)** to reduce the cost of data production and reporting for firrms and deliver more value to both industry and agencies.

- **Retail Banking Business Model Data** that the FCA currently collects across a range of retail banking products and segments, and is critical to support the FCA's competition objective and

is reused by other FCA business units. The data is currently collected 'ad-hoc', without an integrated design that meets the needs of all the data users.

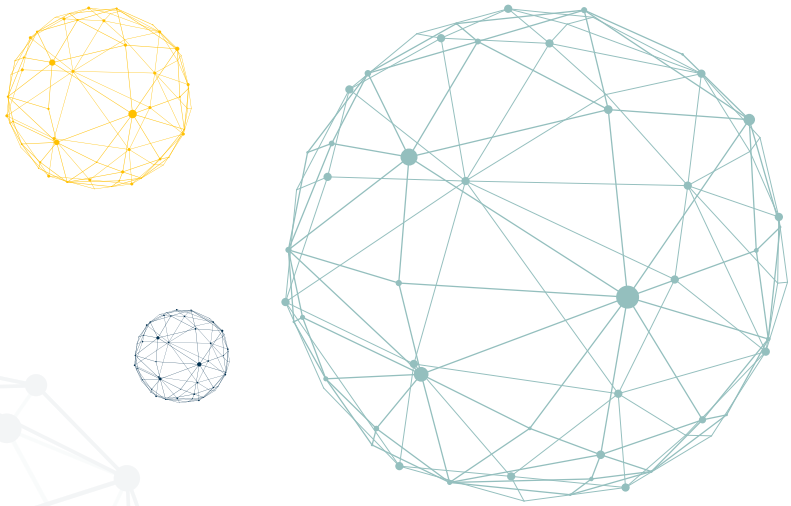
- **Incident, Outsourcing and Third-Party Reporting (IOREP)**, which the BOE, the Prudential Regulatory Authority (PRA) and FCA use to ensure the operational resilience of the financial sector, relying on data with low quality and consistency

In addition, these activities have been undertaken:

- **Data Standards Review (ongoing).** As part of their vision for increased development and adoption of common data standards throughout the financial sector, BOE and FCA set up the Data Standards Committee, which commissioned a review of data standards.

- **Banking Data Review (PRA, ongoing).** The PRA has announced the launch of a 'Banking Data Review' through a discussion paper (BOE 2022b) that outlines the increase in the scope of PRA's policy-making responsibilities and the expected impact of the Financial Services and Markets Bill. This initiative is being run separately by the PRA, however, it is complementary to the work being carried out within the Joint Transformation Programme.

- **Transition to Bank of England Electronic Data Submission (BEEDS) portal.** In 2020, the BOE Data and Statistics Division (DSD) announced a plan to move the collection of statistical data to the Bank of England Electronic Data Submission (BEEDS) portal (BOE 2020). BEEDS is an online application that enables firms to complete and submit data submissions online, and allows firms to view the information held about them by the BOE and keep it up to date. BOE also planned to move the reporting format from XML to XBRL. The transition was completed in December 2022 (XBRL 2022).



5.2. Data processing: Central Bank of the Philippines API-based prudential reporting system and back-office reporting and visualisation application

In 2021, the Bangko Sentral ng Pilipinas (Central Bank of the Philippines, or BSP) issued a memorandum to all banks requiring the submission of prudential reports via API. The regulation is phased to account for the varied stages of information technology maturity in the private sector, and phases are designed based on the readiness of banks. Along with this scaled rollout, the BSP introduced updates to their accounting procedures as well as the collected data elements. The ability to incorporate all the changes in one place for all supervised entities with a push API, rather than making manual updates across dozens of redundant, file-based reporting schemes, minimised what could have otherwise required prohibitively large amounts of time and effort.

This is the latest development of an initiative for the automation of data collection and processing that BSP started in 2017 in partnership with the R²A developing an API and back-office reporting and visualisation application (R²A 2018a) to:

- Allow the digitally diverse and geographically distributed set of financial institutions to submit high-quality, granular data digitally and automatically to the financial authority with higher frequency.
- Enable BSP staff to make data validation faster and analysis sharper by generating customised reports in different formats for supervisory and policy development purposes.

At that time, the BSP Supervisory Data Center (SDC) was receiving incomplete, late, and inconsistent reports from their supervised financial institutions. The process was primarily manual and highly resource-intensive, and data cleaning and

FIGURE 35.
BSP'S NEW API-BASED DATA ARCHITECTURE FOR PRUDENTIAL REPORTING

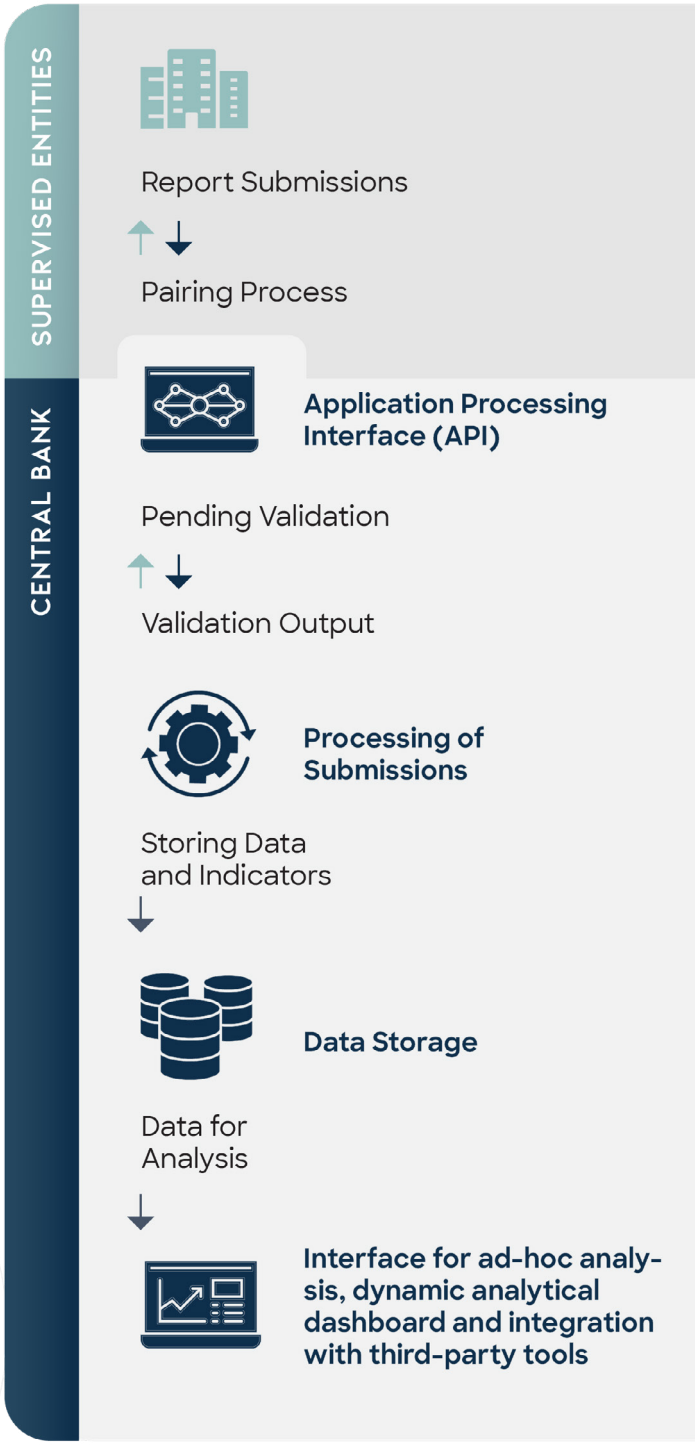
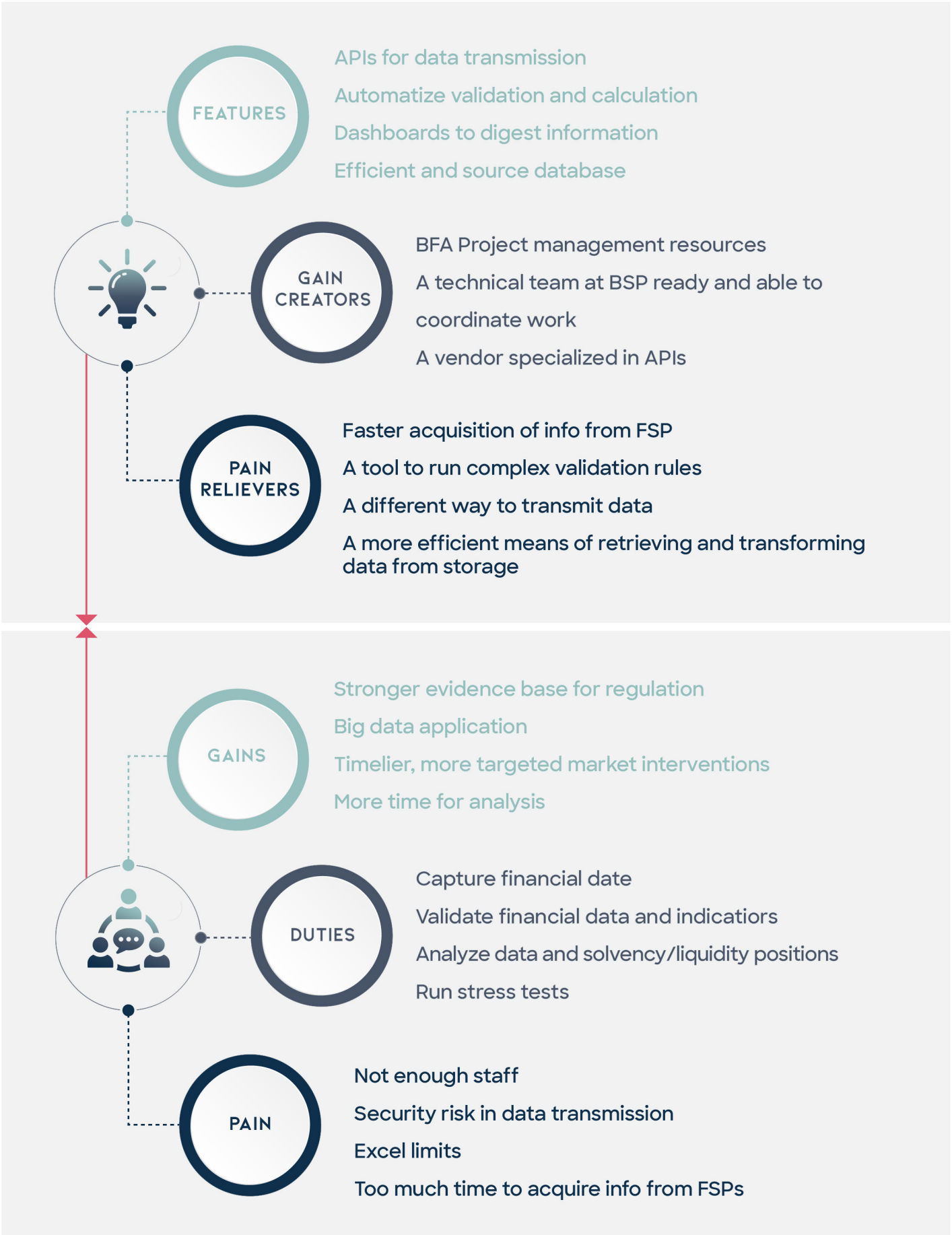


FIGURE 36.

VALUE PROPOSITION FOR BSP'S API AND BACK-OFFICE REPORTING AND VISUALISATION APPLICATION



validation were consuming significant resources.

By improving data quality and access, and developing new tools for data visualisation and analysis, the suptech application helped BSP implement a risk-based supervisory approach that reduces compliance costs and promotes financial inclusion while ensuring financial stability and integrity.

The solution represented a fundamental re-engineering of BSP's prudential reporting system. It consisted of a mix of relatively cost-effective, best-in-class technologies targeted at the various pain points of the existing reporting system. Together they formed a coherent, streamlined, and nimble architecture for (1) transmission, (2) processing, (3) warehousing and (4) analysis of banks' prudential reporting data. Each of these technical layers are described below, highlighting the efficiency gains reaped by diagnosing, automating and digitally transforming key aspects of the reporting process.

Application programming interface

The API establishes a direct line of machine-to-machine transmission between the supervised banks' core banking system and BSP without human intervention. Rather than having supervised entities manually populate multiple spreadsheet-based report templates, raw data is extracted from their databases and converted into a single encrypted XML-based document. This content is then 'pushed' directly to BSP's processing queue, obviating the need for legacy processes such as email or web data portals.

XML-based APIs have several advantages over the existing prudential reporting system:

- **Larger size limits**. The volume (and hence granularity) of data that can be transferred is increased considerably. Not only can the XML format hold far more memory than Excel files, but API transfers have fewer size limitations than email attachments.
- **Increased security**. This mode of transmission is inherently more secure than email as there is less room for human error (for example, a file is accidentally sent to the wrong person or through a compromised email account). Furthermore, BSP's solution contains several industry-standard security protocols to encrypt and authenticate files.
- **Reduced redundancy**. Sending raw data cuts the total number of data points submitted – in BSP's case, from ~107,000 to ~50,000 – since duplicate and calculated fields are eliminated. Likewise, transferring raw data rather than preformatted templates means that amendments to data requests can be accommodated relatively easily.

In addition to the XML-based machine to machine APIs described above, an alternative, phased approach was designed into the system to aid in onboarding supervised entities with less technically advanced staff and infrastructure. For these entities, software was provided that directly converts legacy Excel-based templates to XML documents to be submitted via the API, allowing them to continue to comply via legacy processes during their digital transformation, while still affording the benefits of the API.

Processing engine

A new processing engine performs all data validation and formula calculation in one dedicated location based on rules defined in a dynamic business syntax. The API uses a technological standard called XML Schema Definitions (XSD) to specify strict validation rules for required elements, attributes, and data types and formats allowed to be submitted. Thus, validation tests are run near- instantaneously upon receipt of the file submission, with the reporting entity receiving an automated

response (via API and message) detailing the test results and flagging any abnormal or missing values. Broken rules result in the report being rejected and sent back for review and resubmission. Once validation has been completed, the data is run through the calculation engine, where all the relevant prudential indicators and risk metrics, such as capital adequacy ratios (CAR), are calculated.

With over 7,000 validation rules prescribed by BSP, a single processing engine ensures that all tests and calculations are run uniformly rather than on disparate spreadsheets whose formulae may be inconsistent, broken or outdated. It also allows for more complex number crunching than might be possible in Excel. Automation significantly reduces processing time, especially for datasets that were validated through painstaking manual reconciliation. Excluding these, the average processing time fell from an average of 1,800 seconds to around 10 seconds. Finally, a calculation engine with a user-friendly interface where indicators are defined or amended means that methodological changes – for instance, in calculating risk-weighted assets (RWAs) – are implemented relatively easily.

Data storage

After processing, the cleansed data is passed directly into a centralised, secure, access-controlled database for storage without the need for manual uploads as in the old system. This warehouse facilitates historical data access and more efficient database management. BSP will be able to extract data in different ranges using pre-scripted SQL queries without requiring BSP's intervention.

Analytical tools

A centralised database also expands the array of analytical tools that can be applied, including dashboards and statistical software. The BSP prototype included a web-based pivot table tool as well as charts depicting key performance and risk indicators. More interactive visualisations and sophisticated big data applications such as predictive analytics and machine learning are also under consideration for future deployments.

TABLE 2.
KEY PERFORMANCE METRICS FOR BSP'S PRUDENTIAL REPORTING SYSTEM

PRIOR REPORTING SYSTEM	API PROTOTYPE
29 reporting schemes (14 financial reports + 15 related reports)	**1 unified scheme** (in XSD format)
~ 243 reporting templates	**~ 210 reporting templates**
~ 107,000 data points	**~ 50,000 data points** (other data points are calculated in fields)
~ 7,000 validation rules (in several layers, some spot checked by human)	**~ 7,000 validation rules** (single validation layer in dynamic business syntax)
~ Layered reporting packages (multiple submission of reporting package via email)	**Single reporting package** (automatic submission via API, no human intervention needed
Processing time avg. > 1800 sec (exculding manual workarounds)	**Processing time avg. ~ 10 sec** (excluding manual workarounds)
Multiple processing and analytical layers (some manual)	**Single processing and analytical layers** (processing and validation layer)

Beyond the prototype

In 2021, the BSP issued a memorandum to all banks regarding the submission of prudential reports via API. The regulation is phased to account for the varied stages of information technology maturity in the private sector, and has phases designed based on the readiness of banks.

Along with this scaled rollout, the BSP introduced updates to their accounting procedures as well as the collected data elements. The ability to incorporate all the changes in one place for all supervised entities with a push API, rather than making manual updates across dozens of redundant, file-based reporting schemes, minimised what could have otherwise required prohibitively large amounts of time and effort.

5.3. Data storage: National Bank of Rwanda Electronic Data Warehouse

The Electronic Data Warehouse (EDW) is a supervision information system (SIS) initiative of the National Bank of Rwanda (BNR), based on 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 costed approximately USD$ 1 million to implement. It overhauled previous data-management systems, requiring investments in hardware and software as well as in upgrading staff skills. The BNR also spent signficant time in coordination with over 600 institutions it supervises.

The EDW solution introduced three new dimensions to BNR's regulatory reporting infrastructure:

1. Data-pull technology that allows supervisors to connect directly to databases of the supervised firms and collect data from the source rather than sharing data via Excel spreadsheets
2. The collection of more granular account-level data, provided daily, rather than aggregated by an institution monthly or quarterly
3. Data analytics and reporting that is now automated and linked to interactive dashboards

The BNR uses the EDW to pull data directly from the IT systems of more than 600 supervised financial institutions, including commercial banks, insurance companies, microfinance institutions, pension funds, forex bureaus, telecom operators and money transfer operators. Data is automatically pulled from these institutions at various intervals (even every 15 minutes for mobile money and money transfer operators). For this purpose, a data dictionary was developed, and each

financial institution was required to write data scripts that would map the data dictionary to the information in its own systems. This, however, was a one-off investment.

The mapped information is then put in a 'staging area' where the BNR can pull the information it needs. The encrypted data is transmitted over a VPN channel and through data integrity check mechanisms. In this way, the data pull approach delivers timely, consistent, and reliable data to the BNR. It has also led to improvements in financial institutions' data, which they now also use for internal risk-management purposes. The EDW reduces errors and inconsistencies.

To this end, the BNR has built quality and integrity rules into the system. If data does not meet certain standards, it is rejected, and an automatic email alert is sent to bank examiners within the BNR and the supervised institution. Furthermore, historical data cleaning from the data supply side is currently in progress. Next to improving data quality, the EDW also offers flexibility and the ability to quickly analyse large amounts of data.

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 reporting firms. This is a result of its technical design for software interoperability. Regulated institutions can continue using the same database provider (e.g., Oracle, SQL, MySQL) and connect to the EDW using simple data-transfer protocols. Further, management at the BNR reports that frequent engagement with FSPs, particularly relating to providers' concerns about the level, nature and frequency of supervisors' access to their data, was key to its ultimate widespread adoption.

The BNR can adapt its supervisory processes and methodologies to fully leverage the collected data and allocate supervisory resources more efficiently. The BNR has not ended manual reporting completely due to data gaps on the part of external stakeholders. The BNR is also streamlining internal business processes to ensure that information is completely captured.

Throughout the three-year initiative, management at the BNR indicated the importance of managing change within the financial authority. Supervisory staff accustomed to the 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 mechanical processes like consolidation and cleaning.

FIGURE 37.
DATA FLOW DIAGRAM OF SUPERVISORY INFORMATION SYSTEM (SIS) SOLUTIONS

DATA COLLECTION & VALIDATION

Supervised Institutions
share data via automated date flows (**ADFs**), APIs, or other secured transmission.

System performs **data validations** and **transformations.**

DATA STORAGE & MANAGEMENT

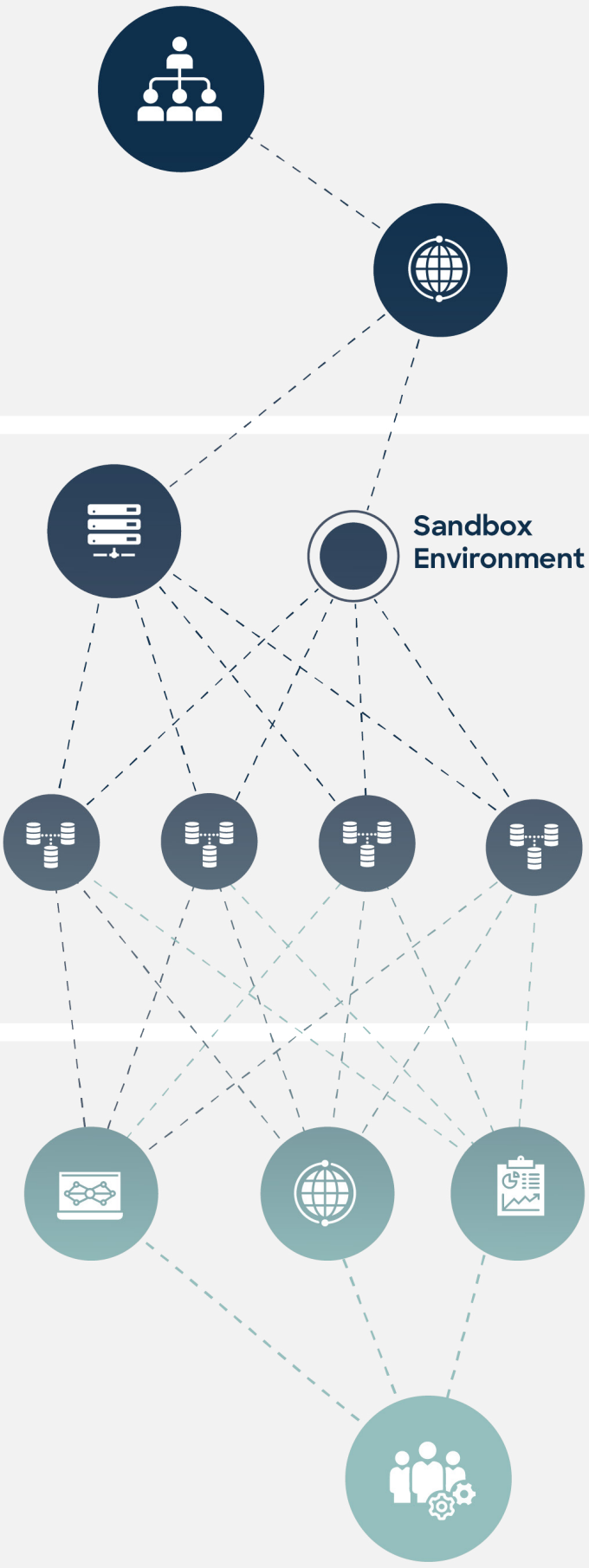
Enterprise Data Warehouse
The validated data is archived and stored in a central data warehouse. Additional datasets are merged and test data is available in the Sandbox Environment

Datamarts
Datamarts manage permission-based access of data to specific departments or teams

DATA ANALYSIS & REPORTING

Analysis Tools
Dashboards, alerts, statistical tables and graphs are created to understand trends and risk within the data.

Data Consumer
Supervisors, risk experts, and other regulatory staff use the date and insights for oversights.



5.4. Data analytics: Central Bank of the Netherlands' outlier detection tool for AML/CFT/PF supervision

The Central Bank of the Netherlands (De Nederlandsche Bank – DNB) has developed a suspicious activity detection application to select those customer files with the highest chance of requiring reclassification to a higher risk level. The development of this capability resulted from the successful cooperation between on-site and off-site supervisors and the DNB Data Science Hub (DSH).

To comply with the Dutch Anti-Money Laundering and Anti-Terrorist Financing Act (WWFT), financial institutions must assign risk categories to their customers (see DNB's Guidance on WWFT for details). To assess whether the institution's risk classification is consistent with the customer profile, supervisors analyse a limited number of individual customer files in their customer-due-diligence (CDD) deep dives.

Traditionally, to select these customer files, supervisors define hard-coded tell-tale identifiers ('business rules') to spot potential suspicious activities from a dataset containing millions of accounts and billions of transactions. These rules are defined based on a combination of account characteristics and transaction patterns, among other conditions. For example, 'multiple accounts on a single address' with 'a single deposit per month and immediate withdrawal' could point to human trafficking of seasonal workers – the hypothesis being that subcontractors organise housing for seasonal workers (a perfectly legal activity), but then withdraw the wages deposited and only pay a fraction to the worker, which is illegal.

In practice, the DNB applies business rules combining more than two conditions to improve accuracy. However, given the large size of this dataset, it is impossible for supervisors to map all the potentially significant business rules, especially those that combine conditions that interact in non-linear and unpredictable ways. So, how to detect potentially fraudulent transactions in a dataset that contains millions of accounts and billions of transactions? Data science has a clear role to play here.

The DNB's Supervision Department and the DSH jointly developed an application to enhance data analytics capabilities and generate richer insights for supervision. After exploring the supervisors' needs and expectations, the features of the application were defined as follows:

- An unsupervised learning model to select customer files with a high chance of requiring reclassification to a higher risk level
- Reusable code to run a similar analysis for future datasets
- Dashboards to visualise the results of the model, highlighting risky customers

The use of an unsupervised learning model was proposed because supervisors wanted to identify new risks but didn't have sufficiently suitable labels for supervised classification. To build the model, the team started with some exploratory data analysis and background research about commonly used analytical methods. Next, they loosely experimented with different methods, such as clustering, t-distributed Stochastic Neighbour Embedding (t-SNE), and multiple algorithms for outlier detection. Due to the lack of experience with this dataset and some data quality issues, the team needed around one full-time employee (FTE) for about five months to build the model.

On top of the explored methods, the team built an outlier detection model using the Isolation Forest algorithm (Isotree R package), which was found to be the

most suitable for the business case. The model was programmed in the R programming language using RStudio, and data was stored in several tables in a Microsoft Server SQL Database. Due to the unsupervised nature of the model, the supervisors were asked to label whether the top outliers were interesting enough to perform a CDD deep dive on and to compare the outcome with the risk category assigned by the financial institution. In the selection of files, supervisors need to analyse data within the client files. To aid this process, dashboards were developed to provide the supervisors with a first insight into the details in the client files.

The results showed that the model worked well, especially for the portfolio of private clients. Iterating over several deep dives at different banks, the team trained the model and created dashboards that allow supervisors to have much more informed discussions with supervised institutions. Dashboards were created using Shiny in R and can be run locally.

Benefits and impact

In the last year, supervisors have employed the outlier detection model, being able to use the most unlikely combinations of conditions to spot potential suspicious activity in half the time previously needed. According to Tim Haarman, a data scientist on the project, the application has proved very useful:

> *"Our colleagues in integrity supervision can now do their work in a more efficient manner by selecting the riskiest files using data science.*

According to the supervisors, the benefits of the model are:

- **New risk identification**: Previously, using manual analysis, supervisors could only check for risks they already knew existed. The risks they were unaware of may remain under the radar. An objective unsupervised analysis, instead, allows new risks to become apparent.
- **Efficiency gain**: The DNB estimated that going through the data manually with queries would cost around two weeks of one FTE. Using the outlier detection model, the same task took around one week. For both the old and new approaches, onboarding and cleaning the data is still necessary. However, a fair comparison is difficult since, in contrast with the old manual approach, where only a tiny fraction of the data is investigated, the new approach considers all available data.

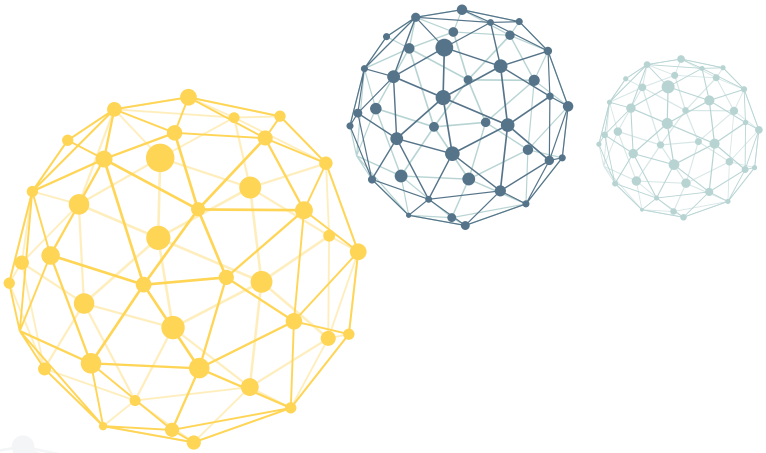


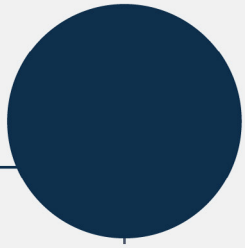
FIGURE 38.

HOW THE DNB APPLIES THE OUTLIER DETECTION MODEL FOR CUSTOMER FILES SELECTION

1

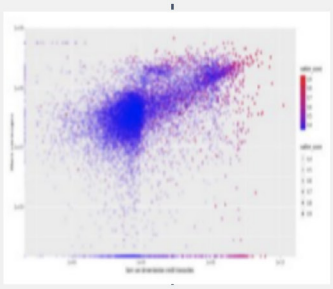
LARGE VOLUME OF CLIENTS AND TRANSACTION DATA RECEIVED FROM SUPERVISED BANK

Mass Retail Portfolio:
Millions of clients
Billions of transactions



2

DATA PARSED BY ALGORITHM PROVIDE LIST OF OUTLIERS



3

FIRST SELECTION

Mass Retail Portfolio:
Billions of transactions

Outlier Detection:
Less than 1000 clients

4

OUTPUT: A LIST OF ALL DOSSIERS ABOVE A PARTICULAR OUTLIER SCORE

R	S	T	U	V	W	X
878240	14849	236	238	111	41	0.907103353
545120.3	58767.9	86	9	763	34	0.898110413
0	17175.08	471	69	17	3	0.893839643
117370	144540.2	139	12	69	2	0.886751398
0	0	157	1	9	0	0.88504812
142350	0	10	0	10	7	0.871525497
0	0	7	0	0	0	0.856794763
0	29425.22	175	155	90	0	0.856768825
0	259768.67	127	29	2	5	0.855067009
0	0	228	0	230	3	0.852238653
50370	590000	596	375	67	4	0.843576917
0	137548.79	3	3	0	0	0.843477995
0	80500	106	11	18	2	0.84039536
0	0	111	5	1	0	0.838696024

5

SELECTED CLIENT FILES ARE CHECKED

Mass Retail Portfolio:
Millions of clients

Outlier Detection:
Less than 1000 clients

Outlier Detection:
Selection of small set of files

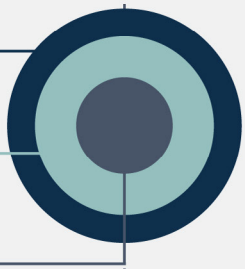


FIGURE 39.

HOW THE DNB OUTLIER DETECTION MODEL IMPROVED CUSTOMER FILES SELECTION

Lessons learned

Working closely with supervisors and data scientists can definitely add value to the supervisory process. The proximity between data scientists and business units allows them to provide input to developing the model, as well as interpreting its outcomes. However, in organisational terms, combining data science knowledge with business knowledge can present challenges. If you embed data science into business units, it is difficult to keep data science knowledge up to date. Conversely, if you set up a central data science office, you would worry about hobby horses and empire-building.

To address these challenges, the DNB chose a hub-and-spoke system. The DNB data scientists work in a hub (Data Science Hub, DSH) but collaborate with the spokes – the various business units of DNB – to develop the tools they need jointly. This comes with the additional benefit of making optimal use of potential spill-overs – using the same code for different applications – while working on a variety of topics.

So far, so good. But if the business units want to embed the data science tools, such as the outlier detection tool, into existing workflows, there is the challenge of correctly implementing and, subsequently, maintaining the tool.

In the DNB's current organisational setup, the supervision department has its own innovation sub-department, which has people both technically skilled and with knowledge of the supervision process. The DSH collaborates closely with this sub-department providing the data science and machine learning expertise. Since their local innovation sub-department is involved in the development process and also technically skilled, supervisors can implement the tool into the workflow and run it themselves for new investigations. Having ownership of the product close to the business is crucial in the long term.

An alternative approach with which DNB is experimenting to guarantee the maintenance of the tool is something called BizDevOps. This is an organisational form where both the development and operation of the tool are taken on by the business unit. It encourages collaboration between three parties – the business, the developers, and the operations team – by combining development and operations into a single team. What this organisational form aims to achieve is that there is ownership of and more engagement with the project on the business side. Otherwise, innovative data pilots tend to show a cool POC and die off afterwards as no one continues maintaining it. Having this engagement from the beginning makes it much more likely that developed products will actually be used in the long term. The DNB is not yet sure whether this provides an optimal solution, but they believe it is worth testing it.

5.5. Data products: Reserve Bank of India's DAKSH

The Reserve Bank of India (RBI) has been taking various measures to strengthen supervision, including adopting the latest data and analytical tools and leveraging technology for implementing more efficient and automated work processes.

RBI's leadership has explicitly called for harnessing the power of suptech for "improving efficiency through the use of automation, introducing new capabilities and streamlining workflows," and enjoyed early successes with their Import Data Processing and Monitoring System (IDPMS), Export Data Processing and Monitoring System (EDPMS) and Central Repository of Information on Large Credits (CRILC).

In October 2022, the RBI announced a new suptech initiative – Reserve Bank's Advanced Supervisory Monitoring System. Referred to as DAKSH – a Hindi term meaning 'efficient' or 'adept' – the tool offers a web-based interface, through which the RBI can monitor compliance requirements more closely.

It has been referred to as an "anytime-anywhere secure" end-to-end workflow application, which offers:

- Field service management software that provides end-to-end solutions for processes from the beginning of a workflow until the end.
- Everything from the preparation stages right through to the post-launch of the application.
- Elimination of any unnecessary steps in the end-to-end process and, in turn, enhanced efficiency of the system.

Through DAKSH, the RBI aims to improve the compliance culture in supervised entities, such as banks and non-bank financial companies (NBFCs) and make the supervisory processes more robust. The tool is also expected to enable seamless communication, inspection planning and execution, cyber incident reporting and analysis, and provision of various MIS reports. Additional advantages include:

- Secure access to regulatory data, anytime from anywhere
- Ability to make informed decisions using specific sets of data
- Complete visibility from start to finish on a single project
- More time to focus on providing cost-effective solutions

In December of 2022, the RBI announced the fraud reporting module of their Central Payments Fraud Information Registry (CPFIR) would be migrated to the DAKSH system effective January 1, 2023.

All Payment System Operators (PSOs) will be required to submit all payment frauds, including attempted incidents, irrespective of value, either reported by their customers or detected by the entities themselves. Historical data previously submitted via the legacy Electronic Data Submission Portal (EDSP) will be migrated to the new system. All data collected will be required to follow a standardised, compressed format that can easily be parsed by validation and processing tools.

This effort is expected to further "streamline reporting, enhance efficiency and automate the payments fraud management process." In addition to migrating this bulk upload functionality from the legacy EDSP, DAKSH will introduce further benefits to the CPFIR system including:

- Maker-checker facility
- Online screen-based reporting
- Option for requesting additional information
- Facility to issue alerts/advisories
- Generation of dashboards and reports.

5.6. Full stack: BIS Project Ellipse, an integrated regulatory reporting and data analytics platform

Project Ellipse is a POC and working prototype launched by the Bank for International Settlements Innovation Hub along with the Monetary Authority of Singapore (MAS), Bank of England (BOE), Financial Network Analytics (FNA), and Accenture to explore how supervision could become insights-based and data-driven using a full-stack, integrated regulatory data and analytics platform.

Project Ellipse's platform aims to empower regulatory authorities with the ability to digitally extract, query and analyse a large quantity of data from diverse sources, such as those included in a demonstration of the initial working prototype:

- Total and secured loan exposures
- Total supervised banks
- Total borrowers
- Breakdown of exposure by bank
- Breakdown of exposure by industry.

These data can then be matched against other data sources in real time, notably including both structured and unstructured data, drawn from both public and proprietary sources. Examples presented in the prototype included:

- Internal audit reports
- External audit reports
- Credit risk
- Liquidity risk.
- Governance risk
- Social network mentions and public sentiment
- Public financial and governance data on supervised entities
- Current news and events relating to the financial system.

Insights derived from these combined datasets are made visible via dashboards, to provide the ability to conduct network analysis and stress test (e.g., by modeling defaults and contagion), and ultimately inform financial authorities of effective early supervisory actions that may need to be taken.

Project Ellipse is a uniquely prominent demonstration of how suptech solutions have the potential to enable supervision to be more forward-looking, insights-based and data-driven using a full-stack, integrated regulatory data and analytics platform. Advanced analytics are then applied to those integrated data sources to provide supervisors with early warning indicators, analytics and prudential metrics.

Challenges of regulatory reporting, and potential solutions through Project Ellipse

Supervisors today rely heavily on regulatory reporting to identify potential risks that may be forming in regulated entities, which could have broader and systemic implications for the financial system. However, there are challenges with the information supervisors receive from regulatory reports compared with the richer data introduced by suptech tools like Project Ellipse.

The core supervisory limitation resulting from these challenges is the ability to form a complete and accurate picture of exposures to have predictive insights into emerging risks using these data sets. Project Ellipse, therefore, demonstrates one means through which supervisors can better identify and assess emerging risks in real time to inform them of early supervisory actions that may need to be taken.

To meet these challenges in the digital age, authorities could benefit from 'on-demand' access to timely and integrated data sources to help support and inform their supervisory assessments. Several possible solutions are therefore explored in this project.

In phase 1, the project investigated how data-driven supervision could be enabled by machine-executable digital reporting using a cross-border common data model. This POC explored the feasibility of cross-border digital reporting, and is intended as a first step towards bringing authorities and stakeholders closer to a common understanding of data that is collected by authorities globally. The exploration confirmed that:

- Regulatory reporting requirements can be expressed in unambiguous machine-readable logical reporting instructions underpinned by a consistent data model.
- Technical standardised programmatic specifications of the steps for generating regulatory reports can be published alongside regulation and ensure clear understanding at the most granular level of the expected data.
- Executable libraries can be automatically generated and published alongside regulations to assist accelerated implementation.

TABLE 3.
CHALLENGES IN REGULATORY REPORTING

CHALLENGE	DESCRIPTION
TEMPLATE-BASED, AGGREGATED	Regulatory requirements are often template-based and ask for aggregated data, meaning that datasets are fixed to a use case, and hence the data received cannot be easily reused for other purposes. New reporting requirements are needed each time additional or ad-hoc information is needed.
DATA IS INCONSISTENTLY DESCRIBED	Reporting data is often sourced from legacy data systems within reporting firms that are not always integrated. This often results in the heterogeneity of data for any given product or transaction – both within a bank and across different banks – as systems will describe these data differently.
INFREQUENT BACKWARD LOOKING	Regulatory reports are submitted to supervisors from reporting entities on an infrequent basis (for example, every month or quarter). At times of heightened risk, the need for up-to-date data increases, but given the static nature of regulatory reports, supervisors may not have the timeliest data to make informed judgments.
DIFFERENT SOURCES OF DATA ARE NOT INTEGRATED	Information contained in regulatory reports is often linked with other types of information that may point to emerging risks, but these sources of information are not connected. For instance, information sourced from market data and news are often the first indications of emerging risks, but it is difficult for supervisors to scan through the vast volumes of market and news data to assess what events should trigger early action.

- If a common data standard was agreed to and implemented, financial institutions may no longer need to interpret reporting instructions and submit aggregated data by use case.
- With additional logical instruction based on the same data model, supervisors may also be able to query the underlying transaction data automatically and generate regulatory metrics referencing that standardised data.

Phase 1 illustrated the possibilities and efficiencies that could be gained if machine-executable reporting using

TABLE 4.
SOLUTIONS TO ADDRESS CHALLENGES IN REGULATORY REPORTING

SOLUTIONS	DESCRIPTION
GRANULAR DATA	The collection of granular data from reporting entities could replace the need for authorities to request information using templates. It could also enable authorities to reuse those data for different use cases. Supervisory metrics could also be derived using granular data, as opposed to requiring reporting entities to aggregate the required data prior to submission.
COMMON DATA MODELS	Differences in the description of data for similar products and transactions across banks can be addressed using data standards and common data models. Granular reporting requires a common understanding by authorities and financial institutions of what those data are so that financial institutions can map their operational data to a common 'input' before the required data can be reported. Supervisory metrics could then be derived using programmable rules referencing machine-readable and machine-executable common data models.
REAL-TIME INFORMATION	Real-time insights using advanced analytics could be derived from large volumes of unstructured data that would supplement the granular reporting available. This would provide supervisors with additional indicators and early warnings of at-risk exposures of reporting entities.
INTEGRATION OF STRUCTURED AND UNSTRUCTURED DATA	Integrating granular data from reporting entities with other sources of unstructured information, such as news and market data onto the same platform means supervisors would not have to manually scan for information. Advanced analytics such as artificial intelligence and machine learning could be used to make risk correlations and analyse sentiment, alerting supervisors in real time of issues that may need further investigation.

common data models were adopted. This could also increase the volume of granular data available to supervisors, as needed to enable the use of advanced analytics.

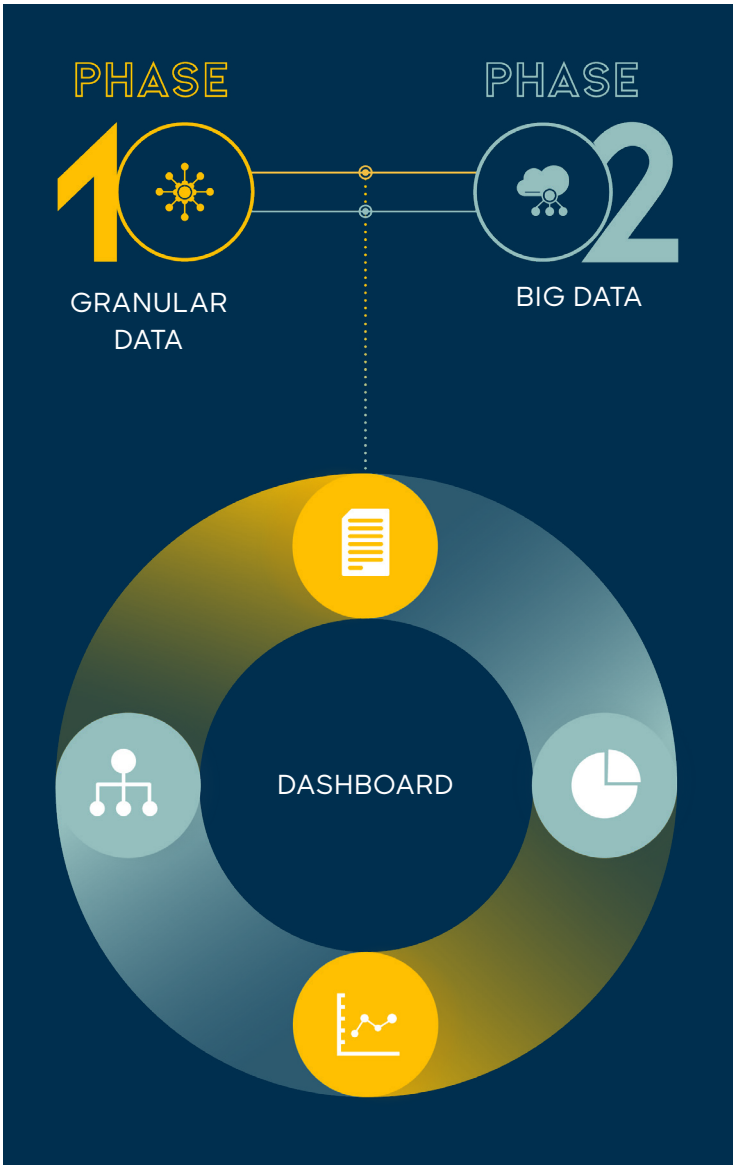
In phase 2, the project explored how a single platform could be built so that authorities could benefit from 'on-demand' access to timely and integrated sources of data to help support and inform their supervisory assessments. An important aim of phase 2 was to show that, as authorities continue to explore digital reporting and the collection of granular data, they could, in the meantime, use existing regulatory data on the Ellipse platform, such as data collected on large exposures. It also wanted to show that a single platform could act as a 'one-stop shop', where supervisors could find regulatory information quickly and be guided by insights generated from the analytics running on the platform.

Outcomes

During the limited POC and working prototype, Project Ellipse demonstrated critical proof points for the use of suptech in addressing supervisory challenges. Insights include:

- Regulatory data that has sufficient granularity, such as large exposures, can be integrated on the same platform with other sources of unstructured information, such as news and market data
- Advanced analytics can automate the assessment of possible impacts on key prudential metrics, providing supervisors with real-time insights and early warnings of the at-risk exposures of reporting entities
- Access to information automated in real time enables supervisors to investigate risks further and challenge reporting entities faster, allowing earlier intervention

FIGURE 40.
THE TWO PHASES OF PROJECT ELLIPSE



- A single-user interface with multiple functionalities can meet the needs of different units with different supervisory or oversight responsibilities, with authentication systems allowing access only to authorised users
- A platform – built to accommodate multiple applications – offers the potential to scale up with larger and more granular datasets as the financial sector and the digital datasets surrounding it continue to evolve.

6.

CONCLUSIONS

The suptech ecosystem is expanding. As from the early 1990s, a wave of innovation has emerged across the globe and is modernising financial supervision, accelerated in recent years when most financial authorities have engaged in the adoption of suptech strategies, roadmaps, and solutions.

Financial supervisory authorities are making headway in their suptech adoption journey leading to desirable outcomes such as improved risk-based supervision and efficiencies in data collection and use of resources. Suptech is enabling the automation of existing processes, as well as brand new use cases, opening completely new opportunities for supervisors. The ability to ingest massive online datasets like social media streams to conduct sentiment analysis, to parse online reviews to assess risks or identify fraudulent fintech apps, and to conduct real-time, on-chain analyses for digital assets supervision are just a few of many examples.

Agencies in AEs are early adopters of suptech, however ,the gap seems relatively small - and certainly there are innovation champions in the global south - and they report challenges that are very similar to the their peers in EMDEs are facing (first and foremost, limitations in budget, data quality and technical skills).

Worldwide, agencies are still addressing issues related to the foundations of their data architecture (data collection as well as descriptive and diagnostic analytics), and many suptech efforts are still in the experimentation stage, lacking the needed resources to be taken into production and scaled.

Although it is encouraging to see more suptech deployments, many suptech vendors are still struggling to get traction, transform or scale, and are calling for legacy procurement processes to be reformed to be able to serve the market.

The survey, questionnaire, and interviews that the Lab implemented for this Report, and evidence from extensive engagement with supervisors and experts through other activities (such as the Lab events and the Innovation Leadership Programme) suggest a few key takeaways on possible trends to monitor in 2023.

Develop a suptech strategy and/or roadmap

While running experiments at an early stage of the modernization journey can be useful to familiarise with foundational innovation methodologies and test new ideas, defining suptech strategies/roadmaps is critical, and only a few financial authorities have done this so far.

Suptech modernisation is a journey, not an event. Some agencies believe that the digital transformation of financial supervision leads with technology, and have not yet developed a strategy/roadmap for incremental, comprehensive transformation - guided by outcomes and data, not tech - and tackled important aspects such as change management and the upgrade of technical skills.

The findings shared in this Report show that financial authorities have adopted different strategic approaches in the deployment and governance of suptech. While there are portable elements (e.g., taxonomies and open software) there is no singular one-size-fits-all; each agency must adapt suptech to strike a balance between the breadth of its transformation goals and the existing context or culture of the agency. When it comes to suptech deployment, some agencies have taken a holistic approach and incorporated suptech as part of an institution-wide digital transformation strategy. In contrast, others have taken a phased approach based on supervisory focus areas or use cases, which is the more popular approach. The governance of suptech deployments also adopts different approaches. Primarily, the suptech deployments are led by either the operational teams or the IT department.

We foresee in the near future a growing number of agencies elaborating suptech

strategies or roadmaps (possibly with their broader digital transformation planning). Three insights point in that direction (the first two from this Report, the third one through the Lab's panel series):

1. Recently, more financial authorities have formed dedicated teams or units to lead suptech deployments.
2. Senior management has been increasingly endorsing the development of suptech initiatives.
3. Agencies are increasingly conducting rapid, iterative test-and-learn cycles on new technologies and process innovation.

Build data capabilities for the supervisors of the future

Rolling out suptech solutions can be an undertaking. And the development and implementation of suptech strategies and roadmaps also present challenges. One element that cannot be overlooked is the need for specific competencies and capabilities in the management and workforce. According to the survey findings:

- Limited data capabilities are among the top five challenges supervisors face in deploying suptech.
- Training in technology and data is on top of supervisors' lists when asked what support they require to successfully implement suptech.
- Supervisory agencies from EDMEs wish to have additional support in the form of technical assistance to further data analytics.

Acquiring and/or training data talent and skills throughout agencies is imperative for the development of suptech strategies and for scoping, prototyping and deploying suptech applications, especially as the advanced technologies that underpin suptech become more pervasive. Supervisors will need to understand how and why these new technologies deliver value, and, in most cases, they may need to work collaboratively with complex applications and data environments, which means they may require not only new skills but also entirely new ways of thinking about data-enabled, data-informed work. Moreover, data taxonomies, glossaries, and governance frameworks remain to be developed.

Throughout 2022, supervisors shared with the Lab's team the need and desire for their agencies to develop a diverse set of skills that - in addition to data science and analytics - includes business analysis, product design, and product and project management. In 2023, the Lab aims to address this demand with various capacity-building and training activities.

Grow a data-driven innovation culture

Cultural resistance is a key inhibitor of the adoption of suptech initiatives and applications. Past efforts in this space (BIS 2015, BIS 2016, R²A 2018d) have established a set of good practices that are applicable in today's suptech context:

- Establish appropriate communication with stakeholders and seek proper institutional endorsement
- Ensure a clear legal and regulatory basis to support data-sharing
- Establish fully fledged cooperation at all levels
- Collect common data using joint methodological and technical standards
- Ensure sound measures to protect confidential information
- Formalise governance and cooperation arrangements.

More recent work (HBS 2020, di Castri 2023 forthcoming) takes it one step further, in an effort to integrate suptech deployments into data-driven culture as a fundamental component of transforming the broader organisation. To this end, recommendations include:

- Updated mission statements and strategic plans, informed by technical experts from both within the agency

and the industry

- Recruiting and retraining, to ensure high-caliber data and tech talent
- Process transformation to break vertical and horizontal siloes
- Focus on field examiners to ensure this shift is not happening in a vacuum, but having the intended impact on the supervisors themselves
- Agile workflow, to allow for more granular and measurable achievements and accelerate progress
- Location and lifestyle, to promote a workforce that attracts high-tech talent
- Realistic standards, that allow for measurement of progress against the status quo and not against perfection

By getting these foundational processes right, financial authorities can unlock the scalability, interoperability and actionable intelligence promised by suptech.

We expect agencies to increasingly realise that suptech goes beyond disruption through technology, and the scope of supervisory innovation initiatives to embrace their overall capability to rapidly evolve and adapt through intelligent use of data, technologies and talent. It is about the agency's ability to develop agile and innovative models to exploit new opportunities to maximise efficiencies and create value.

Scale

As suptech initiatives (strategies, roadmaps, and applications) are accelerated, innovation culture becomes pervasive, and change management is addressed, suptech is starting to scale across two distinct and complementary layers:

- **Intraorganisational scale.** On a 'micro' level, scale occurs within a given financial authority. This type of scale consists of evolving from building a given set of innovative solutions, approaches, or processes produced through the success of a small pilot, into building a larger more strategic initiative across departments or thematic focus areas of an agency, once concrete value has been demonstrated and associated risks provably mitigated (OECD 2017).
- **Ecosystem-wide scale.** On a 'macro' level, scale means an increased ability to leverage data and resources from across a growing set of interrelated initiatives. For financial supervisors, this may include drawing from and combining an expanded set of data sources (R²A 2020), incorporating successes codified in open-source software repositories and digital public goods (DPGA 2022), and replicating/adapting modular components of proven applications (Taterinov et al 2022) such as the type listed in the Lab's SupTech Marketplace.

Through continued investment in addressing the use cases and challenges identified in this Report, suptech applications will become applicable and portable to the needs of a growing set of supervisors within and across agencies and jurisdictions. Thus, financial authorities must prepare themselves to intentionally and strategically harvest the bountiful crop of valuable lessons and resources from today's existing suptech initiatives, to consume the resources to address their own needs by building and procure tomorrow's suptech solutions, and to plant the seeds for new growth in the broader suptech ecosystem by sharing their own successes, lessons and digital resources.

To complement this imperative and accelerate progress toward such ecosystem-wide scale, the development of an intentional strategy for the effective and responsible sharing of code, data, models, and approaches related to suptech will remain a core focus of the Lab (Grasser 2023 forthcoming). Only through such collaboration can suptech as a movement ensure these scalable solutions are built and kept resilient to the nuanced and rapidly evolving needs of financial authorities undergoing digital transformation, as well as of the broader set of players exploring and engaged with the suptech ecosystem.

REFERENCES

1. Asian Development Bank (2022) Building regulatory and supervisory technology ecosystems for Asia's financial stability and sustainable environment. Manila: ADB. Available at: https://www.adb.org/publications/regulatorytechnology-ecosystems-asia-financial-stability

2. Alliance for Financial Inclusion (2020) Guideline Note on Sex-Disaggregated Data Report Templates. https://www.afi-global.org/publications/guideline-note-on-sex-disaggregated-data-report-templates/

3. Bauguess, S (2017) The role of Big Data, Machine Learning, and AI in assessing risks: A regulatory perspective, keynote address. SEC Keynote Address: OpRisk North America 2017, Available at: SSRN https://ssrn.com/abstract=3226514

4. Basel Committee on Banking Supervision (2017) Sound practices: Implications of fintech developments for banks and bank supervisors. Basel: BIS. Available at: https://www.bis.org/bcbs/publ/d415.pdf

5. Bank for International Settlements (2014) Integrated management of micro-databases - Deepening business intelligence within central banks' statistical systems. Basel: BIS. Available at: https://www.bis.org/ifc/publ/ifcb37.pdf

6. Bank for International Settlements (2015a) Data-sharing – issues and good practices. Basel: BIS. Available at: https://www.bis.org/ifc/events/7ifc-tf-report-datasharing.pdf

7. Bank for International Settlements (2015b) Central banks' use of and interest in "big data". Irving Fisher Committee on Central Bank Statistics (IFC): BIS. Available at: https://www.bis.org/ifc/publ/ifc-report-bigdata.pdf

8. Bank for International Settlements (2016) The sharing of micro data – a central bank perspective. Available at: https://www.bis.org/ifc/publ/ifc-report-microdata.pdf

9. Bank for International Settlements (2018) Innovative technology in financial supervision (suptech) - the experience of early users. FSI Insights 9. Basel: BIS. Available at: https://www.bis.org/fsi/publ/insights9.htm

10. Bank for International Settlements (2019) The SupTech Generations. FSI Insights 19. Basel: BIS. Available at: https://www.bis.org/fsi/publ/insights19.pdf

11. Bank for International Settlements (2021) Suptech tools for prudential supervision and their use during the pandemic. FSI Insights 37. Basel: BIS. Available at: https://www.bis.org/fsi/publ/insights37.pdf

12. Bank of England (2015) Big data and central banks. Available at: https://journals.sagepub.com/doi/pdf/10.1177/2053951715579469

13. Bank of England (2015a) Advanced analytics at the Bank of England. Available at: https://www.riksbank.se/globalassets/media/forskning/seminarier/2015/building-data-strategies-for-central-banks-in-light-of-the-data-revolution/advanced-analytics-at-the-bank-of-england.pdf

14. Bank of England (2019) Future of finance review on the outlook for the UK financial system: what it means for the bank of England. Available at: https://www.bankofengland.co.uk/-/media/boe/files/report/2019/future-of-finance-report.pdf?la=en&hash=59CEFAEF01C71AA551E7182262E933A699E952FC

15. Bank of England (2020) Green Notice 2020/02. Available at: https://www.bankofengland.co.uk/statistics/notice/2020/green-notice-2020-02

16. Bank of England (2021a) Transforming data collection from the UK financial sector: a plan for 2021 and beyond. Available at: https://www.bankofengland.co.uk/paper/2021/transforming-

data-collection-from-the-uk-financial-sector-a-plan-for-2021-and-beyond

17. Bank of England (2021b) Transforming data collection - Phase one recommendations with Bank and FCA response. Available at: https://www.bankofengland.co.uk/news/2022/july/transforming-data-collection-phase-one-recommendations-with-bank-and-fca-response

18. Bank of England (2022a) Transforming data collection communication to firms. Available at: https://www.bankofengland.co.uk/news/2022/november/transforming-data-collection-communication-to-firms-28-november

19. Bank of England (2022b) DP4/22 – The Prudential Regulation Authority's future approach to policy. Available at: https://www.bankofengland.co.uk/-/media/boe/files/prudential-regulation/discussion-paper/2022/dp422.pdf?la=en&hash=5F3F2D67F893F3BFAF266F05CFD0BEB736D49F3F

20. Boston Consulting Group and Eastern Foundry (2017) Why Startups Don't Bid on Government Contracts. Available at: https://www.bcg.com/en-us/publications/2017/public-sector-agency-transformation-why-startupsdont-bid-government-contracts.aspx

21. Cambridge Centre for Alternative Finance and World Bank (2020) The Global Covid-19 FinTech Regulatory Rapid Assessment Study. Cambridge: CCAF. Available at: https://www.jbs.cam.ac.uk/faculty-research/centres/alternative-finance/publications/2020-global-covid-19-fintech-regulatory-rapid-assessment-study/

22. di Castri, S. (2022) Powering the paradigm shift in financial supervision: The Cambridge SupTech Lab Innovation Leadership Programme. Available at: https://lab.ccaf.io/blog/powering-the-paradigm-shift-in-financial-supervision-the-cambridge-suptech-lab-innovation-leadership-programme/

23. di Castri, S. (2023, forthcoming) The digital transformation of financial authorities. Making regulatory and supervisory agencies smarter and more agile to tackle the challenges of the digital age. Cambridge: Cambridge SupTech Lab. Available at: https://www.cambridgesuptechlab.org/

24. Digital Public Goods Alliance (2022) State of the Digital Public Goods Ecosystem 2022 Report. Available at: https://digitalpublicgoods.net/DPG-Ecosystem-2022.pdf

25. Dunleavy, P., and L. Carrera (2013) Growing the Productivity of Government Services. Cheltenham, U.K.: Edward Elgar.

26. European Central Bank (2014) ECB Statistic Paper Series – Social Media Sentiment and Consumer Confidence. Available at: https://www.ecb.europa.eu/pub/pdf/scpsps/ecbsp5.en.pdf

27. European Central Bank (2015) Data as a core central banking asset – the strategy of the ECB. Available at: https://www.riksbank.se/globalassets/media/forskning/seminarier/2015/building-data-strategies-for-central-banks-in-light-of-the-data-revolution/data-as-a-core-central-banking-asset--the-strategy-of-the-european-central-bank.pdf

28. European Central Bank (2021) ECB Annual Report on supervisory activities. Available at: https://www.bankingsupervision.europa.eu/press/publications/annual-report/pdf/ssm.ar2021~52a7d32451.en.pdf

29. European Central Bank (2022) Building an innovative supervisory culture. Available at: https://www.bankingsupervision.europa.eu/press/publications/newsletter/2022/html/ssm.nl220216_3.en.html

30. Financial Conduct Authority (2020) Digital Regulatory Reporting: Phase 2 Viability Assessment. Available at: https://www.fca.org.uk/publication/discussion/digital-regulatory-reporting-pilot-phase-2-viability-assessment.pdf

31. Financial Stability Board (2015) Emerging Opportunities and Challenges with Central Bank Data. Available at: https://www.riksbank.se/globalassets/media/forskning/seminarier/2015/building-data-strategies-for-central-banks-in-light-of-the-data-revolution/emerging-opportunities-and-challenges-with-central-bank-data.pdf

32. Financial Stability Board (2020) The Use of Supervisory and Regulatory Technology by Authorities and Regulated Institutions: Market developments and financial stability implications. Available at: https://www.fsb.org/2020/10/the-use-of-supervisory-and-regulatory-technology-by-authorities-and-regulated-institutions-market-developments-and-financial-stability-implications/

33. Grasser, M. (2023, forthcoming) The anatomy of suptech solutions: Dissecting applications to inform global ecosystem acceleration. Cambridge: Cambridge SupTech Lab. Available at: https://www.cambridgesuptechlab.org/

34. Harvard Business School (2020) Modernizing Consumer Financial Regulation for the Digital Age. Available at: https://www.hks.harvard.edu/sites/default/files/centers/mrcbg/files/AWP_152_final.pdf

35. Linköping University (2015) Learning from Big Data. Available at: https://www.riksbank.se/globalassets/media/forskning/seminarier/2015/building-data-strategies-for-central-banks-in-light-of-the-data-revolution/learning-from-big-data.pdf

36. Manon, R. (2017) Financial regulation - the forward agenda. Available at https://www.bis.org/review/r170328a.htm

37. Organisation for Economic Co-operation and Development (2017) Embracing Innovation in Government. Global Trends. OECD Publishing, Paris. Available at: https://www.oecd.org/gov/innovative-government/embracing-innovation-in-government.pdf

38. Organisation for Economic Co-operation and Development (2019) Government at a Glance 2019. OECD Publishing, Paris. Available at: https://dx.doi.org/10.1787/8ccf5c38-en

39. RegTech for Regulators Accelerator (2018a) An API-based Prudential Reporting System for the Bangko Sentral ng Pilipinas (BSP). Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3596276

40. RegTech for Regulators Accelerator (2018b) A Chatbot Application and Complaints Management System for the Bangko Sentral ng Pilipinas (BSP). Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3596268

41. RegTech for Regulators Accelerator (2018c) An AML SupTech Solution for the Mexican National Banking and Securities Commission (CNBV). Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3592564

42. RegTech for Regulators Accelerator (2018d) The RegTech for Regulators Accelerator (R²A) Process: Giving Financial Authorities Superpowers. Available at SSRN: https://ssrn.com/abstract=3250848

43. RegTech for Regulators Accelerator (2020) The 'DataStack': A Data and Tech Blueprint for Financial Supervision, Innovation, and the Data Commons. Available at: https://papers.ssrn.com/abstract=3595344

44. Riksbank (2015) Big Data: Building data strategies for central banks in light of the data revolution. Available at: https://www.riksbank.se/en-gb/press-and-published/conferences/2015/big-data-building-data-strategies-for-central-banks-in-light-of-the-data-revolution/

45. Tatarinov K., Ambos T. C. and F. T. Tschang (2022) Scaling Digital Solutions for Wiked Problems: Ecosystem Versability. Available at: https://link.springer.com/article/10.1057/s41267-022-00526-6

46. Toronto Centre (2020) Cloud Computing: Issues for Supervisors. Available at: https://res.torontocentre.org/guidedocs/Cloud%20Computing%20FINAL.pdf

47. World Bank (2016) Benchmarking Public Procurement 2016. Washington, D.C: World Bank. Available at: https://openknowledge.worldbank.org/bitstream/handle/10986/22649/9781464807268.pdf

48. World Bank (2017) Benchmarking Public Procurement 2017. Washington, D.C: World Bank. Available at: https://documents1.worldbank.org/curated/en/121001523554026106/Benchmarking-Public-Procurement-2017-Assessing-Public-Procurement-Regulatory-Systems-in-180-Economies.pdf

49. World Bank (2020) A Roadmap to SupTech Solutions for Low Income (IDA) Countries. Washington, D.C: World Bank. Available at: https://openknowledge.worldbank.org/handle/10986/34662

50. World Bank (2021) The Next Wave of SupTech Innovation: SupTech Solutions for Market Conduct Supervision. Washington, D.C: World Bank. Available at: https://openknowledge.worldbank.org/entities/publication/79d2a5dc-c88b-58ee-834e-a175f7b165b6

51. World Bank and CCAF (2022) The 3rd Global Fintech Regulator Survey, World Bank Group and the Cambridge Centre for Alternative Finance (CCAF). Available at: https://www.jbs.cam.ac.uk/faculty-research/centres/alternative-finance/publications/3rd-global-fintech-regulator-survey/

52. XBLR (2009a) XBRL Implementation for Capital Markets in Chile. Available at http://archive.xbrl.org/19th/sites/19thconference.xbrl.org/files/SVS-Chile.pdf

53. XBLR (2009b) BOJ's XBRL Project. Towards the next stage. Available at http://archive.xbrl.org/19th/sites/19thconference.xbrl.org/files/BOJXBRL200906inParisFNL2.ppt

54. XBRL (2009c) XBRL-based Basel II Reporting System: Experience of Reserve Bank of India. Available at: http://archive.xbrl.org/19th/sites/19thconference.xbrl.org/files/19th XBRL International Conference - Presentation by RBI.ppt

55. XBRL (2010) Financial Supervision Reporting and XBRL: Benefits of adopting XBRL for financial supervision reporting. Available at: http://www.xbrl.org/bpboarddocs/financialsupervisionreporting.pdf

52. XBLR (2022) OSCA switched off as Bank of England transitions to BEEDS. Available at: https://www.xbrl.org/news/osca-switched-off-as-bank-of-england-transitions-to-beeds/

Appendix 1: List of respondents

Table A1. Financial authority respondents to the primary survey

	Financial authority	Jurisdiction	Income level	Type of financial authority
1	Albanian Financial Supervisory Authority	Albania	Upper-middle income	Capital market, securities and investment instruments (CMSII)
2	Banque d'Algérie	Algeria	Lower-middle income	Central bank
3	Banco Nacional de Angola	Angola	Lower-middle income	Central bank
4	Anguilla Financial Services Commission	Anguilla	High income	CMSII
5	Banco Central de la República Argentina	Argentina	Upper-middle income	Central Bank
6	Australian Securities and Investments Commission	Australia	High income	CMSII
7	Central Bank of the Republic of Azerbaijan	Azerbaijan	Upper-middle income	Central bank
8	Central Bank of The Bahamas	Bahamas	High income	Central bank
9	Bangladesh Bank	Bangladesh	Lower-middle income	Central bank
10	Central Bank of Barbados	Barbados	High income	Central bank
11	National Bank of Belgium	Belgium	High income	Central bank
12	Central Bank of Belize	Belize	Lower-middle income	Central bank
13	Bermuda Monetary Authority	Bermuda	High income	Central bank
14	Banco Central do Brasil	Brazil	Upper-middle income	Central bank
15	British Virgin Islands Financial Services Commission	British Virgin Islands	High income	CMSII
16	Bulgarian National Bank	Bulgaria	Upper-middle income	Central bank
17	National Bank of Cambodia	Cambodia	Lower-middle income	Central bank
18	British Columbia Securities Commission	Canada	High income	CMSII
19	Cayman Islands Monetary Authority	Cayman Islands	High income	CMSII
20	Banco Central de Chile	Chile	High income	Central bank

	Financial authority	Jurisdiction	Income level	Type of financial authority
21	People's Bank of China	China	Upper-middle income	Central bank
22	Superintendencia Financiera de Colombia	Colombia	Upper-middle income	CMSII
23	Banco de la República	Colombia	Upper-middle income	Central Bank
24	Superintendencia General de Entidades Financieras	Costa Rica	Upper-middle income	CMSII
25	Superintendencia General de Seguros de Costa Rica	Costa Rica	Upper-middle income	Other
26	Superintendencia General de Valores	Costa Rica	Upper-middle income	CMSII
27	Croatian National Bank	Croatia	High income	Central Bank
28	Central Bank of Cyprus	Cyprus	High income	Central Bank
29	Cyprus Securities and Exchange Commission	Cyprus	High income	CMSII
30	Banque Centrale du Congo	Democratic Republic of Congo	Low income	Central Bank
31	Central Bank of Djibouti	Djibouti	Lower-middle income	Central Bank
32	Banco Central de la República Dominicana	Dominican Republic	Upper-middle income	Central Bank
33	Banco Central del Ecuador	Ecuador	Upper-middle income	Central Bank
34	Central Bank of Eswatini	Eswatini	Lower-middle income	Central Bank
35	Reserve Bank of Fiji	Fiji	Upper-middle income	Central Bank
36	National Pensions Regulatory Authority	Ghana	Lower-middle income	Other
37	National Insurance Commission	Ghana	Lower-middle income	Other
38	Bank of Ghana	Ghana	Lower-middle income	Central Bank
39	Gibraltar Financial Services Commission	Gibraltar	High income	CMSII
40	Bank of Greece	Greece	High income	Central Bank
41	Banque Centrale de la République de Guinée	Guinée	Low income	Central Bank
42	Comision Nacional de Bancos y Seguros	Honduras	Lower-middle income	Other

	Financial authority	Jurisdiction	Income level	Type of financial authority
43	Magyar Nemzeti Bank	Hungary	High income	Central Bank
44	Securities and Exchange Board of India	India	Lower-middle income	CMSII
45	Otoritas Jasa Keuangan	Indonesia	Lower-middle income	CMSII
46	Bank Indonesia	Indonesia	Lower-middle income	Central Bank
47	Israel Securities Authority	Israel	High income	CMSII
48	Banca d'Italia	Italy	High income	Central Bank
49	Jersey Financial Services Commission	Jersey, Channel Islands	High income	CMSII
50	Central Bank of Jordan	Jordan	Upper-middle income	Central Bank
51	Astana Financial Services Authority	Kazakhstan	Upper-middle income	CMSII
52	Capital Markets Authority	Kenya	Lower-middle income	CMSII
53	Central Bank of Kenya	Kenya	Lower-middle income	Central Bank
54	The Central Bank of Kuwait	Kuwait	High income	Central Bank
55	Banque du Liban	Lebanon	Upper-middle income	Central Bank
56	Central Bank of Liberia	Liberia	Low income	Central Bank
57	Financial Market Authority Liechtenstein	Liechtenstein	High income	Central Bank
58	Lietuvos Bankas	Lithuania	High income	Central Bank
59	The Monetary Authority of Macao	Macao, China	High income	Central Bank
60	Reserve Bank of Malawi	Malawi	Low income	Central Bank
61	Malawi Communication Regulatory Authority	Malawi	Low income	Other
62	Securities Commission Malaysia	Malaysia	Upper-middle income	CMSII
63	Malta Financial Services Authority	Malta	High income	Central Bank
64	Office of the Banking Commission	Marshall Islands	Upper-middle income	Other
65	Bank of Mauritius	Mauritius	Upper-middle income	Central Bank
66	Financial Services Commission	Mauritius	Upper-middle income	CMSII

	Financial authority	Jurisdiction	Income level	Type of financial authority
67	Comision Nacional de Bancos y Valores	Mexico	Upper-middle income	Other
68	Banco de México	México	Upper-middle income	Central Bank
69	National Bank of Moldova	Moldova	Upper-middle income	Central Bank
70	Bank of Mongolia	Mongolia	Lower-middle income	Central Bank
71	Central bank of Montenegro	Montenegro	Upper-middle income	Central Bank
72	Securities and Exchange Commission	Myanmar	Lower-middle income	CMSII
73	Bank of Namibia	Namibia	Upper-middle income	Central Bank
74	Namibia Financial Institutions Supervisory Authority	Namibia	Upper-middle income	CMSII
75	Nepal Rastra Bank	Nepal	Lower-middle income	Central Bank
76	Securities Board of Nepal	Nepal	Lower-middle income	CMSII
77	Reserve Bank of New Zealand	New Zealand	High income	Central Bank
78	Securities and Exchange Commission	Nigeria	Lower-middle income	CMSII
79	National Bank of the Republic of North Macedonia	North Macedonia	Upper-middle income	Central Bank
80	State Bank of Pakistan	Pakistan	Lower-middle income	Central Bank
81	Palestine capital market authority	Palestine	Lower-middle income	CMSII
82	Banco Central del Paraguay	Paraguay	Upper-middle income	Central Bank
83	Comision Nacional de Valores	Paraguay	Upper-middle income	CMSII
84	Superintendencia de Banca, Seguros y AFP del Perú	Peru	Upper-middle income	Other
85	Superintendencia del Mercado de Valores	Peru	Upper-middle income	CMSII
86	Bangko Sentral ng Pilipinas	Philippines	Lower-middle income	Central Bank
87	Securities and Exchange Commission	Philippines	Lower-middle income	CMSII

	Financial authority	Jurisdiction	Income level	Type of financial authority
88	Narodowy Bank Polski	Poland	High income	Central Bank
89	Portuguese Securities Market Commission	Portugal	High income	CMSII
90	Banca Na□ională a României	Romania	Upper-middle income	Central Bank
91	National Bank of Rwanda	Rwanda	Low income	Central Bank
92	Central Bank of Samoa	Samoa	Lower-middle income	Central Bank
93	Banca Centrale della Repubblica di San Marino	San Marino	High income	Central Bank
94	Banco Central de São Tome e Principe	São Tome e Principe	Lower-middle income	Central Bank
95	National Bank of Serbia	Serbia	Upper-middle income	Central Bank
96	Securities Commission	Serbia	Upper-middle income	CMSII
97	Central Bank of Seychelles	Seychelles	High income	Central Bank
98	Bank of Sierra Leone	Sierra Leone	Low income	Central Bank
99	Národná Banka Slovenska	Slovakia	High income	Central Bank
100	Central Bank of Solomon Islands	Solomon Islands	Lower-middle income	Central Bank
101	Central Bank of Somalia	Somalia	Low income	Central Bank
102	South African Reserve Bank	South Africa	Upper-middle income	Central Bank
103	Financial Sector Conduct Authority	South Africa	Upper-middle income	Other
104	Banco de España	Spain	High income	Central Bank
105	Insurance Regulatory Commission of Sri Lanka	Sri Lanka	Lower-middle income	Other
106	Central Bank of Sudan	Sudan	Low income	Central Bank
107	Centrale Bank van Suriname	Suriname	Upper-middle income	Central Bank
108	Central Bank of Syria	Syria	Low income	Central Bank
109	Financial Supervisory Commission	Taiwan, China	High income	CMSII
110	Central Bank of the Republic of China of Taiwan	Taiwan, China	High income	Central Bank
111	Bank of Tanzania	Tanzania	Lower-middle income	Central Bank
112	The Office of Insurance Commission	Thailand	Upper-middle income	Other

	Financial authority	Jurisdiction	Income level	Type of financial authority
113	Bank of Thailand	Thailand	Upper-middle income	Central Bank
114	Central Bank of The Gambia	The Gambia	Low income	Central Bank
115	Banco Central de Timor Leste	Timor Leste	Lower-middle income	Central Bank
116	Trinidad and Tobago Securities and Exchange Commission	Trinidad and Tobago	High income	CMSII
117	Central Bank of Trinidad and Tobago	Trinidad and Tobago	High income	Central Bank
118	Conseil du Marche Financier	Tunisia	Lower-middle income	CMSII
119	Central Bank of Tunisia	Tunisia	Lower-middle income	Central Bank
120	Bank of Uganda	Uganda	Low income	Central Bank
121	National Information Technology Authority Uganda	Uganda	Low income	Other
122	Abu Dhabi Global Market	United Arab Emirates	High income	Other
123	Central Bank of The United Arab Emirates	United Arab Emirates	High income	Central Bank
124	Dubai Financial Services Authority	United Arab Emirates	High income	Other
125	Financial Conduct Authority	United Kingdom	High income	Other
126	Bank of England	United Kingdom	High income	Central Bank
127	Banco Central del Uruguay	Uruguay	High income	Central Bank
128	Consumer Financial Protection Bureau	USA	High income	Other
129	Central Bank of Uzbekistan	Uzbekistan	Lower-middle income	Central Bank
130	Reserve Bank of Vanuatu	Vanuatu	Lower-middle income	Central Bank
131	Securities and Exchange Commission	Zambia	Lower-middle income	CMSII
132	Competition and Consumer Protection Commission	Zambia	Lower-middle income	Other
133	Bank of Zambia	Zambia	Lower-middle income	Central Bank
134	Securities and Exchange Commission of Zimbabwe	Zimbabwe	Lower-middle income	CMSII

Table A2. Additional financial authorities that contributed to the supervisory data questionnaire

	Financial authority	Jurisdiction	Income Level	Type of financial authority
135	Ministry of Finance	Azerbaijan, Republic of	Upper-middle income	Other
136	British Colombia Securities Commission	Canada	High income	CMSII
137	Financial Transactions and Reports Analysis Centre of Canada	Canada	High income	Other
138	Banque des États de l'Afrique Centrale	Cameroon	Lower-middle income	Central Bank
		Central African Republic	Low income	Central Bank
		Chad	Low income	Central Bank
		Equatorial Guinea	Upper-middle income	Central Bank
		Gabon	Upper-middle income	Central Bank
		Republic of the Congo	Lower-middle income	Central Bank
139	Comisión del Mercado Financiero	Chile	High income	Other
140	National Bank of Ethiopia	Ethiopia	Low income	Central Bank
141	Reserve Bank of India	India	Lower-middle income	Central Bank
142	Bank Negara Malaysia	Malaysia	Upper-middle income	Central Bank
143	Banco de Moçambique	Mozambique	Low income	Central Bank
144	Securities and Exchange Commission of Thailand	Thailand	Upper-middle income	CMSII
145	National Bank of Ukraine	Ukraine	Lower-middle income	Central Bank
146	Palestine Monetary Authority	West Bank and Gaza	Lower-middle income	CMSII

Appendix 2: SupTech Taxonomy

Table A3. Supervisory areas and use cases (Cambridge SupTech Lab 2022)

Supervisory area	Supervisory use case
Anti money laundering, counter financing of terrorism and proliferation financing supervision	1. Assisted/automated examination
	2. KYC assessment
	3. Suspicious activity detection
	4. Misconduct analysis
	5. Metadata intelligence
	6. Advanced text analysis
	7. Derisking analysis
	8. Onsite examination
Capital market, securities and investment instruments	9. Market manipulation detection
	10. Insider trading detection
	11. Improved insights
	12. Poor disclosure detection
	13. Onsite inspection
	14. Risk-based prioritisation
Competition monitoring	15. Competition monitoring
	16. Fees and rates monitoring
Compliance assistance	17. Automated guidance
	18. Automated compliance auditing
Climate/ESG risks supervision	19. Risk classification
	20. Green market monitoring
	21. Scenario analysis
	22. Portfolio analysis

Supervisory area	Supervisory use case
Consumer protection and market conduct supervision	23. Consumer fraud detection
	24. Complaints handling
	25. Complaints analysis
	26. Complaints monitoring
	27. Algorithmic auditing
	28. Sentiment analysis
	29. Templates validation
	30. Interdepartmental analysis
	31. Cross-entity analysis
	32. Early warning systems
	33. Onsite examination
	34. Peer-group/risk classification
	35. Misconduct analysis
	36. Predatory pricing detection
	37. Poor disclosure detection
	38. Credit bureau rectification
	39. Alternative dispute resolution
Cyber risk supervision	40. Cybersecurity assessment
	41. Audit trail examination
	42. Compliance monitoring
	43. Onsite inspection
Digital assets/cryptocurrencies oversight	44. Data handling
	45. Automated compliance auditing
	46. Automated data validation
	47. On-chain analysis
	48. Cross-jurisdictional analysis
Financial inclusion policymaking	49. Gender-based analysis
	50. Simplified KYC assessment
	51. Geospatial analysis
	52. Advanced/real-time monitoring
	53. Consumer education
	54. Consumer satisfaction analysis

Supervisory area	Supervisory use case
Insurance supervision	55. Data handling
	56. Automated compliance auditing
	57. ORSAs reporting and analytics
	58. Automated data validation
	59. Stress testing
	60. Registration of intermediaries
	61. Fit & proper assessment
	62. Product registration
	63. Risk assessment
	64. Onsite examination
Licensing	65. Automated guidance
	66. Automated processing
Payments oversight	67. Advanced/real-time monitoring
	68. Network performance monitoring
	69. RTGS stress testing
Prudential supervision of banks and non-bank deposit-taking institutions	70. Data handling
	71. Automated data validation
	72. Interdepartmental analysis
	73. Cross-entity analysis
	74. Early warning systems
	75. Sectorial credit monitoring
	76. Cross-entity rating monitoring
	77. Threshold monitoring
	78. Investment patterns analysis
	79. Automated credit examination
	80. Microprudential supervision
	81. Risk-based prioritisation
	82. Automated report generation
	83. Onsite examination
	84. Stress testing
	85. Scenario analysis
	86. Peer-group/risk classification
	87. Fit & proper assessment

Table A4. Technologies & data science tools for supervision (Cambridge SupTech Lab 2022)

Layer of supervisory stack	Technology	Suptech generation
Data collection	Web portals or other document management	2G
	Application programming interfaces (APIs)	3G
	Advanced collection (e.g., scraping, streaming, AI-based)	4G
Data processing	Automated validation	2G
	Task automation	3G
	Advanced processing (e.g., machine learning)	4G
Data storage	On-premises databases	2G
	Cloud and hybrid computing systems	3G
	Big data tools	4G
Data analytics	Descriptive/Diagnostic analytics tools	2G
	Predictive Analytics	3G
	Prescriptive Analytics	4G
Data products	Static Charts and Metrics	2G
	Interactive visualisations	3G
	Advanced business intelligence tools (e.g., AI-driven)	4G

Appendix 3: Definitions

A

APPLICATION PROGRAMMING INTERFACE (API)	APIs allow software programs to interact by exchanging data which can prompt certain actions, such as making a transaction. This includes payment APIs, data APIs, 'ecosystem expansion' APIs and 'consent and identity' APIs. (World Bank 2020b)
ARTIFICIAL INTELLIGENCE (AI)	Defined as IT systems that perform functions requiring human capabilities. AI can ask questions, discover and test hypotheses, and make decisions automatically based on advanced analytics operating on extensive data sets. Machine learning (ML) is one subcategory of AI. (World Bank 2020a)

B

BIG DATA	High-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation. (Gartner 2022)
BUSINESS INTELLIGENCE (BI)	Software and services to transform data into actionable insights that inform an organisation's strategic and tactical business decisions. BI tools access and analyse data sets and present analytical findings in reports, summaries, dashboards, graphs, charts and maps to provide users with detailed intelligence about the state of the business. The term business intelligence often also refers to a range of tools that provide quick, easy-to-digest access to insights about an organisation's current state, based on available data. (CIO 2019)

C

CHATBOT	A computer program that simulates and processes human conversation (either written or spoken), allowing humans to interact with digital devices as if they were communicating with a real person. (Oracle 2022)
CLOUD COMPUTING	An innovation in computing that allows for the use of an online network ('cloud') to host processors, leading to an increase in the scale and flexibility of computing capacity. (FSB 2020)

COMPUTER VISION	A field of artificial intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs – and take actions or make recommendations based on that information. Subcategories include image segmentation (where items are in an image) and image classification (what the items are). (IBM 2022, R²A 2018)
CONSUMER PROTECTION	The framework of laws, regulations, and institutional arrangements that safeguard consumers by ensuring fair and responsible treatment for them in the financial marketplace. (World Bank 2022)
CYBERSECURITY	Preservation of confidentiality, integrity and availability of information and/or information systems through the cyber medium. In addition, other properties, such as authenticity, accountability, non-repudiation and reliability can also be involved. (FSB 2018)

D

DATA LAKES	A data lake is a centralised repository designed to store, process, and secure large amounts of structured, semi structured, and unstructured data. It can store data in its native format and process any variety of it, ignoring size limits. (Google 2022)
DATA PROCESSING	The collective set of data actions (the complete data life cycle, including, but not limited to, collection, retention, logging, generation, transformation, use, disclosure, sharing, transmission and disposal). (NIST 2020)
DESCRIPTIVE ANALYTICS TOOLS	Interactive applications that are used to search and summarise historical data to identify patterns or meaning, including dashboards, data visualisation tools, and automated statistical summaries. (TechTarget 2022)
DATA VALIDATION	An activity aimed at verifying whether the value of a data item comes from the given (finite or infinite) set of acceptable values. For example, this ensures a postal code is valid or that a numeric value does not include letters or symbols. These rules can be enforced in either a manual or automatic fashion. (OECD 2013)

DATA VISUALISATION	The graphical representation of data for understanding and communication. This typically takes the form of exploratory (trying to explore and understand patterns and trends within your data) or explanatory (surfacing something in your data you would like to communicate to your audience) forms. (Johns Hopkins 2022)
DATA WAREHOUSE	A data management system designed to enable and support business intelligence (BI) activities, especially analytics. Data warehouses are solely intended to perform queries and analysis and often contain large amounts of historical data. The data within a data warehouse is usually derived from various sources, such as application log files and transaction applications. (Oracle 2022)
DESCRIPTIVE ANALYTICS TOOLS	Interactive applications used to search and summarise historical data to identify patterns or meaning, including dashboards, data visualisation tools and automated statistical summaries. (TechTarget 2022)
DIGITAL ASSETS	Digital instruments issued or represented by using distributed ledger or similar technology. This does not include digital representations of fiat currencies, such as e-money. (FSB 2022)
DISTRIBUTED LEDGER TECHNOLOGY (DLT)	A technology – such as blockchain – that records information through a distributed ledger (a repeated digital copy of data at multiple locations). These technologies enable nodes in a network to securely propose, validate and record state changes (or updates) to a synchronised ledger distributed across the network's nodes. (World Bank 2020a)

F

FINANCIAL INCLUSION	The uptake and usage of a range of appropriate financial products and services by individuals and MSMEs (micro, small, and medium enterprises), provided in a manner that is accessible and safe to the consumer and sustainable to the provider. (World Bank 2020b)

FINANCIAL STABILITY	A stable financial system is capable of efficiently allocating resources, assessing and managing financial risks, maintaining employment levels close to the economy's natural rate, and eliminating relative price movements of real or financial assets that will affect monetary stability or employment levels. Financial stability is paramount for economic growth, as most transactions in the real economy are made through the financial system. (World Bank 2016).
FINTECH	An acronym for 'financial technology', it refers to the advances in technology that have the potential to transform the provision of financial services spurring the development of new business models, applications, processes, and products. (World Bank 2020a)

G

GEOGRAPHIC INFORMATION SYSTEMS (GIS)	A computerised system for capturing, storing, checking, and displaying data related to positions on Earth's surface, and enabling analysis and visualisation based on spatial relationships between these data. (NatGeo 2022)

I

IMAGE PROCESSING	The general process of digitising and formatting visual information (e.g., photographs, video) such that useful information can be automatically extracted via technologies such as optical character recognition (OCR), facial recognition, and other computer vision techniques. (R²A 2018)

M

MACRO PRUDENTIAL SUPERVISION	Supervision that considers the interactions among individual financial institutions, as well as the feedback loops of the financial sector with the real economy, including the costs that systemic risk entails in terms of output losses. (ECB 2014)

MARKET INTEGRITY	Concerned with the capacity to pursue the 'dirty money' that flows through the global financial system, imposing a significant cost on national security, economic opportunity, and the rule of law. It is also connected with regulators' ability to uncover, prosecute, and prevent such movements in the future, as well as to restore official funds stolen in corruption to public coffers. (World Bank 2022)
MICROPRUDENTIAL SUPERVISION	Supervision that focuses on safeguarding individual financial institutions from idiosyncratic risks and preventing them from taking too much risk. (ECB 2014)

N

NATURAL LANGUAGE PROCESSING (NLP)	An interdisciplinary field of computer science, artificial intelligence, and computation linguistics that focuses on programming computers and algorithms to parse, process, and understand human language. NLP can be regarded as a form of AI. (FSB 2020)
NETWORK ANALYSIS	The use of quantitative and qualitative data to model and draw insights regarding the formal and less-formal interconnections between a set of related entities, for example, a measure of the degree to which a financial system will be weakened by the cascading transmission of financial distress across institutions (IMF 2010)

O

OPTICAL CHARACTER RECOGNITION (OCR)	A specific form of computer vision that focuses on the transcription of image data into textual data. Examples include licence plate readers, OCR-enabled scanners and mobile apps, passport and other identification card readers, and file conversion tools. (R²A 2018)

P

PREDICTIVE ANALYTICS TOOLS	The advanced analysis of historical data to create statistical models to predict future events, values, facts or characteristics. This process may include recommendation engines (tools where the prediction is an optimal value or action), and employ machine learning (computerised, iterative optimisation of the aforementioned statistical models). (TechTarget 2022)

R

REGTECH

An acronym for 'regulatory technology'. It involves new technologies to help regulated financial service providers streamline audit, compliance and risk management and other back-office functions to enhance productivity and overcome regulatory challenges, such as the risks and costs related to regulatory reporting and compliance obligations. This can also refer to firms that offer such applications. (World Bank 2020a)

ROBOTIC PROCESS AUTOMATION (RPA)

The automation of the basic tasks defined by a user; these tasks can include, e.g., filling forms and checking forms for completeness. (TechTarget 2022)

S

SUPTECH

An acronym for 'supervisory technology'. It is the application of technology and data analysis solutions to complement and enhance a financial authority's financial market oversight capabilities. Suptech applications are used by financial authorities to access more granular, diverse, timely and trustworthy data to improve operational efficiency and generate previously unattainable insights, thus improving decision-making. (Cambridge SupTech Lab 2022)

SENTIMENT ANALYSIS

A specific form of NLP that focuses on inferring the emotional content expressed in a given corpus of text or transcribed speech. Examples include social media data mining to understand public sentiment surrounding a given topic or entity and analysing customer service requests/complaints to inform escalation. (R²A 2017)

T

TEXT MINING

The process of discovering interesting and useful patterns and relationships in large volumes of text. This process uses tools from statistics and artificial intelligence. (IBM 2022)

W

WEB SCRAPING

The process of using software to extract data from websites. (Cambridge SupTech Lab 2022)

CAMBRIDGE **SUPTECH LAB**

www.cambridgesuptechlab.org

Cambridge
Centre
for Alternative
Finance

UNIVERSITY OF
CAMBRIDGE
Judge Business School