SUERF
The European Money and Finance Forum

SUERF Policy Brief
No 1208, July 2025

Central banks' experience in enhancing the access to and sharing of information through data science

Douglas Araujo, Adam Cap, Ilaria Mattei, Rafael Schmidt, Olivier Sirello and Bruno Tissot
Bank for International Settlements (BIS)

Keywords: data science, data access, data sharing, central banks, official statistics, innovation

JEL codes: C80, C83, E58, O31, O33

Abstract

Data science offers significant potential for leveraging both traditional and emerging data sources, improving statistical processes and enabling advanced data analysis in central banking. It also facilitates the secure sharing of granular data sets while protecting sensitive information. However, there are various challenges, including the need for robust IT infrastructure, organisational barriers and limited quality of secondary sources, which can constrain their use for official purposes. Fortunately, central banks are well-equipped to address these issues, not least because of their expertise as both data producers and users. Looking ahead, enhancing data management, promoting interoperability, investing in modern data platforms and fostering structured exchanges of experiences across stakeholders are essential to unlock the full potential of data in today's digital age and effectively support central banks' public mandates.

Disclaimer: This Policy Brief is based on IFC (2025a) published by the Irving Fisher Committee on Central Bank Statistics (IFC) of the Bank for International Settlements (BIS). The authors are at the BIS. The views expressed are their own and do not necessarily reflect the views of the BIS, the IFC or its members.

Introduction

The ongoing data revolution presents central banks with various opportunities to support their statistical function as well as the formulation and implementation of evidence-based policies. The exponential growth in the amount of accessible information enables more real time, detailed and multidimensional insights, particularly through the integration of novel sources such as big data into traditional macroeconomic aggregates (IFC (2024)). However, challenges often arise from the limited quality and standardisation of large or complex data sets, as well as from operational, ethical and legal barriers such as confidentiality restrictions.

Fortunately, **innovative data science techniques** - including big data analytics, machine learning (ML) and artificial intelligence (AI) – **can be part of the solution**. These tools help central banks to better tap into unstructured data, such as text. They can also improve data quality, extract meaningful insights and facilitate data access and sharing while protecting sensitive information. Against this backdrop, data science has emerged as a practical solution for central banks to fully unlock the potential and value of data.

Drawing on the IFC bulletin "Data science in central banks: enhancing the access to and sharing of data", **this policy brief underscores the benefits of data science for central banks as both producers and users of reference information**. It also highlights some key obstacles, particularly with emerging and secondary data sources, and emphasises the importance of international collaboration to address these challenges.

Accessing and making data available for further analysis

Data science can enhance the use of existing as well as novel data sources, with benefits for both producers and users of economic and financial information in central banks.

For producers, data science and ML in particular play a crucial role in enhancing data quality, accuracy and timeliness. Central banks are using these technologies for tasks such as anomaly detection, error identification and uncovering patterns in financial data to support areas such as prudential supervision and anti-money laundering (Shah (2025)). ML also facilitates the production of more granular and timely statistics by leveraging real time sources and automating processes across the data life cycle, including data collection, integration and evaluation (IFC (2022)).

Turning to users, innovative techniques, such as clustering, neural networks and natural language processing (NLP), **can enable the analysis of large, multidimensional and/or complex data sets**. These methods help, for instance, to uncover patterns, assess market sentiment and improve forecasting accuracy for key economic indicators. In particular, NLP and large language models (LLMs) allow central banks to tap into unstructured data, such as text. This can be quite valuable for them, for instance in extracting qualitative insights from their communications to improve forecasting of macroeconomic variables such as inflation (Araujo et al (2024)).

Data science can also improve the access to and sharing of data. Advanced techniques enable information to be shared or analysed while securing an adequate level of protection, eg through privacy-enhancing technologies, such as data obfuscation and encrypted data processing. Perhaps more significantly, data science allows to share information without disclosing it. Techniques such as federated learning, for instance, enable models to be trained without revealing the raw data. Multi-party secure private computing, on the other hand, supports the sharing of computations among data holders, allowing them to retain full control over their information assets without needing access to others' data (Ricciato (2024)).

Key obstacles in accessing, sharing and using data

While innovative data techniques can enable better access to and sharing of vast and diverse amounts of information, various challenges remain.

A key challenge is the **substantial investments in IT systems required to effectively leverage data science**. Long-term strategies may thus be needed to address the growing need for high-performance computing and platforms for managing diverse data types. Fortunately, central banks appear to have significantly upgraded their IT and data infrastructures over the past few decades (IFC (2020)). They are also increasingly interested in adopting cloud solutions because of their scalability and performance. Yet they are also aware of a number of issues posed by the cloud, not least increased dependency on external providers, potentially weaker information protection and concerns about data sovereignty.

The **adoption of data science may also face organisational barriers**. First, data silos can be an issue as they prevent users from effectively tapping into the available information due to isolated data management being split across the various functional domains of central banking. Second, the low transparency of techniques such as ML can also explain organisational risk aversion towards adopting data science, while also highlighting the need for continuous training and human oversight (IFC (2025b)). Finally, limited collaboration between data scientists, IT and subject-matter experts can prevent the full deployment of data science within the organisation. Experiences from central banks suggest that success depends on a clear strategy, sound data governance and close stakeholder collaboration, for instance by setting up organisational structures such as data science hubs (Duijm and van Lelyveld (2025)).

Finally, the **deployment of data science can be hindered by limitations in data quality**. Emerging data sources, such as big data, are often a byproduct of other processes and may feature a low degree of standardisation and consistency, making them unsuitable for statistical purposes without adequate preprocessing (Perrazzelli (2025)). Challenges also arise due to the lack of consistent identifiers, impeding tasks like quality assurance and integration as shown by the BIS' experience with building a database on cryptocurrencies (Illes and Mattei (2025)).

Making data access and sharing more effective

The above considerations highlight that achieving effective and adequate data access and sharing may require further progress in four key areas.

First, **data management and governance** are important for ensuring data quality, usability and security through clear sets of principles and adequate resources (Križman and Tissot (2022)). Key steps might include ensuring the dissemination of high-quality information through traceable documentation and citation standards, enhancing data usability by developing common inventories and adopting a master data management approach to harmonise statistical concepts within a common data dictionary (Gonçalves et al (2025)). Data governance is also a key component of AI governance, not least to ensure adequate provenance and traceability of the data used by AI systems (Figure 1; IFC (2025b)).

Implementing AI governance: four key components Figure 1

Source: IFC, "Governance and implementation of AI in central banks", *IFC Report*, no 18, April 2025.

Second, the development of **interoperable information standards** can also be instrumental in effective data use and reuse, including for training ML models (Araujo (2023)). Standards such as Statistical Data and Metadata eXchange (SDMX) provide a unified governance framework for both macro and micro data exchange and metadata management, supporting smoother integration, reduced reporting burdens and enhanced efficiency (IFC (2025c)). Efforts to ensure interoperability across standards, such as SDMX and the eXtensible Business Reporting Language (XBRL), can further support the handling of diverse data types and accommodate users' needs. Interoperability also facilitates decentralised data architectures, allowing tailored processes to business requirements while maintaining compatibility across systems. Moreover, standards act as AI enablers by improving data discovery, ensuring consistent understanding and enhancing AI accuracy through robust metadata and documentation, thereby mitigating data quality issues (Anvar (2025)).

Third, leveraging large, diverse and complex data sets calls for **modern, metadata-driven and accessible data platforms**. Single access points such as the European Single Access Point and collaborative frameworks such as the Common European Data Spaces can further enhance data access and reuse. Central banks have also developed specialised solutions, such as research data centres and remote execution systems, to securely facilitate academic research on granular data (IFC (2024)).

Finally, **strengthening collaboration** among peers and counterparts plays an important role in advancing data science. This can involve co-developing open-source solutions, such as the BIS OpenTech and sdmx.io initiatives. It can also entail fostering a structured exchange of experiences among stakeholders, as exemplified by the series of workshops on "Data science in central banking" co-organised by the IFC and Bank of Italy (IFC (2022, 2023, 2025a, forthcoming)). Collaboration can also be further strengthened by developing common international data frameworks (Tissot (2025)). In this context, the G20 Data Gaps Initiative represents a significant and promising step towards improving access to alternative data for official statistics and promoting the effective sharing of economic and financial information for the public good (IMF et al (2023)).

References

Anvar, E (2025): "Why SDMX matters? A community journey towards SDMX as an AI-enabler and a data mesh enabler", IFC Bulletin, no 64, May.

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

Araujo, D, N Bokan, F Comazzi and M Lenza (2024): "Word2Prices: embedding central bank communications for inflation prediction", CEPR Discussion Paper, no 19784, December.

Duijm, P, and I van Lelyveld (2025): "Experiences, essentials and perspectives for data science in the heart of central banks and supervisors", IFC Bulletin, no 64, May.

Gonçalves, A R, M Lourenço, D V Sousa and T Verheij (2025): "New strategy of data sharing and data access in statistics: the view from Banco de Portugal", IFC Bulletin, no 64, May.

International Monetary Fund (IMF), Inter-Agency Group on Economic and Financial Statistics and FSB Secretariat (2023): G20 Data Gaps Initiative 3 workplan – People, planet, economy, March.

Illes, A, and I Mattei (2025): "Building a database on cryptocurrencies", IFC Bulletin, no 64, May.

Irving Fisher Committee on Central Bank Statistics (IFC) (2020): "Computing platforms for big data analytics and artificial intelligence", IFC Report, no 11, April.

——— (2022): "Machine learning in central banking", IFC Bulletin, no 57, November.

——— (2023): "Data science in central banking: applications and tools", IFC Bulletin, no 59, October.

——— (2024): "Granular data: new horizons and challenges", IFC Bulletin, no 61, July.

——— (2025a): "Data science in central banking: enhancing the access to and sharing of data", IFC Bulletin, no 64, May.

——— (2025b): "Governance and implementation of AI in central banks", IFC Report, no 18, April.

——— (2025c): "SDMX adoption and use of open source tools", IFC Report, no 17, February.

——— (forthcoming): "Generative artificial intelligence in central banking", IFC Bulletin.

Križman, I and B Tissot (2022): "Data governance frameworks for official statistics and the integration of alternative sources", Statistical Journal of the IAOS, vol 38, no 3, August, pp 947–55.

Perrazzelli, A (2025): "Data science in central banking: enhancing the access to and sharing of data", IFC Bulletin, no 64, May.

Ricciato, F (2024): "Steps toward a shared infrastructure for multi-party secure private computing in official statistics", Journal of Official Statistics, vol 40, no 1, March.

Shah, B (2025): "Project Aurora: the power of data, technology and collaboration to combat money laundering across institutions and borders", IFC Bulletin, no 64, May.

Tissot, B (2025): "Data science in central banking: enhancing the access to and sharing of data", IFC Bulletin, no 64, May.

About the author(s)

Douglas Araujo joined the BIS in September 2018, where he is Adviser for data science in economics since May 2025. Before that, Douglas was an Economist at the BIS' International Data Hub. Previously, Douglas was in the Secretariat of the Basel Committee on Banking Supervision overseeing a range of policy and supervisory topics. Before that, he worked at the Central Bank of Brazil on macroprudential supervision (2011-15) and led the efforts to develop and implement the Brazilian proportionality framework for prudential regulation (2015-18). Douglas worked on financial stability monitoring as a Fellow at the BIS Financial Stability Institute in 2014. From 2015 to 2018, he also supported a number of countries in enhancing their macroprudential frameworks as a member of IMF missions. Until 2011, Douglas worked in the private sector in Brazilian financial markets. His current research focuses on using ML and other econometric techniques to forecast and to identify causal effects. Douglas also contributes to open source software at the intersection of machine learning and economics.

Adam Cap is Senior Macroeconomic Analyst in the Monetary and Economic Department at the BIS. He joined the BIS in July 2019 following his role in the International Policy Analysis division at the ECB. At the BIS, he has supported a variety of research projects related to the financial markets and monetary policy. Following an assignment at the BIS Representative Office for Asia and the Pacific, his focus shifted towards macroeconomic topics. He also has a strong interest in the application of data science and artificial intelligence, having previously supported two editions of the IFC Workshop on Data Science in Central Banking. Adam holds an MSc in Finance from Goethe University Frankfurt.

Ilaria Mattei is a Senior Financial Market Analyst in the Monetary and Economic Department at the BIS. Her expertise lies in financial markets statistics, central bank digital currencies and stablecoins. Prior to joining the BIS, she worked at the European Commission's Directorate General for Economic and Financial Affairs (DG ECFIN) and previously in the Analytical Credit and Master Data Division at the European Central Bank. She holds an MSc in Economic and Social Sciences from Bocconi University.

Rafael Schmidt is the Head of IT in the Monetary and Economic Department at the BIS where he oversees data processing, analytics, and dissemination systems. He also serves as Chairman of the SDMX Secretariat, promoting the ISO SDMX data standard sponsored by numerous international organisations. Previously, he had managerial roles in the Risk Management Department at the BIS. He also held research positions at the University of Cologne and the London School of Economics and Political Science. Rafael has graduated in Mathematics, Economics and Statistics. He holds a PhD and a habilitation in Financial Statistics and Econometrics.

Olivier Sirello is a Senior Statistical Analyst in the Monetary and Economic Department at the BIS. His areas of expertise are central bank statistics, balance of payments, and securities statistics. He is also involved in a number of international initiatives related to official statistics, including the High-Level Group for the Modernisation of Official Statistics of the United Nations Economic Commission for Europe. He previously worked as an economist-statistician at the Bank of France and held post-graduate teaching positions at Sciences Po Paris. He studied at Sciences Po Paris, Princeton University and Bocconi University. He holds two MScs in economics and public policy.

Bruno Tissot is the Head of Statistics and Research Support at the BIS and Head of the Secretariat of the IFC Committee. He is also the BIS Representative in the SDMX Sponsors' Committee and chairs the international Working Group on Securities Databases (WGSD). He has been working at the BIS since 2001, as Senior Economist and Secretary to the Markets Committee of Central Banks and then as the Adviser to the General Manager and Secretary to the BIS Executive Committee. Between 1994 and 2001 he worked for the French Ministry of Finance, having graduated from École Polytechnique (Paris) and the French Statistical Office, INSEE.

SUERF Policy Notes and Briefs disseminate SUERF Members' economic research, policy-oriented analyses, and views. They analyze relevant developments, address challenges and propose solutions to current monetary, financial and macroeconomic themes. The style is analytical yet non-technical, facilitating interaction and the exchange of ideas between researchers, policy makers and financial practitioners.

SUERF Policy Notes and Briefs are accessible to the public free of charge at https://www.suerf.org/publications/suerf-policy-notes-and-briefs/.

Editorial Board: Ernest Gnan, David T. Llewellyn, Donato Masciandaro, Natacha Valla

Designed by the Information Management and Services Division of the Oesterreichische Nationalbank (OeNB)

SUERF Secretariat
c/o OeNB, Otto-Wagner-Platz 3A-1090 Vienna, Austria
Phone: +43 1 40 420 7206
E-Mail: suerf@oenb.at
Website: https://www.suerf.org/