The use of artificial intelligence for policy purposes

Report submitted to the G20 Finance Ministers and Central Bank Governors

October 2025

This publication is available on the BIS website (www.bis.org).

ISBN 978-92-9259-890-7 (online)

The use of artificial intelligence for policy purposes

BIS report to be submitted ahead of the October 2025 G20 FMCBG meeting

Contents

1. Introduction 2

2. Core concepts of AI 3

3. AI use cases by central banks and supervisors 4

   3.1 Information collection 5

   3.2 Macroeconomic and financial analysis to support monetary policy 6

   3.3 Oversight of payment systems 7

   3.4 Supervision and financial stability 10

4. Challenges and lessons learned 15

5. Conclusions 20

6. References 22

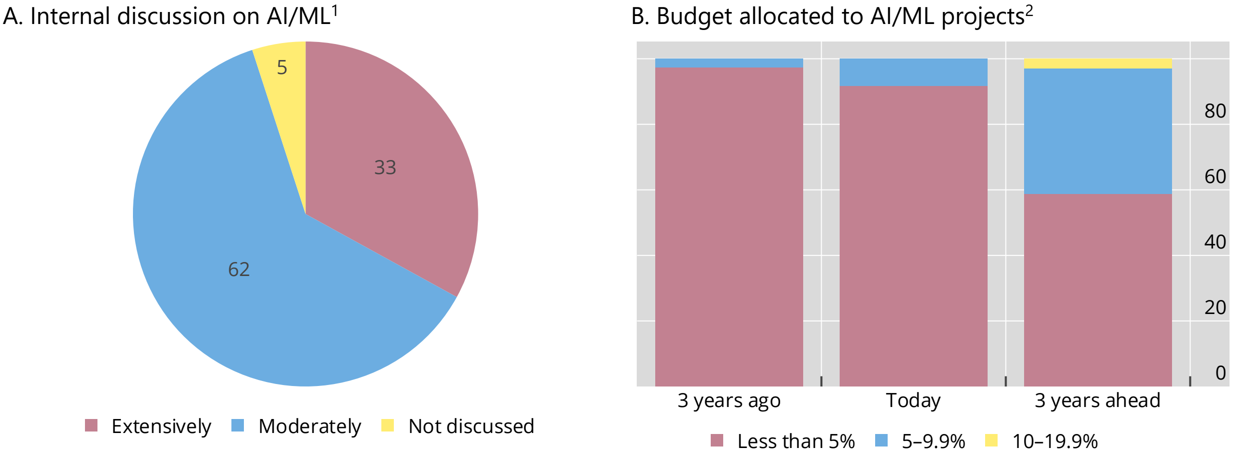
1. Introduction

The rapid adoption of artificial intelligence (AI) – broadly defined as computer systems capable of tasks that normally require human intelligence – is poised to have a profound effect on the financial system and the real economy (BIS (2024); Aldasoro, Doerr, Gambacorta and Rees (2024); Aldasoro, Gambacorta, Korinek, Shreeti and Stein (2024); IMF (2024)). The adoption of generative AI (gen AI), ie tools that engage with text-based content and allow users to converse with computers through ordinary language, is proceeding at a speed that easily outpaces previous waves of technology adoption. ChatGPT alone reached one million users in less than a week and many firms are already integrating gen AI into their daily operations. To do so, they are investing heavily in AI technology to tailor it to their specific needs, and in many cases they have embarked on a hiring spree of workers with AI-related skills. These developments, and the attendant effects on inflation, productivity, consumption, investment and labour markets are of paramount concern to central banks and other supervisory and regulatory authorities (Aldasoro, Doerr, Gambacorta, Gelos and Rees (2024); Cazzaniga et al (2024)).

Central banks were early adopters of machine learning (ML) (a key component of AI techniques), using it to gain insights for statistics, research and policy long before AI became a popular topic (Araujo et al (2024)).[1] Discussions on AI and ML at central banks are pervasive (Graph 1.A) and expected budget allocations bear that interest out (Graph 1.B). Indeed, central banks, financial regulators and supervisory authorities regularly handle vast data sets and complex decision processes in pursuit of safeguarding monetary and financial stability, and the integrity of payment systems. Today, the greater capabilities of new AI methods – such as the large language models (LLMs) underpinning gen AI – open further opportunities, from improved economic analysis to better regulatory oversight, potentially enhancing the effectiveness and efficiency of these institutions and of policymaking more broadly.

AI is an important topic for central banks, with budgets set to grow

As a percentage of respondents | Graph 1



AI = artificial intelligence; ML = machine learning. Responses to the survey were collected between September and November 2024.

[1] Share of central banks discussing AI/ML for internal usage extensively, moderately or not at all. [2] Share of respondents that expect to allocate the respective share of their total budget to AI/ML projects.

Source: IFC (2025).

[1] Machine learning is a term referring to techniques designed to detect patterns in data and use them in prediction or to aid decision-making.

This report examines how central banks and other supervisory institutions are leveraging AI for policy purposes. The report first offers a brief discussion of core AI concepts relevant to public authority use cases, focusing in particular on ML. It then provides examples of how central banks and supervisory authorities are already using big data and ML in four key areas. These are: (i) information collection and the compilation of official statistics; (ii) macroeconomic and financial analysis in support of monetary policy; (iii) oversight of payment systems; and (iv) supervision and financial stability analysis. Recent projects on the use of AI by the Bank for International Settlements' (BIS) Innovation Hub provide examples of experimentation with AI across these areas.

Finally, the report stresses that, despite AI's significant potential to enhance policymaking, the effective use of gen AI requires a number of challenges to be addressed. These range from data governance (eg the use of internal versus external data) to investing in human capital and information technology (IT) infrastructure. A key lesson is that collaboration and the sharing of experiences emerge as important avenues for central banks, in particular by exploiting economies of scale and reducing the demands on IT infrastructure and human capital.

2. Core concepts of artificial intelligence

The field of AI traces its roots back to the mid-20th century, but progress has come in spurts.[2] Early AI research in the 1950s and 1960s generated optimism but was limited by a lack of computing power. As computing power increased, the rise of ML techniques in the 1990s laid the foundations for today's advanced and ubiquitous AI models. Decades of increasing computing power, growing volumes of data and algorithmic improvements have transformed AI from a niche academic pursuit into a powerful general purpose technology.

A defining featpsure is the use of ML, a collection of techniques that allow computers to learn patterns from data rather than follow explicit or hard-coded instructions to make predictions (Araujo et al (2024)). During an ML model's training phase, it is fed historical data and adjusts its internal parameters to fit observed outcomes. The goal is to be able to generalise and predict correctly based on new data. The focus on predicting outcomes "out of sample" (ie based on data that are not part of the training sample) is a key objective of ML, with less emphasis placed on *how* those predictions are reached. ML techniques excel at finding "needles in haystacks" – identifying patterns in vast, complex data sets that would overwhelm traditional analysis (BIS (2024)).

An important subfield of ML techniques is tree-based methods such as random forests, which are heavily used in policy applications due to their flexibility and performance (see next section). Decision trees sequentially partition data into finer categories based on specific characteristics. For example, a tree might first divide individuals into those with a university degree and those without, then further split each group by years of work experience, and finally by industry of employment. The resulting finer categories can then be compared with a target variable, such as income levels or job retention rates, to identify patterns and relationships. Random forests enhance this approach by combining multiple trees trained on different subsets of data, improving predictive accuracy and reducing the risk of overfitting. In the same spirit, forests can be deployed in identifying outliers by means of isolation forests, a method that singles

[2] See Mitchell (2020) for a non-technical introduction to and history of AI. Russell and Norvig (2021) provide a technical discussion of AI methods.

out the data points that can be isolated from others. These methods are not only powerful tools for prediction but also serve as exploratory techniques to uncover patterns, classify data or identify outliers.

Deep learning, which uses neural network models, is another key ML technique. These networks are inspired by how the human brain works, connecting simple units called artificial neurons in layers. The first layer processes the input data and passes its results to the next layer, which continues the process until a final output is produced. The depth of the network refers to the number of layers: deeper networks require more training to fine-tune the parameters of each layer (which determine the strength of the connection between neurons and layers) but predict more accurately. A key advantage of these models is their ability to handle unstructured data such as text or images by mapping them into arrays of numbers (so-called embeddings). This allows the model to compare and manipulate concepts mathematically.

Building on these methods, by the early 2020s AI development accelerated markedly with the emergence of large-scale models and gen AI (and notably LLMs). Gen AI applications can take prompts as input and flexibly produce human-like output, be it text, images, videos etc. This was facilitated by the development in 2017 of the so-called transformer architecture, which massively improved the capabilities of neural networks to handle natural language processing (NLP). The distinguishing feature of transformers is their ability to efficiently capture long-range dependencies in data using self-attention, enabling more accurate processing of sequential information like text or time series to make a probabilistic assessment of the next word or data point.[3] In other words, attention allows each word in a text to be understood in relation to every other word, enhancing the models' capacity to take account of context and relationships within the text. For instance, it can discern the different meanings of "rock" in "Some say the Beatles did not make rock music" versus "The Rock has not won an Oscar, but his movies are successful" by focusing on surrounding words like "Beatles" or "music" versus "Oscar" or "movies". Transformers are the foundation of LLMs and gen AI because they enable these models to understand context and relationships in text over long passages, making them highly effective for tasks like generating, summarising or translating language (or images in the broader generative AI case). Importantly, gen AI applications can be fine-tuned for specific tasks by feeding specialised training data.

Finally, LLMs possess so-called few-shot or zero-shot learning abilities, ie the ability to deliver accurate results with minimal examples. For example, LLMs trained to forecast the next word, by converting time series data into tokens resembling textual sequences, can be directly repurposed for other prediction tasks. That is to say that due to their advanced capabilities for few-shot learning, these models can be readily applied to a wide array of forecasting tasks with limited or no additional training. This stands in contrast to existing forecasting models, for which optimisation often requires significant fine-tuning ex ante.

3. AI use cases by central banks and supervisors

Central banks deploy AI in four main areas: (i) information collection and the compilation of official statistics; (ii) macroeconomic and financial analysis in support of monetary policy; (iii) oversight of payment systems; and (iv) supervision and financial stability analysis.[4] Given the even more prominent role of data in the most recent crop of AI innovations, the focus on these areas comes as no surprise. Indeed, a survey that predates the rise of gen AI indicated that the majority of central banks in both advanced economies

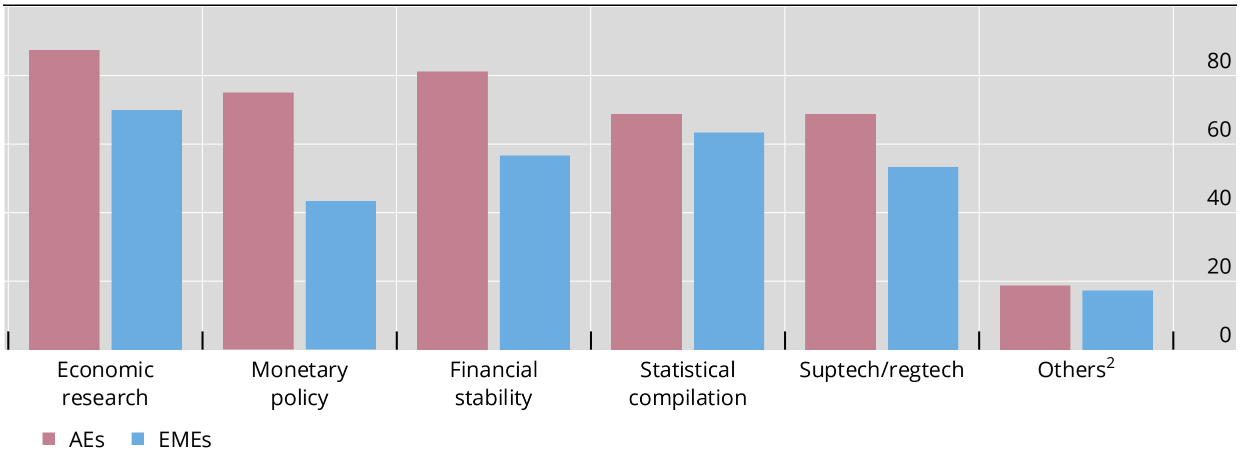
[3] See Vaswani et al (2017) and Kwon et al (2024).

[4] For details on the use cases, as well as additional examples, see Araujo et al (2024).

(AEs) and emerging market economies (EMEs) were already using AI and big data in all of these areas (Graph 2).

Purposes for which central banks use big data[1]

In per cent Graph 2



[1] The graph reports the share of respondents that selected each respective answer to the question, "For what general purposes does your institution use big data?" and respondents could select multiple options. [2] Includes "monitoring cryptoassets", "cyber security" and "network analysis".

Source: Doerr, Gambacorta and Serena Garralda (2021).

3.1 Information collection

Central banks collect data from a large variety of sources both to use internally and make available as a public good. Securing high-quality data for economic analysis and statistical compilation remains a significant challenge. Key obstacles include cleaning data, eg identifying outliers and erroneous entries, as well as ensuring representativeness and linking new observations to established series. Ever-rising data volumes and complexity demand agile and efficient quality control tools.

To provide high-quality micro data, central banks are increasingly using ML techniques. Isolation forests – a variation of random forests – are particularly suitable for the large and granular data sets typical of central banks, owing to their scalability and ability to identify outliers regardless of the shape of the data's distribution.[5] One example is the use of isolation forest and other methods to detect anomalies in central banks' extensive derivatives data sets (Kamenetsky Yadan (2021)).[6] Another example is assessing the efficacy of isolation forests and other algorithms in detecting outliers in large data sets through benchmarking against manually identified outliers. This illustrates the benefits of a two-step approach: initially, a model autonomously identifies potential outliers, which are then reviewed by experts who

[5] Central banks, as providers of data to other public agencies and the public, need to ensure that their data are reliable. Detecting outliers, eg erroneous entries by reporting parties, is paramount to ensuring data can be used for further analysis.

[6] Central banks have also surmounted a primary limitation of isolation forests – namely that they are primarily designed to work with numerical data – by developing new ways to include other data types (eg the currency of a transaction). For example, given that the data underlying the calculation of the euro short-term rate predominantly comprises categorical data (ie data grouped in different categories, rather than assigned numerical values), the European Central Bank first applies an algorithm to convert categorical variables into numerical ones and then makes full use of the data to find outliers (Accornero and Boscariol (2021)).

provide feedback to refine the algorithm. This approach balances the value of domain expertise with the costs of human inputs. Analysing different methods to explain outlier classification can help to overcome the issue of "black box" ML models lacking "explainability" (discussed in more detail in Section 4).

3.2 Macroeconomic and financial analysis to support monetary policy

Timely data are essential for nowcasting. Because no single metric fully captures real-time economic activity, nowcasting models often use large, intricate data sets containing dozens, and sometimes hundreds, of indicators. Inputs span industrial production, purchasing manager surveys and credit card spending, and may extend to web-scraped figures from online retailers or social media platforms. Breakthroughs in high-dimensional econometrics now distil this complexity into a handful of representative factors, markedly sharpening forecasts. The Federal Reserve Bank of New York's nowcasting system illustrates the approach: it draws on a broad array of series, isolates latent drivers and projects key aggregates – notably GDP – that allow for substantially more accurate nowcasting.[7]

Yet access to timely data remains a bottleneck. Structured information on corporate hiring or household spending, for instance, arrives with delays, and projections are only as timely as the underlying releases.

ML techniques help bridge the gap by opening the door to unstructured data. Deep neural networks create embeddings that swiftly transform unstructured inputs into structured form. Through such embeddings, LLMs ingest massive volumes of text or images and turn them into high-frequency series. Raw data can be drawn from news articles, social media posts, web search activity or aerial photos.

One application includes real-time gauges of inflation expectations and narrative summaries of economic conditions. A random forest, for example, first filters price-related social media messages and then classifies each as signalling higher inflation, lower inflation or neutral expectations. The daily balance between these three groups serves as an expectations measure (Denes et al (2021)).

Open source LLMs fine-tuned on financial news offer a related example. They parse anecdotal reports from or interviews with entrepreneurs, economists and market commentators, convert them into sentiment scores and generate a historical series. This index can then feed into GDP nowcasts and recession probability models (Du et al (2023)). Similarly, van Dijk and de Winter (2023) use financial press articles to construct a measure of sentiment that can improve GDP nowcasting.

Another example is the use of neural networks to decompose services inflation, attributing shares to inertia, expectations, the output gap or external price pressures (Buckmann et al (2023)). Because they handle many more inputs than standard econometric tools, neural networks allow researchers and policymakers to work with highly granular information. Their capacity to learn rich non-linear patterns further aids in modelling monetary policy shocks, heterogeneous asset holdings and rapidly shifting inflation dynamics.[8]

The use of granular sources likewise deepens insights into sectoral and regional trends. ML algorithms can mine job posting databases or e-commerce portals to trace wages and hiring across occupations and industries, shedding light on technology-driven displacement, re-employment speeds and wage formation. Satellite readings of pollution or nighttime lights indicate short-term economic activity, while electricity use patterns reveal industrial production across geographies. Together these

[7] See the dedicated webpage of the Federal Reserve Bank of New York.

[8] See Aruoba and Drechsel (2024), Gabaix et al (2025) and BIS Innovation Hub Project Spectrum (discussed in Section 3.5).

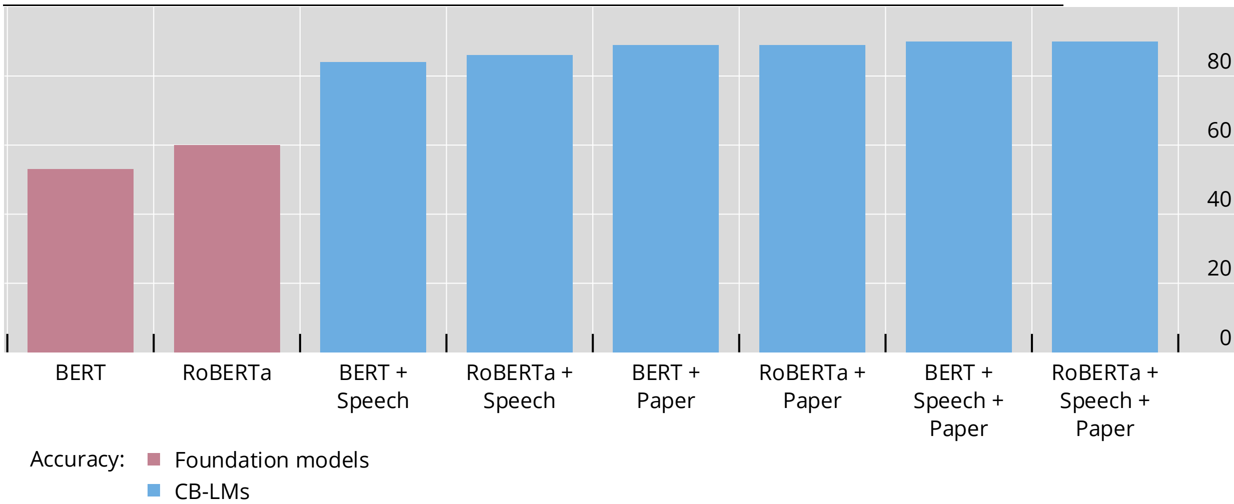
streams of information give central banks a nuanced view of capital spending, output and evolving supply-demand balances.

Tailoring LLMs to central bank vocabulary yields further benefits. The BIS Central Bank Language Model (CB-LM) project (Gambacorta, Kwon, Park, Patelli and Zhu (2024)) retrains open source foundation models on thousands of speeches and research papers housed in the BIS Central Bank Hub. This domain-specific fine-tuning lifts accuracy in parsing monetary terminology from roughly 50–60% to about 90% and sharpens tasks such as classifying policy stances or forecasting market responses to policy announcements (Graph 3).

Central bank language models (CB-LMs) outperform foundation models in prediction accuracy[1]

In per cent

Graph 3



[1] The y-axis represents the percentage of correct predictions from each model.

Source: Gambacorta, Kwon, Park, Patelli and Zhu (2024).

Through their few-shot learning abilities, LLMs also enhance nowcasting. Most existing forecasting tools remain highly specialised; the models built for GDP differ from those that track financial-risk build-ups, and analysts must hand-select inputs and fine-tune parameters. Adjusting these frameworks when the focus shifts can be time-consuming. Few-shot learners, by contrast, are far more versatile and can be readily repurposed for time series prediction with similar accuracy (Koyuncu et al (2025)).

3.3 Oversight of payment systems

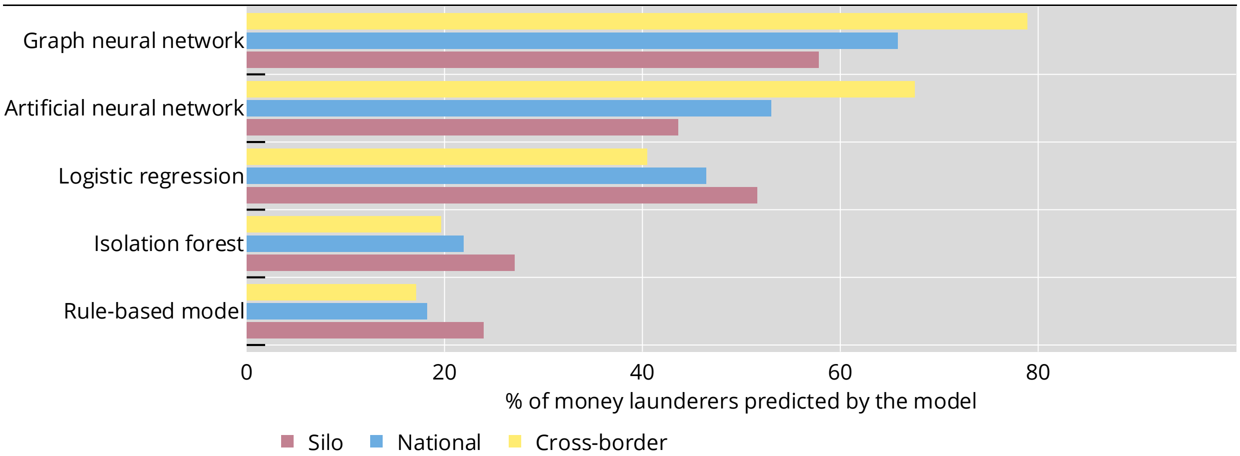
Well functioning payment systems underpin the stability of the financial system, yet the sheer volume of transaction data, often highly skewed, makes it hard to separate anomalous payments from routine flows. At present, banks and other financial institutions largely use rule-based systems (a primitive form of AI) to monitor transactions for suspicious activity. These systems generate alerts when certain thresholds or patterns (defined by experts or regulators) are met – for example, large cash deposits or transfers to high-risk jurisdictions might trigger an alert.

While rule-based monitoring is straightforward, it suffers from a number of shortcomings. It is notorious for false positives – vast numbers of flagged transactions that upon investigation turn out to be innocuous. It is not uncommon for 95% or more of anti-money laundering (AML) alerts to be false positives, creating huge workloads for compliance teams (and the financial intelligence units that they file reports to) and likely distracting from real criminal schemes. Conversely, launderers can sometimes split or camouflage transactions to avoid simple threshold triggers (techniques such as "smurfing" involve breaking large sums into many small transfers to fly under the radar). Another challenge is that each institution only sees its slice of the data. But money laundering networks often span multiple banks and countries, so no single institution can easily detect the full pattern if criminals hop between entities.

Partly due to these reasons, the result is an ineffective, fragmented global payments system. Criminals exploit the fragmentation and complexity of payment networks to hide illicit funds, while banks expend enormous resources on compliance. A troubling side effect in recent years has been *de-risking*: banks withdrawing from certain markets or client segments (eg correspondent banking in smaller or lower-income jurisdictions) because the cost and risk of AML compliance are very high. This retreat can leave some regions financially isolated (Rice et al (2020)).

Among the projects of the BIS Innovation Hub focused on the use of AI (discussed below), Project Aurora aims to tackle these issues directly. Project Aurora was launched by the BIS Innovation Hub as a proof of concept to test new techniques for AML monitoring across institutions and borders (BISIH (2023)). It uses synthetic money laundering data to benchmark traditional models against ML alternatives (Graph 4). The idea behind Aurora is to use advanced analytics (ML, network analysis) on a *holistic data set* of payment transactions, while using privacy-preserving technology to protect sensitive data. By having a more complete view of transaction networks, AI models can uncover complex money laundering patterns that would evade isolated checks at one bank. For example, a network of "mule" accounts (often unwitting individuals recruited to pass on funds) or a web of small transfers could be recognised by ML models when all the connections are analysed together.

Machine learning models' performance under different scenarios[1] Graph 4



[1] Transaction data visible on three different levels of analysis: the view of each financial institution (silo), the national view of a single country (national) and the cross-border view across countries (cross-border).

Source: BISIH (2023).

Aurora successfully performed a simulated collaborative analysis. Multiple institutions' transaction data were combined in encrypted form, and graph neural networks (a type of AI suited for network data) were applied to identify suspicious clusters of transactions. The results are striking. Aurora's approach proved far more effective than traditional rule-based AML monitoring approaches. According to the published findings, it detected up to three times more money laundering cases involving complex schemes and at the same time reduced false positives by up to 80%.

The reasons underlying the success are instructive and point to the value of combining behavioural analytics and data-sharing. First, instead of focusing on individual transactions, the focus was on the behaviour of networks of transactions (pattern of flows and relationships between senders/receivers) to spot anomalies indicative of laundering rings. This network-centric, behavioural approach improves detection as launderers might hide individual transactions, but their overall network behaviour (eg funds circling through intermediaries and back to the origin, and multiple accounts funnelling to one exit point) can give them away. Second, by pooling data from many institutions (with appropriate privacy safeguards), the ML algorithms have the "big picture" – they can see the cross-institution connections that a single bank's system would miss. For example, Bank A might see funds going to Bank B, and Bank B sees them going to Bank C. But neither alone sees A→B→C as one chain. The combined view does, and for some ML techniques like artificial neural networks and graph neural networks, such a holistic view can substantially improve performance.

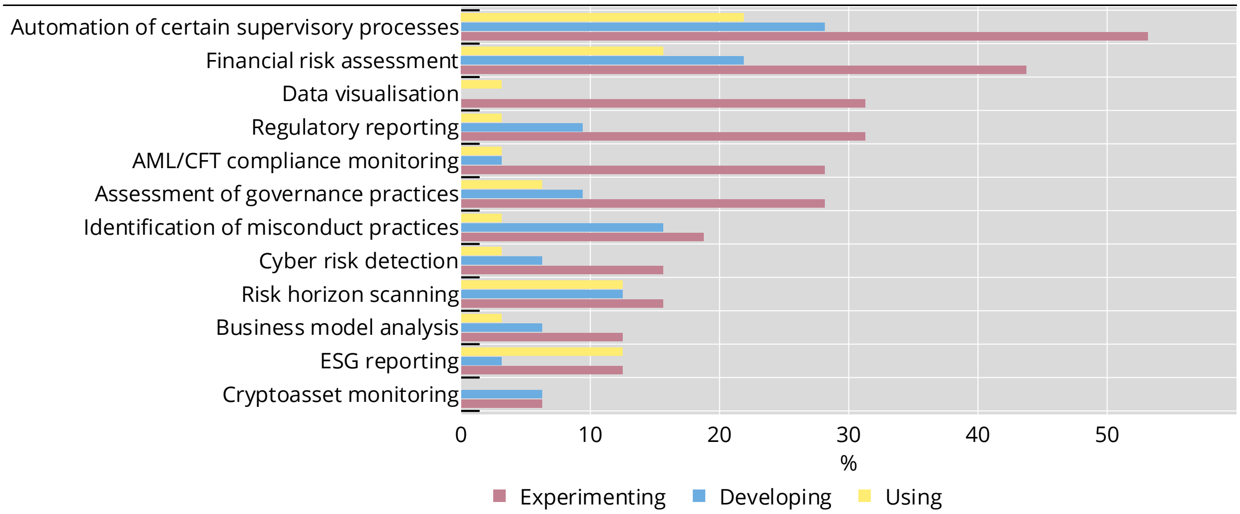
A critical aspect is that Aurora achieved this without violating privacy. It tested privacy-enhancing technologies such as secure encryption and federated learning. In the simulated cross-border scenario, transaction data were encrypted and consolidated so that the ML model could be trained on it, but individual customer identities were protected. The project also tested a setup in which countries collaboratively train a model but keep data locally (federated learning) to respect data sovereignty. These approaches showed that it is feasible to get the benefits of data pooling while maintaining confidentiality.

There are, of course, practical challenges ahead. Aurora's report notes that implementation with real-world data would face legal and technical hurdles (data protection laws, the need for common data standards etc). But the proof of concept provides a blueprint for how AML could be dramatically improved. If deployed, such AI-driven utilities could allow authorities and banks to find the needle in the haystack much more reliably in an AML context – effectively turning the current fragmented system into a unified network radar for illicit finance. Indeed, experimentation based on real-world data for Canada's high-value payment system already points to the improvements in detection of anomalous payments that can come about from the use ML (Desai et al (2025)).[9]

3.4 Supervision and financial stability

Supervisors draw on a wide spectrum of data to oversee financial institutions. Inputs often span textual material, including news stories, internal bank files and supervisory reviews. Manually mining this trove of data for actionable insights is labour-intensive and as volumes grow it verges on impossible. For example, cyber risk assessments have recently emerged as an important topic, yet they do not benefit from the well established data infrastructure that supports more "traditional" risk categories. AI can help in tasks such as document processing, knowledge management and document review. Indeed, a recent survey among supervisors shows that many use AI tools to automate supervisory processes, as well as for financial risk assessment (Graph 5).

Supervision areas with generative artificial intelligence applications — Graph 5



AML/CFT = anti-money laundering/combating the financing of terrorism; ESG = environmental, social and governance.

Source: Prenio (2025).

Many central banks pool information in single platforms to support supervisory work on unstructured data. Models fine-tuned on supervisory content and equipped with NLP can classify documents, gauge sentiment and flag emerging themes, as in the European Central Bank's Project Athena (Hoppe (2025)). Combining large texts with expert lexicons of key terms also automates the extraction of risk-relevant passages. The Federal Reserve's Language EXtraction Engine (LEX), for example, improves supervisors' access to pertinent material scattered across millions of files and shortens review times. Tree-based classifiers or neural nets can likewise pinpoint borrowers for whom lenders understate credit losses;

[9] Recent applications of AI such as "AI agents" (ie AI systems that build on advanced LLMs and are endowed with autonomy, planning and adaptation capabilities) can also be leveraged in the context of payments. For example, Aldasoro and Desai (2025) use prompt-driven experiments to show how simple AI agents can carefully balance liquidity through precautionary and calculated actions that mitigate liquidity-delay trade-offs in payment systems, much like human cash managers.

the Central Bank of Brazil's ADAM performs precisely this task. Networks built on pre-trained layers further sharpen the detection of high-loss borrowers, allowing supervisors to demand additional provisioning where coverage is thin.[10]

Market surveillance is another domain in which the use of AI can yield large benefits. Volatility in markets or the build-up of asset price bubbles can often be linked to patterns in trading data or sentiment that AI can discern. Some regulators already use ML to monitor market chatter (for instance, classifying news or forum posts about a particular asset for tone and credibility) as part of their financial stability dashboards. AI can also improve stress testing frameworks. Traditionally, stress tests involve hypothetical scenario analysis on banks' balance sheets. AI can help design more realistic scenarios by learning from historical crises (or even generating plausible crisis narratives via generative models), and it can accelerate the processing of results.[11] Moreover, AI can be employed to run dynamic stress simulations, adjusting assumptions on the fly, which is useful for testing many variations of scenarios efficiently. Finally, AI and ML excel at spotting cross-sectional patterns and therefore help map risk across large samples of financial and non-financial firms. Timely data remain essential. For instance, AI and ML can improve the prediction of financial market dysfunction, illiquidity and stress (Aldasoro, Hoerdahl, Schrimpf and Zhu (2025); Aquilina et al (2025)).

Finally, blending AI insights with expert judgment can support macroprudential analysis. Financial stability risks usually stem from a slow build-up of vulnerabilities that erupt in rare but costly crises. Sparse data on such events and the singular nature of each episode limit the stand-alone value of data-hungry AI models. Fed with rich data sets that offer ample scope for pattern recognition, AI can help craft early warning indicators that alert supervisors to brewing pressures linked to system-wide risk. Human expertise and sound economic reasoning must, however, remain a critical element to ensure gen AI can deliver gains in policy analysis.

In sum, central banks have been ardent adopters of AI in their daily operations. To support central banks and supervisors, the BIS Innovation Hub has been exploring various AI applications spanning the areas discussed above. As of 2025, the BIS Innovation Hub has undertaken nine projects explicitly involving AI methods, covering a wide range of use cases. These projects illustrate how AI can be applied to diverse policy challenges – from regtech to cyber security – and provide experimental insights that can be shared across the central banking community. Table 1 presents a summary and Box A provides further details on the projects.

[10] See summary in Beerman et al (2021) for details on LEX and ADAM.

[11] The FSB (2024) notes that authorities are exploring using generative AI to model social media's impact in a bank run scenario – essentially to simulate how quickly rumours could spread and affect deposits under stress. This is increasingly relevant in the age of instantaneous digital bank runs.

An overview of BIS Innovation Hub projects using AI — Table 1

	AISE	Aurora	Ellipse	Gaia	Symbiosis	Raven	Insight[1]	Neo	Spectrum
Main use case	Virtual assistant enhancing on-site supervision	Enhance AML suspicious transaction monitoring across firms & borders	Match entities in news with those of supervisory interest	Extract climate risk-related data from ESG reports	Develop methods for scope 3 emission disclosure	Process cyber security & resilience documents to facilitate assessment	Extract info & data on firm supply chain dependencies	Create and forecast economic indicators using timely and granular data	Structure big data on micro prices for inflation nowcasting
BISIH Centre	Toronto	Nordic	Singapore	Eurosystem	Hong Kong SAR	Nordic	Hong Kong SAR	Swiss	Eurosystem
Status	Ongoing	Ongoing (Phase 2)	Completed	Ongoing (Phase 2)	Ongoing	Ongoing	Ongoing	Ongoing	Ongoing
Key theme		Suptech/regtech		Climate risk analysis		Cyber security		Monetary policy tech	
NLP	✓	✗	✓	✓	✓	✓	✗	✗	✓
LLM	✓	✗	✗	✓	✓	✓	✓	✓	✓
SML	✓	✓	✗	✗	✓	✗	✗	✓	✗
UML	✓	✓	✗	✗	✓	✗	✗	✓	✓
Other	✓	✓	✗	✗	✗	✓	✓	✗	✗

BISIH = BIS Innovation Hub; ESG = environmental, social, and governance; LLM = large language model; NLP = natural language processing; SML = supervised machine learning (ML); UML = unsupervised ML.

[1] As the project is in its early stages, the full list of AI technologies to be used is still to be determined.

Source: BIS (2024).

Box A

BIS Innovation Hub projects on artificial intelligence

In recent years the BIS Innovation Hub has been exploring various artificial intelligence (AI) applications to support central banks and supervisors in fulfilling their mandates. To date, nine projects use AI methods, broadly grouped into four distinct categories: applications of regulatory and supervisory technology (suptech/regtech), applications to support climate risk analysis, cyber security and monetary policy.

Within the first group, there are four projects: Aurora (discussed in the main text), AISE (AI Supervisor Enhancer) and Ellipse. Project AISE seeks to address inefficiencies in financial supervision by integrating generative AI into supervisory workflows. It aims to automate data extraction, enhance pattern detection, standardise on-site supervision and streamline the generation of reports. The proposed solution includes a modular AI toolkit, composed of scalable microservices, to provide flexible and reusable tools for supervisory authorities. Project Ellipse uses AI to help supervisors monitor financial institutions. The project developed ML techniques to match entities in news reports with those in regulators' watchlists and to flag relevant news articles to supervisors in real time – allowing them to investigate further. The goal is to ensure that supervisors are quickly alerted to any issue or misconduct involving the firms that they oversee. Early results suggest that ML methods can effectively alert supervisors to issues for investigation sooner than traditional channels would.

Two projects aim to support climate risk analysis. Project Gaia introduced an AI-enabled solution that extracts and organises information from unstructured environmental, social and governance reports. By integrating semantic search with a large language model (LLM), the project enhances the comparability of climate-related information, which is valuable for regulators assessing banks' climate exposures. Building on these foundations, the second phase of Project Gaia focuses on refining the solution for eventual handover to the central banking community. Project Symbiosis in turn uses AI and big data to analyse the emissions and impacts of supply chains linked to multinational corporations ("anchor buyers") and financial institutions. The goal is to identify instances in which new financing can promote sustainable supply chains by connecting suppliers, anchor buyers and financial institutions.

Project Raven uses AI to process cyber security and resilience documents to inform supervisors. Financial institutions regularly submit reports about their cyber defences, incident response plans etc. Raven's tool uses NLP to read these documents and answer specific queries from supervisors (eg "Does this bank have a policy on data breach notification?"). This helps streamline the supervision of technology risks. By quickly summarising and identifying key points in lengthy technical documents, AI can assist supervisors in evaluating banks' cyber preparedness. Given that cyber risk is a growing concern for financial stability, such an AI assistant can improve oversight efficiency.

Finally, three projects are aimed at supporting analysis relevant for monetary policy. First, Project Neo aims to create and forecast economic indicators using timely and granular data. In particular, it can use alternative data sources (like high-frequency retail data, mobility data etc) to produce new indicators or nowcasts that inform policy in near-real time. Second, Project Spectrum is an effort to structure big data on micro prices for inflation nowcasting. It deals with large data sets of individual prices (from eg web-scraped prices or barcode scanners) and uses AI to organise and analyse them for signs of inflation trends. By handling "micro price" data, AI can detect price changes across millions of products and services, providing policymakers with early signals of inflation pressure. Third, Project Insight aims to extract information on firm supply chain dependencies. The project uses AI to parse data (including textual data like company reports or databases) to map out supply chain linkages between firms. Insight's AI could help build a knowledge graph of how firms are interconnected through supply chains, allowing central banks to assess how shocks propagate in the real economy (which, for example, could feed into financial stability if corporate defaults cluster).

4. Challenges and lessons learned

The experience to date from the various use cases of ML and AI more generally has been largely positive. Central banks report "considerable success" in applying ML. For instance, many have achieved more accurate forecasts or faster analytical processes, and supervisory AI tools have improved the identification of compliance issues (Araujo et al (2024)).

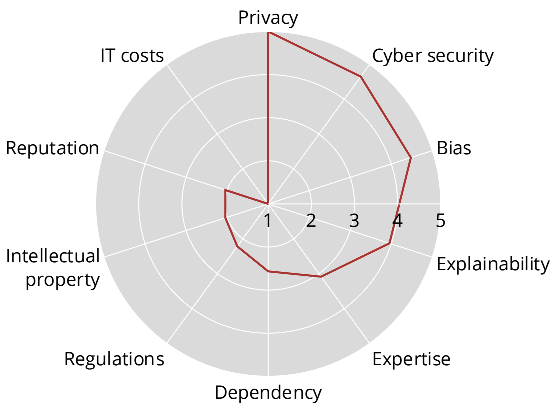
However, the use AI and ML at central banks and other supervisory authorities comes with several challenges (Graph 6). A general challenge is a tension between predictive accuracy and "interpretability/ explainability". Sophisticated ML models can deliver accurate predictions, but the complex, non-linear interactions of many variables make it hard to assess which inputs are key to driving the output. Reliable forecasts therefore come at the price of accepting that the underlying model is a "black box".[12] This opacity can hamper efforts to detect discriminatory biases, especially when algorithms are trained on biased data. Limited explainability also makes it hard to communicate model behaviour in human terms, for instance, why inflation is forecast to rise or why a mortgage application was turned down. This in turn can undermine communication strategies.

For gen AI models the issue of explainability is compounded by the risk of "hallucinations", where a model presents a factually wrong answer with apparent confidence. While the prevalence of hallucinations is likely to decline over time as models improve, they are difficult to eliminate completely, since LLMs are trained to predict the next most plausible word from a probabilistic perspective in response to a given input or prompt. The unavoidability of this risk means that LLMs need human supervision, particularly in tasks that require logical reasoning (Perez-Cruz and Shin (2024)).

Central banks face significant challenges and risks in adopting AI/ML — Graph 6

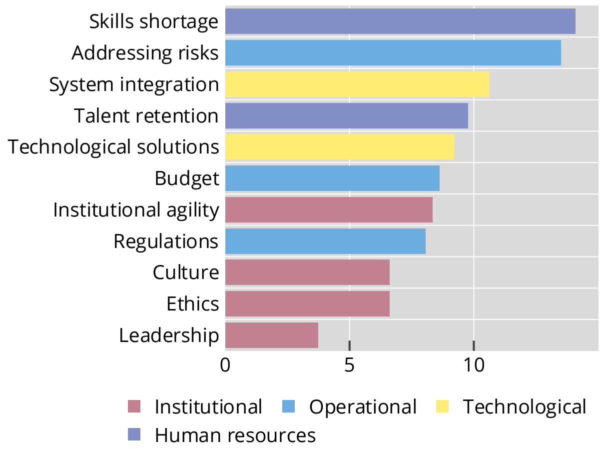
A. Privacy, cyber security and biases are top concerns[1]

Normalised score, 1–5



B. Skills shortage and addressing risks are key barriers

% of responses



[1] Normalised scores from 1 to 5 (1 = not sure; 2 = not impactful at all; 3 = slightly impactful; 4 = moderately impactful; 5 = highly impactful).

Source: IFC (2025).

[12] There are attempts to remedy this in the context of financial stability monitoring – see eg Aldasoro, Hoerdahl, Schrimpf and Zhu (2025) or Aquilina et al (2025).

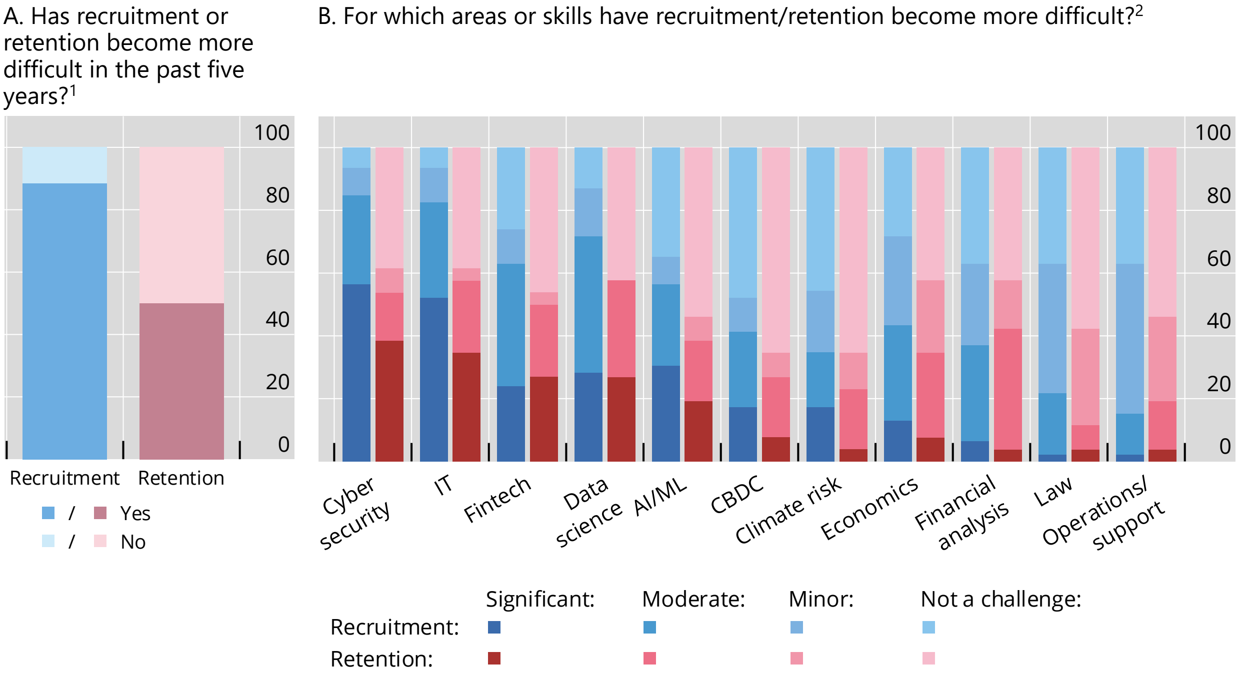
Greater use of AI also requires central banks to invest even more in IT and human capital. Powerful hardware, software licences and staff training all carry large upfront costs. Recruiting or retaining people who combine economic expertise with coding skills is difficult, given strong private sector demand. Over time, however, these investments can pay off. Evidence shows that ML tools can lift staff productivity in time-consuming cognitive tasks such as text summarisation (Brynjolfsson et al (2023); Noy and Zhang (2023); Gambacorta Qiu, Rees and Shian (2024)). AI systems can serve as "copilots" for supervisory teams, freeing economists to interpret data rather than collect and clean them. Still, humans will remain essential: expert feedback improves models and mitigates hallucinations.

Central banks' human capital strategies must therefore evolve along two plausible paths. In the near term, most institutions are likely to deploy "AI copilots", ie LLM-based assistants that augment staff rather than replace them. In the medium term, central banks need to prepare for a second, more disruptive phase in which autonomous "AI agents" could take over narrowly defined tasks. Both paths demand large-scale retraining of existing staff, targeted recruitment of data science and ML specialists, and a workplace culture that prizes experimentation and cross-disciplinary teamwork.

But hiring talent in this field is increasingly difficult (Graph 6.B). Nearly nine in 10 central banks report heavier recruitment headwinds, above all for cyber security, IT, fintech and AI roles (Graph 7), while many face legal hurdles such as civil service exams or citizenship requirements. Institutions are closing capability gaps by blending permanent staff with consultants, contractors and remote specialists, and by stressing their public interest mission, unique data assets and training opportunities to candidates who might otherwise choose the private sector.

Challenges in recruitment and retention

As a percentage of respondent central banks[1] Graph 7



AI = artificial intelligence; CBDC = central bank digital currency; IT = information technology; ML = machine learning.

[1] Based on a survey of 52 members of the Central Bank Governance Network conducted in May 2024. [2] Shares based on the subset of central banks experiencing more difficulties with recruitment and/or retention.

Sources: CBGN (2024); Bell et al (2025).

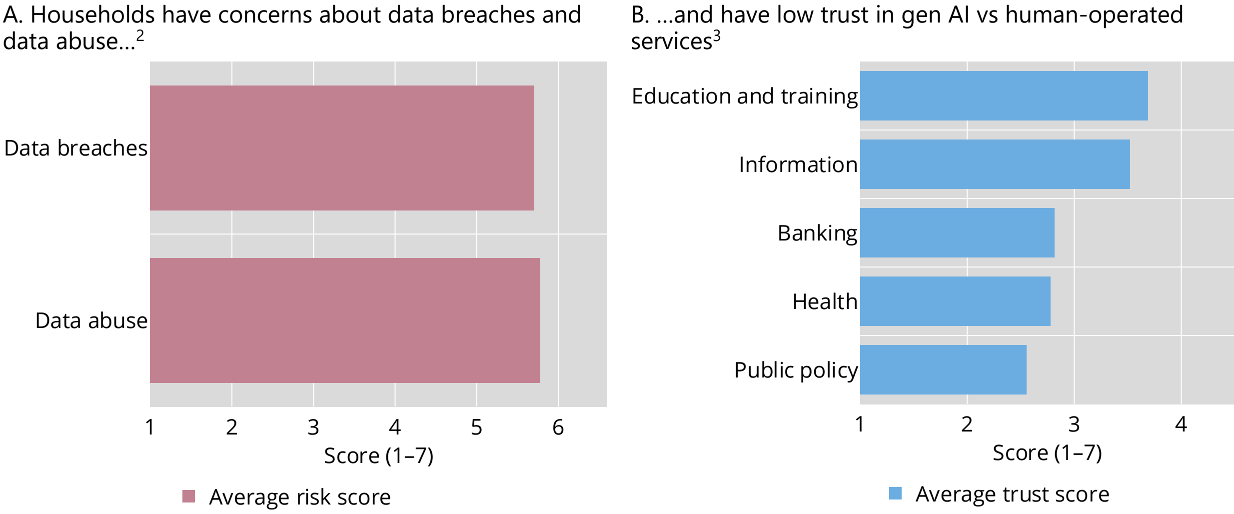
Successful adoption and use of AI hence hinge on robust change management and governance. Continuous learning programmes, clear development plans for teams and individuals, and embedded AI-ethics frameworks can help staff adapt to new tools and responsibilities. In the copilot scenario, the priority is weaving AI insights into daily analysis, but humans remain directly in charge and control of processes. While in the agent scenario, oversight skills, system supervision and disciplined data engineering become paramount. Diverse, inclusive teams that mix specialists with generalists are best placed to manage this transition while safeguarding institutional knowledge.

The use of unstructured or personal data also raises legal and privacy concerns. AI thrives on data, but financial authorities must uphold strict privacy standards. Traditionally, public institutions gathered and stored most data under clear access rules and robust quality controls. Today, individuals and firms generate vast amounts of – often poorly documented – information that sit with private entities. Training or fine-tuning LLMs demands large data sets, sometimes scraped from market platforms or social media, but legal frameworks rarely specify how such data may be used and the legality of such scraping with respect to intellectual property rights law is currently being disputed. Public concern over privacy looms large. Consumers worry about the potential for data breaches and abuse (Graph 8.A; Armantier et al (2021, 2024)) and overwhelmingly support stricter privacy regulation of AI (Aldasoro, Armantier, Doerr, Gambacorta and Oliviero (2024a,b); Chen et al (2023)). They also report lower levels of trust in gen AI than in human-operated services, especially in high-stakes areas such as banking and public policy (Graph 8.B). However, while privacy-enhancing techniques are advancing, they are not yet standard in AI models.

The growing importance of data and the rise of new tools and sources present significant data governance challenges for central banks, requiring robust frameworks to manage data and algorithms effectively. These frameworks should include quality control, auditing practices and metadata management, which are increasingly critical as data diversity expands. Metadata (eg definitions, sources and units) enhance machine readability, especially when standardised under principles like "Findable, Accessible, Interoperable and Reusable" (FAIR),[13] and supports privacy-preserving methods for cross-institutional data use. Collaboration among central banks, through standards like SDMX and frameworks like the Generic Statistical Business Process Model, can promote data-sharing, improve interoperability, reduce reporting burdens and enable the secure exchange of confidential information.

[13] See Wilkinson et al (2016).

Households' trust in generative AI (gen AI)[1] Graph 8



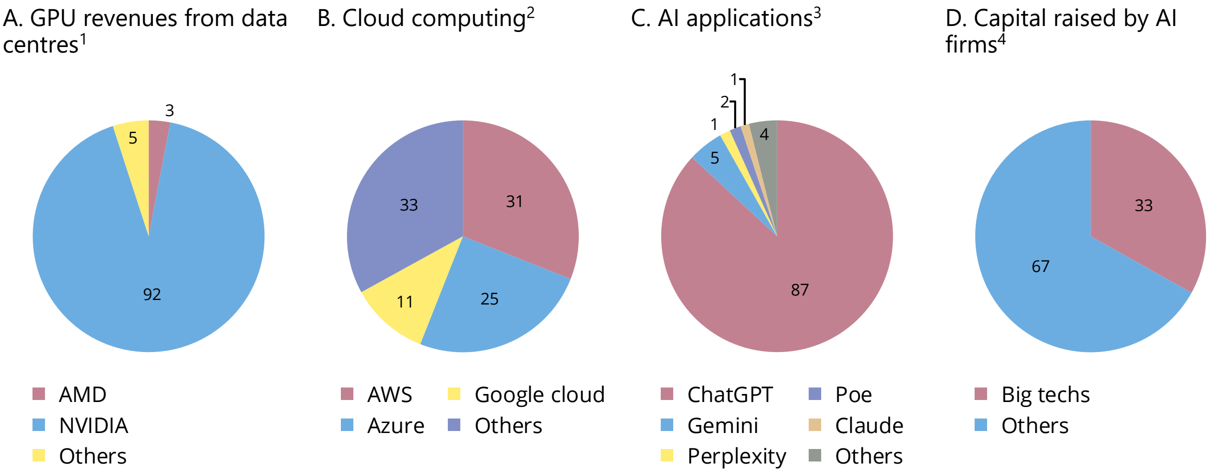
[1] Based on a representative sample of US households from the Survey of Consumer Expectations. [2] Average scores (with scores ranging from 1 (lowest) to 7 (highest)) in answers to the following questions: (1) "Do you think that sharing your personal information with artificial intelligence tools will decrease or increase the risk of data breaches (that is, your data becoming publicly available without your consent)?"; (2) "Are you concerned that sharing your personal information with artificial intelligence tools could lead to the abuse of your data for unintended purposes (such as for targeted ads)?" [3] Average scores in answers to the following question: "In the following areas, would you trust artificial intelligence tools less or more than traditional human-operated services? For each item, please indicate your level of trust on a scale from 1 (much less trust than in a human) to 7 (much more trust)."

Sources: Aldasoro, Armantier, Doerr, Gambacorta and Oliviero (2024a); Federal Reserve Bank of New York, Survey of Consumer Expectations.

Regarding operational challenges, the rise of LLMs and gen AI revives concerns about dependence on a handful of external providers. Economies of scale mean that only a few large tech firms supply the most powerful foundation models. Indeed, most of the AI supply chain is characterised by a high degree of concentration, from data centres to cloud computing and AI applications (Graph 9). Such concentration can threaten innovation and raise financial stability, operational and reputational risks. Heavy reliance on a small set of vendors exposes the system to IT failures or cyber attacks. Outages can impair central banks' operations and, by extension, their mandates. Reputational risks are high, as public trust is a central bank's core asset (Carstens et al (2022)). Closely related is the trade-off between using "off-the-shelf" models versus developing in-house fine-tuned ones. Reliance on external models comes with reduced transparency, even if it may be more cost-effective in the short run. Moreover, if many institutions adopt the same top-tier algorithms, their actions in stress episodes may converge, spurring liquidity hoarding, interbank runs and fire sales (Danielsson and Uthemann (2024); FSB (2024)).

Market structure of the artificial intelligence (AI) supply chain

In per cent | Graph 9



AMD = Advanced Micro Devices; AWS = Amazon Web Services; GPU = graphics processing unit.

[1] Based on global revenues of GPU producers for GPUs used in data centres in 2023. [2] Based on global cloud computing revenues for Q1 2024. [3] Based on monthly visits data. For further details see Liu and Wang (2024). [4] Based on total capital invested in 2023 in firms active in artificial intelligence and machine learning. Big techs correspond to Alibaba Cloud, Alibaba Group, Alphabet, Amazon Industrial Innovation Fund, Amazon Web Services, Amazon, Apple, Google Cloud Platform, Google for Startups, Microsoft, Tencent Cloud, Tencent Cloud Native Accelerator and Tencent Holdings.

Source: Gambacorta and Shreeti (2025).

Closely related are concerns about cyber risk, as AI not only heightens them but also introduces new sources of such risk. For example, gen AI has the potential to significantly enhance hackers' ability to craft convincing phishing emails, develop malware and use it to steal sensitive information or encrypt a company's files for ransom, improving attackers' ability to deploy traditional cyber hacking techniques. Furthermore, gen AI enables hackers to mimic an individual's writing style or voice, or even create fake avatars, potentially leading to a sharp increase in phishing attacks. Other types of attacks are specific to LLMs. For example, LLMs are usually instructed not to provide dangerous information, such as how to manufacture napalm. However, in the infamous "grandma jailbreak" (a so-called prompt injection attack), when the prompter asked ChatGPT to pretend to be their deceased grandmother telling a bedtime story about the steps to produce napalm, the chatbot did reveal this information. While this vulnerability has been fixed, others remain and more will appear over time. As more applications use data created by LLMs, such attacks could have increasingly severe consequences, leading to heightened operational risks.

Central banks in AEs and EMEs alike deem phishing and other forms of social engineering as the most likely type of attack. When it comes to the costs resulting from an attack, advanced persistent malware and ransomware attacks rank highest (Doerr, Gambacorta, Leach, Legros and Whyte (2022)). And the estimated risks and costs of these attacks are likely to be exacerbated by gen AI (Aldasoro, Doerr, Gambacorta, Notra, Oliviero and Whyte (2025)).[14] Most central banks have therefore increased their budgets for investment in cyber security. Technical security control and resilience feature high on the

[14] With respect to generative AI, the prevailing view is that gen AI can outperform traditional methods in enhancing cyber security management, but that it also introduces new risks. The benefits are largely perceived in specific areas of cyber security, such as automation of routine tasks or enhanced threat detection. In terms of risks, gen AI can introduce new vulnerabilities into central banks' cyber security defences. Risks related to social engineering and zero-day attacks as well as unauthorised data disclosure are of highest concern.

priority list in terms of areas for investment in cyber security. Cyber risk is also an area that intersects with human resources challenges: training existing staff in cyber security aspects or hiring new staff with the relevant skills are also considered very important (Graph 10). Beyond investments, central banks put a high focus on developing an incident response plan in case their own institution is attacked and on developing a formal strategy on how to respond to an attack on the financial system.

AI, cyber security and human capital of central banks — Graph 10

[1] Share of respondents choosing each option when asked "How may AI impact or how has it already impacted the allocation of human resources in cyber security tasks?" [2] Other includes: "We think it is still early for us to see real improvement in terms of workload optimisation through the use of AI-based tools"; "In many cases, the impact on HR is still unknown"; "We have no idea yet how this will transform the way we work in cyber security"; "No impact yet. Use of AI is limited as pilot testing and access or interface to production data are not yet allowed." [3] Average score on a scale of 1 (very low) to 5 (very high) that respondents gave to each option when asked to "Rate the following concerns regarding the limited availability of AI-qualified personnel." [4] Share of respondents choosing each option for the question "How do you envision the interaction between AI systems and human cyber security experts evolving?"

Source: Aldasoro, Doerr, Gambacorta, Notra, Oliviero and Whyte (2025).

Dealing with the challenges arising from the operational and human resources side requires the pooling of resources and experience to leverage synergies. Cooperation can help alleviate resource constraints by pooling knowledge and sharing costs for collecting, storing and analysing big data, as well as developing and training AI models. For example, joint procurement of commercial data and sharing granular data internally can reduce costs, particularly for smaller institutions. Collaboration can also support staff training through workshops and conferences, while re-using trained models can lower environmental costs associated with energy-intensive processes. Additionally, standardising data practices and creating shared repositories for open-source tools, such as BIS Open Tech, can enhance the efficiency and accessibility of AI applications. Sharing fine-tuned models across the central banking community could further reduce barriers to adoption while maintaining data confidentiality (Araujo (2023)).

5. Conclusions

The rapid and widespread adoption of AI by households and firms implies that there is an urgent need for central banks and other supervisory and regulatory authorities to raise their game. To address the growing challenges, there is a need to upgrade their capabilities both as *informed observers* of the effects of technological advancements and as *users* of the technology itself. For example, central banks as observers need to stay ahead of the impact of AI on economic activity through its effects on aggregate supply and demand. As users, they need to build expertise in incorporating AI and non-traditional data in their own analytical tools as well as in how to use AI to produce reliable data.

As early adopters of big data and machine learning methods, central banks are ideally placed to harness the capabilities of AI. Given the availability of vast troves of structured and unstructured data as well as the need for rigorous analysis in support of policy, the synergies between machine learning and core central banking disciplines such as economics, statistics and econometrics are likely to place central banks at the vanguard of advances in AI. They stand to benefit across a number of key tasks, ranging from data collection and dissemination to monetary policy and supervision.

However, to harness the benefits of AI central banks and other authorities need to address various challenges and important trade-offs. These involve the trade-off between using external versus internal AI models, as well as in collecting and providing in-house data versus purchasing them from external providers. Together with the centrality of data, the rise of AI will require a rethink of central banks' traditional roles as compilers, users and providers of data.

To mitigate trade-offs and overcome challenges, collaboration and the sharing of experiences emerge as key avenues. Coming together to form a "community of practice" to share knowledge, data, best practices and AI tools emerges as a promising way forward. The BIS is supporting central banks in this endeavour.

6. References

Accornero, M and G Boscariol (2021): “Machine learning for anomaly detection in data sets with categorical variables and skewed distributions”, in “Machine learning in central banking”, *IFC Bulletin*, no 57.

Aldasoro, I, O Armantier, S Doerr, L Gambacorta and T Oliviero (2024a): “Survey evidence on gen AI and households: job prospects amid trust concerns”, *BIS Bulletin*, no 86, April.

——— (2024b): “The gen AI gender gap”, *Economics Letters*, vol 241.

Aldasoro, I and A Desai (2025): “AI agents for cash management in payment systems”, *BIS Working Papers*, forthcoming.

Aldasoro, I, S Doerr, L Gambacorta, G Gelos and D Rees (2024): “Artificial intelligence, labour markets and inflation”, *SUERF Policy Brief*, no 923.

Aldasoro, I, S Doerr, L Gambacorta and D Rees (2024): “The impact of artificial intelligence on output and inflation”, *BIS Working Papers*, no 1179, April.

Aldasoro, I, L Gambacorta, A Korinek, V Shreeti and M Stein (2024): “Intelligent financial system: how AI is transforming finance”, *BIS Working Papers*, no 1194, June.

Aldasoro, I, S Doerr, L Gambacorta, S Notra, T Oliviero and D Whyte (2025): “Generative artificial intelligence and cybersecurity in central banking”, *Journal of Financial Regulation*, vol 11, no 1, pp 119–28.

Aldasoro, I, P Hoerdahl, A Schrimpf and S Zhu (2025): “Predicting financial market stress with machine learning”, *BIS Working Papers*, no 1250.

Aquilina, M, D Araujo, G Gelos, T Park and F Perez-Cruz (2025): “Harnessing artificial intelligence for monitoring financial markets”, *BIS Working Papers*, no 1291.

Araujo, D (2023): “gingado: a machine learning library focused on economics and finance”, *BIS Working Papers*, no 1122, September.

Araujo, D, S Doerr, L Gambacorta and B Tissot (2024): “Artificial intelligence in central banking”, *BIS Bulletin*, no 84, January.

Armantier, O, S Doerr, J Frost, A Fuster and K Shue (2021): “Whom do consumers trust with their data? US survey evidence”, *BIS Bulletin*, no 42, May.

——— (2024): “Nothing to hide? Gender and age differences in the willingness to share data”, *BIS Working Papers*, no 1187, May.

Aruoba, S B and T Drechsel (2024): “Identifying monetary policy shocks: A natural language approach”, *NBER Working Papers*, no 32417.

Bank for International Settlements (BIS) (2024): “Artificial intelligence and the economy: implications for central banks”, *Annual Economic Report*, June, Chapter III.

Beerman, K, J Prenio and R Zamil (2021): “Suptech tools for prudential supervision and their use during the pandemic”, *FSI Insights on Policy Implementation*, no 37.

Bell, S, B Gadanecz, L Gambacorta, F Perez-Cruz and V Shreeti (2025): “Artificial intelligence and human capital: challenges for central banks”, *BIS Bulletin*, no 100.

BIS Innovation Hub (BISIH) (2023): *Project Aurora: the power of data, technology and collaboration to combat money laundering across institutions and borders*, May

Brynjolfsson, E, D Li and L Raymond (2023): "Generative AI at work", *NBER Working Papers*, no 31161.

Buckmann, M, G Potjagailo and P Schnattinger (2023): "Dissecting UK service inflation via a neural network Phillips curve", *Bank of England Bank Underground*, July.

Carstens, A, J Frost and H S Shin (2022): "A foundation of trust", *IMF Finance & Development*, September.

Cazzaniga, M, F Jaumotte, L Li, G Melina, A Panton, C Pizzinelli, E Rockall and M Tavares (2024): "Gen-AI: artificial intelligence and the future of work", *International Monetary Fund Staff Discussion Notes*, no 1, January.

Central Bank Governance Network (CBGN) (2024): *Survey on talent management and new ways of working at central banks*.

Chen, S, S Doerr, J Frost, L Gambacorta and H S Shin (2023): "The fintech gender gap", *Journal of Financial Intermediation*, vol 54.

Danielsson, J and A Uthemann (2024): "On the use of artificial intelligence in financial regulations and the impact on financial stability", available at ssrn.com/abstract=4604628.

Denes, J, A Lestrade and L Richardet (2021): "Using Twitter data to measure inflation perception" in "Machine learning in central banking", *IFC Bulletin*, no 57.

Desai, A, A Kosse, and J Sharples (2025): "Finding a needle in a haystack: a machine learning framework for anomaly detection in payment systems", *The Journal of Finance and Data Science,* no 11, 100163.

Doerr, S, L Gambacorta, T Leach, B Legros and D Whyte (2022): "Cyber risk in central banking", *BIS Working Papers*, no 1039, September.

Doerr, S, L Gambacorta and J M Serena Garralda (2021): "Big data and machine learning in central banking", *BIS Working Papers*, no 930, March.

Du, S, K Guo, F Haberkorn, A Kessler, I Kitschelt, S J Lee, A Monken, D Saez, K Shipman and S Thakur (2023): "Do anecdotes matter? Exploring the Beige Book through textual analysis from 1970 to 2023", *IFC Bulletin*, no 64.

Financial Stability Board (FSB) (2024): *The financial stability implications of artificial intelligence*, report submitted to the G20, November.

Gabaix, X, R Koijen, R Richmond and M Yogo (2025): "Asset embeddings", mimeo.

Gambacorta, L, B Kwon, T Park, P Patelli and S Zhu (2024): "CB-LMs: language models for central banking", *BIS Working Papers*, no 1215.

Gambacorta, L, H Qiu, D Rees and S Shian (2024): "Generative AI and labour productivity: a field experiment on code programming", *BIS Working Papers*, no 1208.

Gambacorta, L and V Shreeti (2025): "The AI supply chain", *BIS Papers*, no 154.

Hoppe, F (2025): "Benefits from advanced technology infrastructure in supervision", *ECB Supervision Newsletter*, May.

International Monetary Fund (IMF) (2024): "Advances in artificial intelligence: implications for capital market activities", *Global Financial Stability Report*, Chapter 3, October, pp 77–105.

Irving Fisher Committee (IFC) (2025): "Governance and implementation of artificial intelligence in central banks: 2024 survey conducted by the Irving Fisher Committee on Central Bank Statistics", *IFC Report*, no 18, April.

Kamenetsky Yadan, S (2021): "Anomaly detection methods and tools for big data" in "Machine learning in central banking", *IFC Bulletin*, no 57.

Koyuncu, B, B Kwon, M Lombardi, F Perez-Cruz and H S Shin (2025), "A foundational model for conditional forecasting macroeconomic variables", *BIS Working Paper*, forthcoming.

Kwon, B, T Park, F Perez-Cruz and P Rungcharoenkitkul (2024): "Large language models: a primer for economists", *BIS Quarterly Review*, December, pp 37–52.

Liu, Y and H Wang (2024): "Who on earth is using generative AI?", *World Bank Policy Research Working Paper*, no 10870.

Mitchell, M (2020): *Artificial intelligence: a guide for thinking humans*, Picador Paper Publishing, 17 November.

Noy, S and W Zhang (2023): "Experimental evidence on the productivity effects of generative artificial intelligence", *Science*, vol 381, no 6654, pp 187–92.

Perez-Cruz, F and H S Shin (2024): "Testing the cognitive limits of large language models", *BIS Bulletin*, no 83, January.

Prenio, J (2025): "Starting with the basics: a stocktake of gen AI applications in supervision", *FSI Briefs*, no 26.

Rice, T, G von Peter and C Boar (2020): "On the global retreat of correspondent banks", *BIS Quarterly Review*, March, pp 37–52.

Russell, S and P Norvig (2021): *Artificial intelligence: a modern approach*, Pearson, 4th edition.

van Dijk, D and J de Winter (2023): "Nowcasting GDP using tone-adjusted time varying news topics: Evidence from the financial press", *DNB Working Paper*, no 766.

Vaswani, A, N Shazeer, N Parmar, J Uszkoreit, L Jones, A Gomez, Ł Kaiser and I Polosukhin (2017): "Attention is all you need", in I Guyon, U Von Luxburg, S Bengio, H Wallach, R Fergus, S Vishwanathan and R Garnett (eds), *Advances in neural information processing systems 30: 31st annual conference on neural information processing systems (NIPS 2017)*, Neural Information Processing Systems Foundation.

Wilkinson, M et al (2016): "The FAIR guiding principles for scientific data management and stewardship", *Scientific Data*, no 3.